

RUNNING HEAD: MOUSE TRACKING IAT

Beyond Reaction Times:  
Incorporating Mouse Tracking Measures into Implicit Association Test to  
Examine Its Underlying Process

Zonghuo Yu, Fei Wang, Dengfeng Wang

Peking University

Mike Bastin

University of Nottingham

(This paper has been accepted as a Research Article for publication in Social Cognition.)

Author notes:

\* The first two authors contributed equally to this work.

\* The E-mail addresses for future contact: wdfeng@pku.edu.cn ,or

yuzonghuo@pku.edu..

\* We express our gratitude to Jon Freeman for his help on Mouse Tracker software and valuable comments on an earlier version of the manuscript. We also express our gratitude to Dr. Jeff Sherman, Dr. Karl Christoph Klauer and another anonymous reviewer for their valuable review comments on this paper.

### Abstract

Although the Implicit Association Test (IAT) has been widely used as an implicit measure over the past decade, research into its underlying mechanism remains woefully insufficient, partly due to the limitation of reaction time measures it uses. In two experiments, we modified the procedures of flower-insect IAT and two implicit self-esteem IATs by instructing participants to respond with a computer mouse instead of pressing keys. Analysis of motor trajectories showed that, although participants chose the correct response button in most of the trials, their mouse movement was continuously attracted towards the alternative response button, suggesting that both response representations are partially and simultaneously activated during the process. Furthermore, analysis of velocity profiles indicated that mouse movements towards the correct response button were slower in incompatible trials than in compatible trials, especially for attribute stimuli. Theoretical and methodological implications of these results are discussed.

Keywords: Implicit Association Test (IAT), continuous and dynamic real-time cognition, mouse tracking

## **Beyond Reaction Times: Incorporating Mouse Tracking Measures into Implicit Association Test to Examine Its Underlying Process**

### Introduction

Implicit Association Test (IAT; Greenwald et al., 1998) is one of the most influential and popular research paradigms in the increasingly popular area of implicit social cognition during the past decade (De Houwer et al., 2009). However, our understanding about its underlying process remains very limited. In our studies, we aimed to investigate the possibility of incorporating mouse-tracking measures into standard IAT procedure and use it to examine further the nature of the underlying process of IAT.

### A BRIEF INTRODCUTION TO IAT

In a typical IAT task there are four categories of stimuli, two of which are *concept stimuli* (e.g. names of flower/insect) while the other two are *attribute stimuli* (e.g. negative/ positive words). Participants are instructed to categorize the stimuli by pressing one of the two response keys and the assignment of the keys is varied across different experiment blocks. For example, in some trials participants are instructed to press the left key for flower names and positive words, and press the right key for insect names and negative words. These trials are called *compatible trials*, because valance of the concept stimuli (*flower* which is positive) are similar to valance of the attribute stimuli (positive words) which are mapped onto the same response key.

However, in some other trials the assignment of response keys is different, with insect names and positive words assigned to the left key and flower names and negative words assigned to the right key. These trials are called *incompatible trials*, since valances of concept stimuli and attribute stimuli which share the same response key are opposite.

The basic assumption of IAT is that participant's performance in the categorization task is a function of association strength between the two categories that are assigned to the same response key. Therefore, if there is a significant *compatibility effect* (i.e. reaction time is longer in incompatible trials than in compatible trials), we may conclude that the flower-positive/insect-negative associations are stronger than the insect-positive/flower-negative associations. If we further assume that the association strength between concept and attribute stimuli is a reflection of one's attitude, then we can treat the magnitude of compatibility effect as an indicator of participants' implicit attitude towards flower and insect.

Following this logic, IAT has been used as an implicit measure in many research areas over the past years, such as racial attitude (Greenwald et al., 1998), implicit self-esteem (Greenwald & Farham, 2000), gender stereotype (Rudman et al., 2001), and implicit personality (Asendorpf et al., 2002).

#### THEORETICAL ACCOUNTS OF THE UNDERLYING PROCESS OF IAT

Despite IAT's popularity, however, there have been heavy debates on its validity and how to interpret its results, even since its inception (for reviews, see: Nosek et al.,

2007; De Houwer et al., 2009). One question related to this issue, but one which has received less attention, is the mechanism of IAT. What is the underlying process that causes the compatibility effect? What happens during the period after the onset of stimuli and before a participant makes a response? To date, researchers have proposed several theoretical accounts. For example, Brendl et al.'s (2001) random walk model account proposes that every stimulus in IAT generates information towards the response representations, and when accumulated information exceeds a threshold the participant will make an overt behavioral response. A concept stimulus generates two kinds of information: *identity* (e.g. flower or insect) and *valance* (e.g. positive or negative); in compatible trials these two sources provide information in the same direction and move the information accumulation process toward the correct response representation, while in incompatible trials these two provide information in opposite directions, resulting in slower activation of representation and delayed response. Attribute stimuli, on the other hand, can only generate valance information. For example, the word *happiness* generates similar information towards the positive response representation, no matter whether in compatible or incompatible trials. Therefore, the attribute stimuli information accumulation rate should be the same in compatible and incompatible trials.

DeHouwer's (2001; 2003) relevant features account also interprets the compatibility effect based on the response activation process. It proposes that the extent to which a response representation would be activated automatically depends on the degree of similarity between the stimulus and the response, and in IAT such

stimulus-response compatibility varies across different trials. In compatible trials, concept and attribute stimuli of same valance are assigned to the same response button, therefore the valance of the response button is consistent with both kinds of stimuli, facilitating the response process. In incompatible trials, however, the concept and attribute stimuli of opposite valances are assigned to the same response button, therefore the valance of response button is less similar to either kinds of stimuli, resulting in lower stimulus-response compatibility and slower representation activation.

Both of the random walk model account and the relevant feature account attribute the compatibility effect to the process of response representation activation. Besides these two, there are also other accounts which attribute the IAT effect to other processes (for a review, see: De Houwer et al., 2009), such as the task-switching account (Klauer & Mierke, 2005), the salience asymmetry account (Rothermund & Wentura, 2001), the diffusion model (Klauer et al., 2007), and the quad model (Conrey et al., 2005). Despite of this abundance of theoretical accounts, as De Houwer et al. (2009) pointed out, our understanding about how the IAT effect comes about is still very limited and there is a clear need for more research into this effect. In current studies, we aim to contribute to this literature by introducing a new measure of IAT into our research toolbox, which may in turn generate useful data on the underlying IAT process.

## FRAME WORK OF HUMAN COGNITION

Indexes generated by the standard IAT procedure are reaction times and error rates, both of which are *end products* of the underlying process of IAT, not *online measurements* of it. They can be used as indicators of the participant's overall performance in IAT, but what really happens during an IAT trial is beyond their scope.

To supplement these existing measurements, we want to adopt a new measure which is designed to track mental processes on a real-time basis, namely the mouse tracking paradigm (Spivey et al., 2005; Dale et al., 2007; Farmer et al., 2007; Freeman et al., 2008; Freeman & Ambady, 2009; Wojnowicz et al., 2009; Freeman et al., 2010; Freeman & Ambady, 2010; Freeman & Ambady, 2011), into the standard IAT procedure. In mouse tracking studies, participants are instructed to make responses using a computer mouse, and the trajectories of mouse movement are recorded and analyzed. Because these motor trajectories are captured during the entire response process, analyzing them may reveal real-time course characteristics of this process (Freeman & Ambady, 2010).

As an emerging paradigm, mouse tracking has been used to demonstrate the underlying process of a variety of phenomena in recent years. For example, in a pioneering study, Spivey et al. (2005) showed that when asked to choose between different objects according to a spoken word (e.g. candy), participants' motor trajectories were curved towards a similarly sounded distractor (e.g. candle), even if the responses were correct. Freeman et al. (2008) demonstrated that mouse

movements of sex judgments for atypical faces (e.g. male with long hair) were spatially attracted towards the opposite sex (e.g. female). Furthermore, Wojnowicz et al. (2009) found that when participants were asked to report their explicit attitude (like vs. dislike) towards Black and White people, the motor trajectories of the “like” response for Black people were continuously attracted by the “dislike” option.

Underlying these studies is a theoretical framework of continuous and dynamic real-time cognition (Spivey & Dale, 2004, 2006; Spivey, 2007; Freeman & Ambady, 2011). Traditional models in cognitive psychology often regard human cognition as discrete-stage based processes. On the other hand, the continuous and dynamic framework of cognition proposes that real-time cognitive process is best described as a continuous and dynamic changing course of neuronal activities. The cognitive system spends most of its time in intermediate and ambiguous states before reaching a stable and explicit response; during this process, multiple representations may be partially activated. Online measures such as mouse tracking, as demonstrated in the examples described above, may be used to explore such a mixture of mental states. The continuous spatial attraction effects yielded in mouse tracking studies were considered to demonstrate the continuous and dynamic nature of such a process.

#### INCORPORATING MOUSE TRACKING MEASURES INTO IAT PROCEDURES TO EXAMINE ITS UNDERLYING PROCESS

We believe that the response representation activation process proposed by several theoretical accounts of IAT is in line with Spivey et al.’s continuous and

dynamic framework. At the announcement of a stimulus word, information generation and accumulation begins, while levels of activation of response representations change over the time period. Both response representations are partially activated and the system is in the intermediate area for the majority of the time, before finally making a behavioral response (Brendl et al., 2001; Klauer, et al., 2007). The traditional measures of reaction time are limited in demonstrating this process directly, because they do not provide real-time measurement. Mouse tracking measures, on the other hand, may help us to track such a process due to its online nature.

The aim of the present research is to explore the possibility of incorporating mouse tracking measures into the existing IAT procedure and demonstrate its potential in exploring the underlying process of IAT. The only change in the mouse tracking version of IAT is that participants choose the response buttons with a computer mouse rather than pressing keys. More specifically, our goals are threefold:

Firstly, we want to make sure that the compatibility effect in reaction time can be replicated in the new version. Although the participant's method of response is changed, the stimuli, structure and characteristics of the task remain intact. Therefore, we can expect to witness the classical IAT effect in our experiments.

Secondly, we want to use mouse tracking data to examine the continuous and dynamic nature of the underlying process. Although continuity of process is a common presupposition in many theoretical accounts of IAT, either explicitly or implicitly, on the empirical level few studies have demonstrated such process in a real-time manner. Following the same logic of previous mouse tracking studies, if the

underlying process of IAT is indeed continuous and dynamic and the two response representations are indeed simultaneously activated, then we can expect that participants' motor trajectories would manifest a continuous spatial attraction effect towards the incorrect response button, especially in incompatible trials. What's more, as shown in Freeman et al. (2009, Study 3), discrete-staged vs. continuous process can be differentiated by analyzing the distributions of curvature of mouse trajectories. In this research, we will also conduct such analysis to provide further evidence of the continuous and dynamic view.

Thirdly, to demonstrate further the usefulness of adding mouse tracking measures into IAT, we will explore the velocity profiles of mouse movements. In previous mouse tracking studies, mouse velocity has been used as an indicator of response activation speed (Wojnowicz et al., 2009); faster movement towards the final response button was considered to indicate higher net information accumulation rate towards this response representation. Based on the same presupposition, in present research we will examine mouse's velocity over the whole course of its movement to reveal the online nature of the underlying process of IAT effect, which is largely beyond the scope of traditional measures of reaction time and error rate.

## Experiment 1

In Experiment 1, we began with a mouse tracking conversion of the classical flower-insect IAT (Greenwald et al., 1998, Study 1). As noted by Greenwald et al. (1998), most people hold positive attitudes towards flowers and negative attitudes

towards insects, which represent an excellent starting point. If implantation of mouse tracking measures into IAT were to be successful, the IAT effect in reaction times would be replicated in this modified version, while analyses on mouse trajectories and velocity profiles would provide further insight into the underlying process.

## METHOD

*Participants* A total of 38 Peking University students (17 males, 21 females) participated and received ¥10 (approximately \$1.5) for their participation.

*Stimuli* 40 Chinese words adapted from previous studies (Greenwald et al., 1998; Cai, 2003) were used as stimuli. Concept stimuli consisted of 10 flower names, e.g. *rose* and 10 insect names, e.g. *ant*. Attribute stimuli consisted of 10 positive nouns, e.g. *peace*, and 10 negative nouns, e.g. *death*<sup>1</sup>.

*Procedure* The procedure was almost identical to the standard IAT procedure (Nosek, et al., 2007), except that the response choice was made by mouse movement and click rather than simply pressing different keys. There were 7 blocks and 20-40 trials in each block (see Table 1). The blocks 4 and 7 were the critical blocks. For half of the participants block 4 is the compatible block (flower+positive/insect+negative) and

<sup>1</sup> We asked 25 students from Peking University to rate the valence of all of the stimuli used in the two experiments, using a 9-point Likert-like scale from “totally negative” to “totally positive”. Valence strength for each stimulus word was computed as participant’s rating score’s deviation from midpoint 5. Across all of the 3 IATs in the two experiments, attribute stimuli had significantly greater valence strength than concept stimuli (for flower-insect IAT in Experiment 1,  $t(24)=9.24, p<.001, d=1.85$ ; for the affective IAT in Experiment 2,  $t(24)=7.94, p<.001, d=1.58$ ; for the affective IAT in Experiment 2,  $t(24)=17.66, p<.001, d=3.53$ ). However, we believe that such inequality in valence strength across stimuli type is not a severe problem in current study, as our major focus is on how compatibility of IAT task would affect participants’ performance for each of the stimuli types, not the performance differences across them.

block 7 is the incompatible block (insect+positive/flower+negative), while for the other half the reverse was true.

Figure 1 demonstrates the course of a typical trial. At the beginning of each trial, there was a small box labeled “start” at the bottom of the screen. After participants clicked the box, a stimulus word appeared, and participants moved the mouse cursor to one of the response buttons in the top-left or top-right corner of the screen, and clicked it to end the trial. Inter-trial interval was 750 milliseconds (ms). The instructions for the classification task were displayed at the beginning of each block.

To capture the online features of the IAT process, it is important to make sure that mouse movement starts before an explicit response is already generated. Therefore, we instructed participants to begin to move the mouse right after the stimulus word appears, even if they were not sure which response button to choose at that time. Furthermore, if the mouse didn't move for longer than 500ms, a warning would appear on the screen at the end of the trial.

We presented the stimuli and recorded the mouse trajectories using the Mouse Tracker software package, developed by Freeman and Ambady (2010). In this software, the streaming x, y coordinates of the computer mouse was recorded at a sample rate of approximately 70 hertz.

## RESULTS AND DISCUSSIONS

*Data screening and preparation* Our analyses focused on the motor trajectories in blocks 4 and 7. For each subject, there were totally 80 trials in these two critical

blocks, half of which were compatible trials while the others were incompatible trials. Trials with incorrect responses, or with too long a reaction time ( $>4000\text{ms}$ ) or initial time ( $>500\text{ms}$ ), were discarded. For the 38 subjects, the discard rates were 1.9% for compatible trials and 3.4% for incompatible trials.

Before submission to data analysis, all trajectories were remapped rightward if needed, and rescaled into a standard coordinate space with  $[-1, 1.5]$  at the top left and  $[1, 0]$  at the bottom right, leaving  $[0, 0]$  at the start point of movement. Because the trajectories varied in duration, they were normalized into 101 time steps using linear interpolation. These data preparations were also done in Mouse Tracker.

Every subject's trajectories were collapsed into two averaged trajectories, one for compatible trials and one for incompatible trials. The following analyses, except the bimodality analysis, were based on averaged trajectories.

*Reaction time analysis* To begin with, we must insure that the classical IAT effect in reaction time was replicated in our modified paradigm. This was indeed the case, as a  $2$  (trial type: compatible / incompatible)  $\times$   $2$  (stimulus type: attribute / concept) within-subject ANOVA yielded a significant main effect for trial type,  $F(1,37)=11.90$ ,  $p=.001$ , partial  $\eta^2 = .24$ . Therefore, the reaction times of incompatible trials ( $M=1139\text{ms}$ ,  $SD=169\text{ms}$ ) were significantly longer than those of compatible trials ( $M=1067\text{ms}$ ,  $SD=150\text{ms}$ ), resulting in an IAT effect of approximately 72ms and an averaged D score (Greenwald, Nosek & Banji, 2003) of 0.26. Neither the main effect of stimulus type ( $F(1,37)=0.04$ ,  $p=.84$ , partial  $\eta^2 < .01$ ) nor the interaction effect of trial type and stimulus ( $F(1,37)=0.19$ ,  $p=.66$ , partial  $\eta^2 = .01$ ) was significant,

suggesting that the compatibility effect in reaction time was identical for attribute or concept stimulus.

*Shape analysis* As shown in Figure 1, Panel A, mouse trajectories in incompatible trials were attracted more towards the incorrect response button than those in compatible trials. To verify this initial observation, we examined the trajectories' curvature by computing their maximum deviations (MD), defined as the trajectory's maximum perpendicular deviation from a straight line connecting the trajectory's starting point and endpoint. Larger MD would indicate more attraction towards the alternative. As expected, a 2 (trial type: compatible / incompatible)  $\times$  2 (stimulus type: attribute / concept) within-subject ANOVA yielded a significant main effect for trial type,  $F(1,37)=6.08$ ,  $p=.02$ , partial  $\eta^2 = .14$ . Therefore MDs were significantly larger in incompatible trials than in compatible trials, indicating a greater spatial attraction effect in incompatible trials. Neither the main effect of stimulus type ( $F(1,37)=0.65$ ,  $p=.43$ , partial  $\eta^2 = .02$ ) nor the interaction effect of trial type and stimulus ( $F(1,37)=1.00$ ,  $p=.32$ , partial  $\eta^2 =.03$ ) was significant.

Correlational analyses indicated that MD were not significantly correlated with reaction time in both compatible ( $r(37)=-.02$ ,  $p=.90$ ) and incompatible trials ( $r(37)=.09$ ,  $p=.60$ ). However, the compatibility effect on MD was positively correlated with the compatibility effect on reaction time,  $r(37) = .56$ ,  $p<.001$ , and with D score,  $r(37) = .67$ ,  $p<.001$  Therefore, the slower a participant performed in incompatible trials than in compatible trials, the more his or her trajectories were attracted towards the incorrect response button.

*Distributional Analysis* However, there are also alternative accounts for this spatial attraction effect. One possibility is that in some trials the mouse moved directly towards the correct response button, while in the other trials the mouse firstly moved directly towards the incorrect response button and then made a sharp midflight correction. Averaging these two kinds of trajectories would also lead to an overall curved trajectory. Freeman et al. (2009, Study 3) suggested that one way to eliminate such a possibility is to examine the bimodality of the curvature distribution. The logic behind such examinations is that sharp midflight-correction trajectories would have very large MD, while nearly straight trajectories would have very small MD. If the trajectories indeed consisted of these two types, then we could expect that the distribution of MD should be bimodal; otherwise, the distribution should be normal.

We computed the  $b$  coefficient (SAS Institute, 1989) of the  $Z$ -score-transformed MD distributions for each kind of trial using trial-level data. The  $b$  coefficient was 0.401 for compatible trials, and 0.435 for incompatible trials, both of which were smaller than the standard cut-off point 0.555, suggesting that bimodality was not the case for current data. Furthermore, the Kolmogorov-Smirnov test suggested that the shapes of the distributions are statistically indistinguishable for compatible and incompatible trials,  $D = .04$ ,  $p = .294$ . These distributional analyses challenge the bimodality explanation of the spatial attraction effect and further supported the continuous and dynamic underlying process of IAT.

*Velocity analysis* To examine the information accumulation rates in different trials, we constructed velocity profiles based on normalized mouse trajectories. We

break the whole movement into 5 time bins and computed horizontal velocity in each time bin as incremental X-coordinate between the current and previous time bins.

Positive velocity indicated movement towards the correct response button, while negative velocity indicated movement towards the incorrect response button. For each participant, velocity profiles were created separately for incompatible and compatible trials as well as for attribute stimuli and concept stimuli (Figure 1, Panel B).

We conducted a 2 (trial type: compatible / incompatible)  $\times$  2 (stimulus type: attribute / concept)  $\times$  5 (time) within-subject ANOVA on these velocity profiles. The factor of time yielded the only significant main effect,  $F(4,148) = 201.76$ ,  $p < .001$ , partial  $\eta^2 = .85$ ; post-hoc analysis suggested that velocities differed in each of the 5 time bins and peaked in time bin 3. However, the main effect of trial type was not significant,  $F(1,37) = 0.72$ ,  $p = .79$ , partial  $\eta^2 = .002$ . The main effect of time was qualified by a significant two-way interaction effect between trial type and time,  $F(4, 148) = 16.12$ ,  $p < .001$ , partial  $\eta^2 = .30$ . Post hoc analysis suggested that the compatibility effect on velocities reached significance in time bin 3,  $F(1, 37) = 18.31$ ,  $p < .001$ , partial  $\eta^2 = .33$ , suggesting that the peak velocities were lower in incompatible trials. However, there was also a reversed compatibility effect in time bin 4,  $F(1, 37) = 20.46$ ,  $p < .001$ , partial  $\eta^2 = .36$ , indicating that velocities were greater in incompatible trials in the latter stage of mouse movement. This is understandable, as in compatible trials, compared with incompatible trials, the mouse would move nearer to the destination at this stage and therefore already begin to slow down.

Moreover, the factor of stimulus type didn't yield any significant main effect

( $F(1,37) = 0.003, p = .96, \text{partial } \eta^2 < .001$ ) or interaction effect with other factors (with trial type:  $F(1,37) = 0.88, p = .36, \text{partial } \eta^2 = .02$ ; with time:  $F(4,148) = 0.89, p = .47, \text{partial } \eta^2 = .02$ ; with trial type and time:  $F(4,148) = 0.85, p = .50, \text{partial } \eta^2 = .02$ ). Therefore, between the two stimulus types there was no significant difference on the compatibility effect on velocity profiles.

To further examine the relationship between velocity profiles and the IAT effect, we computed the correlation between compatibility effects on velocity and participant's D score, in time bin 3 and time bin 4. For both type of stimuli, they were significantly positively correlated in time bin 3 (for attribute stimuli,  $r(37) = .49, p = .002$ ; for concept stimuli,  $r(37) = .35, p = .03$ ), and significantly negatively correlated in time bin 4 (for attribute stimuli,  $r(37) = -.36, p = .03$ ; for concept stimuli,  $r(37) = -.43, p = .01$ ).

In summary, velocities of mouse movements didn't differ for attribute and concept stimuli, and the peak velocities in time bin 3 were lower in incompatible trials than compatible trials for both types of stimuli. If we accept the presupposition that mouse velocity indicates information accumulation rate, then these results suggest that response representation was activated slower in incompatible trials for both concept and attribute stimuli. We will discuss implications of these results on existing theoretical accounts of IAT in the section of General Discussion.

## Experiment 2

In Experiment 2, we tried to replicate the results of Experiment 1 in another two

widely used IAT tasks — the *affective* and *evaluative* implicit self-esteem IAT (Greenwald & Farham, 2000). Previous studies have shown that people tend to associate self with positive words more than with negative valance words, and these IAT effects were robust in different cultural contexts as well (Cai, 2003). Depending on the attribute words used (pleasant/unpleasant meaning words or desirable/undesirable trait words), the implicit self-esteem IAT could be affective or evaluative. We expected that the continuous spatial attraction effect and patterns in velocity profiles would be replicated in Experiment 2.

## METHOD

*Participants* 56 Peking University students (28 males, 28 females) participated and each received ¥15 (approximately \$2.2) for their participation.

*Stimuli* For each IAT, 40 Chinese words adapted from previous implicit self-esteem research (Cai, 2003) were used as stimuli. Concept stimuli were identical in both IATs, consisting of 10 words referring to the self (e.g. *mine*) and 10 words referring to others (e.g. *theirs*). Attribute stimuli in evaluative self-esteem IAT consisted of 10 positive trait words (e.g. *clever*) and 10 negative trait words (e.g. *ugly*). Attribute stimuli in affective self-esteem IAT were identical with Experiment 1, consisting of 10 positive words (e.g. *peace*) and 10 negative words (e.g. *death*).

*Procedures* Every participant performed mouse tracking versions of affective and

evaluative implicit self-esteem IAT consecutively, with a 5 minutes break in between<sup>2</sup>. The procedure of each IAT was identical to Experiment 1. The order of the two IATs and block order within each IAT were counter-balanced between participants. Mouse trajectories were recorded using the Mouse Tracker software package (Freeman & Ambady, 2010).

## RESULTS AND DISCUSSIONS

*Data screening and preparation* The data screening criterion was the same as Experiment 1. For the affective IAT, the discard rate was 0.7% for compatible trials and 1.7% for incompatible trials. For the evaluative IAT, the discard rate was 1.0% for compatible trials and 1.8% for incompatible trials. For each IAT, every subject's trajectories were remapped, rescaled, and collapsed into two average trajectories, one for compatible trials and one for incompatible trials.

*Reaction time analysis* We conducted 2 (trial type: compatible / incompatible) × 2 (stimulus type: attribute / concept) within-subject ANOVA on reaction time data for both IATs. For both IATs, the main effect of trial type was significant (affective IAT:  $F(1,55)=29.53, p<.001$ , partial  $\eta^2 = .35$ ; evaluative IAT:  $F(1,54)=29.95, p<.001$ , partial  $\eta^2 = .36$ ), suggesting that participants spent a longer time in incompatible trials than in compatible trials. However, neither the main effect of stimulus type (affective IAT:  $F(1,55)=0.002, p=.96$ , partial  $\eta^2 < .001$ ; evaluative IAT:  $F(1,54)=0.79, p=.38$ , partial  $\eta^2 = .01$ ) or the interaction effect between these two factors was significant

---

<sup>2</sup> One subject's data of evaluative IAT was lost due to computer malfunction, so the sample size of evaluative IAT was 55.

(affective IAT:  $F(1,55)=2.74$ ,  $p=.10$ , partial  $\eta^2 = .05$  ; evaluative IAT:  $F(1,54)=2.13$ ,  $p=.15$ , partial  $\eta^2 = .04$ ). Therefore, the classical IAT effect in reaction times was replicated in both mouse tracking IATs. For the affective IAT, the magnitude of IAT effect was 71ms, while for the evaluative IAT, the magnitude of IAT effect was 75ms. The averaged D scores for these two IATs were 0.29 and 0.31, respectively. The two IAT effects were marginally significantly correlated,  $r(54) = .25$ ,  $p = .062$ .

*Shape analysis* As shown in Figure 2 & 3, Panel A, there was a spatial attraction effect towards the incorrect response button for both IATs. We conducted 2 (trial type: compatible / incompatible)  $\times$  2 (stimulus type: attribute / concept) within-subject ANOVA on MD for both IATs. For both IATs, the main effect of trial type was significant (affective IAT:  $F(1,55)=7.88$ ,  $p=.01$ , partial  $\eta^2 = .13$ ; evaluative IAT:  $F(1,54)=7.90$ ,  $p=.01$ , partial  $\eta^2 = .13$ ), suggesting that MDs were larger in incompatible trials for both IATs. The main effect of stimulus type was not significant (affective IAT:  $F(1,55)=0.83$ ,  $p=.37$ , partial  $\eta^2 = .02$  ; evaluative IAT:  $F(1,54)=1.31$ ,  $p=.26$ , partial  $\eta^2 = .02$ ).

Surprisingly, the interaction effect between these two factors was significant for both IATs (affective IAT:  $F(1,55)=4.75$ ,  $p=.03$ , partial  $\eta^2 = .08$  ; evaluative IAT:  $F(1,54)=7.37$ ,  $p=.01$ , partial  $\eta^2 = .12$ ). Post hoc analysis on this interaction effect showed that the compatibility effect on MD was significant only for *attribute* stimuli (affective IAT:  $F(1,55)=14.35$ ,  $p<.001$ , partial  $\eta^2 = .21$  ; evaluative IAT:  $F(1,54)=11.56$ ,  $p=.001$ , partial  $\eta^2 = .18$ ), but non-significant for *concept* stimuli (affective IAT:  $F(1,55)=1.26$ ,  $p=.27$ , partial  $\eta^2 = .02$  ; evaluative IAT:  $F(1,54)=0.99$ ,

$p=.33$ , partial  $\eta^2 = .02$  ).

In both IATs, MD were not significantly correlated with reaction time in both compatible (affective IAT:  $r(55)=.04$ ,  $p=.79$ ; evaluative IAT:  $r(54) = .07$ ,  $p = .61$ ) and incompatible trials (affective IAT:  $r(55)=.18$ ,  $p=.19$ ; evaluative IAT:  $r(54) = -.003$ ,  $p = .98$ ). However, the compatibility effect on MD was positively correlated with the compatibility effect on reaction time ( $r(55) = .33$ ,  $p = .01$ ) and D score ( $r(55) = .39$ ,  $p=.003$ ) in the affective IAT. In the evaluative IAT, both of the correlations were marginally significant, with  $r(54) = .24$ ,  $p = .082$  for the compatibility effect on reaction time, and  $r(54) = .22$ ,  $p = .10$  for D score.

*Distributional Analysis* To rule out the discrete-stage account of this attraction effect, we computed the  $b$  coefficient of the Z-score-transformed MD distributions for the two kinds of trials. In the affective IAT, the  $b$  coefficient was 0.378 for compatible trials, and 0.416 for incompatible trials; in the evaluative IAT, the  $b$  coefficient was .420 for compatible trials, and 0.403 for incompatible trials. All of them were smaller than the standard cut-off point 0.555. The Kolmogorov-Smirnov (K-S) test indicated that the shapes of the two distributions did differ statistically for both IATs (affective:  $D = .07$ ,  $p<.001$ ; evaluative:  $D = .08$ ,  $p<.001$ ). However, the K-S test is a null hypothesis significance test (NHST), and the large sample size in these trial-based analyses (>2200 for each type of trial) would make NHST reach significance too easily. The magnitude of the  $D$  value we witnessed (.07-.08) is at a comparable level with those in previous mouse tracking studies (e.g. Freeman et al., 2008). Overall, from distributional analyses we might conclude that bimodality is not

a serious problem in our current data.

*Velocity analysis* Velocity profiles were constructed for the affective IAT and evaluative IAT as in Experiment 1. As shown in Figure 2 & 3, Panel B, in both IATs the time course pattern of velocity profiles in two IATs were similar to Experiment 1. We used  $2$  (trial type: compatible / incompatible)  $\times$   $2$  (stimulus type: attribute / concept)  $\times$   $5$  (time) within-subject ANOVA on two IATs separately. For both IATs, the main effect of time (affective IAT:  $F(4,220) = 274.23, p < .001$ , partial  $\eta^2 = .83$ ; evaluative IAT:  $F(4,216) = 291.19, p < .001$ , partial  $\eta^2 = 0.84$ ) and its two-way interaction effect with trial type (affective IAT:  $F(4, 220) = 9.95, p < .001$ , partial  $\eta^2 = .15$ ; evaluative IAT:  $F(4,216) = 10.33, p < .001$ , partial  $\eta^2 = .16$ ) reached significance. Post hoc analysis yielded patterns similar to Experiment 1. The compatibility effects on velocities reached significance in time bin 3, (affective IAT:  $F(1, 55) = 10.88, p = .002$ , partial  $\eta^2 = .17$ ; evaluative IAT:  $F(1, 55) = 17.04, p < .001$ , partial  $\eta^2 = .24$ ), suggesting that the peak velocities were lower in incompatible trials. However, there were also reversed compatibility effects in time bin 4, (affective IAT:  $F(1, 54) = 11.68, p = .001$ , partial  $\eta^2 = .18$ ; evaluative IAT:  $F(1, 54) = 16.54, p < .001$ , partial  $\eta^2 = .23$ ), indicating that velocities were greater in incompatible trials in the latter stage of mouse movement.

However, the three-way interaction effect between trial type, stimulus type and time also reached significance for both IATs (affective IAT:  $F(4,220) = 4.90, p = .003$ , partial  $\eta^2 = .08$ ; evaluative IAT:  $F(4,216) = 5.73, p < .001$ , partial  $\eta^2 = .10$ ). Further examination into these interaction effects revealed an interesting pattern: there were

significant or marginally significant two-way interaction effects between stimulus type and trial type in time bin 2 ( $F(1,54) = 5.31, p=.03$ , partial  $\eta^2=.09$ ) of evaluative IAT and in time bin 3 (affective IAT:  $F(1,55) = 5.13, p=.03$ , partial  $\eta^2=.09$ ; evaluative IAT:  $F(1,54) = 3.23, p=.08$ , partial  $\eta^2=.06$ ) and time bin 4 (affective IAT:  $F(1,55) = 14.99, p<.001$ , partial  $\eta^2=.21$ ; evaluative IAT:  $F(1,54) = 11.06, p=.002$ , partial  $\eta^2=.17$ ) of both IATs, such that in these time bins, the compatibility effects for concept stimuli were either *smaller* than those for attribute stimuli, or didn't reach significance at all.

As in Experiment 1, we also computed the correlation between compatibility effects on velocity and participant's D score, in time bin 3 and time bin 4. For attribute stimuli, they were significantly positively correlated in time bin 3 ( $r(55)=.39, p=.003$ ), and significantly negatively correlated in time bin 4 ( $r(55)=-.32, p=.02$ ) in affective IAT; in evaluative IAT, the correlations were in the same directions as in affective IAT, but failed to reach significance (time bin 3:  $r(54)=.20, p=.14$ ; time bin 4:  $r(54)=-.21, p=.13$ ). For concept stimuli, however, the correlations were not significant in both IATs in either time bin 3 (affective IAT:  $r(55)=.10, p=.47$ ; evaluative IAT:  $r(54)=.04, p=.79$ ) or time bin 4 (affective IAT:  $r(55)=.12, p=.39$ ; evaluative IAT:  $r(54)=-.06, p=.66$ ).

In summary, for attribute stimuli we observed patterns similar to experiment 1, which support the relevant feature account. However, for concept stimuli the difference between compatible and incompatible trials on velocity profiles was smaller or even absent, which replicated the pattern in MD analysis. We will address this reduced compatibility effect for concept stimuli in the General Discussion

section.

## General Discussion

### SUMMARY OF RESULTS

In 2 experiments, we altered 3 different IAT procedures by instructing participants to respond with a computer mouse rather than pressing keys. While the classical reaction-time IAT effects were replicated in these mouse tracking IATs, analysis of motor trajectories further demonstrated its potential in revealing the underlying process of IAT.

Across the 3 IATs, the motor trajectories were spatially attracted towards the incorrect response button, especially in incompatible trials, even though the behavioral responses were always correct. Evidence from distributional analyses supported the view that this attraction effect was not due to the averaging of straight and midflight-correction trajectories, but was continuous over the time course. Analysis of velocity profiles further indicated that mouse movements towards the correct response button were slower in incompatible trials, especially for attribute stimuli.

### IMPLICATIONS TO THE CONTINUOUS AND DYNAMIC REAL-TIME COGNITION FRAMEWORK

On a broad level, these results are consistent with the continuous and dynamic real-time cognition framework (Spivey & Dale, 2006; Spivey, 2007). According to this view, human cognition rarely operates in a discrete-stage based manner. Rather,

the cognitive system spends most of its time traveling within the intermediate region between different mental states, before settling on a stable and explicit response.

Mouse tracking measures have been used to demonstrate such a process in different areas, including language processing (Spivey et al., 2005; Dale et al., 2007; Farmer et al., 2007), social categorization (Freeman et al., 2008; Freeman & Ambady, 2009; Freeman et al., 2010), and social judgment (Wojnowicz et al., 2009), and others as well.

The current research contributes to this literature by showing that such a continuous and dynamic process may also enlighten understanding of underlying IAT tasks. As discussed in the Introduction, although many theoretical accounts of IAT more or less explicitly assume that the underlying process of IAT is a continuous one, there has been little research which aims to demonstrate the process directly, partly due to the limitation of reaction time measures.

Both of our Experiments yielded significant compatibility effects on trajectory curvature, such that mouse movements were more curved towards the incorrect response button in incompatible trials than compatible trials. One may argue that such effects were due to averaging staged-based trajectories rather than continuously attraction from response representations. We ruled out one of such stage-based accounts by examining the curvature distributions. What's more, Freeman and Ambady (2011b) demonstrated that, in a sex-categorization task, difference in trajectories curvature after hearing sex-typical and sex-intypical voices was absent when the alternative response button was replaced with a neutral control (e.g.

replaced “*female*” with “*farm*”), further supporting the continuous spatial attraction account. If we accept the presupposition of mouse tracking measures, then the continuous and spatial attraction effect we observed here may be considered as a demonstration of the continuous and dynamic view of the IAT underlying process. Although mouse movement is generally towards the correct response button, the two response representations are simultaneously and partially activated during the whole process.

#### IMPLICATIONS FOR THEORITICAL ACCOUNTS OF IAT EFFECT

The traditional IAT procedure only uses reaction times and error rate measures. Both of them have disadvantages in capturing the online characteristics of the underlying process of IAT, because they only capture the end product, not online measurement of the process. Therefore, researchers have to rely on techniques such as computational modeling to speculate on the nature of the mechanism of IAT (e.g. Conrey et al., 2005; Klauer et al., 2007). In this research, we provide an alternative solution to this problem by incorporating the mouse tracking measures into the standard IAT procedure. Because mouse tracking is a measure that records real-time activity, it can help us to examine the continuous and dynamic characteristics of the underlying process of IAT.

To demonstrate the potential of using mouse tracking measures, we examined the velocity profiles of mouse movement. Across 2 Experiments, results indicated that mouse movement towards the correct response button was slower in incompatible trials than compatible trials, especially for attribute stimuli. We believe these results

have implication for the random walk model account (Brendl et al., 2001) and the relevant feature account (DeHouwer, 2001; 2003) of IAT effect.

As described in Introduction, the random walk model account proposes that concept stimuli would provide two kinds of information (*valance* and *identity*), and these two sources would generate information in opposite directions in incompatible trials, therefore slow down the response process. However, attribute stimuli could only generate valance information, so the information accumulation rate should be the same in compatible and incompatible trials. The relevant feature account, on the other hand, proposes that stimulus-response compatibility should be lower in incompatible trials for both concept and attribute stimuli, therefore the information accumulation process should also be hindered for attribute stimuli as well, which is opposite to the random walk model's account.

Previous empirical studies showed that participants do spend more time in incompatible trials than compatible trials for attribute stimuli (Brendl et al., 2001). However, Brendl et al. attributed this to heighten response threshold rather than slower activation of correct response representation. Klauer et al. (2007) produced a diffusion model analysis of IAT, and found that parameter estimation of drift rate  $v$  (i.e. information accumulation rate) for attribute stimuli was lower in incompatible trials than in compatible trials, which was consistent with the relevant feature account's prediction.

Compared with these studies, mouse tracking measures used in our experiments allowed us to examine the information accumulation process for attribute stimuli in a

more direct way. Across 3 IATs, analysis of velocity profiles showed that for attribute stimuli, the mouse indeed moved slower in incompatible trials than in compatible trials. If we accept the presupposition that mouse velocity reflects information accumulation rate, then these results are consistent with the relevant feature account's predictions.

For trials with concept stimuli, we observed a novel finding for the two implicit self-esteem IATs: the compatibility effect on velocity for these stimuli was smaller than for attribute stimuli, or even absent. This is predicted by neither the random walk model account and the relevant feature account. We speculate that this reduced compatibility effect may be related to the concept category used in these implicit self-esteem IATs. As noted by Pinter and Greenwald (2005), IAT involves two contrasting concept categories and they may differ in the strength of their associations with the IAT's attribute categories. For flower-insect IAT, the category *flower* is of positive valance while the category *insect* is of negative valance, and such associations are universal for most people. However, in self-esteem IAT although the category *self* is likely to be positive, the category *other* is rather neutral in valance (for debates on this issue, see: Karpinski, 2004; Pinter & Greenwald, 2005). The process proposed by the relevant feature account may not be relevant in this case, as for these stimuli the most relevant feature would be their identity information. More theory-building and empirical studies which compare concept categories of different valance are needed to examine and interpret this novel phenomenon.

## LIMITATIONS AND CONCLUSION

We must make it clear that our results reported here are not conclusive but explorative. While the usage of mouse tracking measures has been growing rapidly over the past few years, it is still a new research field and our understanding about the nature of this measurement is still very limited at present. For example, although we treated the mouse trajectories as a reflection of representation activation process underlying IAT, it is highly possible that other processes also have impact on these trajectories as well, and empirical study on influence factors of mouse movement is still lacking.

Furthermore, we used the spatial attraction effect to support the continuous and dynamic view of the IAT process, but there are also alternative interpretations for this effect. Even though we ruled out the midflight-correction account by examining the bimodality of curvature distribution, there are also other possibilities. For example, the mouse might initially stay close the midline for a period, and then departed from it towards the correct response, and the compatibility effect on MD would be due to that incompatible trajectories lingered around the midline for a longer time than compatible ones. The absence of compatibility effect on velocities before time bin 3 seems also consistent with this account, and we were not able to rule out it.

As for the velocity profiles analysis, although we believe that our results are more consistent with the relevant feature account, there are also ambiguity around this claim. For example, previous research used mouse velocity an index of information accumulation rate (Wojnowicz et al., 2009), but more studies are needed to justify such usage.

What's more, Brendl et al. (2001) used heightened threshold to explained the compatibility effect on reaction times for attribute stimuli. As they have pointed out, response threshold is often difficult to examine empirically, and currently we don't have a clear idea on how to rule out this possibility with mouse tracking data.

Lastly, besides these two accounts there are other theoretical accounts for IAT which attribute the IAT effect to processes other than response representation activation (e.g. Klauer & Mierke, 2005, Rothermund & Wentura, 2004). Whether and how these processes may be mapped onto mouse tracking data is still an open question.

In summary, the way we interpreted mouse tracking data in our present research involves many presuppositions, and more evidence is needed to validate these. Nonetheless, the current research explores the possibility of incorporating mouse tracking measures into IAT procedure, and demonstrated its potential in revealing the underlying process of IAT. Because of its advantage in recording the process over real-time course, we believe that it is a useful addition to our research toolbox. As De Houwer et al. (2009) have pointed out; our understanding about the mechanism of the IAT effect is still very limited. Further research which employs mouse tracking measures could potentially provide more insight into this issue, as well as broaden our understanding about the basic principle of human cognition in general.

## References

- Asendorpf, J.B., Banse, R., & Muecke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, 83, 380–393.
- Brendl, C.M., Markman, A.B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology*, 81, 760-773.
- Cai, H.J. (2003). The effect of implicit self-esteem and the relationship of explicit self-esteem and implicit self-esteem. (in Chinese). *Acta Psychologica Sinica*, 35, 796-801.
- Conrey, F.R., Sherman, J.W., Gawronski, B., Hugener, K., & Groom, C.J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. *Journal of Personality and Social Psychology*, 89, 469-487.
- Dale, R., Kehoe, C., & Spivey, M.J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory & Cognition*, 35, 15-28.
- De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology*, 37, 443-451.
- De Houwer, J. (2003). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Kaluer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219-244). Mahwah, NJ: Erlbaum.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit

- Measures: A Normative Analysis and Review. *Psychological Bulletin*, 135, 347-368.
- Farmer, T.A., Anderson, S.E., & Spivey, M.J. (2007). Gradiency and visual context in syntactic garden-paths. *Journal of Memory and Language*, 57, 570-595.
- Freeman, J.B. & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science*, 20, 1183-1188.
- Freeman, J.B. & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, 42, 226-241.
- Freeman, J.B. & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, 118, 247-279.
- Freeman, J.B. & Ambady, N. (2011b). When two become one: Temporally dynamic integration of the face and voice. *Journal of Experimental Social Psychology*, 47, 259-263.
- Freeman, J.B., Ambady, N., Rule, N.O., & Johnson, K.L. (2008). Will a category cue attract you? Motor output reveals dynamic competition across person construal. *Journal of Experimental Psychology: General*, 137, 673-69.
- Freeman, J.B., Pauker, K., Apfelbaum, E.P., & Ambady, N. (2010). Continuous dynamics in the real-time perception of race. *Journal of Experimental Social Psychology*, 46, 179-185.
- Greenwald, A.G., & Farnham, S.D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social*

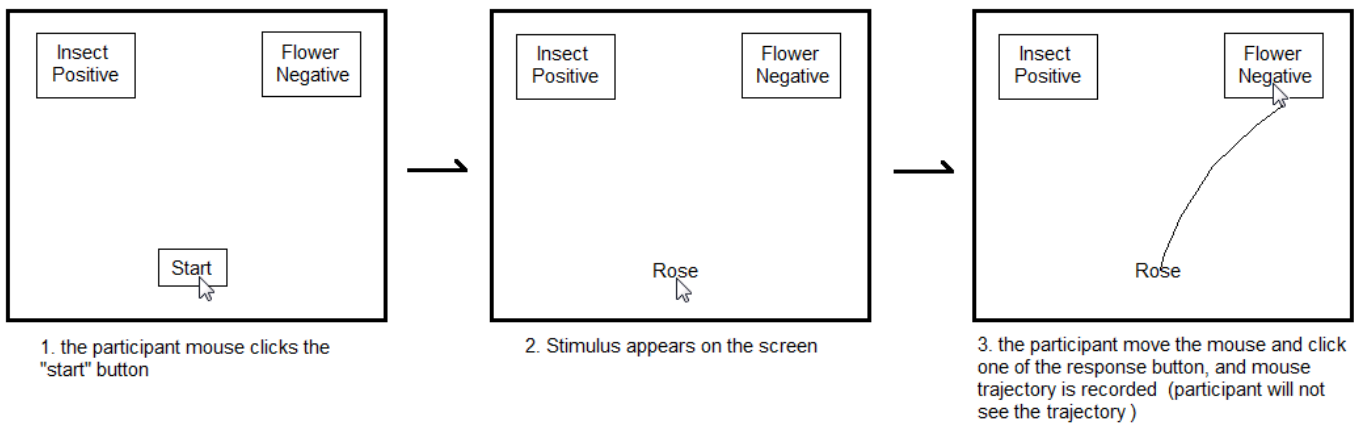
- Psychology*, 79, 1022–1038.
- Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464-48.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197–216.
- Karpinski, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. *Personality and Social Psychology Bulletin*, 30, 22-34.
- Klauer, K.C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the Implicit Association Test effect. *Personality and Social Psychology Bulletin*, 31, 208-217.
- Klauer, K.C., & Voss, A., Schmitz, F., & Teige-Mocigemba S. (2007). Process Components of the Implicit Association Test: A Diffusion-Model Analysis. *Journal of Personality and Social Psychology*, 93, 353-368.
- Nosek, B.A., Greenwald, A.G., & Banaji, M.R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review (pp. 265–292). In Bargh J.A. (Ed.), *Automatic processes in social thinking and behavior*. Psychology Press.
- Pinter, B., & Greenwald, A. G. (2005). Clarifying the role of the “other” category in the self-esteem IAT. *Experimental Psychology*, 52, 74-79.
- Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit

- Association Test (IAT): Dissociating salience from associations. *Journal of Experimental Psychology: General*, 133, 139–165.
- Rudman, L.A., Greenwald, A.G., & McGhee, D.E. (2001). Implicit self-concept and evaluative implicit gender stereotypes: Self and ingroup share desirable traits. *Personality and Social Psychology Bulletin*, 27, 1164–1178.
- SAS Institute (1989). *SAS/STAT user's guide*. Cary, NC: Author.
- Spivey, M.J. (2007). *The continuity of mind*. New York: Oxford University Press.
- Spivey, M. J., & Dale, R. (2004). On the continuity of mind: Toward a dynamical account of cognition. In *The psychology of learning and motivation: Advances in research and theory* (Vol. 45, pp. 87-142). San Diego, CA: Elsevier.
- Spivey, M.J., & Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science*, 15, 207-211.
- Spivey, M.J., Grosjean, M., & Knoblich, G.. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, 102, 10393-10398.
- Wojnowicz, M.T., Ferguson, M.J., Dale, R., & Spivey, M.J. (2009). The self-organization of explicit attitudes. *Psychological Science*, 20, 1428-1435.

**Table 1. Sequence of blocks of the modified flower-insect IAT in experiment 1**

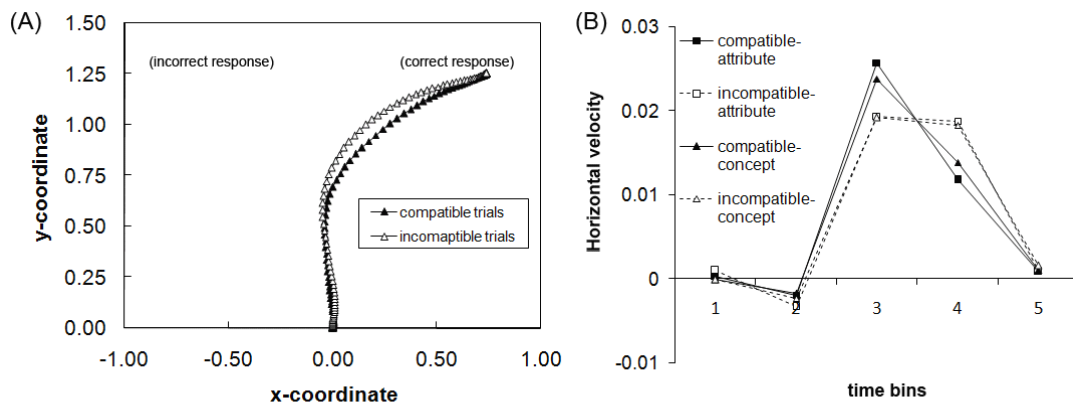
<b>Block number</b>	<b>Stimuli</b>	<b>Response button 1</b>	<b>Response button 2</b>
1	20 concept words	flower	Insect
2	20 attribute words	positive	negative
3	10 concept words + 10 attribute words	flower + positive	insect + negative
4	20 concept words + 20 attribute words	flower + positive	insect + negative
5	20 concept words	Insect	flower
6	10 concept words + 10 attribute words	insect + positive	flower + negative
7	20 concept words + 20 attribute words	insect + positive	flower + negative

Block 3 & 6 were training blocks, and block 4 & 7 were critical compatible & incompatible blocks. For half of the participants, the order of compatible & incompatible blocks was reversed, so their sequence of blocks was 5-2-6-7-1-3-4. Block order did not yield any significant effect on the following data analysis.



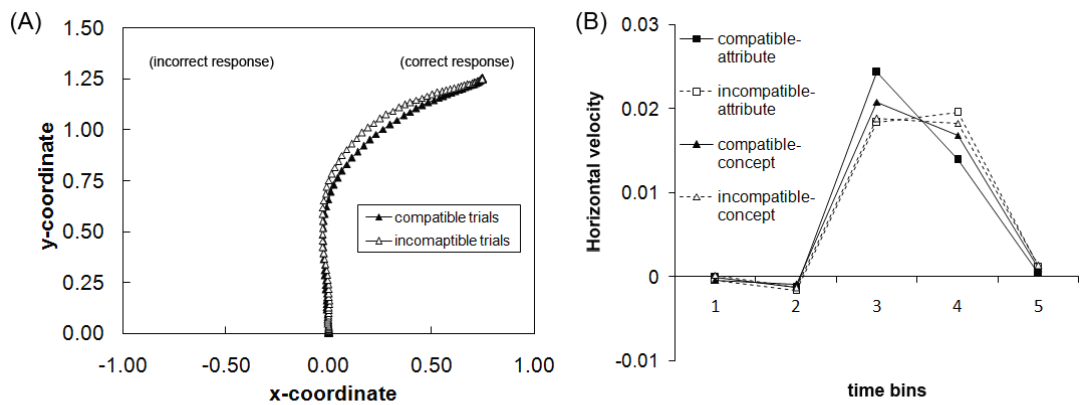
**Figure 1.** Illustration of a typical trial of mouse tracking IAT.

(A)

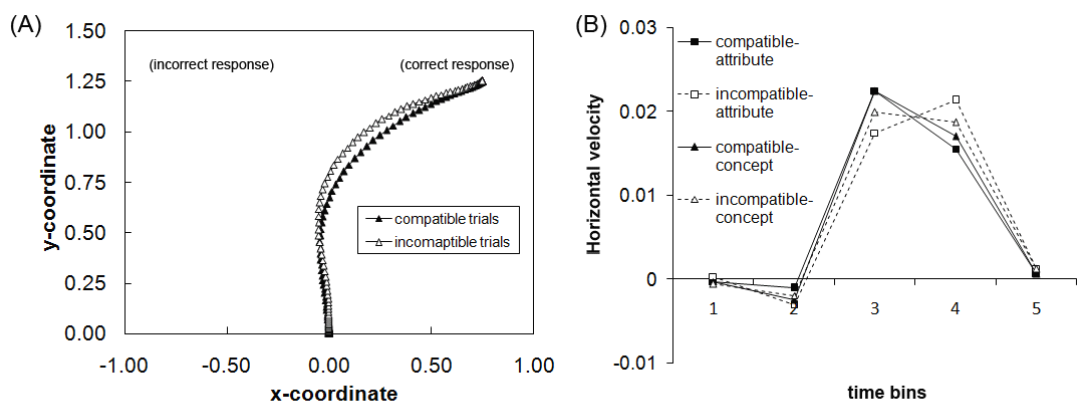


**Figure 2.** (A) Averaged motor trajectories of compatible & incompatible trials in Experiment 1.

(B) Velocity profiles by trial type and stimulus type in Experiment 1.



**Figure 3.** (A) Averaged motor trajectories of compatible & incompatible trials in affective IAT in Experiment 2. (B) velocity profiles by trial type and stimulus type in affective IAT in Experiment 2.



**Figure 4.** (A) Averaged motor trajectories of compatible & incompatible trials in evaluative IAT in Experiment 2. (B) velocity profiles by trial type and stimulus type in evaluative IAT in Experiment 2.