System response delay and user strategy selection

Alexander I. Rudnicky

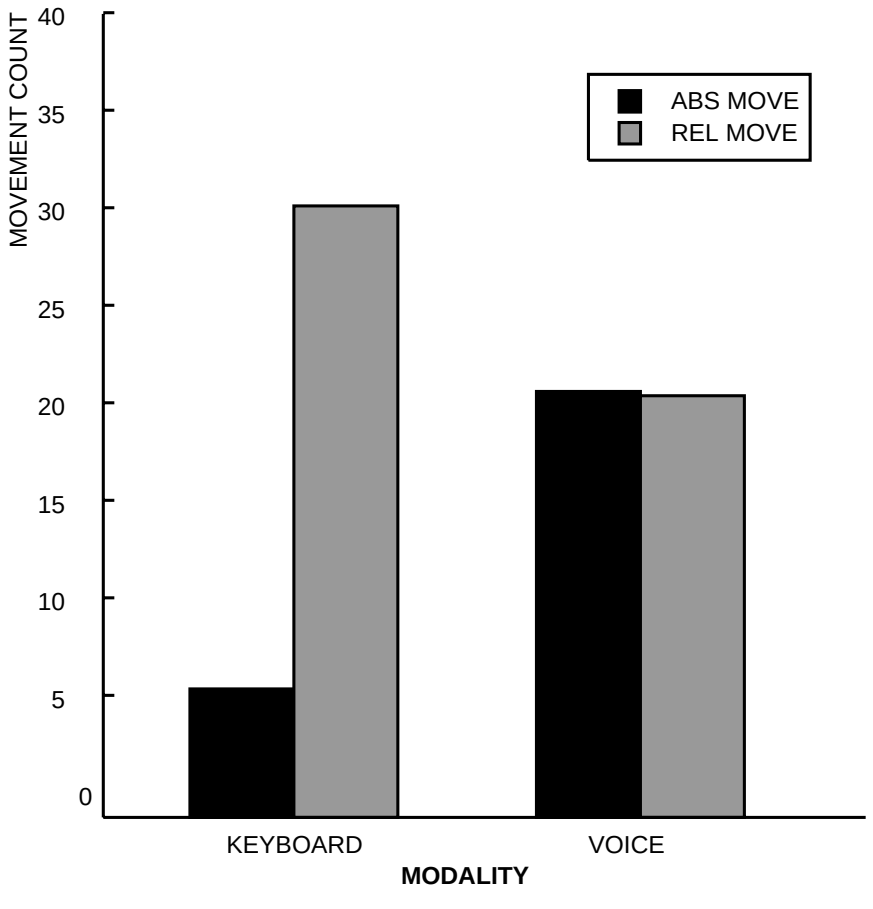
Carnegie Mellon University
Pittsburgh, Pennsylvania
USA 15213

Invited Poster
Chi'90
Conference on Human Factors in Computing Systems
2 April 1990

The Problem

When given a choice between different methods that achieve a given goal, on what basis do people make a choice? Does input modality (voice, keyboard) play a role?

Subjects in an experiment comparing voice and keyboard interfaces to a spreadsheet program [Rudnicky *et al.*, 1989] used different strategies for moving around. In the keyboard mode, subjects preferred to use arrow keys (*relative* movement). In the voice mode, subjects preferred to use ‘‘go to’’ commands (*absolute* movement).

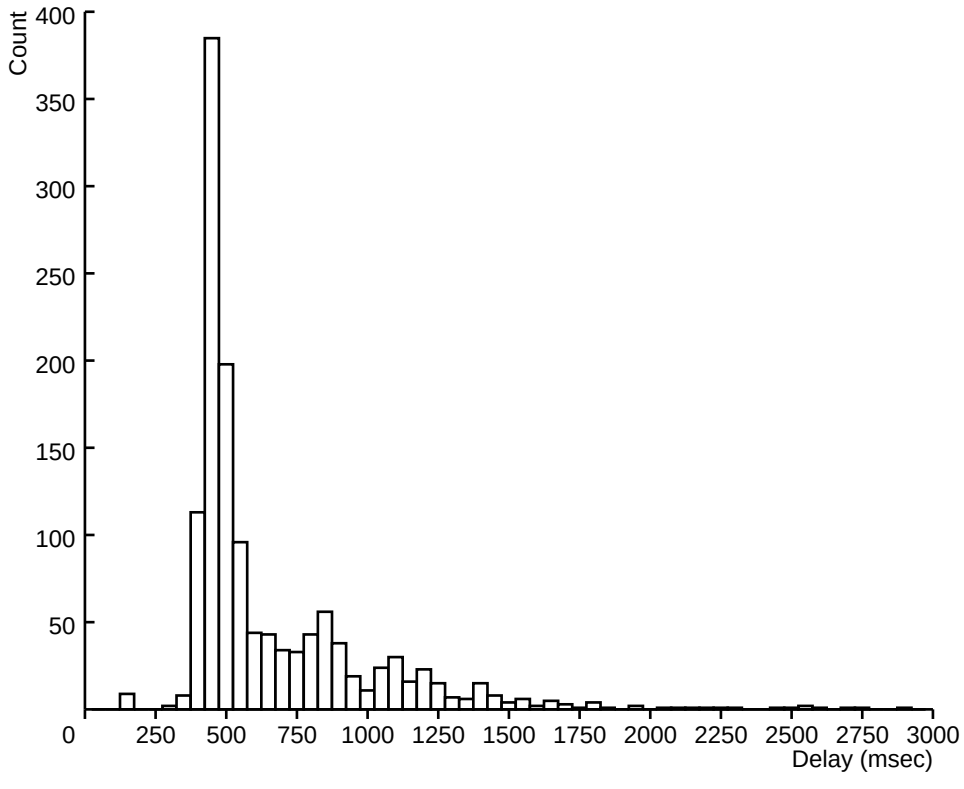


It might be argued that speech encourages more natural expression, by allowing the use of commands that closely reflect goals (e.g., ‘‘*GO TO SALARY*’’), in contrast to commands that specify component acts (e.g., ‘‘*GO DOWN TWELVE*’’).

Alternate interpretations

Before such a conclusion is warranted, we need to take a closer look at the circumstances of the Rudnicky *et al.* (1989) study.

A prominent feature of the voice interaction cycle in that study was a variable *response delay*. This delay was due to properties of the recognition system, specifically its lack of a ‘‘true’’ real-time response.



Questions

- Is it speech or delay that causes people to shift their strategy?

- Would an equivalent fixed-duration delay have the same effect as the random-duration delay?

Experiment

Users performed the same spreadsheet tasks as in the Rudnicky *et al.* (1989) study. This consisted of entering 40 items of financial information into a pre-programmed spreadsheet. Five tasks, blocked, were performed under each condition, by each subject. A total of 9 users participated.

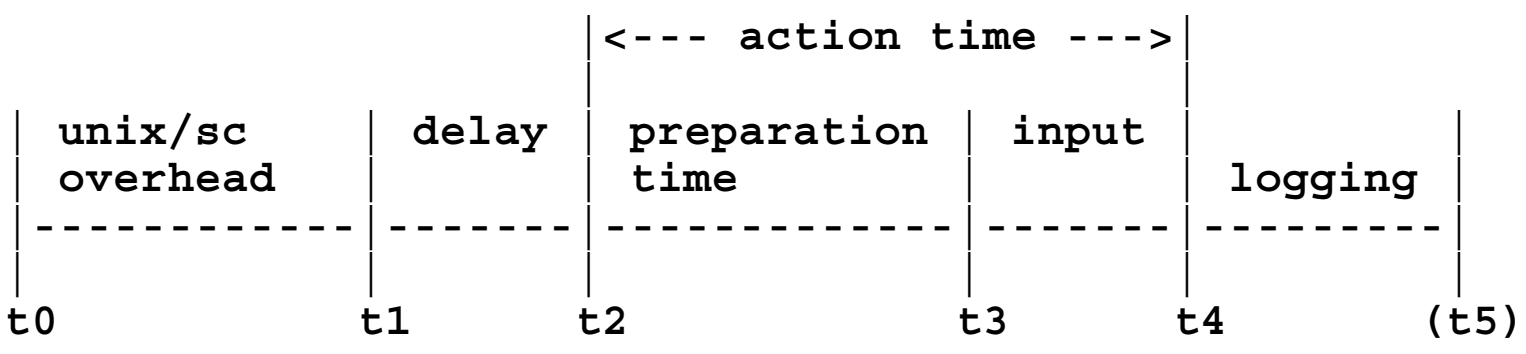
Keyboard input was used exclusively in the current experiment. There were three different delay conditions:

- **No Delay:** Unmodified spreadsheet interaction cycle.

- **Random Delay:** Delays sampled from the same distribution as produced by the speech system.

- **Fixed Delay:** Delay set at 820 msec (at the 75th percentile of the Random distribution).

The delay was made part of the interaction cycle:



Result

The distribution of movement action types for the delay conditions matches that observed for voice interaction.

It can therefore be inferred that *strategy shift is due to the presence of delay rather than to any special properties of the speech modality*.

Analysis

The mean total time to complete a single task (in seconds):

Condition	Raw Time	Action Time
No delay	479.9	468.5
random	573.9	481.8
fixed	578.3	469.2

- There are no significant differences between total actions times. Subjects spent about the same amount of time performing movements in each condition, although using very different strategy mixes.

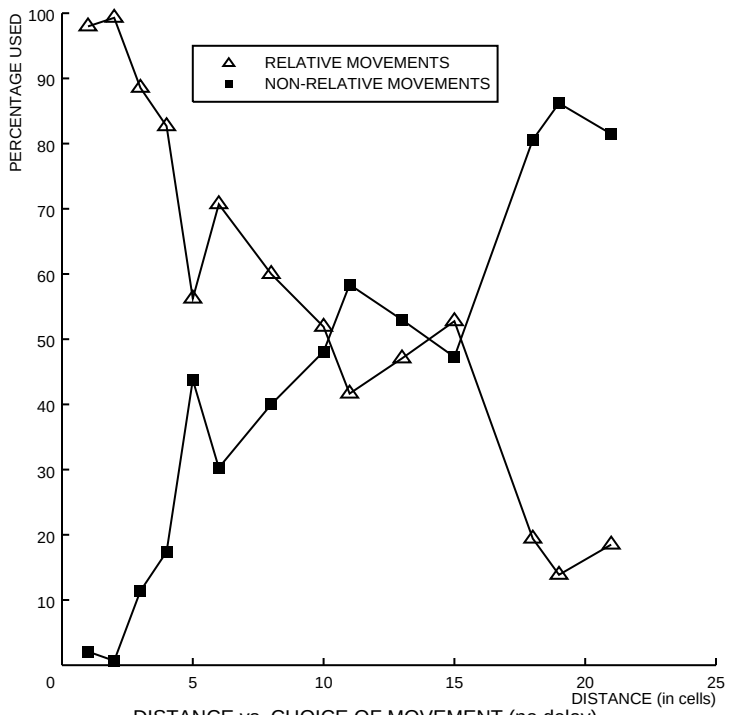
To better understand the user's strategy selection process, we can examine how choice of method varies with the distance that needs to be traversed.

The ***left column*** in the next panel shows the proportion of Relative and non-Relative moves made, as a function of distance. For the **No Delay** condition, subject prefer Relative moves for distances of up to about 10–15 cells. For the **Delay** conditions, subjects shift to non-Relative moves for distances greater than 2 cells.

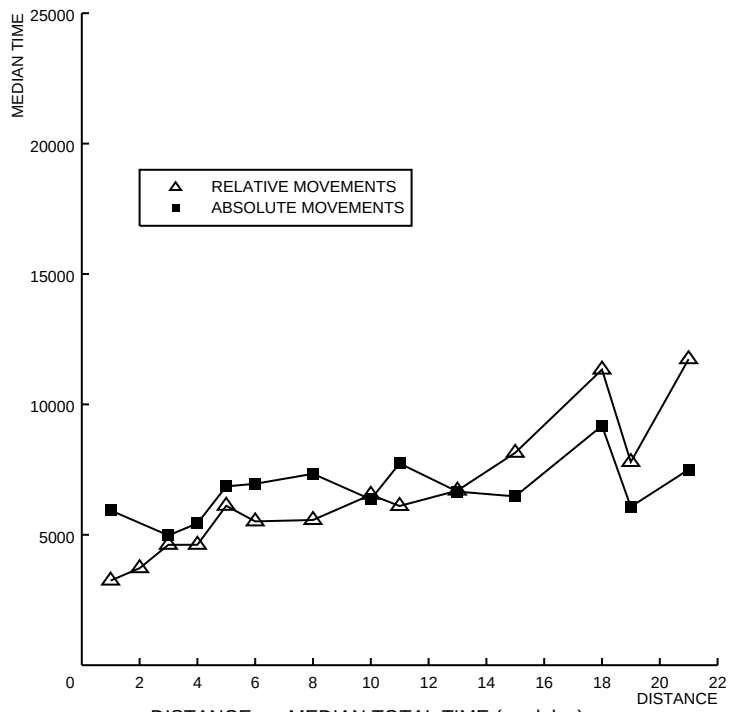
The ***right column*** of the next panel shows, for a given distance, the time it took to make a move using each method.

- Comparison of the two columns shows that subjects chose a strategy that allows them to complete the move *in the shortest time possible*, shifting methods at a point corresponding to the duration cross-over.

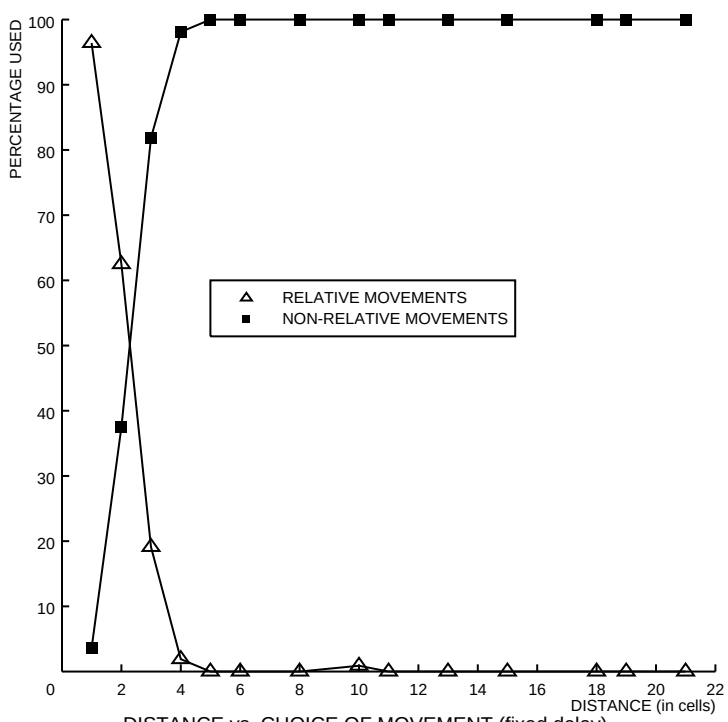
Choice, Action Time and Movement Distance



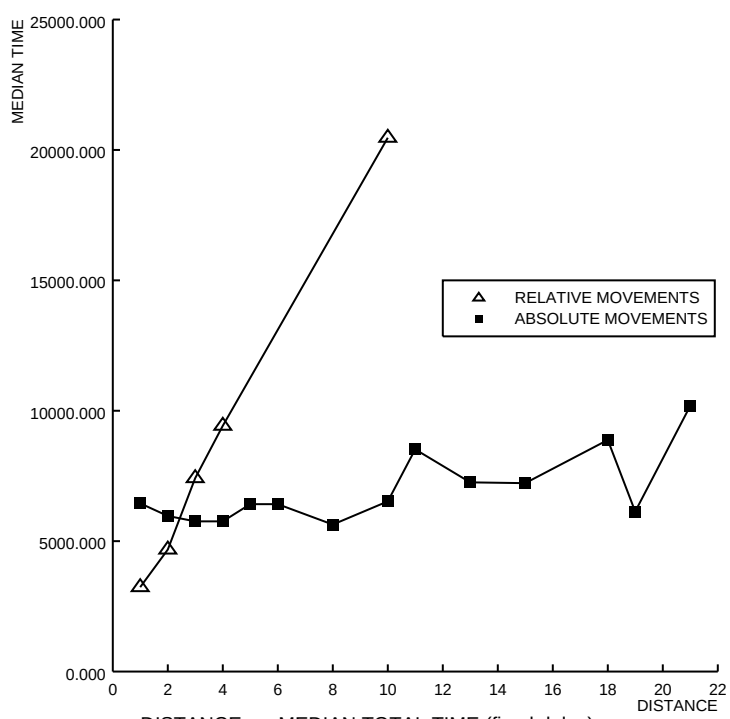
DISTANCE vs. CHOICE OF MOVEMENT (no delay)



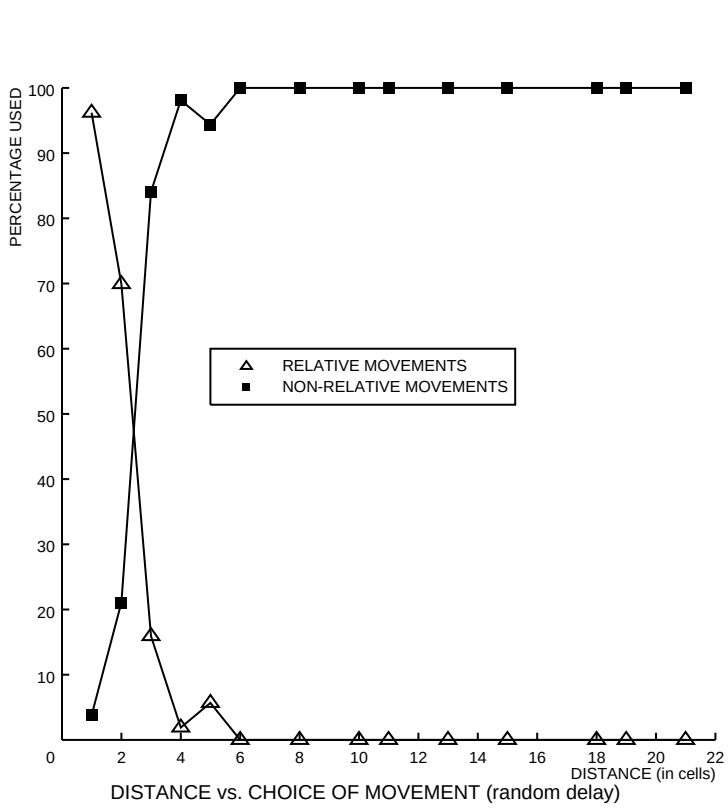
DISTANCE vs. MEDIAN TOTAL TIME (no delay)



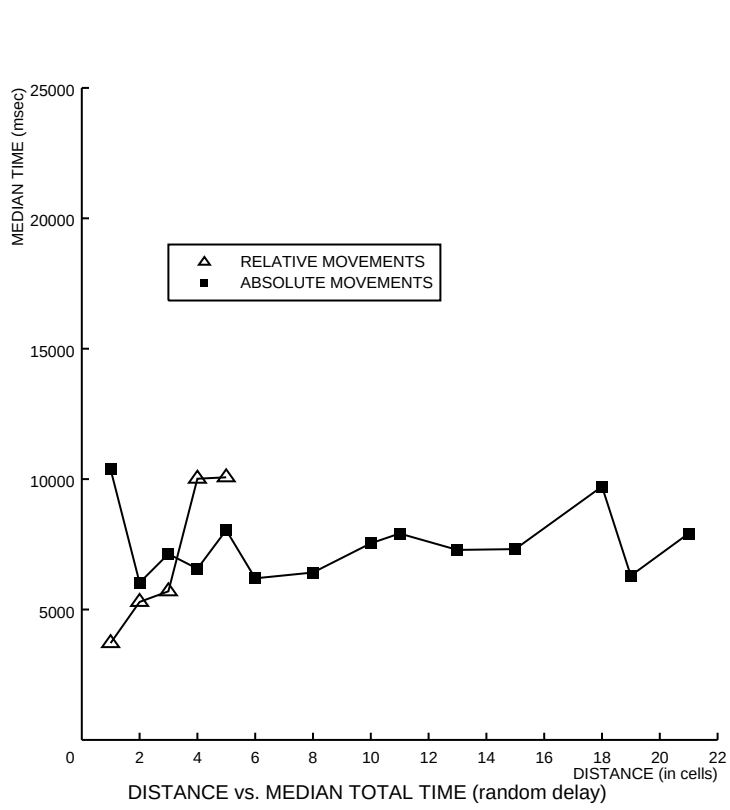
DISTANCE vs. CHOICE OF MOVEMENT (fixed delay)



DISTANCE vs. MEDIAN TOTAL TIME (fixed delay)



DISTANCE vs. CHOICE OF MOVEMENT (random delay)



DISTANCE vs. MEDIAN TOTAL TIME (random delay)

Conclusions

Users can accurately assess the cognitive and practical costs of alternative actions. Their choice of strategy will reflect this cost.

The current experiment shows:

- Subjects are able to accurately estimate the expected duration of an action, even for actions that require an extended keystroke sequence.

- Subjects use the expected total duration for an action to chose the most efficient strategy.

This result can be applied to speech interface design through the following observation:

- Speech systems that do not provide true real-time response carry a built-in penalty, since the resulting delay will distort the true cost of different actions.