To appear in the Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, CA: Human Factors and Ergonomics Society, 2003.

TIMING OF DISCRETE ACTIONS: EVIDENCE FOR PERCEPTUAL DOMINANCE IN HUMAN PERFORMANCE

Xidong Xu and Esa M. Rantanen
University of Illinois at Urbana Champaign

An experiment involving a street-crossing task is described. The subjects were required to release a pedestrian within a narrow window of opportunity to cross the road without being run over by vehicles. To test a number of hypotheses on whether timing of actions is driven by perceptual cues or some internal clocking mechanisms, the experiment involved conditions where the street-crossing task was either visible throughout trials or where the vehicles were removed, and coupled with a concurrent mental arithmetic loading task with either visually or auditorily presented stimuli. The data support the hypothesis that when the stimuli remain visible, people predominantly rely on perceptual cues for timing discrete actions. Significant performance decrements were observed in the removal conditions. Importance of the visual component in the task performance was supported by the larger performance decrement in the unmasked condition when the subjects' attention had to be divided between two sets of visual stimuli. When the primary task stimuli were removed, the clock counting seems to be the preferred strategy for the task under study.

INTRODUCTION

In various domains such as aviation and surface transportation, it is essential that people can predict the future status of moving objects. For example, pilots need to predict the future trajectories and positions of their own positions and other aircraft to avoid collisions. Drivers and pedestrians need to estimate their relative positions to prevent accidents. There are several potential strategies to use for this kind of prediction tasks. For example, one can use the distance divided by speed method to "compute" the time an object will travel a certain distance and then use a clocking or timing mechanism to count down the time. That is, the response will be initiated when the duration reaches the estimated time (Tresilian, 1995). Another possible strategy is referred to as cognitive motion extrapolation (CME) (DeLucia & Liddell, 1998; Lyon & Waag, 1995). According to this hypothesis, an observer may develop a cognitive model or representation of the object's motion at the initial stage of the motion, and then project the object's future motion based on that model, possibly involving visual imagery. DeLucia and Liddell's (1998) results suggest that, at least for lateral motion on a two-dimensional surface, observers use the CME method rather than the pure timing or clocking process in predicting the time an object would reach its destination.

In Rantanen and Xu (2001), the task involved a situation of crossing a busy street simulated on a computer screen. Symbols of vehicles traveled across a traverse two-dimensional surface in a continuous stream and the subject was required to let a "pedestrian" cross the street between vehicles. The independent variables that were manipulated were the size of gaps between consecutive vehicles and the speed of the vehicles, creating three temporal windows of opportunity. Among other findings, the results showed that subjects adjusted their "pedestrian" release times appropriately to the increasing accuracy requirements by becoming more consistent in timing and earlier release times. Of particular interest was the result that although the large window/fast speed and small window/slow speed conditions resulted in identical durations of the temporal windows of opportunity, subjects' timing performance was different in these two conditions, suggesting the visual elements of the task (i.e., the physical size of the windows) dominated the perceived accuracy demand of the task and thus the timing of the "pedestrian" releasing.

The question addressed by the present study, following up study Rantanen and Xu (2001), was how the removal of the stimulus (after being visible for a few seconds) and a concurrent task presented either visually or auditorily influence human performance in timing of discrete actions. This design allowed for examination of several hypotheses: If the primary mechanism behind the timing of the pedestrian release is perceptual, then the removal of the stimuli would be detrimental to the performance. Visual loading task should also degrade subjects' performance more than the auditory loading task while the stimuli remained visible throughout the trial, because the visual loading task would compete with the pedestrian releasing task for the visual resource. If the subjects rely on an internal clock in timing the pedestrian release, however, their performance should be disrupted, with equal performance decrement, by the loading task in both visual and auditory conditions. The removal of the stimulus could supposedly force the subjects to rely on a mental clocking mechanism instead of perceptual cues available on the screen. On the other hand, it is plausible that the subjects would continue tracking the suddenly invisible stimulus and release the pedestrian based on their imagery of the task geometry, which is consistent with the CME model.

The present study allowed us to test whether the subjects in the removal conditions relied on mental imagery in tracking the now invisible vehicle and releasing the pedestrian based on imagined positions of the pedestrian and the vehicle. If this were the case, visual presentation of the loading task should disrupt their task more than auditory loading task. Alternatively, if subjects relied on the clocking mechanism for timing the pedestrian release in these conditions, equal amount of interference would be expected from the visual and the auditory loading tasks.

METHOD

Participants

Ten undergraduate and graduate students (two female and eight male) from the University of Illinois at Urbana-Champaign participated in the present experiment and were paid $8.00/h. All subjects had normal or corrected to normal vision and normal hearing.

Experimental Apparatus and Task

The apparatus and the task for this experiment were basically the same as those used in Rantanen and Xu (2001). A series of simulated vehicles were moving on a two-dimensional traverse surface and the task for the participant was to release a pedestrian waiting at the road side to cross the street, without being run over by the vehicles, by pressing a key. The participant took a God's-eye view of the pedestrian and the vehicles. In Rantanen and Xu (2001), successful completion of the task first required correct judgment of a sufficient interval between successive vehicles (a temporal window of opportunity) among non-window intervals to allow the pedestrian safely cross the road. This was not necessary in the present study, since it was explicitly explained to the participant that he or she needed to cross at intervals formed by vehicles 1 and 2, 3 and 4, etc. in a block. As in Rantanen and Xu (2001), the release of the pedestrian required taking into account the time required for crossing and correct timing of the initiation of the crossing. Since the pedestrian was waiting at a given distance from the road, crossing had to be initiated before the vehicle preceding the window of opportunity had passed the crossing path (i.e., before the window actually opened), requiring the subject to anticipate both the time required for the pedestrian to reach the road and to cross it. Because the pedestrian could not be stopped, rushed, or called back once crossing had been initiated, a failure to ascertain a safe window for crossing resulted in the pedestrian being hit by a car. Timing errors had similar consequences: Too early initiation of crossing resulted in the pedestrian walking in front or into the side of the first car and too late initiation in being hit by the second car (see Figure 1).

In some conditions, the symbols of the two vehicles disappeared or were removed from the screen after a viewing time of four seconds, and the subject's task remained the same as when the vehicles remained visible throughout a trial. In these removal conditions, as soon as a key was pressed to release the pedestrian, the vehicles were visible again to allow for performance feedback.

The width of the gap between two vehicles that constituted a window was 85 mm. All the vehicles were moving at the same speed of 13.6 mm/sec. These variables resulted in a temporal window of opportunity of 480 ms. The pedestrian/vehicle speed ratio was four. Each condition of the crossing task contained one block with 32 windows of opportunity (i.e., replicates).

When the concurrent mental arithmetic task—a summation task presented either visually or auditorily—was added, the participant would need to perform the two tasks at the same time. For the summation task, the participant needed to press a key whenever two consecutive digits (0-9), presented at a rate of two digits/sec, would add up to 9. For both the visual and the auditory summation tasks, 30% of all the possible summations constituted correct answers (i.e., adding up to 9), and the participant was required to achieve 90% or above of correct performance for the summation task. In other words, while they were told to perform both tasks to their best ability, they should treat the summation task as the primary task and the crossing task the secondary one. The visual arithmetic stimulus (a digit between 0 and 9 presented one at a time) was presented 10 mm above the road of the

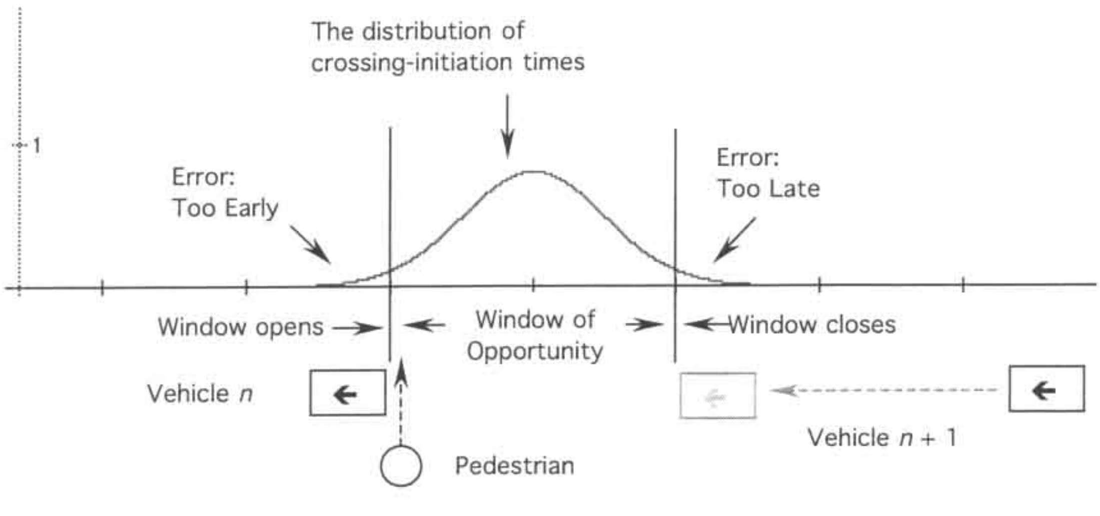

Figure 1. A schematic diagram of the experimental task and the relation of timing data and subjects' performance.

vehicles and 30 mm right above the pedestrian, and the auditory stimuli were presented from two speakers adjacent to the computer display.

Independent Variables

The independent variables were loading task condition, absent, or single task, auditorily presented stimuli, and visually presented stimuli, and the viewing condition, with vehicle symbols present throughout the trial or with the symbols removed after being visible for four seconds.

Dependent Variables

For the crossing task, five dependent measures were analyzed. The first was the overall error rate (either too early or too late release, resulting in the "pedestrian" being hit by a vehicle). The second was the percent release time relative to the window duration, defined by 100 × (release time) / (window duration), where the release time was measured from the "opening" of the window of opportunity (a negative value would indicate a too early release and a release time exceeding the duration of the window a too late release). The standard deviations of the percent release time distribution was also used as a dependent variable, as small standard deviation would indicates consistent and usually better performance. For the mental arithmetic task, percentage of correct responses (i.e., hit rate) false alarm rate, and response time were measured.

Design

A within-subjects 2 × 3 factorial design was used, with two types of viewing conditions (vehicle symbols present vs. removed) and three concurrent task conditions (single task and visual and auditory loading task). The order in which participants completed the experimental conditions was randomized.

Procedure

Participants individually completed a practice session to familiarize themselves with the displays and task. The practice session contained both the crossing and the arithmetic tasks and included both the single and the dual-task conditions. The practice session was followed by one experimental session and participants were allowed to have short breaks between experimental conditions. The computer automatically recorded release times, whether the crossing was successful or not, and for the mental arithmetic task, percentage of correct responses, false alarm rate, and response time. The participants were given feedback regarding their successes and failures on both tasks. For the summation task, whenever it was a success, the participant would hear a beep from the speakers. The experimental session lasted approximately one hour and 20 min.

RESULTS

Because the performance on the arithmetic task was above 90% across all conditions where the task was performed, and there were no significant differences in performance across the conditions, the results presented here are only for the crossing task. In other words, subjects were instructed to place priority on the arithmetic task and the data suggest that performance on this task was protected. This allowed for the examination of the arithmetic task's interference with the crossing task. The analyses were conducted at several levels for the crossing task performance. The first level of analysis was on the overall error rate of crossing. The second level of analysis was performed on the proportion of too early and too late releases. The third level of analysis was on percent release time, and the final level was on the standard deviation of percent release time.

Error rate was analyzed by Chi-squared tests on proportions of errors between the experimental conditions. The overall error rate was significantly higher when the vehicles were removed than they were visible, for single task, $\chi^2 = 17.45, p < .001$, auditory loading task, $\chi^2 = 14.73, p < .001$, and visual loading task $\chi^2 = 5.86, p < .05$, conditions. The difference in error rate between loading task conditions was significant when the vehicles remained visible, $\chi^2 = 7.69, p < .05$, but not when they were removed, $\chi^2 = 1.7, p = .428$ (see Figure 2).

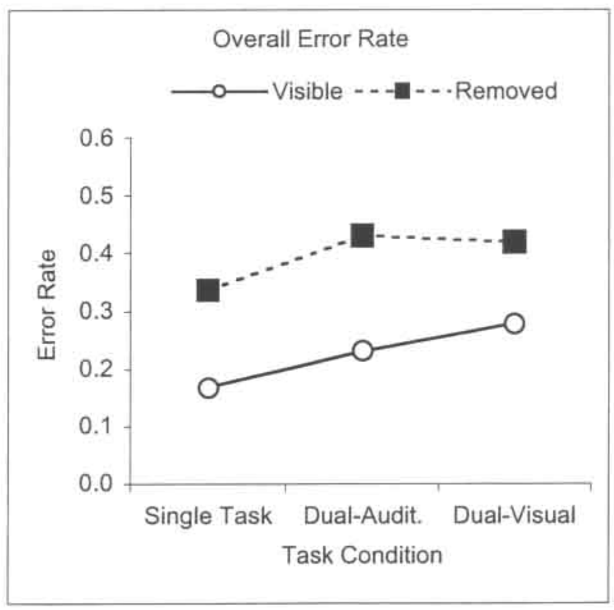

Figure 2. The effect of viewing condition on overall error rate was significant for all task conditions. The difference between task conditions was significant for visible stimuli.

The percent release time indicates how well the subjects were able to time the pedestrian release within the window of opportunity. Ideally this value should be close to 50%, minimizing the probability of any variation resulting in too early or too late release. There was a significant main effect of viewing condition for both the single- and the dual-task conditions, $F(1, 1793) = 35.06, p < .001$, clearly indicating that subjects released the pedestrian earlier when the vehicles were removed than when they were visible throughout a trial (Figure 3). There was also a main effect of task condition, $F(2, 1793) = 20.65, p < .001$, but only the differences between single task and visual loading task and auditory loading task and visual loading tasks were significant, $T = 5.71, p < .0001$ and $T = 5.66, p < .0001$, respectively (Tukey's test). That is, subjects released the pedestrian later when they were performing the two consecutive tasks than they were only performing one task of the crossing. Further, when performing the visual arithmetic task, subjects released the pedestrian later than the task was auditory in modality.

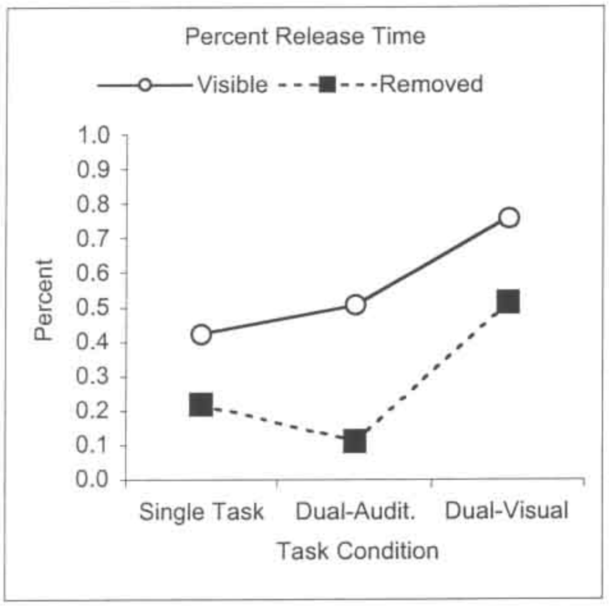

Figure 3. Percent release time. There was a significant effect of viewing condition for both the single- and dual-task condition..

We also analyzed the standard deviation of percent release time as another dependent variables (Figure 4). In addition to the error rate, this is probably the best indicator of subject performance, small SD indicating consistent and overall good performance and large SD depicting performance decrement (i.e., large variable error). There was a significant effect of viewing condition for both the single- and the dual-task conditions, $F(1, 43) = 11.57, p < .005$. When the vehicles were removed, subjects' release time was more variable than in the case of visible condition. There was also a main effect of task condition both for the visible and the removal conditions, $F(2, 43) = 7.88, p < .005$. When the subjects were performing the two tasks at the same time, the release time showed less consistency than when they were only performing the crossing task. The difference between the two loading task conditions was not significant (Tukey's test). Showing the same pattern as the overall error rate, the standard deviation in the dual-task condition for the visible viewing condition was greater when the arithmetic task was presented visually than when presented auditorily. This is consistent with the result that the performance decrement in the dual-task condition was greater for the visually presented arithmetic task than when presented auditorily. However, for the removal condition, there was no significant difference in performance decrement between the visual and the auditory modalities for the arithmetic task.

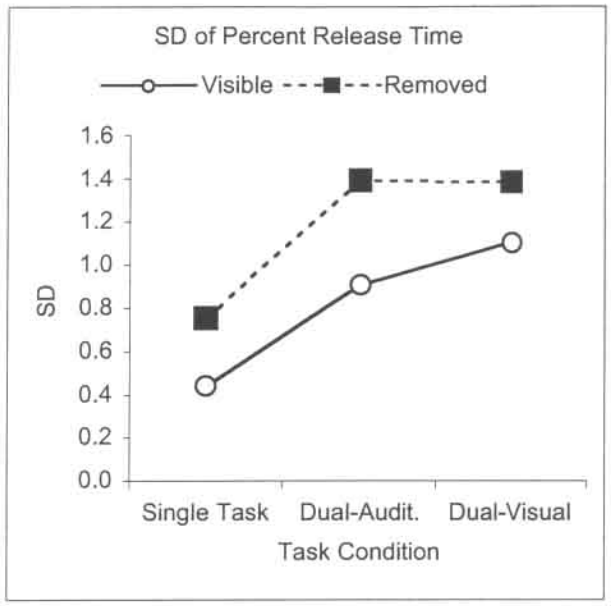

Figure 4. Standard deviation of percent release time. There was a significant effect of viewing condition both for the single- and the dual-task conditions. The effect of task condition was also significant.

DISCUSSION

The central question for this study was concerned with the perceptual and cognitive mechanism involved in the simulated task of a pedestrian crossing a busy street. In Rantanen and Xu (2001), we found the dominance of the (visual) perception over the temporal perception in releasing the pedestrian. The present study allowed us to further explore the perceptual and cognitive process underlying the task. The analyses of the data lead to some conclusions with respect to the hypotheses. When the stimuli were visible throughout a trial, there were two alternative strategies with which the subject could perform the task: pure timing (or clock counting) and visual perceptual. Had the subjects employed the counting method, they could have used the speed information obtained prior to the removal of the vehicles, along with the distance information, which was available throughout the whole trial, to perform the task. The removal of the stimuli would have not made a difference, relative to when they were visible all the time. The fact that performance suffered when the stimuli were removed strongly suggests that the task of releasing the pedestrian within a relatively narrow window of opportunity largely relied on perceptual cues available in the task or at least the counting strategy was not preferred when the stimuli were directly visible. This notion was supported by the greater performance decrement (when the vehicles were visible) regarding the overall error rate with the visual arithmetic task than with the auditory one, suggesting that the former was competing with the crossing task for the visual resource. The data regarding the standard deviation also support this notion of visual dominance in that it was greater in the removal than in the visible condition, and in the dual task condition, it was greater for the visual arithmetic task than for the auditory one. These results are consistent with the visual dominance found in Rantanen and Xu (2001). It is interesting to note that in the removal condition, subjects released the pedestrian earlier, indicative of underestimation of the time or an overestimation of the vehicle's speed, for a reason unknown to us.

The finding of visual dominance in the visible condition does not reveal what strategy was used when the stimuli were removed from the screen, however. It is conceivable that subjects could rely on the CME method (DeLucia & Liddell, 1998; Lyon & Waag, 1995) or used the counting method. If the subjects had relied on the CME using visual imagery in the removal condition, then the performance decrement during the dual-task paradigm would have shown the same pattern as in the visible condition; that is, the performance would have dropped more with the visual arithmetic than the auditory task. But this did not happen to the overall error rate (see Figure 2) or the standard deviation (see Figure 4), suggesting that in the removal condition, the subjects might have relied on the counting method. Indeed, post-experiment interviews with the subjects revealed that the majority of the subjects reported using this method. It seems that this was the preferred method when the stimuli were not present. Clearly, this is not consistent with the results found by DeLucia and Liddell (1998), and Lyon and Waag (1995). Thus, given the discrepancy, it is important to examine under what circumstances people use the counting method and the CME method differentially.

The paradigm used in the present study and our previous study may have important implications for other areas of research. Indeed, predicting the future status of people and moving objects is a common issue in many aspects of our lives and vital to safety. Therefore, our future studies will aim for the scale-up to those more applied areas such as driving and aviation. Meanwhile, it is also possible for the paradigm to scale down to the more basic research for theoretical contributions.

ACKNOWLEDGEMENTS

This research was supported by a grant form the University of Illinois Research Board. We wish to express our many thanks to Sharon Yeakel for the development of the experimental software.

REFERENCES

DeLucia, P. R., & Liddell, G. W. (1998). Cognitive motion extrapolation and cognitive clocking in prediction motion tasks. *Journal of Experimental Psychology: Human Perception and performance, 24,* 901-914.

Lyon, D. R., & Waag, W. L. (1995). Time course of visual extrapolation accuracy. *Acta Psychologica, 89,* 239-260.

Rantanen, E. M., & Xu, X. (2001). Human performance in timing of discrete actions. *Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 527—531). Santa Monica, CA: HFES.

Tresilian, J. R. (1995). Perceptual and cognitive processes in time-to-contact estimation: Analysis of prediction-motion and relative judgment tasks. *Perception & Psychophysics, 57*(2), 231-245.