Measuring computer use behavior: patterns of variability, within and across days and between different users

H.P. Slijper[1], S.E. Mathiassen[2] , E.A.B. Over[1], J. Richter[1], and M.A. Frens[1]

[1] *Department of Neuroscience, Erasmus MC, Rotterdam, The Netherlands, harm.slijper@gmail.com*

[2] *Centre for Musculoskeletal Research, University of Gävle, Sweden, svenderik.mathiassen@hig.se*

Introduction

Computer users are notoriously bad at estimating how much time they have spend using the computer. Levels of overestimation are large (up to 120%) and depend only slightly on the amount of time actually worked with the computer [1]. Thus, in order to arrive at an objective and precise description of computer use behavior we developed registration software which unobtrusively recorded input device use in a large group (N=500) of heavy computer users. The computer users were followed up to several years to characterize the **temporal** (pause behavior) and **spatial** (mouse use) patterns that governed their computer use behavior. With regard to the temporal pattern, we described the episodic nature of work and non-work episodes both *within* days and *across* days. We asked the following questions: 1. How do work/ pause patterns change when the computer is more intensely used? 2. Can the characteristics of a working day be predicted on the basis of recordings from previous days? 3. How well can the average computer use (across a year) be predicted from recorded computer use during only a few days (this determines ergonomist's measurement strategies)? 4. To what extend is computer usage subject specific? With regard to this last question we also looked at the spatial pattern of mouse movements to see whether the characteristics of mouse movements varied systematically between users.

Methods

The recorded time traces of computer use (N>72.000) included timestamps of keyboard strokes and cursor changes (10 Hz). In such a time line of events, which have no duration themselves, "computer-work" and "non-computer work" episodes need to be discriminated. To this end, we implemented a temporal criterion (non computer threshold, NCT) that specifies the amount of time two subsequent events could be separated in time, while the time in between would still be classified as continuous work [2]. On the basis of this NCT, the duration of the working day, the summed duration of the work and non-work episodes and the number of work episodes, were used to describe the pattern of computer use within a work day. By varying the NCT value we investigated how the values of these variables systematically changed across different time scales (i.e. a small NCT will identify a lot of short duration episodes, while a large NCT will identify few long duration episodes). The shape of the relationship between NCT will thus be able to detect whether there is regularity in the usage behavior. To analyze whether the characteristics of a workday could be predicted based on previously recorded days, we calculated day-to-day autocorrelation functions for the above variables. To calculate the error in estimating "the one-year-average duration of computer use" we used empirical re-sampling of data. That is, by randomly drawing subsets (certain number of sample days) of data from all available data from a subject, and repeating this procedure thousands of times (bootstrap-like procedure), we could calculate how reliable the mean work duration across a work year could be estimated. For all variables, we performed a variance component analysis to see how much variability could be explained by differences between subjects and how much by differences between work days. In a separate analysis, we studied the kinematics of mouse movements [4]. Using a velocity threshold, begin and endpoint of individual mouse movements were identified. Consequently, we focused on characterizing directional distributions, since there was a strong bias of users to make movements in particular directions. We investigated whether these directional distributions were invariant across days and computers (used by the same worker).

Results

Q1: We found that distributions of episode (durations) were highly skewed (al lot of small episodes and few large ones). These distributions depend on the chosen NCT according to a log-linear relationship [3] (figure 1). While the time classified as work increases with a mere 3.5% per doubling of the NCT, the number of computer work episodes decreased by 40%, the duration of computer work episodes increased by 90% and the duration of non-computer work episodes increased by 60%. This means that when the duration of a non-work episode is doubled it is 1.6 times less likely to occur. The slopes of these relationships remained invariant while the intercept changed when workers used the computer more intensely (for example see Figure 1).

Q2: Since the day-to-day autocorrelations (r = 0.1-0.2; non significant) were low across a large number of lags (days) and all variables, the characteristics of a working day could not be

Figure 2: Error in time working with the computer for different sample sizes

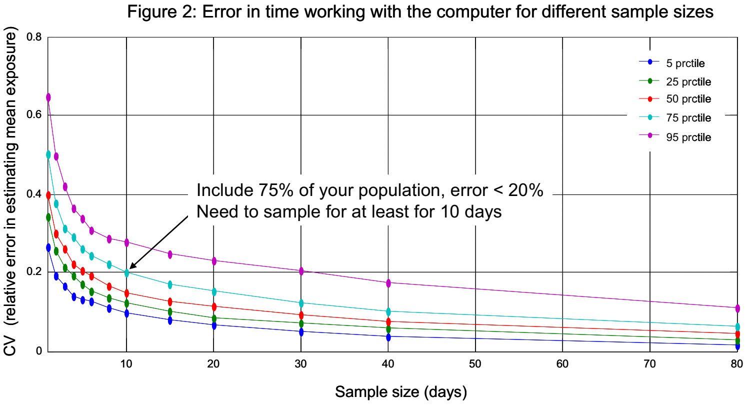


Figure 1: time classified as work depending on NCT

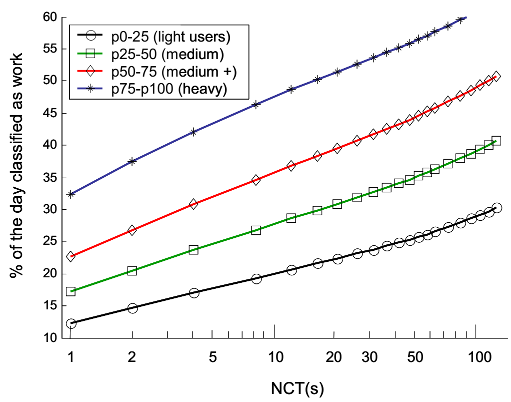
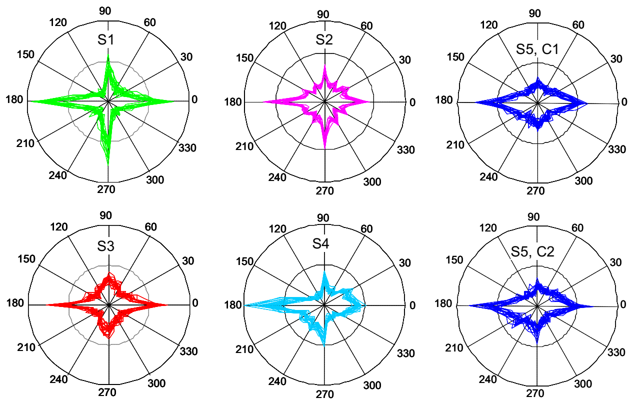


Figure 3: directional distributions in 5 users across 25 days



predicted precisely from the recordings of previous days.

Q3: In order to estimate the reliable assessment of mean computer use duration across a 1-year period, we calculated relative errors (CV) for different sample sizes, as shown in figure 2. This shows that considerable days need to be sampled in order to estimate 1-year average values reliably.

Q4: Variance component analysis showed that between 20-50% of the variance could be explained by differences between workers, depending on the variable. A striking example of a subject specific pattern was found in the directional distribution of mouse movements (Figure 3). Shown is data from 5 users (S1-5), across 25 workdays (different lines). User 5 worked on two computers (C1, C2). Note the preference for movements in cardinal directions, the invariance of the distribution across days and computers, and the idiosyncratic differences between users.

Conclusions

Although computer users are free to choose the onset and duration of their work episodes, the (spontaneous) pattern of work and non-work episodes is highly structured Day-to day variability in computer use is unpredictable which necessitates long recordings to estimate 1-year averages (unfortunate for ergonomists). Subject-specific patterns in the directional distribution of mouse use show that users have a 'mouse-signature'.

References

1. Faucett, J., Rempel, D., 1996. Musculoskeletal symptoms related to video display terminal use: an analysis of objective and subjective exposure estimates. Aaohn J, **44(1)**: 33-9.
2. Slijper, H. P., J. M. Richter, et al. The effects of pause software on the temporal characteristics of computer use. *Ergonomics*, 2007, **50(2)**: 178-91.
3. Richter JM, Slijper HP, Over EAB, Frens MA. Computer work duration and its dependence on the used pause definition. *Applied ergonomics*. 2008, **Jan 3** (online first).
4. Slijper, H.P., Richter, J.M., Over, E.A.B., Smeets, J.B.J., Frens, M.A (2008) Statistics predict Kinematics of Hand Movements during Everyday Activity, *Journal of Motor Behavior* (in press).