Response trajectories reveal conflict phase in image–word mismatch

Floris T. van Vugt · Patrick Cavanagh

Published online: 5 January 2012
© Psychonomic Society, Inc. 2012

Spatial prepositions (words such as *above*, *below*, and *inside*) have been studied extensively in sentence–picture verification studies (Carpenter & Just, 1975; Chase & Clark, 1971, 1972) in which the task was to report whether the spatial preposition accurately described a picture. All of these studies used reaction time measures. In the present study, we investigated response trajectories to gain insight into the time course for the processing of spatial prepositions that is not available in reaction time measures (Brenner & Smeets, 2004; Schmidt & Schmidt, 2009; Song & Nakayama, 2009). In several recent studies, researchers have analyzed the trajectory of a participant's response when selecting the answer with the touch of a finger (Boulenger et al., 2006; Finkbeiner, Song, Nakayama, & Caramazza, 2008; Schmidt & Seydell, 2008; Song & Nakayama, 2008a, 2008b), a saccade (Smit & Gisbergen, 1990) or a computer mouse (Freeman & Ambady, 2009; Spivey, Grosjean & Knoblich, 2005). The response choices are presented at different locations in space so that an in-flight deviation toward the competing, incorrect answer can be revealed.

We studied a processing conflict involving spatial prepositions in which a marker was placed above or below the

word ABOVE or BELOW, and in which the participant reported the location of the marker relative to the word, ignoring the meaning of the latter. A very similar task was originally studied by Paley and Olson (1975), who found no significant difference between the reaction times in the congruent and incongruent conditions. Logan and Zbrodoff (1979) also did not find a significant difference in their similar “spatial task.” Recently, the same task has been used in fMRI studies (Banich et al., 2000) and ERP studies (Stern & Mangels, 2006), revealing only marginal effects in reaction times. Therefore, in order to accentuate the conflict between spatial position and word meaning, we intermixed this *location* task with a second task—a *word* task—in which the participant had to respond to the word meaning ABOVE or BELOW, ignoring its position relative to the marker. The type of trial was indicated by nature of the marker: X for a location trial but O for a word trial (Fig. 1). This procedure, inspired by Harvey (1984), made word meaning relevant on some trials, increasing the probability of its processing even when it was to be ignored.

Using this interleaved task, we found clear evidence of a conflict period in which the incongruity between the word and its location either delayed the start of the trajectory to the correct answer or interrupted it. The response trajectory measure yields insights into the processing stages of decision making (e.g., see Resulaj, Kiani, Wolpert & Shadlen, 2009) and offers measures of processing times for location and word meaning. To examine the timing of the conflict, we varied the SOA of the word and the location marker (Glaser & Glaser, 1982); the marker also indicated the type of task (for a similar approach, see the speed–accuracy trade-off method of McElree & Griffith, 1995). Crucially, the word always appeared before (or simultaneously with) the marker that indicated the type of task.

Electronic supplementary material The online version of this article (doi:10.3758/s13414-011-0261-0) contains supplementary material, which is available to authorized users.

F. T. van Vugt · P. Cavanagh
Laboratoire Psychologie de la Perception, CNRS UMR 8158,
Université Paris Descartes,
Sorbonne Paris Cité,
Paris, France

F. T. van Vugt (✉)
IMMM HMTMH,
Emmichplatz 1,
Hannover, Germany
e-mail: f.t.vanvugt@gmail.com

Fig. 1 **a** Stimulus–response contingencies in the location trials in which the marker indicates the trial type: X reports the location of the marker relative to the word; O reports the meaning of the word. **b** Stimulus–response contingencies in the word trials. **c** the order of presentation

The particular analysis that we used for the trajectory (direction vector) was very sensitive to processing stages, more than curvature (e.g., Spivey et al., 2005) or other analyses (see the [Supplementary Materials](#)) and, as we will show, was more revealing than the final response time, the time at which the mouse-directed cursor reached one or the other of the two answer locations.

Method

Participants

Six right-handed healthy participants (three male, three female) with normal or corrected-to-normal vision participated in our study.

Stimuli

The participant was seated at approximately 50 cm from a 19-in. CRT monitor that presented the $42.3^\circ \times 32.3^\circ$ displays with a screen resolution of $1,024 \times 768$ at 100 Hz controlled by a Mac G4. The responses were directed with the right hand using a computer mouse whose position was sampled at 125 Hz, and this position trace was resampled by

linear interpolation to 40 Hz. The words ABOVE and BELOW were displayed in the center of the screen in white uppercase Verdana font subtending $6.7^\circ \times 2.6^\circ$. The answer words were displayed at $6.7^\circ \times 6.7^\circ$ from the screen corners and subtended $7.6^\circ \times 3.0^\circ$. The markers “X” and “O” appeared in the screen center, subtending $1.3^\circ \times 1.3^\circ$. The answer areas were all positions in these corners that were more than 23.2° from the screen center. Once a participant entered these areas, the corresponding answer was recorded.

Procedure

The task had two conditions: location trials and word trials. The shape of the marker indicated which response was required on each trial: An X indicated that participants were to report the location of the marker relative to the word and to ignore the meaning of the word, whereas an O indicated that they were to report the meaning of the word and ignore the location of the marker (see Fig. 1a, b).

The SOA between the word onset and the marker onset was varied from 0 ms to 200 ms. Consequently, the word appeared either before or simultaneously with the marker so that the word may have triggered significant processing before the marker indicated whether it would be task relevant or not.

To initiate each trial, the participant clicked with the mouse on a button in the bottom of the screen (see Fig. 1c). The word and marker then appeared, their onsets separated by 200, 150, 50, or 0 ms. Both the word and the marker (X or O) remained on screen until the end of the trial. The marker always appeared 200 ms after the participant’s click. The response corners were always the same throughout all sessions, but as a reminder, the two corner labels appeared in the top left and top right corners of the screen on each trial 300 ms after the marker. The participant responded by moving the mouse to the screen corner corresponding to the answer, and the trial ended as soon as the pointer entered the corner answer area. This arrival time will be referred to as the *movement finish time*.

Participants were required to start moving the mouse very quickly. If the mouse pointer had not left a circular area around the start button within 400 ms after initiating the trial, the response was discarded, a warning sounded, and the trial was repeated some later time during the experiment. The participants learned to initiate their responses quickly within a few blocks of training trials at the beginning of a session. Crucially, since they initiated their movement before they had made their decision, they started out moving straight up, approaching both answers without yet choosing either one of them. This initial, neutral upward motion was critical for capturing the moment at which the trajectory first veered off toward an answer corner.

Each of the four stimuli in the location and word conditions was presented for each of the four SOAs (200, 100, 50, and 0 ms). These 32 conditions were repeated 15 times to yield a total of 480 trials per block. Participants began with a training session that first introduced the two conditions (location and word responses) separately, followed by a mixed block, and then only in the end was the early movement requirement introduced. The two blocks together with the training lasted a little over an hour.

Results

Incorrect trials were removed from further analysis (7.7%) as were trials in which the participant reached the answer corner more than 4 *SDs* earlier or later than their average (0.9%).

We analyzed a number of properties of the response trajectories and, rather than measures of curvature used in several articles (e.g., Finkbeiner et al., 2008), we found the moment-to-moment direction of the trajectory to be the most sensitive measure (see the [Discussion](#) section and the [Supplementary Materials](#)). Our analysis therefore focuses on this measure, defined as the tangent to the path at each point in time, with 0° being the vertical tangent and positive values assigned to the direction toward the correct corner (see also Scherbaum, Dshemuchadse, Fischer, & Goschke, 2010).

We investigated the evolution of the movement direction over time for each participant separately using their average movement trace in each of the eight experimental conditions: 2 subtasks × 4 SOAs. As an example, we will present our analysis here for one participant in two such cases for response traces in location trials (Fig. 2).

To describe our analysis, we use the average direction trace from one participant for whom the SOA is 200 ms (Fig. 2, left graphs). The congruent direction curve starts with a consistent direction of 0°, reflecting the participant's initial motion straight upward prior to any deviation toward an answer corner. After around 200 ms, the path starts to arc toward the correct answer, stabilizing at a heading of around 60° until reaching the correct answer. The curve has this shape for all participants in all congruent conditions. We therefore fit a straight line to the upward trend for each participant (see the [Supplementary Materials](#) for details of the fitting procedure). We label the intersection of the linear fit with the baseline the *decision moment* (blue arrow in Fig. 2). At that point, the participant has gathered sufficient information to move toward the answer corner.

The trajectory in the incongruent case is similar, showing an initial launch toward the correct answer corner: There were no instances of an initial motion towards the wrong

corner followed by a correction toward the correct corner. However, the incongruent trace, as here, often shows an interruption. Most likely, once the word's meaning is processed, the conflict between its meaning and the location response interferes with the answer in progress. To capture this interference, we fit a broken line to the incongruent curve (as shown by the orange curves in Fig. 2). Its initial take-off point is set to the same value as in the congruent case, but the curve can be interrupted by a horizontal plateau at any time before resuming its path to the correct answer corner. This gives us a double step clearly seen in the bottom right-hand panel of Fig. 2. The plateau at which the trajectory pauses defines two time points: a conflict onset and an offset (the green area in Fig. 2). Having performed this analysis for all participants and conditions, we find three time points for each participant in each SOA and each task type (*word* or *location*). These three are: (a) the decision moment (common to congruent and incongruent trials), and for the incongruent trials, (b) the conflict onset, and (c) the conflict offset. This analysis was robust enough to reveal a conflict in all incongruent conditions for all participants except two conditions (out of the eight) for one participant (out of six).

In the location task, a repeated measures ANOVA revealed that the conflict duration in incongruent conditions (conflict offset – conflict onset) was significantly greater than zero, $F(1, 5) = 48.31$, $p < .001$, and did not vary significantly with SOA, $F(1, 5) = 2.66$, $p = .16$. In the word task, the conflict duration was also significantly greater than zero for all SOAs, $F(1, 5) = 65.60$, $p < .001$, and did not interact with SOA.

On the basis of the distinctive double-step pattern in the incongruent traces, we were able to choose a simple model to estimate the time at which the word and location encodings were available (Fig. 3). Specifically, the initial response was always to the correct answer corner on incongruent trials, even though the response was later interrupted by the conflict. This suggests that the response began once the task marker was decoded in both congruent and incongruent trials and that on incongruent trials, the conflicting information would interrupt once it was available. The double step pattern also allows us to rule out two other models. If the response began as soon as either the word or relative location was available, without waiting to determine which task was required, there would necessarily be some frequency of initial responses to the wrong corner that would be corrected later on. This was never seen. Or, if the response began when both word and relative location were available, congruent trials could begin immediately without waiting to interpreting the task marker, but all incongruent trials would have to be delayed until the marker was interpreted. In this case, there would never be an initial motion toward the correct corner, starting at the same time as it would in a congruent

Fig. 2 Direction over time for one participant (MZ) for an SOA of 200 ms (left) and 0 ms (right) for location trials. The traces show the point-by-point mean direction of the 60 trials (minus error trials) of that participant in those conditions. Upper graphs: only the congruent/incongruent traces. Lower graphs: our trajectory analyses applied to

the same two direction curves (see the [Supplementary Materials](#)’ details of the fit). The *blue arrow* indicates the decision moment in the congruent case. The time window between the conflict onset and conflict offset is marked in green

condition that was later paused and then resumed. However, over 50% of the average incongruent traces showed exactly this double step.

In order to model the response trajectories, we assumed that the word and location signals arrive with fixed delay after their onsets, and that the response begins once the task is decoded and the relevant signal has arrived. We further assumed that a conflict emerges on incongruent trials once both signals are available. We used only three free parameters (decision moment, conflict onset, and offset) to fit results on both location and word trials (the two tasks were also fit separately, see below). The least-squares fit of our model (see the [Supplementary Materials](#)) to the time points from two tasks ($R^2 = .91$, $\chi_{\text{red}}^2 = 0.48$) gives us an estimate of the processing time of the position and of the word meaning as 251 ms and 325 ms, respectively, and a conflict

duration of 138 ms. We plot this fit in [Fig. 3](#) (green lines). To obtain an estimate of the reliability of this fit, we then fit the same model to the data points of each participant individually and found in the cross-participant averages very similar values: position and word processing time of 260 ± 15 ms and 321 ± 16 ms, respectively, and a conflict duration of 131 ± 15 ms.

We used only three free parameters and found a quite respectable fit. Clearly we could have allowed different values of these three parameters for the two tasks and different values at each SOA. We had no indication from the data that conflict duration, for example, should vary as a function of SOA, but perhaps it might be different for the two tasks. We therefore fit the model to the two tasks separately. Fitting the location task on its own increased the goodness of fit ($R^2 = .98$, $\chi_{\text{red}}^2 = 0.07$), but the

Fig. 3 Signal processing and conflict times for the location (left) and word task (right). The red and gray lines in the bottom indicate the onset of the word and position stimuli, where the task-relevant feature is red and the task-irrelevant feature is gray. The green triangles indicate the mean conflict onset and offset, determined as described previously. The black triangles indicate the decision moment in the congruent and incongruent conditions. Finally, we indicate the movement finish times. The error bars indicate the standard errors of the participant means

parameters showed little change: The word and position processing times were 331 ms and 224 ms respectively, and the conflict duration was 146 ms. A similar separate fit for the word task ($R^2 = .86$, $\chi_{\text{red}}^2 = 0.32$) also had little effect on the best-fitting parameters: word and position processing times of 327 ms and 274 ms respectively and a conflict duration of 118 ms. Fitting each task separately with individual participant data, we found no significant differences between the three parameters for the word and location tasks, nor between the separate and conjoint fits.

These independent fits are a test of the robustness of the simple model applied to these data. However, the weakest point in the fit is the assumption that the word meaning should be available at a fixed duration following its presentation. Specifically, in the word task, the decision moment must increase with SOA with a slope of 1. This part of the fit is less successful than elsewhere, and the deviations may be accounted for by some delay in the processing of the word meaning while waiting for the appearance and decoding of the task cue. We could add this extra parameter to our model, but we felt that there were not enough data points to support this more complex interaction and that the simple

model, despite this deviation for the meaning decision moment, was adequate for our present purposes.

Figure 3 also shows the response finish times for the different conditions, the moment at which the trajectory entered the correct answer corner. These show response times between 700 and 800 ms, typical of many reaction experiments for similar conflict tasks (see the Discussion section). There is a significant delay of 112 ms for the incongruent versus congruent trials in the location task, $F(1,5) = 22.71$, $p = .005$, and of 71 ms for the word task, $F(1, 5) = 21.76$, $p = .005$, but no interaction with word-position SOA that would reveal any details of the word and location processing.

Discussion

Using response trajectories in word and location judgment tasks, we find a remarkably distinctive and stable decision moment at approximately 250 ms when the participant has enough information to begin to respond. In incongruent trials, we also find clear evidence of a conflict that delays or interrupts the response and lasts about 130 ms. This very large congruency effect was not a simple delay but often appeared as a pause in the trajectory well after the initial, correct response had already begun. The timing of these trajectory events also allowed us to derive the processing delay for the word and location signals. As compared with the response finishing time, our response trajectory measures of the conflict show a larger effect and a clearer interaction with the onset delay. Our simple model of the conflict does not capture all of the data with equal accuracy, but it does show a significant measure of success, providing far more information than the final reaction times. Two aspects of our response trajectory measure are critical in this success. The first is that the participant is moving the mouse during the entire trial, beginning before the critical stimuli appear. Therefore, when the word and position are displayed, the participant's trajectory is already underway. As such, we are able to measure the effects of the conflict on the trajectory as it happens instead of deducing that there must have been one from a delayed reaction time registered much later. The second critical aspect is the direction measure that we have used for the trajectory rather than the more typical curvature measures. We found that this measure reveals discrete changes in response trajectory that the curvature measure cannot localize as well or at all (see the Supplementary Materials).

One could argue that we slowed the participant's response time down by making him or her cross the entire screen with the mouse pointer. But previous studies using similar spatial Stroop tasks and various other response modalities have found response times in the same range as the

movement finish times in our experiment. For instance, researchers in several studies (Banich et al. 2000; Seymour, 1973; Walley, McLeod, & Weiden, 1994) had their participants say their responses out loud and found smaller congruency effects of between 15 and 45 ms (with response times in the range of 600 to 900 ms). However, exact comparisons are difficult since some studies used more than two spatial words. Palef and Olson (1975), who had participants respond by pressing a button to only what we called the location task, found earlier reaction times of around 360 ms (or converging to that value across practice). However they did not find a conflict effect at all. This may be because they presented the two tasks (word and location) in separate blocks, thus allowing the participants to switch strategies between blocks.

Conclusion

In the present study, we provided evidence of a reliable Stroop-like effect with spatial prepositions in participant's movement trajectories. Instead of the reaction time (i.e., the moment the participant registers his or her response) we investigated the response tendencies by analyzing the movement direction of the trajectory.

By proposing that the position and meaning information are processed in parallel and that the conflict occurs when both become available, we can deduce that the meaning of spatial prepositions “above” and “below” are processed in approximately 325 ms. Relative position is processed in a much shorter time of approximately 250 ms. The conflict they give rise to lasts for some 138 ms.

Author Note This work was supported by a Chaire d'Excellence grant from the ANR (to P.C.) and an EDF scholarship (to F.T.V.). Correspondence concerning this article should be addressed to Floris van Vugt, IMMM HMTMH, Emmichplatz 1, Hannover, Germany (e-mail: f.t.vanvugt@gmail.com).

References

Banich, M. T., Milham, M. P., Atchley, R., Cohen, N. J., Webb, A., Wszalek, T., et al. (2000). fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *Journal of Cognitive Neuroscience*, 12, 988–1000.

Boulenger, V., Roy, A. C., Paulignan, Y., Deprez, V., Jeannerod, M., & Nazir, T. A. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *Journal of Cognitive Neuroscience*, 18, 1607–1615.

Brenner, E., & Smeets, J. B. J. (2004). Colour vision can contribute to fast corrections of arm movements. *Experimental Brain Research*, 158, 302–307.

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 82, 45–73.

Chase, W., & Clark, H. (1971). Semantics in the perception of verticality. *British Journal of Psychology*, 62, 211–216.

Chase, W., & Clark, H. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, 3, 472–517.

Finkbeiner, M., Song, J.-H., Nakayama, K., & Caramazza, A. (2008). Engaging the motor system with masked orthographic primes: A kinematic analysis. *Visual Cognition*, 16, 11–22.

Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science*, 20, 1183–1188.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 875–894.

Harvey, N. (1984). The Stroop effect: Failure to focus attention or failure to maintain focusing? *The Quarterly Journal of Experimental Psychology. A*, 36, 89–115.

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, 7, 166–174.

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 134–157.

Palef, S. R., & Olson, D. R. (1975). Spatial and verbal rivalry in a stroop-like task. *Canadian Journal of Psychology*, 29, 201–209.

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*, 461, 263–266.

Scherbaum, S., Dshemuchadse, M., Fischer, R., & Goschke, T. (2010). How decisions evolve: The temporal dynamics of action selection. *Cognition*, 115, 407–416.

Schmidt, T., & Schmidt, F. (2009). Processing of natural images is feedforward: A simple behavioral test. *Attention, Perception, & Psychophysics*, 71, 594–606.

Schmidt, T., & Seydell, A. (2008). Visual attention amplifies response priming of pointing movements to color targets. *Perception & Psychophysics*, 70, 443–455.

Seymour, P. H. (1973). Stroop interference in naming and verifying spatial locations. *Perception & Psychophysics*, 14, 95–100.

Smit, A. C., & Gisbergen, J. A. V. (1990). An analysis of curvature in fast and slow human saccades. *Experimental Brain Research*, 81, 335–345.

Song, J.-H., & Nakayama, K. (2008a). Numeric comparison in a visually-guided manual reaching task. *Cognition*, 106, 994–1003.

Song, J.-H., & Nakayama, K. (2008b). Target selection in visual search as revealed by movement trajectories. *Vision Research*, 48, 853–861.

Song, J.-H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks. *Trends in Cognitive Sciences*, 13, 360–366.

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, 102, 10393–10398.

Stern, E. R., & Mangels, J. A. (2006). An electrophysiological investigation of preparatory attentional control in a spatial Stroop task. *Journal of Cognitive Neuroscience*, 18, 1004–1017.

Walley, R. E., McLeod, B. E., & Weiden, T. D. (1994). Increased attention to the irrelevant dimension increases interference in a spatial Stroop task. *Canadian Journal of Experimental Psychology*, 48, 467–494.