Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society.
Santa Monica, CA: Human Factors & Ergonomics Society. 2001.

HUMAN PERFORMANCE IN TIMING OF DISCRETE ACTIONS

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

An experiment on human timing performance is described. A scenario of street crossing without the aid of traffic lights was presented on a computer screen. The subjects' task was to let a pedestrian to cross the street between successive vehicles by pressing a key. Independent variables were the size of the gap between vehicles and the speed of the traffic stream. The subjects' performance was evaluated by the distributions of the crossing initiation times and the proportions of two types of timing errors in the experimental conditions: Too early and too late releases of the pedestrian. Increasing the accuracy requirement by reducing the duration of the window of opportunity resulted in a shift to earlier release times, smaller standard deviations of the release time distributions, and increased proportions of the "too early" errors. Our results show a way to include also the tails of the response distributions in the analyses.

INTRODUCTION

This paper outlines an experimental paradigm designed to study human performance in timing of discrete actions. It is argued that this type of performance has become increasingly important with the progression of automation, automated systems predominantly performing direct control of many processes. Human operator is nevertheless usually tasked with general supervision of the processes and either routine or emergency interventions to support automation. The number of human actions associated with the control of systems commonly form a relatively limited and well-defined set. Consequently, such actions are regularly time-critical and, once initiated, often irreversible. Correct timing of actions is therefore imperative.

A number of relatively recent studies have recognized the role of timing in control of complex systems and ergonomics research (Carmichael, 1997; Decortis, De Keyser, Cacciabue, & Volta, 1991; De Keyser, 1990, 1995; Sougné, Nyssen, & De Keyser, 1993;). Many human errors are also temporal in nature. DeKeyser (1995) classified temporal errors as incorrect estimate of sequence of actions, incorrect estimate of durations, failure in the evaluation of the right time to act, failure in anticipation of an event, and failure in synchronization of collective actions. In addition to the importance of temporal performance of operators in many safety-critical systems, time may offer a useful domain for research of a multitude of human factors aspects. Time may offer attractive methods for the measurement of covert mental models. Time has a long history as a means to investigate cognitive processes, manifested by extensive reaction time research. Timing data (e.g., reaction times) are relatively easy to obtain under both experimental and naturalistic conditions, and time is a variable that is common to the human, the task, and the environment. Time offers thus a common unit of measurement of human performance in the context of the task, and can be used to infer the goodness of the temporal dimension of the operator's mental model of the task or system being controlled.

A second objective of this study was to develop methods of measurement that would expand the analyses beyond those of the mean (e.g., t-tests, ANOVA). The events important to human factors research typically involve the tails of the response distributions rather than the average responses (Wickens, 2001). The tails often represent the exceedances of tolerances or criteria (e.g., too fast, too slow, too early, too late, etc.) and are rightly classified as errors (Hollnagel, 1998). Because of the inherent difficulties in the study of typically very rare errors (Wickens, 2001), we—instead of focusing solely on the means of responses—derived the dependent variables in our experiment from the response distributions in an attempt to capture also the important and interesting "tail ends" of human performance.

METHOD

This experiment emanates from the desire to extend the existing experimental paradigms of timing (e.g., finger tapping and reaction time experiments) to more complex and realistic settings and thus enhance the applicability of results to naturalistic tasks and problems. Because the experiment was designed to be conducted in a laboratory using naive subjects, the task had to be sufficiently abstract to minimize skill and training requirements. At the same time it should retain the demands and characteristics that are necessary to bring about the desired effects in the subjects' performance.

The experimental apparatus and task.

A suitable task involved the familiar situation of crossing a busy street without the aid of traffic lights, presented on a computer screen. Symbols for vehicles traversed across the screen in a continuous stream and the subjects' task was to let a pedestrian, waiting on the roadside, to cross the street between successive vehicles by pressing a key. Successful completion of the task required correct judgment of a sufficient interval between successive vehicles (a temporal window of opportunity) to allow the pedestrian safely cross the road. This required taking into account the time required for crossing and correct timing of the initiation of the crossing. Since the pedestrian was waiting 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).

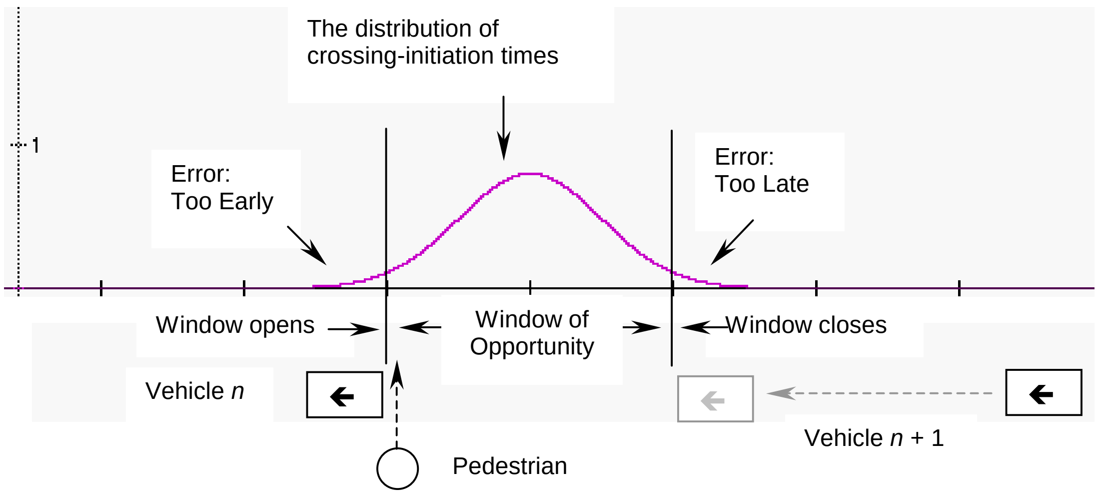

Figure1: A schematic diagram of the experimental task and the relation of timing data and subjects' performance.

Participants.

Eight subjects (three females and five males) volunteered to participate in the experiment. The subjects were members of the Perception and Performance group at the University of Illinois at Urbana-Champaign. All subjects had normal or corrected to normal vision and normal hearing.

Independent variables.

A theoretical minimum gap between vehicles (i.e., a gap within which it was possible to cross the road with perfect timing) was established as a baseline, and insufficient gaps were created by subtracting a number of pixels from this value. Sufficient gaps, or windows of opportunity, were created by adding to the baseline gap. This procedure also allowed for control of the conspicuity of windows of opportunity in the traffic stream; in this experiment these were easily identified. The independent variables were the size of a window of opportunity and the speed of the traffic stream. Two levels of each were utilized. A small window corresponded to a gap of 23.1 mm and the large window to a gap of 29.7 mm between vehicles. The slow speed was 31.6 mm/s and the fast speed 63.2 mm/s. All vehicles moved at the same speed. The pedestrian/vehicle speed ratio was four. These variables resulted in three temporal window of opportunity sizes: 110 ms (small gap, high speed), 210 ms (large gap, high speed and small gap, slow speed), and 420 ms (large gap, slow speed).

Dependent variables.

The primary dependent variable was the crossing initiation time, measured from the "opening" of the window of opportunity. Hence, negative release time would indicate a too early release, and a release time that exceeded the duration of the window too late release. Several other dependent variables were derived from these data, however. In addition to the raw release times, percent release times relative to the window duration were computed, as well as the standard deviations, skewness, and kurtosis values of the release time distributions. The proportions of too early and too late releases were also computed.

Design.

This was a within-subjects 2×2 factorial design, with two levels of window size and two levels of traffic stream speed. The experiment was run in blocks of 96 vehicles in a continuous stream. Each block included 95 gaps between the vehicles, of which 19 were windows of opportunity, dispersed randomly in the traffic stream. Five replicates of blocks in each condition were run in a random order, resulting in a total of 20 blocks.

Procedure.

Participants completed several practice session to familiarize themselves with the task. The data collection session, in which each participant completed the 20 experimental blocks, followed the practice session. The computer automatically recorded sufficient gaps, release times, and whether the crossing was successful or not. The subjects were given feedback on their successes and failures. The data collection session lasted approximately one hour.

RESULTS

The analyses were conducted at three levels. The first level of analysis was on the crossing initiation times (the release time and percent release time relative to the window duration). The second level of analysis was performed on the parameters of the timing distributions (standard deviation, skewness, and kurtosis) for the release time and percent release time. The final level of analysis was on the proportions of errors (i.e., too early and too late releases).

At the first level, there was a significant main effect of speed, $F(1, 145) = 67.63$, $p < .001$, indicating that subjects responded to increased accuracy requirement (increased speed and hence shorter window duration) by releasing the pedestrian earlier. The main effect of speed on the percent release time was also significant, $F(1,145) = 7.32$, $p < .01$, indicating relatively earlier release times as well. This suggests that the subjects did not only release the pedestrian earlier in terms of absolute release times in response to the faster speed, but also did so proportionally relative to the window duration. Similar results were found for the window size: For a small window, the subjects released the pedestrian earlier at both speeds, and the effect of window size was significant, $F(1,145) = 11.61$, $p < .005$ (see Figure 2).

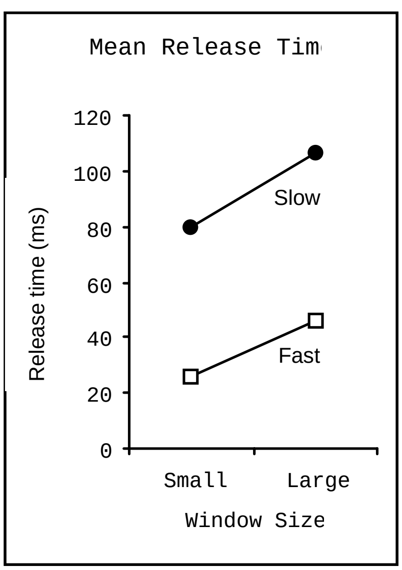

Figure 2. Mean pedestrian release time. The effect of both traffic speed and window size were highly significant.

The effect of window size on the percent release time was also significant, $F(1,145) = 4.9$, $p < .05$, but opposite, that is, the subjects released the pedestrian relatively earlier for larger windows. As was mentioned before, two of the four experimental conditions resulted in the same duration of the window of opportunity. Yet, the subjects performance in these two conditions was markedly different and, analyzed separately, the difference was significant, $F(1,78) = 6.31$, $p < .05$.

An ANOVA on the raw release time standard deviations showed that both the main effects of speed and window size were significant, $F(1, 145) = 4.68$, $p < .05$, and $F(1, 145) = 41.96$, $p < .001$, respectively. Increasing the speed of the vehicles and reducing the window size both resulted in reduced standard deviation of the timing distribution, which is an appropriate response to the increased accuracy demands (Figure 3). However, effects of window size and speed on the standard deviations of percent release times were significant, $F(1, 145) = 4.68$, $p < .05$, and $F(1, 145) = 41.96$, $p < .001$, respectively, but opposite to the raw release times. Hence, the subjects became less variable in their timing performance as the window size increased or the speed decreased relative to the temporal window duration. Analyzed separately, the standard deviation of release times were significantly higher in the small window/slow speed (mean 152, SD 38) than in the large window/fast speed condition (mean 126, SD 49), $F(1, 78) = 6.96$, $p < .01$, despite the fact that the temporal durations of the windows were the same.

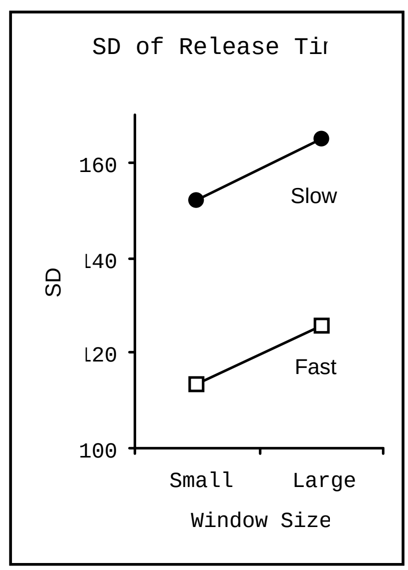

Figure 3. Standard deviation of pedestrian release times. Fast vehicle speeds and small window size reduced the standard deviation of the timing distributions, an appropriate response to increased accuracy demands.

Skewness values of the timing distributions were generally very small (mean –0.063, SD 0.69) and not affected by the experimental conditions. However, the kurtosis values showed some departures from normality (mean –0.113, SD 1.36). The release time distributions associated with small window were also significantly more leptokurtic (i.e., thicker-tailed) than those resulting from large windows, $F(1, 145) = 4.46$, $p < .05$.

Probabilities of two types of error—too early or too late release of the pedestrian—were derived by calculating the area under the tail of the release time distribution, to the left of the opening of the window or to the right of the closing of the window, respectively. The proportion of too early errors increased as the speed of the vehicles increased, $F(1, 145) = 30.13$, $p < .001$, and when the window size decreased, $F(1, 145) = 11.23$, $p < .005$ (Figure 4.)

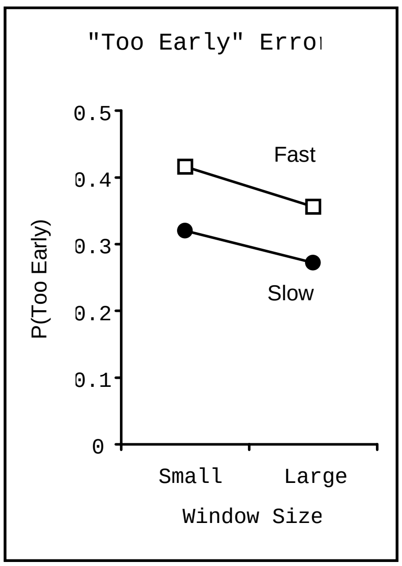

Figure 4. Fast vehicle speeds and small window size increased the probability of too early release times.

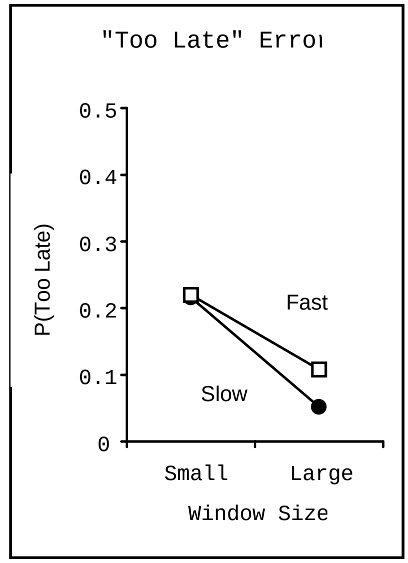

Figure 5. The probability of too late release times was increased by fast vehicle speeds and small window size as well

Small window size resulted in significantly increased proportion of too late errors as well, $F(1, 145) = 110.23$, $p < .001$, as did the fast speed of the vehicles, $F(1, 145) = 5.2$, $p < .05$. The interaction of vehicle speed and window size was also significant, $F(1, 145) = 4.69$, $p < .05$ (Figure 5).

Subject was included as a factor in the ANOVA model, and it was significant in all analyses. However, despite the apparently substantial individual differences between the subjects in the experimental task, the trends described above were obvious.

DISCUSSION

The subjects adjusted their release times appropriately to the increasing accuracy requirements by becoming more consistent in their timing, as indicated by decreasing standard deviations. There was also a clear shift to earlier release times, both in terms of absolute times and times proportional to the window duration. This shift resulted in significant increase in "too early" errors. The shift to earlier release times and the reduction of standard deviations, however, were not sufficient to match the increased accuracy demands of smaller window durations, evident from the increase in too late errors associated with fast speed and small windows.

It was particularly interesting to note that although the large window/fast speed and small window/slow speed conditions resulted in identical durations of the temporal windows of opportunity, the results from these conditions were significantly different. It is apparent that the perceived accuracy demand caused by the faster speed of the traffic stream was responsible for the shift in the subjects' timing performance, not the actual temporal size of the window. This suggests that the visual elements of the task (i.e., the speed of the vehicles and the physical size of the window) dominated the perception of the time available in each experimental condition. Further research is needed to investigate the locus of attention in this type of timing task in detail.

This research also allowed for examination of some existing models of timing performance. Peterken, Brown, and Bowman (1991) explored two alternative explanations for the speed effect on time performance. One was that speed per se was the factor contributing to prediction accuracy. If this were the case, then there should be an effect of speed on both time estimation error (release time in our case) and time estimation error proportion (percent release time in our experiment). Alternatively, faster speed translated into shorter viewing time before releasing the pedestrian and the affected performance (earlier release time and more “too early" errors in our case) might have resulted from the reduced viewing time of the window of opportunity in the traffic stream. If the second scenario was the case, then analyzing the time proportion would eliminate the time factor and significant effect would have been found only for time error in absolute terms (raw release times here). Peterken, Brown, and Bowman’s (1991) results support the second explanation, whereas ours lend support to the first explanation. That is, we found the effect of speed to be significant both for raw release time and percent release time, thus suggesting that speed per se did seem to be the factor influencing the performance. This is consistent with the finding that although the large window-fast speed and small window-slow speed conditions resulted in identical durations of the temporal windows of opportunity, the results from these conditions were quite different. That is, the viewing time of the window was not an important variable.

In addition to analyzing just the means of the release times by ANOVA, this investigation represents an attempt to study the distributions of human performance measures by analyzing the dependent variables derived from them. Our results suggest that this is a valuable method of examining human performance, illustrated by the effects of the traffic speed and the window size on the standard deviations of the timing distributions. By calculating the error proportions we were also able to examine the "tail ends" of these distributions (the “too early “ and “too late” release errors in this study) and thus extrapolate the impact of variability in subjects' performance to practically significant probabilities of the two types of timing errors. These analyses could potentially reveal the most important aspects of human performance in many safety-related domains, such as surface transportation and aviation, that is, accident probabilities (Wickens, 2001).

Much work remains to be done, however. The error probabilities were based on normal distributions and the timing distributions in our study were substantially leptokurtic, or thick-tailed. Hence, the error probabilities were likely underestimated. Improved methods for accurate representation of the distributions of performance measures must be developed. This study also manipulated a very small subset of possible factors affecting human performance (the size of the window of opportunity and preview time). Nevertheless, the study was valuable as a proof-of-concept of novel performance measurement methodologies, and offers a solid foundation for future work. Our next experiments will concentrate on examination of the impact of secondary tasks on human timing performance in light of previous work by Brown (1997) and Zakay, Block, and Tsal (1999).

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

Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Perception and Psychophysics, 59*(7), 1119-1140.

Carmichael, A. R. (1997). Duration estimates: a potentially useful tool for cognitive ergonomists. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics, Vol. 2: Job design and product design* (pp. 267-273). Brookfield, VT: Ashgate.

Decortis, F., De Keyser, V., Cacciabue, P. C., & Volta, G. (1991) The temporal dimension of man-machine interaction. In R. S. Weir and J. L. Alty (Eds.), *Human-Computer Interaction and Complex Systems* (pp. 51-72). San Diego, CA: Academic Press.

De Keyser, V. (1990). Temporal decision making in complex environments. *Philosophical Transactions of the Royal Society of London, B 327,* 569-576.

De Keyser, V. (1995). Time in ergonomics research. *Ergonomics, 38*(8), 1639-1660.

Hollnagel, E. (1998). *Cognitive reliability and error analysis method: CREAM.*. Oxford: Elsevier Science Ltd..

Peterken, C., Brown, B., & Bowman, K. (1991). Predicting the future position of a moving target. *Perception, 20,* 5-16.

Sougné, J., Nyssen, A-S., & DeKeyser, V. (1993). Temporal reasoning and reasoning theories, a case study in anaesthesiology. *Psychologie Belgique, 33-2,* 311-328.

Wickens, C. D. (2001). Attention to safety and the psychology of surprise. *Keynote address presented at the 11th International Symposium of Aviation Psychology*. Columbus, OH: Ohio State University.

Zakay, D., Block, R. A., & Tsal, Y. (1999). Prospective duration estimation and performance. In D. Gopher and A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 557-580). Cambridge, MA: The MIT Press.

Paper submitted for the Human Factors and Ergonomics Society 45th annual meeting in Minneapolis/St Paul, MN, October 8-12, 2001.