Tracing the Process of Rating Decisions through Cursor Movements

Kenpei SHIINA (shiinaATwaseda.jp)
Department of Educational Psychology, Waseda University, Tokyo, Japan

Abstract

To study the decision process during rating tasks, PC cursor trajectories were recorded and analyzed. The trajectories were often successions of rapid saccadic-like movements that are called *strokes* in this paper. The analysis of strokes revealed that the distribution of strokes differed across tasks as a function of task difficulty. A simple number matching task elicited fewer strokes, shorter response times, and velocity patterns resembling simple ballistic reaching movements. A personality rating task tended to elicit multiple strokes and longer RTs, which caused a typical inverted-U RT effect. The shape and speed of tangential velocity of trajectories may reflect participant's internal states, especially when cognitive loads are high.

Keywords: Rating decision; inverted-U effect; Response time; decisional fluctuation.

Rating Scales and Trajectories

Despite the advances in new technologies and modern methods, rating scales are still a mainstay of data acquisition in psychology and cognitive science. Unfortunately, there remain many unresolved fundamental questions concerning the nature of rating scale methods and the data arising from them. It appears that the rating scale is a "black box" that no one is willing to open. One of these problems is that we do not know very well how the rating decisions are physically performed, let alone the internal processes influencing choice and response time. Within the framework of the process tracing paradigm often employed in behavioral decision making research, the present study analyzed the trajectories of cursors in rating decisions using a computer-based decision interface (Figure 1) and contrasted responses to a digit matching Benchmark Task and a more cognitively engaging Big 5 personality questionnaire Task.

Trajectory monitoring of this type has been used in many situations to infer the internal states of decision makers (Baccino, 1994; Baccino & Kennedy, 1995; Arroyo, Selker, & Wei, 2006; Spivey, Grosjean, & Knoblich, 2005; Dale, Kehoe, & Spivey, 2007; Farmer, Cargill, Hindy, Dale, & Spivey, 2007; Freeman, Ambady, Rule, & Johnson, 2008; Shiina, 2008). Recently, Shiina (2011) pointed out that the cursor trajectories in rating tasks were not always smooth continuous curves as in the previous studies, but were often successions of rapid brief movements similar to saccadic eye-movements. They are called *strokes* in this paper.

The purpose of the present paper is two-fold: investigation into the characteristics of such strokes including the relationship between strokes and response times, and the search for the evidence that stroke frequencies and shapes can serve as indices of some decisional characteristics, decisional fluctuation in particular. Finally, implications for decision research are highlighted in Discussion.

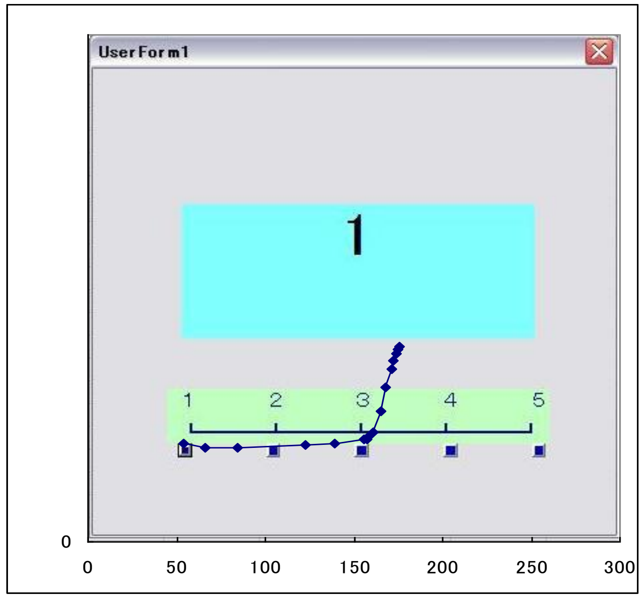


Figure 1: The form used in the number matching Benchmark Task and an example of a cursor trajectory that traveled from "Start" button, which disappeared from the screen after the initial click, to Category 1 (lower left small square button). There are two strokes (rapid movements) in this trajectory. See text and Figure 4 as well.

We used two tasks that call for distinct cognitive processes. In Benchmark Task, the goal of cursor movements was set by the experimenter and the respondents were simply following the order of the experimenter, whereas in the second Big 5 Task, the goal of cursor movements should be set by the participants' spontaneous judgment, which was the reason why a Big 5 personality assessment task with well-studied items was employed. It seems very plausible that the differences between the two tasks reflect the internal states of the participants.

Experiment

In this trajectory monitoring experiment, the form shown in Figure 1 was used. Figure 1 also shows an example of a cursor trajectory. There were 5 ordered categories with a set of numerical labels from 1 to 5 in Benchmark Task and a set of verbal labels in Big 5 Task.

Procedure Using the form shown in Figure 1, a trial was initiated with the presentation of a center button labeled "Start". After the initial click of the start button, the button disappeared from the screen. The start button is not shown in Figure 1 but the origin of the trajectory example shows its approximate location. Immediately after the initial click of the start button, a problem appeared in the stimulus display box in the center of the form and the participants were asked to click a "correct" or "most suitable" category button (small squares in the bottom of the form) as quickly as possible. The time and trajectory of the cursor between the initial and last clicks were recorded. The experimental program was written in VBA for Microsoft Excel and the experiment was run on Excel.

Tasks In Benchmark Task, a random digit from 1 to 5 was presented in the display box and the participant's task was to click the corresponding response button as quickly as possible. The trajectories arising from this task served as baselines with a minimum of cognitive components. The 5 digits were randomly presented 5 times each.

In Big 5 Task, participants were asked to rate their personality by clicking one of the 5 buttons. In the stimulus display box, adjectives or sentences based upon Big 5 theory were presented. In this task, instead of the 5 numerals "No", "Don't know", and "Yes" labels (in Japanese) were placed in the positions of "1", "3", and "5", respectively and there were no labels for "2" and "4". There were 30 personality items written in Japanese (Shimizu and Yamamoto, 2007) and these verbal labels are typically used in questionnaires of this type in Japan. All participants were first given Benchmark Task and then Big 5 Task.

Participants The participants were 483 undergraduates of Waseda University. They were native or quasi-native Japanese speakers.

Analysis of Mean RT and Inverted-U Effect

The time from the initial click of "Start" button to the final click of a rating category button was defined as response time (RT). In Figure 2 the mean response times as a function of the final category clicked (answer) are shown.

In Benchmark Task, average response times were U-shaped, reflecting the physical distance between the central start button and the lower, horizontally arrayed response buttons. In Big 5 Task, in contrast, a typical inverted-U effect was observed. Inverted-U effects have been found in a variety of tasks and stimuli that use response scales with polytomous ordered categories (e.g., Likert scale). In general, the Inverted-U effect refers to the fact that responses are more unstable and require more processing time in the middle than in the extremes of the response scale (Kuiper, 1981). This multi-faceted effect is observed in the forms of response times, errors, and response variabilities, called, inverted-U response time, error, and uncertainty effects, respectively (Mignault, Marley, & Chaudhuri, 2008). There was no inverted-U RT effect in Benchmark Task while Big 5 Task showed an inverted-U RT pattern. Therefore, it is reasonable to say that inverted-U RT effect did occur in Big 5 Task and the difference between the two lines in Figure 2 was due to cognitive loads in Big 5 Task. Further, it might be predicted that inverted-U RT effect does not occur when the cognitive load is low (Shiina, 2011).

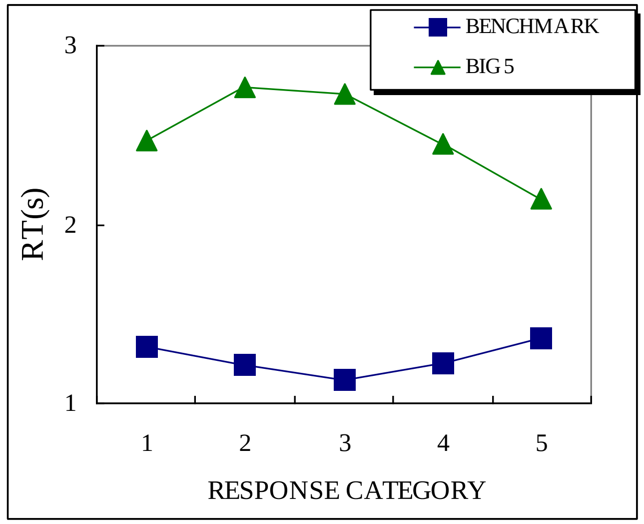


Figure 2: Mean RTs as a function of final category chosen.

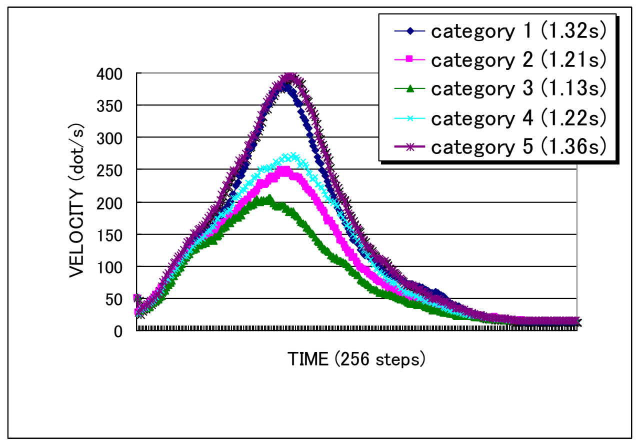


Figure 3a: Average tangential velocities in Benchmark Task.

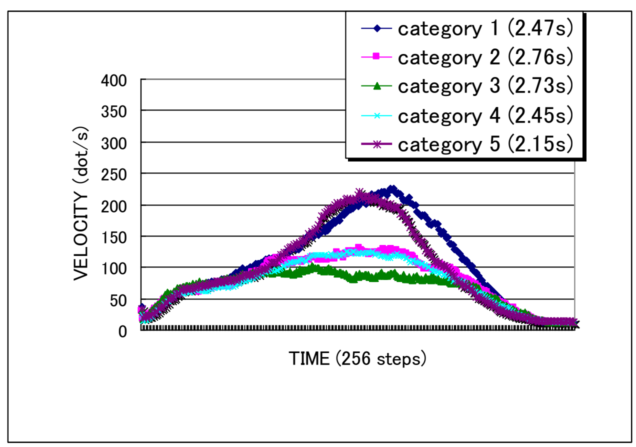


Figure 3b: Average tangential velocities in Big 5 Task.

Analysis of Average Tangential Velocities

A trajectory is a time-indexed 2-dimensional function: $(x(t), y(t)),\ 0 < t < RT$. Because each trajectory had a different RT that was measured from the initial to the final clicks, some standardization was needed. To do this, we first divided the RT by 256 to define a discretized time step and then estimated the locations $(x_i, y_i),\ i = 0, 255$ by linear interpolation. This procedure maintains the length and the shape of a trajectory while equalizing the number of time steps. A convenient way of comparing trajectories without losing information in time domain is to analyze tangential velocity of trajectories (Shiina, 2008). Tangential velocity of a trajectory at time *i* is defined by

$$TV_i = \frac{\sqrt{(x_{i+1}-x_i)^2+(y_{i+1}-y_i)^2}}{RT/256} = \frac{\textit{Traveled Distance}}{\textit{time step}} \quad (1)$$

Then average curves of tangential velocities of the trajectories that arrived at the same rating category (Figure 3) were computed. In Figure 3a, there are 5 average tangential velocity curves corresponding to the 5 final rating categories clicked in Benchmark Task. The curves are showing the "speed" of the cursor as a function of standardized time and the numbers in parentheses are mean RTs. In the same way, the tangential velocity curves in Big 5 Task are shown in Figure 3b. Obviously, one can observe that tangential velocities a) in Big 5 Task were lower, b) toward the middle categories were lower in both tasks, c) toward the middle categories in Big 5 Task were flat, were d) positively correlated with RT in Benchmark Task, and were e) negatively correlated with RT in Big 5 Task. Observation a) implies the effect of cognitive load, b, d) the effect of physical distance, c) decisional hesitation and vacillation (Shiina, 2008) , and e) a realization of inverted-U variability effect (Shiina, 2008, 2011).

The average curves in Figure 3 gives a strong impression that the raw trajectories must also be smooth single movements as in the case of reaching movement (Kelso, Southard, & Goodman, 1979, for example). In reality, however, most of individual trajectories were not continuous smooth curves. Figure 4 presents examples of such velocity curves to be analyzed extensively in the next section. There were several peaks and hills derived from rapid movements of the cursor that are referred to as *strokes* (see Figure 1). Single, twin, triple or more stroke curves were abundant and there were drift patterns as well that are depicted in the last row of Figure 4. Moreover, the strokes do not appear to be homogeneous: we can see pulses, peaks, and hills, although there are no rigorous definitions of them as in the real world. At any rate, the emergence of strokes in decision making is a novel finding, if not an artifact inherent in mouse interface, showing a new aspect of real time processing of decision making.

Tangential Velocity of Individual Raw Trajectories

The first step to analyze individual raw trajectories should be the counting of strokes as a function of final categories clicked. Because it is almost impossible to visually count pulses and hills in the thousands of velocity graphs, a stroke detection filter was designed.

Stroke Detection Filter Finding out strokes from cursor trajectories is equivalent to searching for peaks and hills in

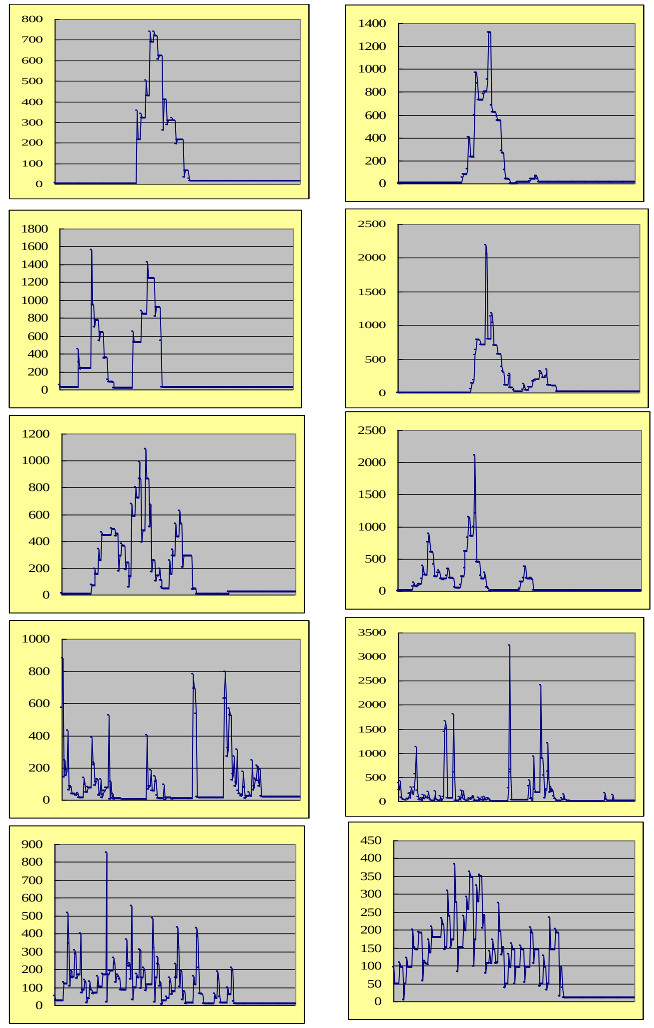


Figure 4: Raw tangential velocity curves (Benchmark Task) showing peaks and hills (strokes) and drifts (the last row). Second row left corresponds to the example trajectory in Figure 1. Abscissa indicates standardized time (256 time steps) and ordinate indicates tangential velocity (dot/s).

tangential velocity graphs shown in Figure 4. In general, it is a hard problem to extract signals from noises. Moreover, the definition of peaks and hills in the present situation is difficult to make partly because there are no previous studies that help us to set guidelines and partly because there cannot be a *correct* definition of a hill or a peak.

With these problems in mind, the filter was designed in the following way. First, because tangential velocities as defined by Equation (1) cannot be negative, detection and filtering methods from digital image analysis can be used. In digital image analysis, edge detection is a fundamental technique with known algorithms and outcomes. A standard procedure is to compute second derivatives of an image and search for zero-crossings. Employing this idea for our stroke detection filter, first derivatives of tangential velocity (acceleration) were computed and then zero-crossings corresponding to hill tops (no acceleration points) were detected. We need not compute second derivatives in our case because we are searching for a hill top and the first derivative at the hill top should be 0, the first derivatives to the left of the top should be positive (upward slope), and the first derivatives to the right of the top should be negative (downward slope). Zero-crossings obtained from this procedure arise from both true strokes and noise: By applying a moving average filter, it was hoped that noise-generated zero-crossings would be eliminated almost entirely. A moving average filter (Smith, 1997, p.277) was used because it is optimal for a common task of reducing random noise while retaining a sharp step response. Finally, a stroke (a hill on velocity graphs) was detected by setting two parameters, *width* and *steepness*. The first parameter defines a minimum width of a knob, and the second a minimum steepness of a knob, to be called a hill. More specifically, the width of a knob is defined on the first derivative time-series as:

Width= Time difference between a valley bottom to the right of a zero-crossing and a hill top to the left of a zero-crossing,

Steepness=Height difference between the hill top and valley bottom divided by *Width* as defined above.

The number of peaks and hills detected is a function of these two parameters, because changing the parameters gives different definitions of a peak or hill.

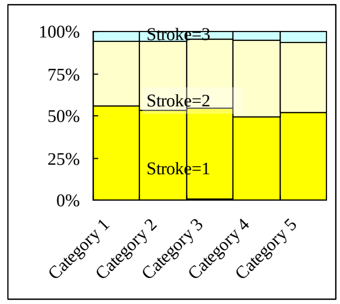
To summarize, strokes in original 2D trajectories are translated into peaks and hills in tangential velocity graphs. Hill tops are translated into zero-crossings in first derivative time-series graphs after application of a 14-point moving average filter. Finally, the zero-crossing detection algorithm counts the number of hills (strokes) using the two parameters that define what a hill should be.

Results of Stroke Analysis Figure 5 shows the results of stroke counting. Define Stroke Ratio or conditional probability:

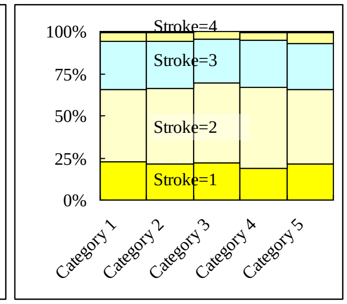
$$P(i|C) \equiv SR_C(i)$$

$$\equiv \frac{\text{The number of trajectories arriving at Category } C \text{ and has } i \text{ strokes}}{\text{The number of trajectories arriving at Category } C}$$

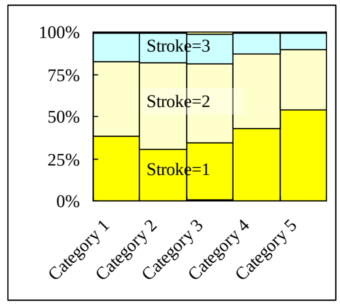
where C is a rating category and i is the number of strokes. In Figures 5a and 5b, Stroke Ratios in Benchmark Task for the 5 response categories are depicted. Figure 5a shows the Stroke Ratios using a strict criterion (*width*=31, *steepness*=0.5) and Figure 5b using a lax criterion (*width*=20, *steepness*=0.5). Similarly, Figures 5c and 5d show results for Big 5 Task. Figure 5c is using the strict criterion (*width*=31, *steepness*=0.5) while Figure 5d is using the lax criterion (*width*=20, *steepness*=0.5). Under the strict criterion, only larger hills and peaks were counted as strokes while under the lax criterion smaller hills and peaks were also detected and counted as strokes. Note that a lax criterion tends to pick up noise-generated hills while a strict criterion tends to reject true signals: This trade-off between noise rejection and signal detection is well known in signal analysis and cannot be eliminated completely. This is the reason for considering several criteria at the same time.



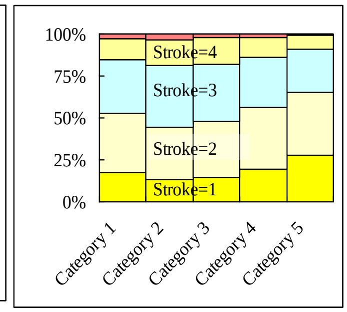
a) Benchmark Strict criterion



b) Benchmark Lax criterion



c) Big 5 Strict criterion



d) Big 5 Lax criterion

Figure 5: The ratios of stroke numbers for each rating category. Very small ratios (more than 4 or 5 strokes) are invisible in this figure. a) Benchmark Task under strict criterion (12,075 total responses) , b) Benchmark Task under lax criterion. In Benchmark Task, Categories 1, 2, 3, 4, and, 5 were clicked 2417, 2414, 2415, 2414, and 2415 times, respectively. c) Big 5 Task under strict criterion (14,490 total responses), d) Big 5 Task under lax criterion. Categories 1, 2, 3, 4, and, 5 were clicked 2642, 1987, 2289, 3367, and 4205 times, respectively.

Major observations are as follows:

1) In Benchmark Task (Figures 5a and 5b), the stroke ratios were about the same across categories under both criteria, meaning that the probability of strokes occurring was not related to the final categories clicked, while in Big 5 Task (Figures 5c and 5d) , we find rather systematic differences in the stroke ratios.

2) Figures 5a and 5c tell us that, under the strict criterion, $SR_C(1)$ (single-stroke trajectories ratio) decreased and multiple-stroke trajectories (2 or more) increased with the cognitive load of personality judgment except Category 5. Similarly, under the lax criterion, comparison of Figures 5b and 5d tells us that $SR_C(2)$ (double-stroke trajectories) decreased and triple-or-more stroke trajectories increased with the cognitive load imposed by personality judgment except again Category 5. Taken together, it is safe to say that the personality judgment increased the number of strokes, with the exception of Category 5.

3) In Big 5 Task (Figures 5c and 5d), it is apparent that the average number of strokes was larger for Categories 2 and 3 under both criteria.

4) In Big 5 Task, shorter averaged RTs (Figure 2) were associated with high SR_C (1)'s, that is, RT was a decreasing function of single-peak ratio. The correlation between them is r = - 0.98 under the strict criterion and r = - 0.96 under the lax criterion, although the values were calculated from only 5 points and are not too reliable.

5) The responses to Category 5 were unique. When participants clicked Category 5 or "Yes" category in Big 5 Task, the stroke ratios were very similar to those in Benchmark Task as if there were no cognitive loads. This *"Yes" effect* merits further investigation in the future.

From these observations, it is apparent that the number of strokes is related to rating judgments and RTs. The next task should be to clarify how and why they are linked. Although it is impossible to fully extend such an analysis, connections to motor movement research and inverted-U effects are described in the next section.

Ballistic Movements and Decisional Hesitation The ratio of single stroke trajectory has a theoretically important meaning because it relates to the ballistic movements. It is well known in motor control studies that simple reaching movements involve an initial *ballistic* phase followed by a second *corrective control* phase (Elliot, Helsen, & Chua, 2001). Because of the speed limitations of neurotransmission, ballistic movements are under feed forward control. This means that in a ballistic movement the initial velocity and direction should be determined before the initiation of the movement and thus should be unaffected by cognitive processes during the moving. In other words, ballistic movements are "thoughtless" after departure, and can serve as an index that the participants are not in the states of hesitation or deliberation. Therefore, if a single stroke movement in the present study is a ballistic movement, it implies that there was little, if any, decisional hesitation. A typical feature of ballistic movement is relatively high and bell-shaped (Gaussian) tangential velocity. Therefore, deviation of tangential velocity curve from bell-shape coupled with velocity levels can help determine whether the trajectory is ballistic or non-ballistic (Shiina, 2008).

In the previous section, we conducted stroke-wise categorization of trajectories. Using these results, average tangential velocities of single stroke trajectories toward the 5 categories in Benchmark Task are drawn in Figure 6a, because it was very plausible that Benchmark Task produces simple reaching movements that contain ballistic components and thus has rapid bell-shaped tangential velocities. The curves are obviously bell-shaped, indicating that Benchmark Task evoked simple reaching movements that contain ballistic components. The curves in Figure 6a also show that the overall tangential velocities were relatively high, another characteristic of ballistic movements.

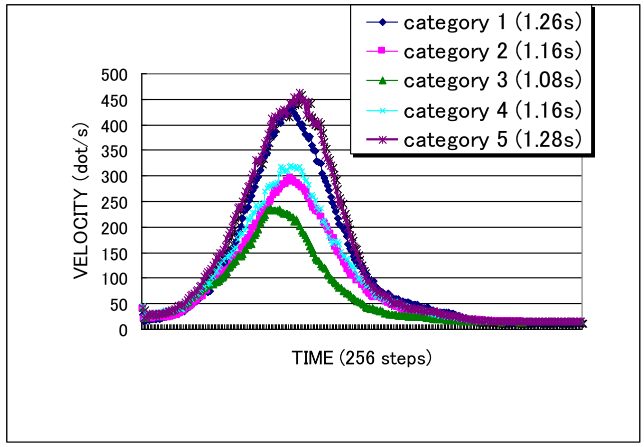


Figure 6a: Average tangential velocities of single stroke trajectories toward the 5 categories in Benchmark Task. Bell-shape is a feature of ballistic movement. Lax criterion.

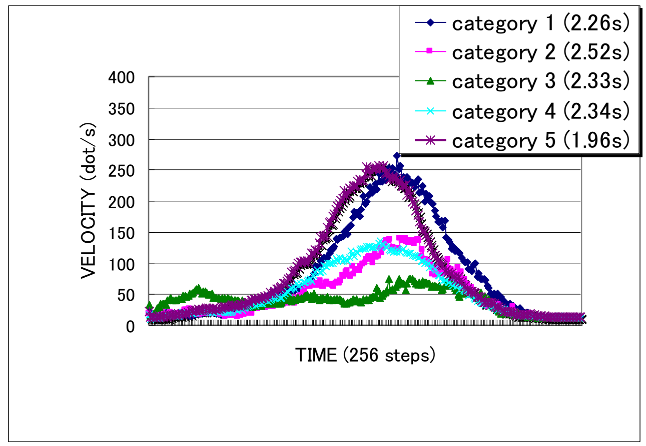


Figure 6b: Average tangential velocities of single stroke trajectories toward the 5 categories in Big 5 Task. Deviation from Bell-shape indicates deviation from ballistic movement. Lax criterion.

In Figure 6b, average tangential velocities of single stroke trajectories toward the 5 categories in Big 5 Task are shown. Deviation from Bell-shape indicates deviation from ballistic movement and then suggests decisional hesitation. According to this rule, we can conjecture that Categories 1 and 5 were chosen easily while the other categories were chosen after deliberation and hesitation. Of course, this type of multi-stage inference is sometimes dangerous and it is too early for drawing a conclusion.

Inverted-U Effect and RT In Figure 2, a typical inverted-U RT effect was found. Then the stroke analysis revealed that the raw tangential trajectories were rather heterogeneous, having several strokes (peaks or hills). Finally, in Figure 6b (and Figure 3b), it was revealed that the single stroke trajectories toward the middle 3 categories deviated from ballistic movements in Big 5 Task. In this task, Double-or-more stroke average tangential velocity curves were much more deviant from bell-shaped and thus different from ballistic movements (although cannot be shown explicitly due to space limitation).

Figure 6 shows the purest average velocities in the sense that they are drawn using filtered velocities and their shapes in Benchmark Task do look bell-shaped. Therefore, a direct cause of inverted U-effect in Big 5 Task would be that the tangential velocities of single-stroke movements toward the middle 3 categories were not bell-shaped and too low that many of them were not ballistic at all.

The most parsimonious overall explanation for the present results is that trajectories toward the middle categories in Big 5 Task had more strokes and were non-ballistic (vibrant and instable) so that their average speed slowed down and thus the inverted-U RT effect occurred. A more bold psychological interpretation would be that participants' hesitation and deliberation caused internal fluctuations and the strokes and trajectory vibrations were the manifestation of these internal fluctuations (Shiina, 2011).

Because average tangential velocity curves strongly reflect stroke-onset frequencies averaged over time, so other interpretations are possible. We are not in the position to make a definite conclusion and the second psychological interpretation, although suggestive, should be justified in the future research.

Discussion

This study is basically an exploratory one and there are many important variables that this study did not deal with: individual differences, direction of velocities, distribution of RT, and distribution of stroke duration and length, to name but a few. Nevertheless, some interesting relationships emerged that suggest the interplay between internal decisions and physical movements (strokes in particular). Major results of this paper can be summarized as follows: 1) The distribution of strokes differed across tasks, and the number of strokes increased with task difficulty. 2) Strokes in a trajectory slowed down RT, causing the inverted-U effect. 3) The shape and speed of tangential velocity of a trajectory may suggest the type of movement (ballistic vs. non-ballistic) that in turn suggest the participant's internal states, especially when the cognitive loads are high. 4) Finally, positive responses did not seem to elicit significant internal conflict. This "Yes" effect appears interesting in connection with recent unconscious decision making research. These findings will be fully explored in the future research.

References

Arroyo, E, Selker, T, & Wei, W. (2006). Usability tool for analysis of web designs using mouse tracks. *CHI '06 extended abstracts on Human factors in computing systems*, 484-489.

Baccino, T. (1994). Contrôle d'un pointage avec souris informatique et codage spatial. *L'Année Psychologique*, **94**, 11-24.

Baccino, T., & Kennedy, A. (1995). MICELAB: Spatial processing of mouse movement in Turbo Pascal. *Behavior Research Methods, Instruments, & Computers,* **27**, 76–82.

Dale, R., Kehoe, C., & Spivey, M.J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory and Cognition*, **35**, 15-28.

Elliott, D., Helsen, W. F., & Chua, R. (2001). A century later: Woodworth's (1899) two-component model of goal-directed aiming. *Psychological Bulletin*, **127**, 342-357.

Farmer, T.A., Cargill, S.A., Hindy, N.C., Dale, R., & Spivey, M.J. (2007). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science*, **31**, 889-909.

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**, 673-690.

Kelso, J. A.S., Southard, D. L., & Goodman, D. (1979). On the nature of human interlimb coordination. *Science*, **203**, 1029-1031.

Kuiper, N. A. (1981). Convergent evidence for the self as a prototype: the "inverted-U RT effect" for self and other judgments. *Personality and Social Psychology Bulletin*, **7**, 438-443.

Mignault, A., Marley, A. A. J., & Chaudhuri, A. (2008). Inverted-U effects generalize to the judgment of subjective properties of faces. *Perception & Psychophysics*, **70**, 1274-1288.

Shiina, K. (2008). Vacillation and Hesitation in Category Rating: Evidence from PC Cursor Trajectories 2008 *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp.1894-1899)**.** Washington D.C.: Cognitive Science Society

Shiina, K. (2011). Trajectory-Variability and Inverted-U Effects in Rating Decisions. *Proceedings of* EUROCOGSCI 2011.

Shimizu, K., & Yamamoto, R. (2007). Modeling of the Relationships among the Constructs of Personality using Parceling method: Big Five, State-Trait, Anxiety, and Mood States. *Bulletin of the Faculty of Sociology (Kansai University)*, **38**, 61-96.

Smith, S. W. (1997). The Scientist and Engineer's Guide to Digital Signal Processing. California Technical Pub.

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