Electronic monitoring systems: an examination of physiological activity and task performance within a simulated keystroke security and electronic performance monitoring system

RON HENDERSON,† DOUG MAHAR,‡ ANTHONY SALIBA,§ FRANK DEANE,¶ RENEE NAPIER

The Trust Project, || RSISE, Australian National University, ACT 0200, Australia, e-mail: ron@science.canberra.edu.au

(Received 11 March 1996 and accepted in revised form 12 September 1997)

Electronic monitoring systems are becoming a prominent feature of the modern office. The aims of the present study were three-fold. First, to assess the effects electronic security monitoring systems (ESM) have on the user’s physiological state. Second, the research aimed to examine the effects explicit security challenges have on both user behaviour and physiological state when using an ESM system. Finally, the research aimed to examine the effects one form of electronic performance monitoring system may have on the user’s physiological state. To this effect, the present study examined the physiological and performance effects of two simulated electronic monitoring systems (security/performance). The computer task required 32 subjects to enter mock clinical case notes under various conditions. In the first session subjects were only required to enter the case notes while keystroke data were collected. The second session was divided into three discrete stages. In the “security baseline” condition, subjects were informed that a keystroke security monitoring system had been instituted, but no security challenges occurred. In the “security challenge” condition, however, a number of explicit security challenges occurred. In the final “performance monitoring” condition, subjects were informed that their data entry speed was monitored and they were placed on a response-cost schedule for poor performance. Blood pressure and continuous inter-heartbeat latency were recorded for the security and performance conditions. Results indicated that monitoring systems have the potential to evoke altered arousal states in the form of increased heart rate and blood pressure. Contrary to expectations, the hypothesized improvement in task performance within the performance monitoring condition was not observed. The implications of these results for the design and implementation of electronically based behavioural-based security and performance monitoring systems are discussed.

© 1998 Academic Press Limited

1. Introduction

Employee monitoring systems may take many forms, from manual supervisory monitoring of on-task behaviour and output to electronic monitoring systems recording data...
entry speed and caller response time, and keystroke biometric verifications systems intended to enhance computer security. Indirect monitoring of employee performance has been given the generic title “Electronic Performance Monitoring”, or EPM for short. The Office of Technology Assessment (OTA), of the US Congress (1987) has defined EPM as the use of computer technology to collect, store, analyse and report information about a worker’s activities. In 1987 it was estimated that approximately six million clerical office workers within the USA had their performance monitored with such systems (OTA, 1987) and by 1990 this figure had been reported at more than 10 million workers in the USA alone (Aiello & Kolb, 1995). Furthermore, due to potential productivity benefits it was also expected that these systems would spread rapidly throughout the world (Schleifer, 1992).

Until recently, EPM has mostly been used with clerical workers and others who perform simple, repetitive tasks. Trends indicate, however, that the technology will expand to cover professional, technical and managerial employees (Garson, 1988; Aiello & Kolb, 1995). Such trends indicate the rapidly expanding diversity of EPM-type systems.

1.1. ELECTRONIC SECURITY MONITORING (ESM) SYSTEMS

One monitoring system not widely examined is the biometric security system. Biometric systems are automated methods of verifying or recognizing the identity of a person based on some physiological characteristic, like a fingerprint or some aspect of behaviour such as keystroke patterns (Miller, 1994). Techniques incorporated within the rubric of biometric systems include: finger prints, hand geometry, keystroke dynamics, retina patterns, signature dynamics, voice verification and face recognition (Sherman, 1992; Miller, 1994) and more recently, pointing device verification (Barrelle, Laverty, Henderson, Gough, Wagner & Hiron, 1996).

A broad distinction between behavioural and physiological biometric systems has been drawn. Face recognition, fingerprint, hand recognition and retina patterns are classified as physiologically based, while signature, voice recognition and keystroke dynamics are considered to be behaviourally based (Sherman, 1992; Miller, 1994). Physiological biometrics can be viewed as being relatively stable and not under user control; whereas behavioural techniques, although ultimately physiologically based, tap a learned psychological component, and to some extent therefore are under user control—thus creating a potential for variation.

There are several purported advantages to employing behaviourally based approaches to user verification. For example, it has been advanced that behaviourally based biometric systems are more acceptable to users (Sherman, 1992; Miller, 1994). This is usually suggested because of the natural and socially accepted conventions of speech, signatures and typing in everyday life. Survey research has questioned the validity of this assertion, however, with Deane, Barrelle, Henderson and Mahar (1995a) finding that within a sample of banking professionals, behaviourally based systems were perceived as less acceptable than physiological systems. It has also been suggested that behaviourally based biometric security systems can operate unobtrusively to the user, making the security verification process independent of operational tasks and therefore not impeding the user’s job performance.
Deane, Henderson, Mahar and Saliba (1995b), however, have argued that such behaviourally based systems have the potential to be used, or perceived as being able to be used, as EPM systems. They have also argued that increased stress levels under a security challenge may elicit atypical typing patterns, resulting in the two-fold effect of increased false rejection rates and a subsequent decrease in system performance, with a concomitant increase in user stress levels.

1.2. ELECTRONIC PERFORMANCE MONITORING (EPM) SYSTEMS

Since their inception EPM systems have been seen as a boon by some and an affliction by others (see Irving, Higgin & Safayeni, 1986). Schleifer and Shell (1992) report that advocates of EPM systems have asserted that the continuous tracking of work activity is essential for high performance and productivity in the modern office. Furthermore, that such systems enable the manager to control human, material and financial resources more effectively. Also, such advocates cite employee benefits in the form of more timely, and perhaps more appropriate, work performance feedback, allowing employee corrective action to be initiated earlier than when non-continuous systems are employed. Such notions have lead the Computer Business and Equipment Manufacturers Association (CBMA) to declare that they strongly support the legitimacy of EPM-type systems and that such systems should be widely used (CBMA, 1986).

Others, however, have asserted that although there exists a potential to increase productivity and reinforce desired worker behaviour, there also exists the potential to decrease worker job control, increase time pressure and increase worker stress levels (e.g. see Smith & Amick, 1989; Amick & Smith, 1992; Carayon, 1993). Carayon (1993) has provided a conceptual model of the proposed effects of electronic performance monitoring (EPM) on worker stress. This model introduces wider work environment factors such as management control principles and work measurement concepts that can influence psychosocial conditions such as social support or job control in a workplace. Central to the model is the assumption that EPM has both direct and indirect impacts on worker stress. Carayon states that: “Technology could be added to the list of work stressors. EPM can be considered a technological work stressor. Theoretical logic backed by some evidence suggests that EPM can increase worker stress” (Carayon, 1993, p. 386). The model proposed that EPM is a work stressor that has the potential to influence other work stressors, operating as a “socio-technical trigger” capable of creating adverse working conditions and increasing worker stress. Later field studies reported by Carayon (1994) have provided support for this notion that EPM systems have an indirect effect on worker stress levels through the job design process.

Consistent with Carayon’s (1993) model, Amick and Smith (1992) maintain that continuous electronic monitoring providing negative feedback to an employee (performance below acceptable level) may lead to increases in workload pressure. To the extent that these monitoring systems increase work pace or time pressure, they can also increase the chances of the worker experiencing stress-related problems as a result of increased workload. Despite this theorizing, empirical evidence to support these claims is limited. Friedman, Rosenman and Carroll (1985; cited by Amick & Smith, 1992) reported that intermittent deadlines on accountants led to increases in stress levels as deadlines drew near. Smith, Carayon, Sanders, Lim and LeGrande (1992) conducted a questionnaire
survey of 745 employees in telecommunications companies to assess the effects of electronic monitoring on job performance, satisfaction and employee health. They found that employees who had their performance electronically monitored perceived their working conditions as more stressful and reported higher workload, more difficulties with supervisor relations, more boredom, tension/anxiety, depression, anger, fatigue and health complaints than unmonitored employees. More monitored employees reported musculoskeletal health complaints such as finger, wrist, arm, shoulder, neck and back problems than unmonitored employees. Psychological problems of high tension, severe fatigue or exhaustion, extreme anxiety and depression were also more frequently reported in monitored workers. A survey study by DiTecco, Cwitco, Arsenault and Andre (1992) partially confirmed these findings and found that telephone operators most strongly linked job stress to call-time pressure items (i.e. trying to keep call time down).

In a laboratory-based study (Schleifer & Amick, 1989; Schleifer & Okogbaa, 1990) 45 female volunteers performed a standard data entry task while system response time (slow/fast) and method payment (incentive/non-incentive) were manipulated. In the study, both cardiovascular activity and somatic discomfort were recorded. Of interest was the finding that incentive pay conditions diminished heart rate variability and increased blood pressure compared to the non-incentive pay, highlighting that EPM-type systems have the potential to evoke physiological reactions consistent with increased stress levels.

Lund (1992) suggested that organizational climate as well as the technology itself might have significant impacts on stress-related outcomes independent of the monitoring itself. Lund stated that, “The existence of statistical association between EPM and stress-related symptoms or other outcomes as a result of monitoring does not necessarily establish a causal relationship nor does it explain the mechanism by which monitoring exerts these influences” (Lund, 1992, p. 58), highlighting the need for more systematic study of the effects of EPM systems.

1.3. THE CURRENT STUDY

The aims of the present study were three-fold. First, to assess the effects electronic security monitoring systems (ESM) have on the user's physiological state. Specifically, Deane et al. (1995a) have argued that ESM systems have the potential to monitor user behaviour and therefore have the potential to elicit physiological arousal effects similar to those that have been hypothesized to be elicited under EPM systems. Second, the research aimed to examine the effects explicit security challenges have on both user behaviour and physiological state when using an ESM system. Drawing from classical stress theory, Deane et al. (1995b) had postulated that explicit security challenges might have the potential to elevate the physiological state of the user. Such an elevated physiological state would subsequently result in atypical typing patterns emerging. As typing verifications systems match the current typing pattern with a previously stored typing pattern, any deviation in typing would subsequently result in an increase in false rejection rate or rejecting the valid user. Finally, the research aimed to examine the effects that one form of electronic performance monitoring system may have on the user’s physiological state. Past research had speculated that the EPM system had the potential to evoke physiological responses consisting of elevated stress levels, but experimental
research literature has been minimal. This research provided the opportunity to evaluate this hypothesized relationship within a laboratory setting.

2. Method

2.1. SUBJECTS
The subjects were 32 student volunteers from an introductory psychology course. Age ranged from 18 to 50, with a mean of 23.64 (S.D. = 8.72) with 21 subjects being female. Each subject was paid a total of $25 for his or her participation in the study.

2.2. APPARATUS AND MEASURES

2.2.1. Physiological measures
Inter-heartbeat latency was recorded continuously during the ESM and EPM. Blood pressure was monitored during a pre-session baseline and after each discrete stage of the study. Non-invasive determination of heart rate was achieved using an Associative Measurement Device (AMLAB) housed in a 486DX50-based computer system. AMLAB measured inter-heartbeat duration (R-to-R wave intervals) and further calculations converted this to inter-beat variability during each condition. A Finapress Continuous Non-invasive Blood Pressure Monitor, interfaced with AMLAB through an analogue output port, measured the blood pressure. AMLAB recorded the raw pressure wave, which was sampled at each heartbeat and provided time stamps for each event.

Inter-heartbeat duration was determined from media-trace (Ag–AgCl₂) pellet electrodes placed on abraded sites. Standard reference sites were observed, with inter-heartbeat latency referenced at the side of the chest cavity (fourth intercostal space; Wilson, Lovallo & Pincomb, 1989).

2.2.2. Data entry task and experimental software
The software used in this study was purpose-built and consisted of two data entry screens. The first screen consisted of four fields (First name, Last name, Address and Case notes). Movement between fields consisted of a two-character sequence consisting of the CTRL key plus the first letter of the desired field. Movement to the second screen required the activation of the three-key sequence:—CTRL-SHIFT N. The second screen consisted of one field, requiring the sequence “the code is trust” to be entered correctly. The three-key sequence CTRL SHIFT N then reactivated a blank copy of screen one, initiating the next data entry task.

In session one, the software was set to allow the subject to enter a total of 30 mock clinical case notes with no software interruptions. In the second session, the same case notes were re-entered in the same order, but the software was set to halt the subject at the end of each block of ten trials. Pre-set prompts were also activated during the second and third block of trials (see Procedure).

In both sessions, a purpose-built program unobtrusively recorded a transcript of key-up and key-down times to millisecond accuracy.
2.2.3. Laboratory setting
All the experimental sessions were conducted in a small air-conditioned room, with a desk positioned in one corner. The data entry computer was positioned centrally on the desk, with the finipress device on one side and the computer housing AMLAB on the other. The finipress and AMLAB screen were always out of the view of the subject, ensuring that no biofeedback occurred. Signals from both the AMLAB and data entry computer output screens were passed through an analogue monitor splitter and patched through to an adjacent room. This enabled an experimenter to view the physiological responses being recorded along with the data entry screen that the subject was viewing. Any security/performance prompts along with any instructive windows were therefore observed by the experimenter.

2.3. PROCEDURE
Prior to undertaking the study, each subject was informed that: the aim of the study was to examine user behaviour and physiological activity whilst operating a computer system; that there were two separate sessions not less than 2 days apart, (min 2 days, max 7 days); that each session would take approximately 1 hour; that they would be required to enter a total of 30 mock clinical case notes into a clinical database; that session one was a pure data entry task, but in session two physiological recordings would be measured. They were also told that for their involvement they would receive a total of 20 dollars, and that during the second session there would be a timed component where they had the opportunity to earn a further 5 dollars if they could enter the data within a pre-set time period.

Upon arrival at the experimental centre each subject was briefed on the task and how to operate the experimental software. Following briefing the experimenter demonstrated the software. For session one, the subject then entered the clinical notes in their own time.

At the beginning of session two, subjects were again told that their heart rate would be monitored continuously throughout the experiment, while blood pressure would be determined at separate designated points. Each piece of apparatus was then explained. Next, the concept of behaviourally based biometric security systems were explained and, in particular, keystroke monitoring. It had been explained to the subjects that during the first session a reference profile of their typing behaviour had been derived and that during this second session their keystrokes would be monitored and compared to the previously derived reference profile. Moreover, if their keystroke patterns did not correspond with the pre-set standard, a security warning by way of a prompt would be activated. Subjects were then shown a hardcopy printout of the security prompt that read “Security error: Please re-enter your personal identification sequence”. They were informed that if such a prompt occurred, they must re-enter the security code “the code is trust”.

Subjects were then instructed as to the placement of the heat monitor electrodes and asked to abrade the site and secure the electrode on themselves. An experimenter of the same sex then checked the placement of these electrodes and applied an appropriately sized photoplethysmographic cuff on the middle finger for determination of the pre-session baseline blood pressure. The finipress device was started and a blood pressure
reading obtained. Subjects were told that there would be a 2 min baseline, then they would be required to enter the clinical notes until prompted to await further instructions. The AMLAB device recorded the inter-heartbeat latency in virtual time from the start of the baseline. At the end of the 2 min pre-session baseline the photoplethsmographic cuff was removed from the finger and the virtual time on the database program was reset. The program recorded the time of each security prompt, which was synchronized with the events recorded by AMLAB. The experimenter then shut the room’s only door and entered the adjacent room containing the signals patched through from the AMLAB and data entry computer screens.

Subjects entered 10 of the 30 clinical notes with no security prompts. Upon completion of the tenth trial a window appeared on the screen requesting the subject to await further instructions. The experimenter then entered the experimental room and recorded blood pressure for 1 min. Subjects were reminded that the security system was still in operation and were told to continue entering clinical notes until they received further instructions. During the entry of the next 10 clinical notes, security prompts were programmed to appear after trials 14, 17 and 20. At trial 14, subjects received only one prompt, where they were asked to re-enter their personal identification sequence. The system then allowed subjects to continue data entry. At trial 17 subjects were prompted to re-enter their personal identification sequence, and upon doing this, given another security prompt and requested to re-enter the sequence again. At trial 20, subjects were requested to re-enter their sequence on three separate occasions, each followed by a security prompt. After the third attempt to re-enter the sequence a window appeared informing the subjects that there had been a security error and to await further instructions. The aim was to simulate being denied access to a system. At this point, the experimenter again entered the room and attached the cuff and recorded the subjects’ blood pressure for 1 min.

Subjects were then informed that the keystroke security system had been deactivated and an electronic performance monitoring (EPM) system had been installed. They were told that the EPM system compared their typing speed to a group-determined level and that they would be notified if their speed fell below this pre-determined level. Subjects were then shown a copy of the notification that read “Caution you are now falling below the required performance standard”. Subjects were told that $1 would be deducted from their $5 honorarium for each performance prompt they received.

The subject then entered the final 10 clinical case notes. EPM prompts were inserted at trials 24, 27, 28 and 30, informing subjects that their performance was falling below the group-determined standard. At the completion of trial 30, a performance prompt appeared, followed by a window thanking subjects for participating in the experiment and to await further instructions. The experimenter entered the room and obtained a final 1 min blood pressure measurement.

To conclude, extensive subject debriefing occurred where the concept of biometric security systems and electronic performance monitoring systems were explained. In addition, subjects were informed that presentation of prompts and warnings was entirely pre-determined and independent of user performance, and they were also informed that they would receive the full $25 for their participation in the study. Figure 1 provides a diagrammatic outline of the procedure, measures and interventions in the current study.
3. Results

3.1. PHYSIOLOGICAL EFFECTS

3.1.1. Blood pressure
The obtained mean and standard deviation blood pressure levels for the subjects over the two pre-session baseline conditions and at the completion of the security baseline condition, the security challenge condition, and the performance monitoring condition were calculated (Table 1). ANOVA revealed that there was a significant main effect for condition $F(4,120) = 8.87, p < 0.001$, and planned comparisons revealed there was a significant decrease in blood pressure level between pre-session one ($M = 116.86, \text{S.D.} = 12.60$) and pre-session baseline two ($M = 112.85, \text{S.D.} = 11.63$), $F(1, 30) = 9.71, p < 0.01$. This second reading was used as the comparison figure with the other conditions.

Comparisons revealed that blood pressure level did not significantly alter between the pre-session reading and the reading observed at the completion of the security baseline condition ($M = 115.60, \text{S.D.} = 11.18$), $F(1, 30) = 2.81, p > 0.05$. There was, however, a significant increase in blood pressure level between the pre-session baseline and the measure recorded at the completion of the security challenge condition ($M = 124.48, \text{S.D.} = 13.30$), $F(1, 30) = 30.10, p < 0.01$ and that measure recorded at the completion of the performance monitoring condition ($M = 121.57, \text{S.D.} = 15.35$), $F(1, 30) = 9.93, p < 0.01$.

One final comparison was undertaken to examine the blood pressure at the completion of the security challenge and the performance monitoring conditions. When this was done, it was found that there was a significant difference between the blood pressure level at the end of each of these conditions at the 0.05 level of confidence, $F(1, 30) = 7.04, p < 0.05$.

3.1.2. Cardiovascular activity
Inter-heartbeat interval over the separate conditions were next derived and examined (Table 2). ANOVA again revealed that a significant main effect was present,
TABLE 1
Mean and standard deviation blood pressure for each condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline 1</td>
<td>116.86</td>
<td>12.60</td>
</tr>
<tr>
<td>Baseline 2</td>
<td>112.85</td>
<td>11.63</td>
</tr>
<tr>
<td>Security baseline</td>
<td>115.60</td>
<td>11.18</td>
</tr>
<tr>
<td>Security challenge</td>
<td>124.48</td>
<td>13.30</td>
</tr>
<tr>
<td>Performance monitoring</td>
<td>121.57</td>
<td>15.35</td>
</tr>
</tbody>
</table>

TABLE 2
Mean and standard deviation inter-heartbeat interval (milliseconds) for each condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline 1</td>
<td>776.97</td>
<td>100.33</td>
</tr>
<tr>
<td>Baseline 2</td>
<td>773.43</td>
<td>105.92</td>
</tr>
<tr>
<td>Security baseline</td>
<td>739.67</td>
<td>92.37</td>
</tr>
<tr>
<td>Security challenge</td>
<td>738.10</td>
<td>108.24</td>
</tr>
<tr>
<td>Performance monitoring</td>
<td>746.97</td>
<td>84.08</td>
</tr>
</tbody>
</table>

\[ F(1, 124) = 3.53, p < 0.01. \]
Comparisons revealed no significant difference between inter-heartbeat latency over the first minute of the pre-session baseline \((M = 776.97, \text{S.D.} = 100.33)\) and the inter-heartbeat latency of the second minute of the pre-session baseline \((M = 773.43, \text{S.D.} = 105.91)\), \(F(1, 31) = 41, p > 0.05. \) For consistency of analysis the mean inter-heartbeat latency of the second baseline session was used in further comparisons.

Contrasts revealed that there was a significant decrease in inter-heartbeat latency between the baseline session and the security baseline condition \((M = 739.67, \text{S.D.} = 92.37)\), \(F(1, 31) = 7.66, p < 0.01. \) Although the inter-heartbeat latencies were lower in both the security challenge \((M = 738.10, \text{S.D.} = 108.23)\) and the performance monitoring \((M = 746.97, \text{S.D.} = 84.08)\) conditions than in the pre-session baseline condition, they were not significantly different from the baseline. Again, it was also observed that there was no difference between the security challenge and performance monitoring conditions, \(F(1, 30) = 1.11, p > 0.05. \)

As the first challenge in both the security challenge and electronic performance monitoring conditions was some way through the respective conditions, it was speculated that this apparent anomalous finding may be an artefact of combining the data as one set within the condition. That is, it was speculated that although the mean inter-heartbeat latency may have decreased after each challenge, any effect would be masked by the pre-challenge data.

To examine this hypothesis both the security challenge and performance monitoring sessions were divided into two components. Pre-challenge consisted of the inter-heart-
beat latency prior to the first challenge, and post-challenge representing the inter-heartbeat latency after the onset of the first challenge.

Pre-challenge inter-heartbeat latency for the security challenge condition ($M = 741.87$, S.D. = 110.88) was longer than the post-challenge latency ($M = 734.33$, S.D. = 112.71), although this difference was not statistically significant. In the performance monitoring session, however, the pre-challenge inter-heartbeat interval was significantly longer ($M = 760.22$, S.D. = 82.01) than the post-challenge latency ($M = 733.71$, S.D. = 97.40), $t(31) = 2.34$, $p < 0.05$.

3.2. PERFORMANCE EFFECTS

3.2.1. Security verification

The efficiency of a behaviourally based security system is based upon the accuracy of its verification system. Accuracy is defined in terms of false acceptance (accepting the invalid user) and false rejection (rejecting the valid user) rates. To test the effect that the security challenges had on user typing behaviour, two tests were conducted.

The first test examined the security code both prior to and after the security challenges in the security challenge condition of session two. This test used a standard keyboard verification algorithm (Mahar, Napier, Henderson, Lawerty, Lawrie, Hiron & Wagner, 1995; Napier, Laverty, Mahar, Henderson, Hiron & Wagner, 1995). Reference profiles were derived from the typed security strings in session one and the security baseline condition of session two. The false rejection rates were then calculated from separate security strings typed in session one and in the security baseline condition of session two, and the security string typed immediately after a security challenge condition of session two.

The obtained false rejection rate for the security strings from session one was 6.1% and 6.4% from the security baseline condition of session two. In contrast, the false rejection rate for the security strings typed immediately after a challenge in the security challenge condition of session two was 16.9%. A one-way ANOVA was then used to test if the differences observed were statistically significant. In this case the main effect was significant, $F(2,284) = 5.89$, $p < 0.01$, and comparisons revealed there was a statistically significant difference between the non-challenged and challenged data, $F(1,142) = 6.76$, $p < 0.01$, but not between the two non-challenged data sets.

To examine the impact the security challenge had on verification rates of data entered prior to and post security challenge, a verification was undertaken on the clinical abstract data. Using the same algorithm (Napier et al., 1995) it was found that the false rejection rates for session one was 12.92%, session two prior to challenge was 15.27%, and session two post-challenge was 15.9%.

3.2.2. Data entry

The impact that security and performance monitoring system may have on task performance was examined by calculating time to enter each clinical case note for each trial and each individual. The mean data entry time for the clinical case notes in session one was 32 min (S.D. = 9.3) with a minimum and maximum time of 18 and 51 min, respectively. The mean data entry time for the clinical case notes in the second session was 26 min (S.D. = 6.9), with a minimum and maximum time of 15 and 41 min, respectively.
For analysis purposes the case notes were grouped into blocks of five in the sequential order they were entered (e.g. 1–5, 6–10, 11–15, etc.) and comparisons were made both between each session and block of five. These comparisons revealed that there was a general increase in performance over sessions and blocks. To examine if the observed increase in performance within the EPM condition was a part of the general trend in increased performance a factorial ANOVA was undertaken. The analysis confirmed that there were significant main effects for both sessions, $F(1,58) = 7.53, p < 0.01$, and block, $F(5,290) = 25.73, p < 0.001$. As a significant interaction was present, $F(5,290) = 4.16, p < 0.01$, however, these main effects cannot be interpreted by themselves. Planned comparisons were conducted to identify the source of the interaction term. It was found that there was a significant difference between sessions on the first blocks only, $F(1,29) = 4.52, p < 0.05$. No other statistically significant differences were observed. Figure 2 presents the mean data entry times for each block of five trials over the two sessions in the current study.

Corrected typing errors were also examined over the two sessions and between the blocks. The corrected errors were operationally defined as the number of DELETE key presses activated whilst entering the clinical case note information. This examination revealed no statistically significant effects. That is, the typing error correction rate was not different over the first and last session, or in any of the three conditions in session

Figure 2. Mean time (milliseconds) to complete each block of five trials over the two sessions in the current study. ▶️ Session one; □ session two.
two. Such a result lends support to the notion that although task performance was increasing this was not at the expense of increased error rates. That is, a speed/accuracy trade-off did not occur.

4. Discussion

Electronic monitoring systems, whether they be security or performance, would seem to have the potential to be stress evoking. This was evidenced by the decreased inter-heartbeat latency observed during the security baseline condition. The altered cardiovascular state was elicited just by the knowledge that such a monitoring system was in place. By the end of the security baseline condition, however, where no prompts had been presented, blood pressure had reduced to the baseline level. This suggests that although knowledge of such monitoring systems may impact negatively on user stress levels, this effect will be transitory if no explicit consequence is observed.

During the security challenge and performance monitoring conditions the effects of explicit challenges were observed. When examining the session mean inter-heartbeat latency, it was observed that although the inter-heartbeat latency was lower than the baseline measure, they were not significantly so. However, when the data was split prior to, and after, the first challenge in each separate condition, it was observed that inter-heartbeat latency significantly decreased in the performance monitoring condition. Furthermore, by the end of both the security challenge and performance monitoring conditions, subjects demonstrated blood pressure levels that were significantly elevated compared to the baseline.

Unexpectedly, there was no increase in task performance during the performance monitoring condition. Rather, a general trend for increasing task performance was observed throughout the study, with the only interaction present occurring between the first block of trial of session one and session two. The implications of this general trend are two-fold. First, it can be argued that the motivational level of the subject group did not decrease during the study sessions. If motivation had decreased we would not have expected to see the continual increase in task performance. Also, it may be argued that a fully typing baseline was never achieved and that learning was still being observed. Interestingly though, contrary to predictions, the performance monitoring condition did not achieve extra task performance over and above what was part of a general trend. One could argue that this is simply an artefact of a weak operational definition in that the prospect of losing 1 dollar for each challenge was not a strong enough stimulus to elicit the desired increase in task performance. The findings of decreased inter-heartbeat latency and increased blood pressure levels do, however, seem to counter this argument. Moreover, the finding that error correction rates did not alter over the sessions or between the conditions of session two indicates that the general trend in increasing task performance was not at the expense of error rates, again supporting the notion that the motivational level of the subject group did not deteriorate over the sessions of the study.

The industrial implications of this research with regard to security monitoring systems are two-fold. First, system managers need to be aware that such systems have the potential to evoke similar stress reactions to performance monitoring systems. It appears that the way a security monitoring system may evoke a stress response is through explicit
security challenges. Therefore, such systems should be designed to minimize the false rejection of the user. Moreover, when security confidence levels fall below the required level, it may be wise to employ a non-obtrusive challenge, as opposed to an explicit security challenge, as explicit challenges may not only evoke increased stress reactions, but also an abnormal typing pattern that will further degrade the system. It must also be noted that an individual’s typing immediately after an explicit security challenge will be significantly different from the typing prior to the challenge. Therefore, such systems must be designed to accommodate such atypical typing patterns after an explicit security challenge. One way would be to examine characteristic typing patterns immediately after the explicit security challenge and develop a separate algorithm to accommodate the atypical typing produced in these circumstances or move to some other non-invasive security prompt like the word-associative or cognitive password (Bunnell, Podd, Henderson, Kennedy-Moffat & Napier, 1997). Conversely, some human-factors user training system may be developed to minimize the impact of such challenges.

Implications of these findings with regard to electronic performance monitoring systems may be far reaching. It appears that in a laboratory-based simulated system, altered physiological states similar to stress can be obtained. This finding supports the past survey research that suggests EPM systems may be stress evoking (e.g. Smith & Amick, 1989; Amick & Smith, 1992; Carayon, 1993). From this present study, however, we cannot ascertain if such reactions are transitory in nature or if, in fact, such reactions would persist for the operational life of the system. From first examination we may infer that the effects would be short lived. Unfortunately, an examination of physiological data within the electronic performance monitoring condition would suggest that this may be not so. On the completion of the electronic performance monitoring condition the user group exhibited significantly elevated blood pressure and reduced inter-heartbeat latency, contrary to the general trend.

The present finding of elevated blood pressure and reduced inter-heartbeat latency are disturbing. Within the confines of this study the operational definition involved was weak (maximum reduction of 5 dollars). Despite this weak operational definition consistent effects were observed.

Supporters of the use of EPM systems may argue that increased blood pressure levels and heart rate simply reflect an increase in arousal, and that this result supports the use of such EPM systems for enhancing on-task performance. This argument may have some merit in the case of appropriately designed and implemented systems. Past research (Westin, 1992) has suggested that appropriately designed systems may fulfil the desired objectives of management, whilst also accommodating the psychological needs of the workers. This study neither confirms nor denies such claims. This current research does, however, support the notion that EPM systems do have the potential to be a work stressor and as such do warrant careful design and implementation.

The limitations of this study are obvious. The study employed student participants and the experimental procedure was short. Future research should attempt to examine the impact of EPM in a more naturalistic setting and should address what factors may moderate the EPM/stress relationship. It may be speculated that both task and individual difference factors may moderate this EPM/stress relationship. For example, Carayon (1994) has provided evidence that job design moderates this relationship, and Davidson and Henderson (1997) have examined the impact of task difficulty on both
performance and self-reported stress. The impact of monitoring at the individual and group level should also be considered.

In summary, this research aimed to examine user physiological states and task performance under two simulated electronic monitoring systems. The results suggest that electronic performance monitoring system do have the potential to elicit arousal states consistent with stress, and that under set circumstances, elicit atypical interactive typing behaviour leading to system degradation. Contrary to expectations, the hypothesized performance increases within the EPM condition was not observed.

References


Paper accepted for publication by Associate Editor, Dr. Ian Witten.