Lighting and Sustained Performance

N.H.Eklund, P.R.Boyce and S.N.Simpson,
Lighting Research Center
Rensselaer Polytechnic Institute
Troy, New York 12180-3590

Abstract

This experiment was undertaken to determine the effect of different lighting conditions on the sustained performance of a repetitive, self-paced, data-entry task. Fifteen subjects worked for one eight-hour day under each of three different lighting installations in a private, windowless office. All three lighting installations provided a similar illuminance on the task without veiling reflections or disability glare, so for the same task they provided similar visibility. However, the three lighting installations were very different in their light distribution over the room; ranging from very uniform indirect lighting to very concentrated overhead lighting. The three lighting installations were identified by experts in lighting as very different in lighting quality. The measurements taken failed to show any difference in task performance, cognitive performance, mood or alertness between the three lighting installations. However, there was a 13% difference in the amount of work done over the day for 6-point print versus 14-point print, and there was a trend in task performance over the day consistent with the human circadian cycle. These results suggest that, for normal office lighting practice, the only aspects of lighting that affect sustained task performance over an eight-hour day are those that directly affect task visibility.

1. Introduction

Most models of how lighting conditions affect task performance are based on abstract tasks designed to have a large visual component (Weston, 1935, 1945; Blackwell, 1982; Rea and Ouellette, 1988). The experiment described here was undertaken to determine the effect of different lighting conditions on the performance of a realistic office task over an eight-hour working day.

There are two aspects of lighting that may affect sustained performance; the lighting of the task and the lighting of the room. The interactions between the lighting of the task and the characteristics of the task are important because they change the stimuli the task presents to the visual system and hence the task visibility. Greater task visibility is expected to lead to better sustained performance. As for the lighting of the room, this does not directly affect task visibility but it may affect mood and hence motivation to perform the task.

2. Method

2.1 An overview

Temporary office workers were recruited to perform a variety of tasks, representative of those done by a low-level clerical worker, for eight hours, each day, for four days, on a fixed schedule. The tasks varied in both visual difficulty and cognitive difficulty. The first day was used for training, and the remaining three days for data collection, one day being spent in one of three rooms, each room having a different lighting installation. Measurements of performance on a data-entry task and cognitive performance tests, subjective feelings of mood and alertness and observations of behavior were made to determine the effects of prolonged working under the different lighting conditions.

2.2 Setting

Figure 1 shows a plan of the setting used. The three private rooms were the same size, shape and layout. Each room was 3.35 m (11 ft) by 2.74 m (9 ft), with a ceiling height of 2.74 m (9 ft). All the rooms were windowless. The walls of each room were painted gray with a matte finish (r = 0.36), the ceiling was standard white, textured acoustical tile (r = 0.70), and the floor was covered in gray carpeting (r = 0.14). The low wall reflectance was chosen to reduce the inter-reflected light and hence to enhance the differences in light distribution for the three lighting installations.

Each room had the same L-shaped workstation, with the same model computer positioned on the corner of the L, with a document stand on the left-hand side (see Figure 1). The rest of the office was furnished with a bookcase, a print and a clock. These items were all located at the same positions within each room and all were a mixture of black, white and gray colors. The differences in the furnishings of the three rooms were limited to the materials used to fill the bookshelves and the content of the print.

2.3 Lighting conditions

A different lighting installation was used in each room.

Room 1 used two, 3-lamp, 2 ft x 4 ft, 27-cell, parabolic luminaires recessed into the ceiling on the diagonal of the office so that one of the luminaires was over each arm of the L-shaped workstation. The lamps used were 32W, T8, 3500 K linear fluorescents with a CIE General Color Rendering Index of 75. The whole installation was fitted with a dimming circuit so that the illuminances produced by the installation could be varied. Figure 2 shows the mean rating of lighting quality given by twelve subjects experienced in lighting design and twelve subjects naïve in lighting. Each subject spent 30 minutes doing office work in each of the three rooms before answering a questionnaire about the lighting. It is clear from Figure 2 that Room 1 was considered to be intermediate in lighting quality (Johnson, Eklund and Boyce, 1997). Figure 3 shows a photograph of Room 1 lit by this installation.

Room 2 used a task-ambient lighting system. The ambient lighting was provided by two, 1200mm (4 ft) long, linear, two-lamp, pendant luminaires, suspended 380 mm (15") below the ceiling, and parallel to each other. The lamps used in the luminaire were the same as in Room 1. The task lighting was provided by an adjustable task-light containing an 9W, 3500 K, compact fluorescent lamp with a CIE General Color Rendering Index of 82. Figure 2 shows this lighting was considered to have the highest lighting quality of the three rooms. Figure 4 shows a photograph of Room 2 lit by this task-ambient lighting installation.

Room 3 used a single, three-lamp, 2 ft x 2 ft, 12-cell, parabolic luminaire recessed in the center of the ceiling, directly over the subject's head. The lamps used were 40W, 3500K, 22.3 in long, compact fluorescents with a CIE General Color Rendering Index of 75. Figure 2 shows this installation was considered to have the lowest lighting quality, by both experts in lighting and people naïve in lighting Figure 5 shows a photograph of Room 3 lit by this lighting installation.

Table 1 shows the mean illuminances produced by these three lighting installations on the workstation surface, the document holder and the screen of the computer monitor. Examination of Table 1 shows that the illuminance on the document holder was similar in all three rooms. Further, the small size of the rooms ensured that the luminaires were not in positions where they would cause discomfort glare, as conventionally defined, nor would they be seen as reflected images in the computer monitor. Also, by using matte materials for the visual tasks, veiling reflections were eliminated. Table 2 shows the range of luminances of the walls that can be seen by the subject when seated at the workstation, the range of luminances on the ceiling and the maximum luminance of the luminaires. Examining Figures 3, 4, and 5 and Tables 1 and 2 confirms that while the lighting conditions in the three rooms were very different in terms of the distribution of light onto the room surfaces and the luminance of the luminaires, in terms of the visibility of a given task, the lighting of the three rooms was very similar.

2.4 Tasks

2.4.1 Data-entry task

The task performed for the bulk of the time spent in the office was the data-entry of alphanumeric codes. This is a simple, self-paced, repetitive task and one that is actually done in offices. The data-entry task required the subject to enter random ten-character alphanumeric strings read from paper copy. A single data entry record consisted of five strings grouped on a page. Each form to be copied was presented on an 11" x 8.5" page, the printing being done by a laser printer to a high contrast, on matte paper. The strings were printed in Geneva font, in three sizes, 6, 10 and 14-point print (see Figure 6).

The form onto which the alphanumeric codes were to be copied was presented in high luminance contrast and negative polarity (dark graphics on a bright background) on the computer monitor (Figure 7). The high luminance contrast ensured the display was easily read. The negative polarity display ensured that there were no high luminance reflected images visible on the screen. The software monitoring the data-entry identified any errors made and required the subject to correct the error before proceeding. Thus, any errors made in the data-entry increased the time taken to complete the entry.

In addition to the different lighting installations, three sizes of print (6, 10, and 14-point) were used in the hard copy from which the subjects read the data. The three print sizes were presented cyclically (page 1: 10 pt, page 2: 6 pt, page 3: 14 pt, page 4: 10 pt, etc.) throughout the data-entry set. The data-entry sheets were bound in a standard three-ring binder, which was placed on a fixed document holder. The three different print sizes were used to provide different levels of task visibility.

Performance on the data-entry task was measured in two ways, the time to enter each set of five alphanumeric strings (work time), and the time between finishing one set and starting the next (rest time) Work time allows the speed of data-entry changes over time to be measured and hence any performance decrement to be identified. Rest time measurements determine if the length of the subject's "micro rests" change over time. Boyce and Eklund (1996) found this measure to be sensitive to lighting conditions.

2.5.2 Cognitive tasks

In addition to the data-entry task, once every hour the subjects performed a short duration battery of cognitive performance tests. The tests used were grammatical-reasoning, addition, short-term memory and coded-lock. All these tests were presented on, and the responses measured by, the computer in each room. The stimuli presented to the subject were of large size and high contrast so that they were highly visible. The display used a negative polarity format (dark graphics on a bright screen) so there were no high luminance reflected images visible on the screen.

The grammatical-reasoning task (Baddeley, 1968) presents the subject with a statement (e.g., "A does not follow B") and a letter pair (e.g., "AB"), and asks the subject to evaluate the truth of this statement (the above example is "true"). This measure looks at logical reasoning. Campbell and Dawson (1990) and Boyce, et al. (1997) have found this measure to be sensitive to lighting conditions during night-shift work Performance was measured as the number of such problems correctly answered in two minutes.

The addition task presents the subject with a pair of two, three or four digit numbers to be mentally added. This measure looks at mathematical ability. Similar tasks have been found to be sensitive to sleep loss (Williams and Lubin, 1967) and noise (Holding, Loeb, and Baker, 1983), and have been used in long term studies of work (Morgan and Alluisi, 1972). Performance was measured by the mean time taken to solve a problem.

In the short-term memory task, subjects are presented with a set of five digits (e.g., 92574), followed by a series of single digit probe numbers. The subject's task is to say if the single digit was one of the set of five digits. After five single digit probes have been presented a new set of five digits is given and the process is repeated. Boyce et al. (1997) found a measure of short-term memory to be sensitive to lighting conditions during night-shift work. Performance on the short-term memory task is measured as the number of correct answers given in two minutes.

The coded-lock task examines memory and logical reasoning. Subjects are presented with a series of five buttons that must be pressed in a certain order that is unknown to the subject. So long as the buttons are pressed in the correct order, they remain depressed. If a button is pressed out of order, then all of the buttons are reset. The subject's task is to find the correct order. Morgan and Alluisi (1972) used a similar task to investigate long term performance. Performance on the coded-lock task is measured as the number of problems solved in two minutes.

These cognitive tasks were introduced to measure the aftereffects of work on the data-entry task.

2.5.3 Mood and alertness

The subject's mood was measured using the Positive and Negative Affect Scale (PANAS) (Watson et al., 1988). This survey measures positive affect ("good feelings") and negative affect ("bad feelings"), and has been shown to be sensitive to work on monotonous tasks in previous studies (Boyce and Eklund, 1996). Ratings of the extent to which the subject has a named feeling are given on a 1 - 5 scale. Positive affect is measured by the response to feeling interested, excited, strong, enthusiastic, proud, alert, inspired, determined, attentive, and active. Negative affect is measured by the responses to felling distressed, upset, guilty, scared, hostile, irritable, ashamed, nervous, jittery and afraid.

Feelings of alertness were measured using the Activation-Deactivation Adjective Checklist (ADACL) (Thayer, 1986). This checklist measures the subject's general activation, deactivation-sleep, high activation and general deactivation. The difference between these terms can be understood from the scales used for each form of activation. General activation is determined by the mean of the ratings on the active, vigorous, energetic, lively and full-of-pep scales. Deactivation-sleep is determined by the mean of the ratings on the drowsy, sleepy, tired, wide-awake and wakeful scales, the last two being reverse-coded. High activation is determined by the mean of the ratings on the tense, jittery, clutched-up, intense and fearful scales. General deactivation is determined by the mean of the ratings on the placid, at-rest, calm, still and quiet scales. This instrument has also been shown to be sensitive to work on monotonous tasks in a previous study (Boyce and Eklund, 1996).

2.5.4 Subjective impressions

At the end of each day, subjects were debriefed using a simple questionnaire. The questionnaire was concerned with several aspects of the subjects' experience. One was the subjects' physiological state at the end of the day, specifically whether they had a headache or eyestrain or if their muscles felt tired. Another addressed the subjects' opinions about the air quality in the room. Yet another asked about the ergonomics of the workstation. The only question specifically about the lighting was one that asked if the room looked gloomy. This questionnaire was structured in this way to support the deception that was part of the experiment (see section 2.6), yet to obtain information on the consequences of working under the different lighting installations.

2.5.5 Post-study lighting evaluation

After all the subjects had completed their work in all three rooms, as many of the subjects as were available were brought back to provide a detailed assessment of the lighting in the three rooms. To do this, each subject worked on the data-entry task in each room for one hour, at the end of which he/she answered a questionnaire about the lighting of the room.

2.6. Deception

Any experiment in which the only obvious differences in the environment are the lighting conditions carries the risk that the subjects will modify their performance to match what they think are the experimenter's expectations. To reduce this possibility, an attempt was made to deceive the subjects as to the purpose of the experiment. To carry through this deception, the subjects were told to report to the Air Quality Laboratory. Besides housing the Lighting Laboratory, the facility where the experiment took place is an incubator center where small companies start their life. The Air Quality Laboratory was located at the opposite end of the building from the Lighting Laboratory. The subjects were told the purpose of the experiment was to examine the effect of the physical environment on the sustained performance of office work. There were different lighting installations in the three rooms but these were explained as being left over from when the space was used by the Lighting Laboratory. Although the workstation furnishings were the same in all three offices, the offices differed in what sort of portable air quality device they had. One office had no air quality device. Another was equipped with a humidifier. The third had an air filter and ionizer. Both these air-quality devices were wired so that the operating light was lit when power was connected but in all other respects they were inoperative, i.e., the humidifier was without water and a filter and the ionizer was without a filter and was not electrically connected.

2.7. Time order of experimental week and day

Each subject was hired for four successive days. On the first day, the subject was screened and trained on the tasks. The training was done in the open reception area. The subject did thirty pages of the data-entry task and answered the Room Assessment Questionnaire twice. The subject also did the cognitive tasks until they were able to achieve the desired level of performance, twice. Specifically, the subject had to get more than 57 out of 64 grammatical reasoning tests correct in 7 minutes; do more than 9 addition measurements in 2 minutes; get 86% correct responses for more than 16 short term memory tests in 2 minutes; and solve at least 5 coded-lock problems in 90 seconds. Any subject who failed to achieve these performance levels after four hours was rejected. The remaining three days were spent in the private rooms, one day in each room.

Each day spent in one of the rooms was divided into eight, one-hour periods, with a half-hour break for lunch from 12:30 to 13:00 hours. During each hour, subjects performed (at most) 44 minutes of data entry, and exactly 2 minutes of grammatical-reasoning, 2 minutes of addition, 2 minutes of short-term memory, and 2 minutes of coded-lock. The subject's mood and feelings of alertness were measured four times a day. The data-entry task occupied the first part of each hour, followed by the cognitive performance tests and, in every other hour, the questionnaires for mood and alertness. After this cycle of activities had been completed, the data-entry task was started again.

The subjects could leave the room at any time to go to the restroom, take a break or get a drink, as they wished, but they were encouraged not to leave the room while doing the cognitive performance tests. The time the subject was out of the room was recorded as "Break time".

2.8. Statistical controls

A Latin Square was used to allocate subjects to rooms over the three days. This balances the order in which the subjects, as a group, experience the lighting installations. To limit the influence of any psychological effect of the presence of the inoperative air-quality devices, another Latin Square was used to balance their presence across all three rooms.

2.9. Subject Recruitment

Subjects were recruited from a temporary service agency. All the subjects who took part in the experiment were paid at the agency rate for their work. Demographic details of the subjects are given in Table 3.

3. Results

3.1 Data treatment

In total, nineteen subjects completed the experiment. A check on the consistency with which the statistical design had been followed revealed that due to dropouts during the three day periods, the balancing of the order in which the subjects experienced the three rooms was not perfect. To correct this potential source of bias, the data from four subjects were dropped, the subjects chosen for dropping being those with the longest break times, subject to the need to perfectly balance the order of experiencing the three rooms. Subjects with the longest break times were dropped because there were fewer data available for these subjects. In fact, it made little difference which subjects were dropped. The same statistical analysis applied to the work time data (see section 3.2) after dropping four subjects with short break times produced the same pattern of statistical significances as occurred after dropping four subjects with long break times.

3.2 Work time on the data-entry task

The statistical significances of the factors print size, hour, and lighting on work time were determined using a repeated-measures analysis of variance (ANOVA). This analysis showed that the only statistically significant factors were print size ($F=130.5$, $df = 2, 28$; $P<0.0001$) and hour ($F=2.6$, $df = 7, 98$; $p<0.05$). There was no statistically significant effect of lighting, nor any statistically significant interactions, although the interaction between print size and hour approached statistical significance ($F=1.77$, $df = 14, 196$; $p=0.07$).

The mean work times, averaged over both lighting and hour factors for the three print sizes, are shown in Figure 8. It is clear that as the print size gets smaller, the mean work time gets longer. In fact the mean work time for the 6-point print is 13% longer than for the 14 point print. Also shown in Figure 8 are the mean work times obtained for the data-entry task in two other studies. The first was a pilot study, in which four subjects did the data entry task for four hours (Eklund, Boyce and Johnson, 1997). The other was an earlier study where slightly different print sizes were used and the total time spent doing the data entry task was 160 minutes (Boyce and Eklund, 1996). These results, obtained from different subjects and, in one case, different lighting installations producing a similar illuminance, show effects of print size on work times of similar magnitude to those obtained in the present experiment.

The mean work times, averaged over both lighting and print size factors, for the eight hours are shown in Figure 9. There is a trend with time that can be associated with the daily performance rhythm for working memory tasks (Folkard and Monk, 1985). Over the morning there is a steady reduction in work time, i.e. faster work, particularly between the first and second hours, After lunch there is an increase in work time, i.e., slower work until the final hour when the speed of work increases again. Supplementary tests showed that the mean work times for hours 2, 3 and 4, as a group, were statistically significantly shorter than for the hours 5, 6 and 7, as a group ($F = 5.70$, df = 1,126; $p < 0.05$).

As for the almost statistically significant interaction between print size and hour, Figure 10 shows the mean work times, averaged over all three lighting conditions, for each print size and hour. It is apparent that if there is a real interaction it occurs because the variation in mean work time over the eight hours is greater for the 6-point print size than for the 10 and 14-point print size. Specifically, there is a greater increase in work times in the afternoon for the 6-point print than for the 10- and 14-point print.

3.3 Rest time on the data-entry task

The statistical significances of the factors print size, hour, and lighting on rest time, were determined using a repeated-measures ANOVA. This analysis showed that the only significant factor was hour ($F=3.83$, df = 7, 98; $p<0.01$). There were no statistically significant effects of lighting or print size nor any statistically significant interactions.

The mean rest times, averaged over both lighting and print size factors, for the eight hours, are shown in Figure 11. Again, there is a trend with time that can be associated with the daily rhythm of performance for working memory tasks. At the beginning of the morning, there is a marked reduction in rest time until mid-morning, after which there is a slight increase until late in the afternoon, when rest time decreases again.

3.4 Break time

The total break times for the whole day were collected for each subject, for each room. A repeated measures ANOVA failed to show any statistically significant differences in mean total break times for the three rooms. The mean total break times were Room 1 = 1074 s, Room 2 = 960 s, and Room 3 = 1011 s.

3.5 Cognitive performance

The statistical significances of the lighting conditions and hour factors on the performance of the grammatical reasoning, addition, short-term memory and coded-lock tasks were examined using repeated-measures ANOVA. These analyses revealed only one statistically significant effect, that being on the time per addition on the addition test ($F = 3.05$, df = 7, 98; $p<0.01$). Figure 12 shows the mean time per problem for each hour, averaged over all three lighting conditions. The trend is one where the time taken to do an addition problem increases over the day until the last two hours.

4.6 Positive and negative affect

The statistical significance of the lighting conditions and hour factors on negative and postive affect were examined using repeated-measures ANOVA. These analyses revealed only one statistically significant effect, that of time, for both negative and positive affect ($F = 7.61$, df = 3, 42; $p<0.005$; and $F = 4.52$, df = 3, 42; $p< 0.01$, respectively). Figure 13 shows the mean positive and negative affect scores, averaged over all three lighting conditions, for the four times the subject's mood was measured. It is clear that there is a tendency for the positive affect to decrease and the negative effect to increase over time. Similar trends in positive and negative affect following prolonged work on the data-entry task have been found in an earlier study (Boyce and Eklund, 1996).

4.7 Alertness

The statistical significance of the lighting conditions and hour factors on the general deactivation, deactivation-sleep, high activation and general activation scores obtained from the ADACL scale, were examined for all four measures using repeated-measures ANOVA. These analyses revealed no statistically significant effects. The mean ratings, averaged over all lighting conditions and the four times the scale was given, are general deactivation = 13.6; high activation = 17.3; deactivation-sleep = 15.4; general activation = 14.1. These values are about the middle of the range of possible values (5 - 20), which indicates that the subjects had no strong feelings about their general deactivation, high activation, deactivation-sleep or general activation.

4.8 Subjective impressions

The statistical significance of the lighting conditions on the strongly disagree / strongly agree scale for each statement in the Room Assessment Questionnaire was examined for the statements using a repeated-measures ANOVA. These analyses revealed only one statistically significant effect, that being on the statement "The temperature is not too high" ($F = 5.75$; $df = 2,28$; $p<0.01$), for which Room 2 is considered warmer than the other two rooms (Figure 14). Given that all three rooms are supplied from the same air-conditioning system, there were no statistically significant differences between the rooms on the opposing statement "The temperature is not too low", nor for the statement "I am comfortable working in this room". It seems likely that the statistical significance is a type 1 error. It is also worth noting that there was no statistically significant difference in the rated air quality of the three rooms.

4.9 Post-study lighting evaluation

The statistical significances of the lighting conditions on the strongly disagree / strongly agree scale for each statement in the post-study Room Lighting Questionnaire were examined for the statements using repeated-measures ANOVA. There was no statistically significant effect of lighting condition for any statement (Figure 15).

5. Discussion

There is both bad news and good news in these results. The bad news is that the different lighting conditions have no effect on the sustained performance of the data-entry task and the cognitive tasks, nor on the mood and alertness of the subjects. The good news is the magnitude of the difference in work times for the different print sizes used in the data-entry task.

The absence of any effect of the different lighting conditions on the sustained performance of the data-entry task implies that only aspects of lighting which directly affect task visibility affect task performance. The three lighting installations examined provided similar illuminances on the document holder carrying the material used in the data-entry task. The material was of matte reflectance so there were no veiling reflections present to reduce the luminance contrast of the data to be entered. The rooms were small, the luminaires were mounted in the ceiling and the room surfaces were matte, so there was no possibility of disability glare occurring. Taken together, these characteristics mean that visibility of the data-entry material of the same print size was very similar in all three rooms, regardless of the differences in lighting installations. The only difference in visibility occurs for the different print sizes used in the data-entry task. There was a highly statistically significant difference in work time for the different print sizes, the smallest print size with the lowest visibility having the longest work time (Figure 8). This effect is what would be expected from models of relative visual performance (Rea and Ouellette, 1988)

Although the lighting installations in the three rooms produce very similar visibilities for the same print size of the data-entry task, they do differ markedly in the light distribution on the room surfaces and in the luminance of the luminaires. The indirect lighting in Room 2 produces a uniformly lit ceiling and walls, from a low luminaire luminance. The lighting in Room 3 produces very dark ceiling and walls, and strong shadows, from a single luminaire with a very high luminance directly above the subject. The two recessed parabolic luminaires in Room 1 produce ceiling and wall luminances and luminaire luminances somewhere between those in Rooms 2 and 3. It was assumed that these lighting installations would be considered different in lighting quality and this would affect the subject's mood over the day which, in turn, would influence the amount of work they did. The results have shown this assumption to be unwarranted. The question now is why?

The first thing to say is that the ten subjects who were available for the post-study lighting evaluation did not consider the three lighting installations to differ in lighting quality. Figure 16 shows the mean rated quality of the lighting in the three rooms, given by subjects expert in lighting and naive in lighting (Johnson, Eklund and Boyce, 1997) and the temporary office workers available for the post-study lighting evaluation in this experiment. Figure 16 reveals that the subjects expert in lighting consider the three lighting installations to differ markedly in quality as do the naïve subjects used in the pilot, but the subjects in this experiment saw no statistically significant difference in quality for the three lighting installations. A similar pattern is evident in the responses to almost all the other statements used in the questionnaire. Figures 17 to 22 show the mean responses by these groups of subjects to the statements " I like the lighting used in this room", "The lighting in this room is not glaring", "The lighting does not cause pronounced shadows", "The lighting of this room allows me to see well", "This is better than average lighting for an office" and "I am comfortable working in this room". For all six statements, the subjects used in the post-study lighting evaluation showed much less difference in their responses for the three lighting installations than did the expert and naïve subjects used in the pilot study.

Assuming that the ten subjects making the post-study lighting evaluation are representative of the subjects who completed the study, and Table 4 suggests they are; then their failure to discriminate between the lighting installations provides a plausible explanation of the failure to show any effect of the different lighting conditions on the data-entry task performance. All three lighting installations provided similar visibility of the task and the subjects did not consider the aspects for the lighting which did differ to be of any importance. The question now becomes, why did they consider the very obvious difference in light distribution irrelevant to their work? There are at least four possible answers. The first is that the data-entry task tends to concentrate the subject's attention on the task area and this was lit very similarly in all three rooms. It should also be noted that the subjects making judgements about the lighting in the pilot study (Johnson, Eklund and Boyce, 1997) had their attention directed to the lighting while the subjects in this experiment did not. The second is that the subjects in this study had spent a much longer time in each room than did the subjects in the pilot study (8 hours versus 30 minutes). This longer time may have allowed habituation to occur. The third is that the subjects used in this study really were insensitive to lighting conditions. It seems plausible that people expert in lighting design would be more sensitive to lighting differences than people who were naive in lighting. The fourth is that the temporary office workers who were used in this study may have seen a much wider range of office lighting conditions than those presented in the three offices and this may have compressed the range of their responses.

It is not possible to separate these possibilities with the data available. However, given that the three lighting installations represent the maximum range of light distributions likely to be found in office lighting practice, that the subjects are representative of office workers, and the task is representative of some types of office work, then, in practice, there is little likelihood that lighting conditions, other than those which alter task visibility, will have any effect on task performance.

The statistically significant trends that do occur in the performance of the data-entry task, the cognitive tasks, and the subjects' mood are related to the passage of time. For the data entry task, the mean work time and mean rest times decrease over the morning and increase slightly after lunch. This pattern is consistent with the changes in performance for other tasks involving short-term storage and processing of information and is associated with human circadian rhythms (Folkard and Monk, 1985). The change over time for the addition task shows a steady increase in the mean time to do mental arithmetic, a change which could be related to fatigue. As for the changes in mood, the trend is for the subject to show a slight increase in negative feelings and a more marked decrease in positive mood over the course of the day, probably because of the boring nature of the data-entry task.

The largest effects found were the differences in work time for the different print sizes. These differences confirm that task visibility does matter for sustained performance. The mean cycle time on the data-entry task, i.e., the time from the start of one form to the start of the next, is 80.1s for the 6-point print and 70.4s for the 14-point print. Assuming that these mean times are maintained throughout the 8-hour day, then changing the printing from 6-point to 14-point would allow the number of forms entered in an 8-hour day to be increased from 360 to 409, an increase of 14% or 49 more forms per 8-hour day. It would be interesting to determine if differences of this magnitude occur when the lighting conditions are used to change the task visibility and whether such changes in task performance are predictable from existing models of relative visual performance.

7. Acknowledgements
This work was undertaken with financial support and technical review from the Lighting Research Office of the Electric Power Research Institute. Co-funding for this research was provided by: Osram Sylvania Inc.; Consolidated Edison of New York; Lighting Corporation of America and the Electric Power Research Institute.

In addition, the Environmental Protection Agency (EPA) provided a grant to support the funding for this research. The material in this report is based upon work partially supported by EPA under Grant No. 825727-01-0. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of EPA.

8. References
Baddeley, A., A 3 min reasoning test based on grammatical transformation, Psychon. Sci., 10 (10), 343, 1968.

Blackwell, H.R., A comprehensive quantitative method for prediction of visual performance potential as a function of reference luminance, Journal of the Illuminating Engineering Society, 12, 40-63, 1982.

Boyce, P.R., Beckstead, J.R., Eklund, N.H., Strobel, R.W., and Rea, M.S., Lighting the graveyard shift: the influence of a daylight-simulating skylight on the task performance and mood of night-shift workers, Lighting Research and Technology, 29, 105, 1997.

Boyce, P.R., and Eklund, N.H., An evaluation of office lighting options, Proceedings, IES Annual Conference, Illuminating Engineering Society of North America, Seattle, 1996.

Campbell, S.S., and Dawson, D., Enhancement of nighttime alertness and performance with bright ambient light, Physiology and Behavior, 48, 317, 1990.

Eklund N.H., Boyce, P.R., and Johnson, A.R., Pilot Study 2: Effects of three lighting installations on four-hour sustained task performance, Lighting Research Center, August 1997.

Folkard, S., and Monk, T.H., *Hours of Work*, John Wiley and Sons, Chichester, 1985.

Holding, D., Loeb, M., and Baker, M., Effects and aftereffects of continuous noise and computation work on risk and effort choices, Motivation and Emotion, 7(4), 331, 1983.

Johnson, A. R., Eklund, N.H., and Boyce P.R., Pilot Study 1: Perception of three lighting installations by people experienced and naïve in lighting, Lighting Research Center, July, 1997.

Morgan, B. and Alluisi, E., Synthetic work: Methodology for assessment of human performance, Perceptual and Motor Skills, 35, 835, 1972.

Rea, M.S., and Ouellette. M.J., Relative visual performance using reaction times. Lighting Research and Technology, 20, 139, 1988.

Thayer, R., Activation-Deactivation Adjective Checklist: current overview and structural analysis, Psychological Reports, 58, 607, 1986.

Watson, D., Clark, L.A., and Tellegen, A., Development and validation of brief measures of positive and negative affect: the PANAS Scales, Journal of Personality and Social Psychology, 54, 1063, 1988.

Weston H.C., The relation between illumination and industrial efficiency: The effect of size of work, Joint Report of the Industrial Health Research Board and the Illumination Research Committee, His Majesty's Stationary Office, London, 1935.

Weston H.C., The relation between illumination and industrial efficiency: The effect of brightness contrast, Industrial Health Research Board of the Medical Research Council. Report 87, His Majesty's Stationary Office, London, 1945.

Williams, H., and Lubin, A., Speeded addition and sleep loss, Journal of Experimental Psychology, 73(2), 313, 1967.

Table 1: Mean illuminances (fc) measured on the workstation, document holder and monitor screen for the three rooms.

Measurement position	Room 1	Room 2	Room 3
Workstation surface	45.5	23.9 (without task light)	45.8
Document holder	43.2	49.8	45.2
Monitor screen	17.6	12.8	11.0

Table 2: Range of luminances (cd/m^2) on the walls and ceiling and the maximum luminance of the luminaire, for the three rooms.

Measurement position	Room 1	Room 2	Room 3
Walls in front of subject when seated at the workstation	9 – 48	18 - 20	9 – 17
Ceiling	8 – 26	24 – 348	6 – 28
Luminaire	7100	990	18,300

Table 3: Demographic details of subjects

Age	Mean = 27.2 years, standard deviation = 12.3 years
Gender	Male = 5, Female = 10
Optical correction	No correction = 7; Spectacles = 3; Contact lenses = 5
Handedness	Right handed = 11; Left handed = 4
Visual acuity	Acuity 20/20 = 10; Acuity 20/25 = 3; Acuity 20/30 = 2

Table 4: Mean rating of agreement (and associated standard error of the mean) with the statement "This room does not look gloomy" for the fifteen subjects who answered the Room Assessment Questionnaire and the ten subjects who answered the post-study Lighting Evaluation Questionnaire. The range of possible values is -5.0 to +5.0.

Room	Room Assessment Questionnaire		Post-study Lighting Evaluation Questionnaire	
	Mean	Standard Error	Mean	Standard Error
1	-0.97	0.81	-0.25	0.96
2	-1.70	0.70	-0.90	0.78
3	-1.00	0.85	-0.20	0.95

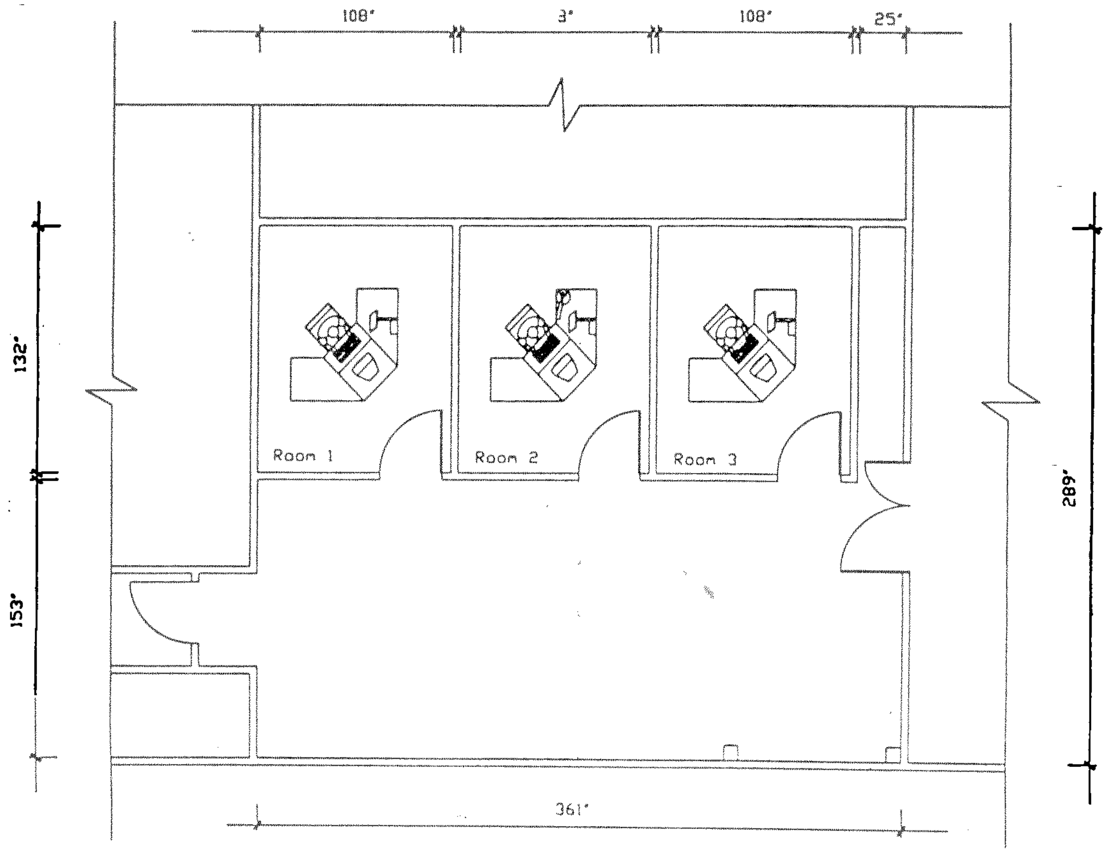


Figure 1 A plan of the three rooms and the furniture in them

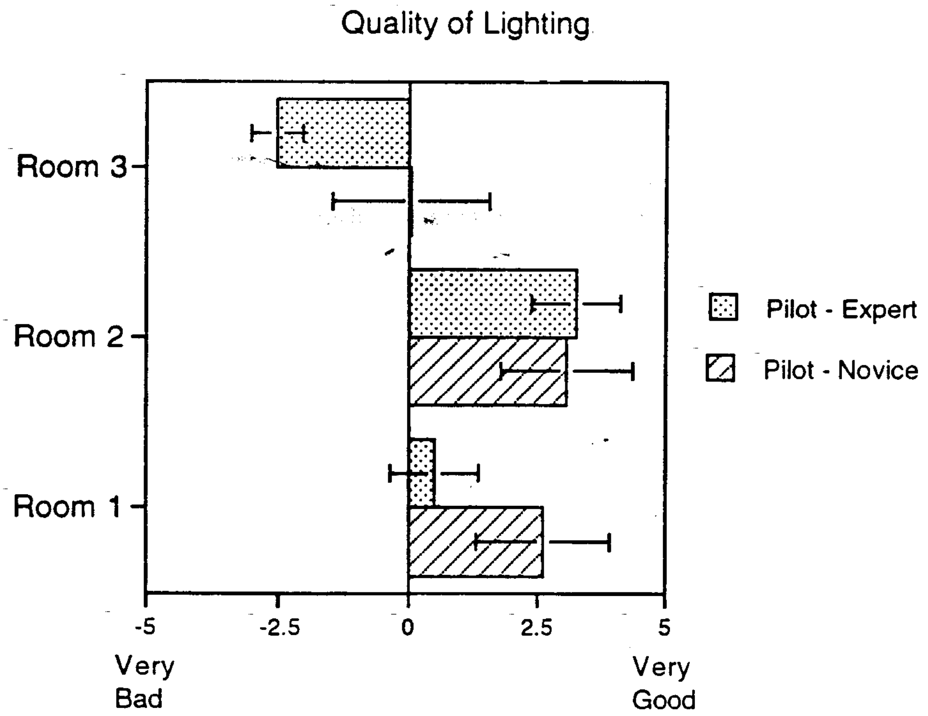


Figure 2 Mean ratings of lighting quality for the three rooms, given by subjects expert in lighting and naïve in lighting. The error bars are the 95% confidence intervals, (Johnson, Eklund and Boyce, 1997)

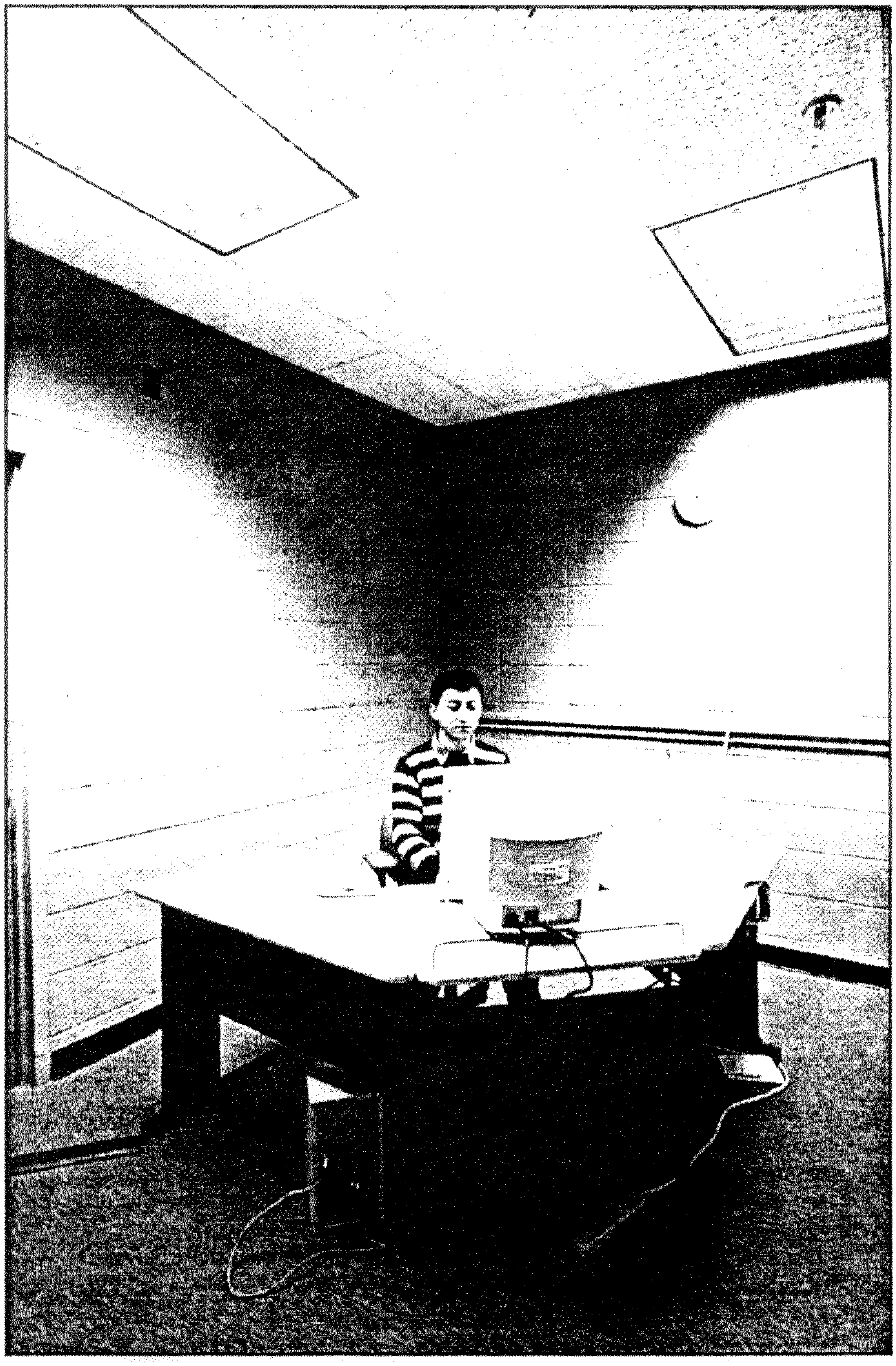


Figure 3: A view of Room 1 taken from the doorway.

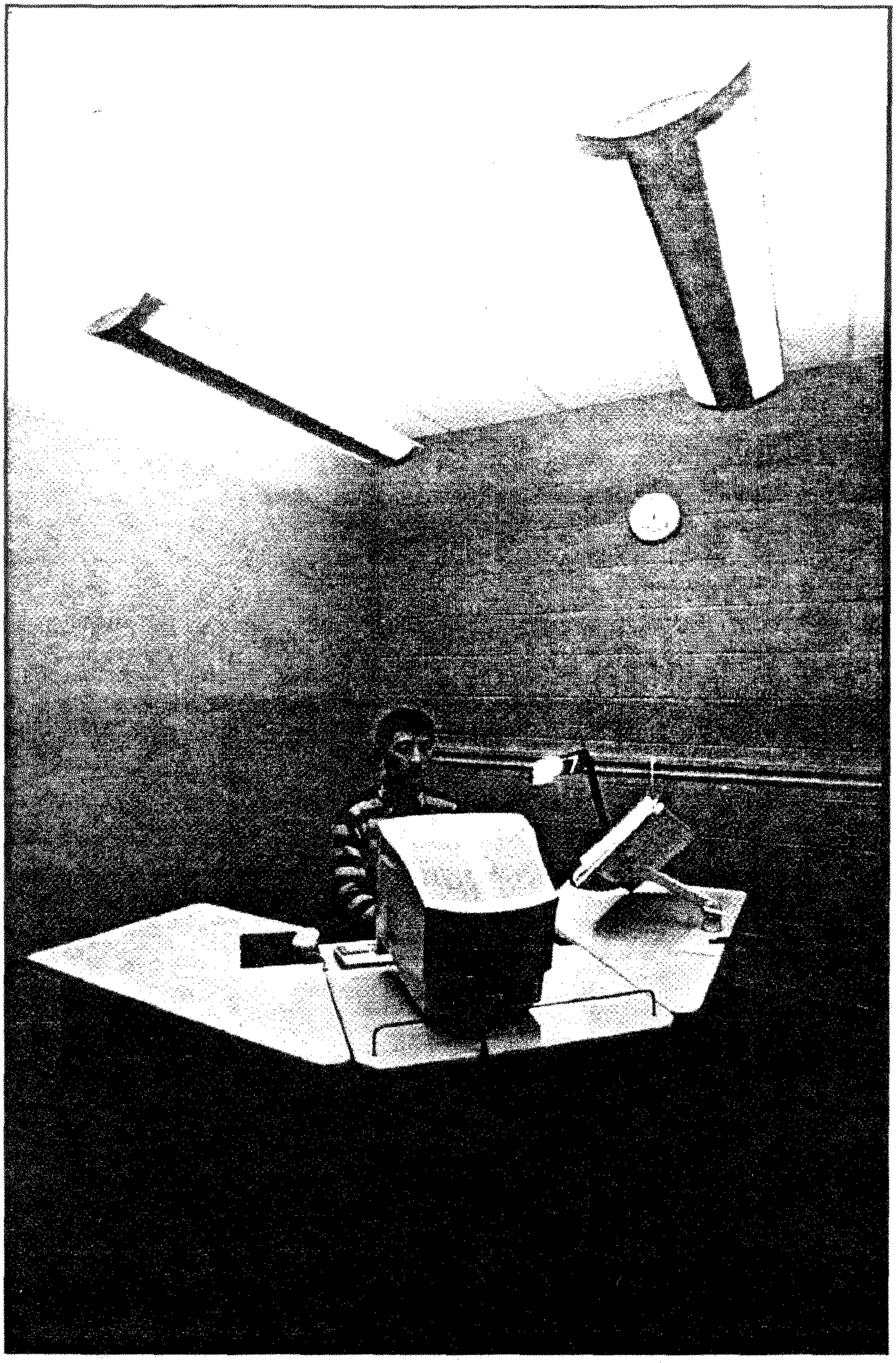


Figure 4: A view of Room 2 taken from the doorway.

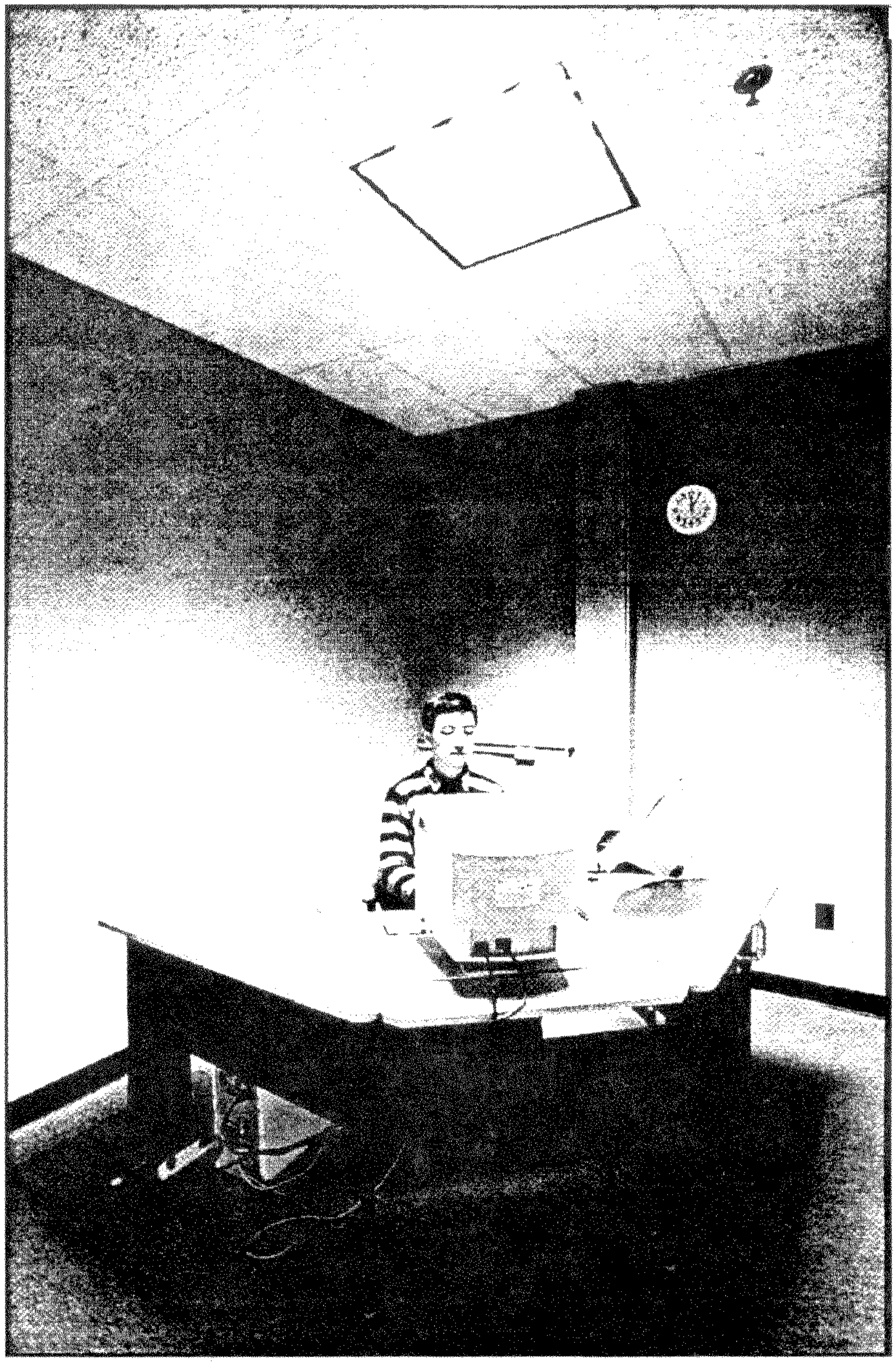


Figure 5: A view of Room 3 taken from the doorway.

Health Plan	3UU63J8962
Region	PK07MSXGY1
Dental Plan	6A60M4AYJ5
Security Code	9973508814
Certificate Number	BNCTJYJXYE

Health Plan	UZHNJEVMPB
Region	6186917896
Dental Plan	5S4F526620
Security Code	MQZ2EYUCHG
Certificate Number	B47G2TNQV6

Health Plan	TV0MLYVLDD
Region	4863Z33804
Dental Plan	7664110064
Security Code	3E669URRAG
Certificate Number	EMAECYBAKN

Figure 6 An example of the data-entry task in 6-point, 10-point and 14-point print. Note that in the experiment only one print size occurred on each page.

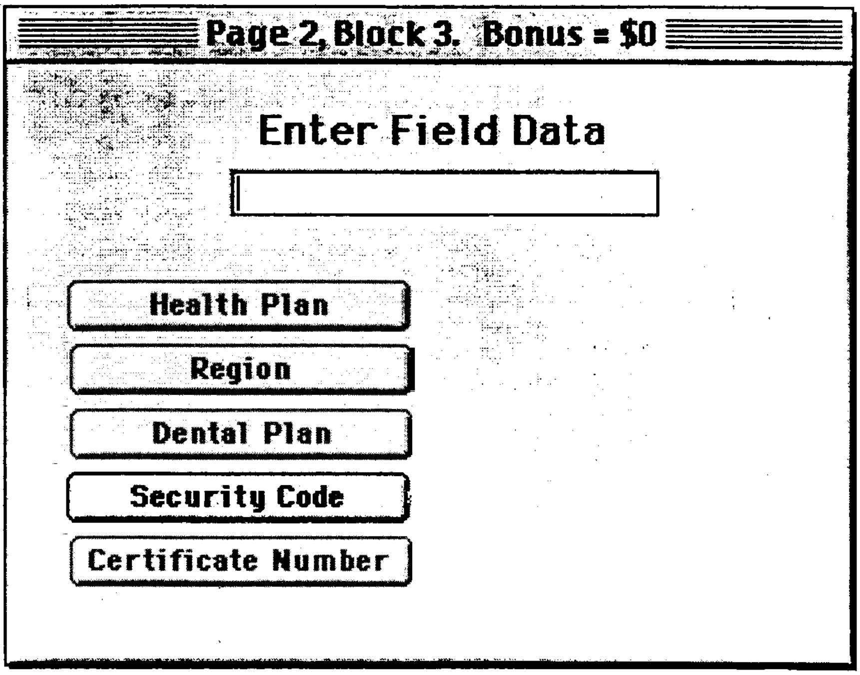


Figure 7: The computer display onto which the data were entered

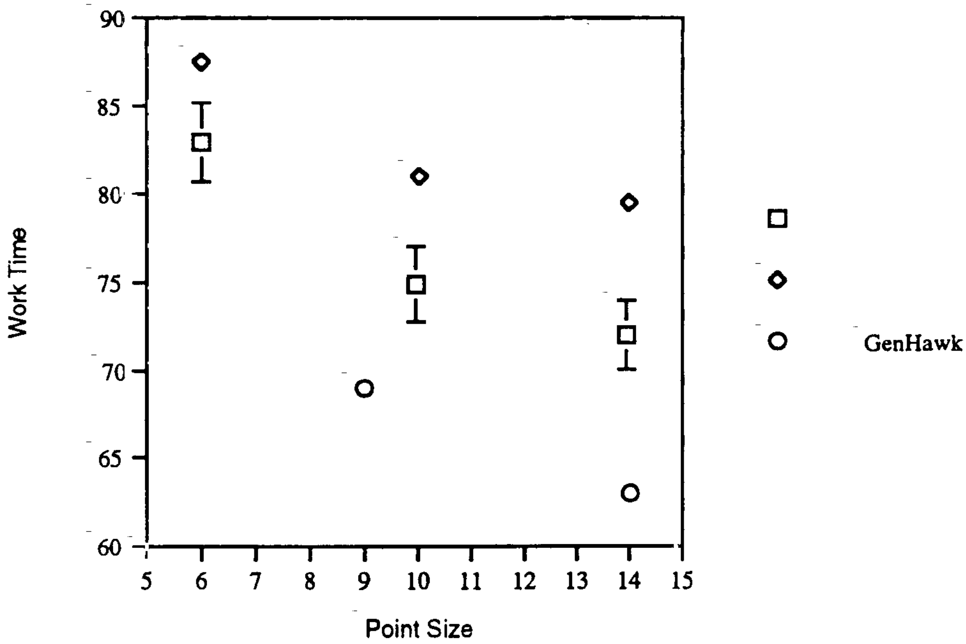
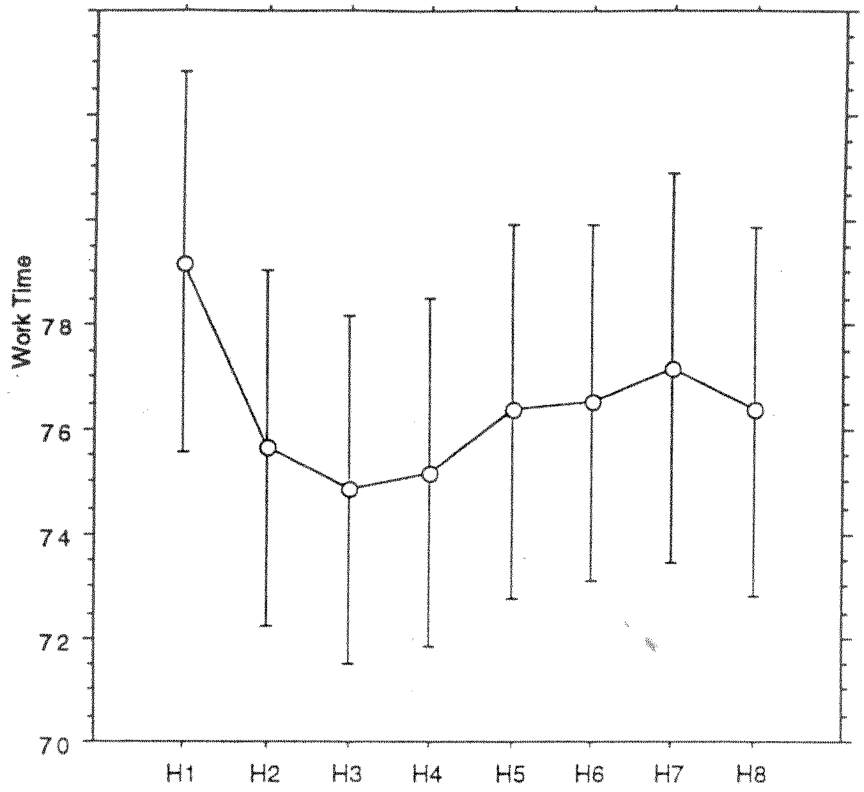


Figure 8: Mean work time (s) for the data-entry task for the three print sizes. The error bars indicate the 95% confidence interval. Also shown are the mean work times for the same data entry task taken from the second pilot study (Eklund, Boyce and Johnson, 1997) and from an earlier study using a different lighting installation (Boyce and Eklund, 1996).



igure 9: Mean work time (s) for the data-entry task for each hour of work. The error bars indicate the ;% confidence interval.

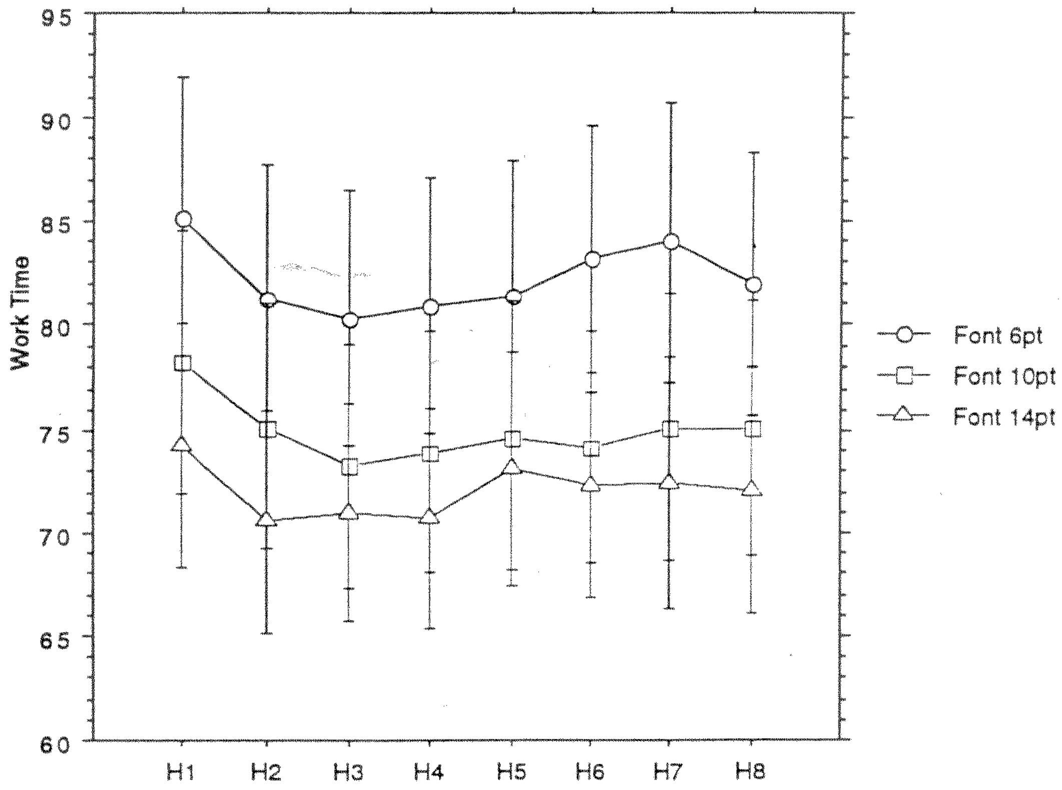


Figure 10: Mean work time (s) for the data-entry task, for the three print sizes, for each hour of work. The error bars indicate the 95% confidence interval.

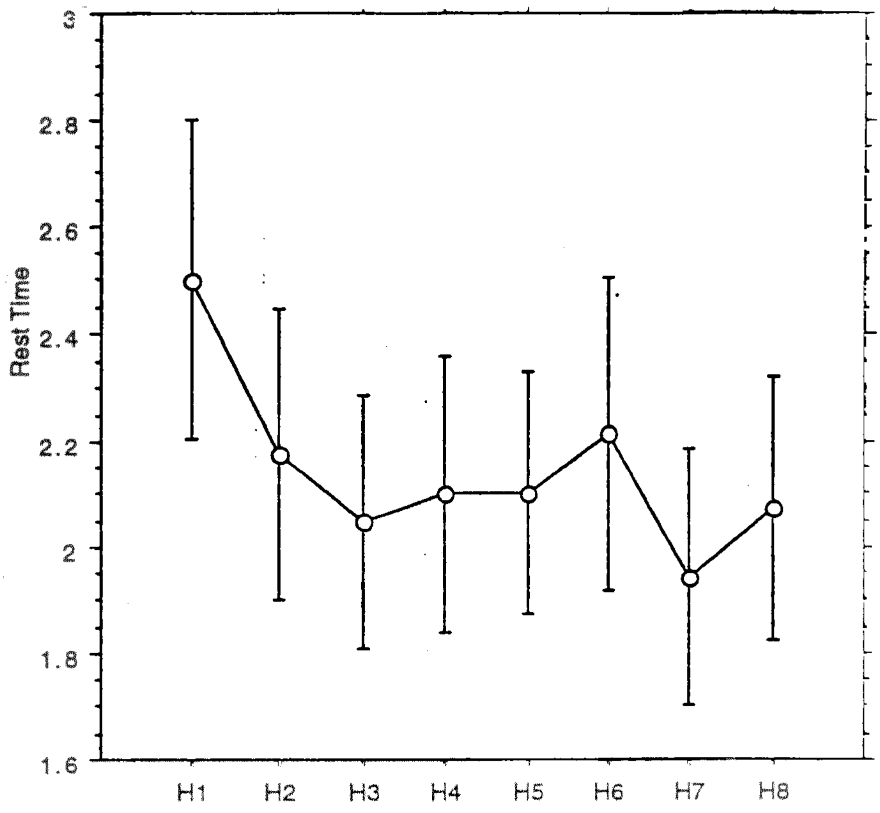


Figure 11: Mean rest times (s) for the data-entry task, for each hour. The error bars indicate the 95% confidence interval.

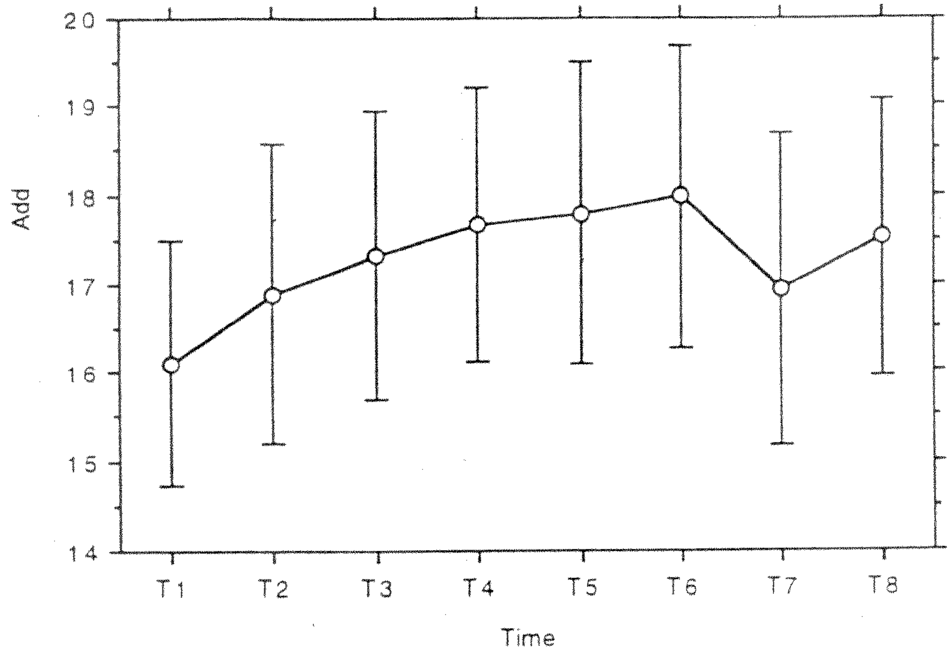


Figure 12: Mean time to solve a problem in the addition test, for each hour. The error bars indicate the 95% confidence interval.

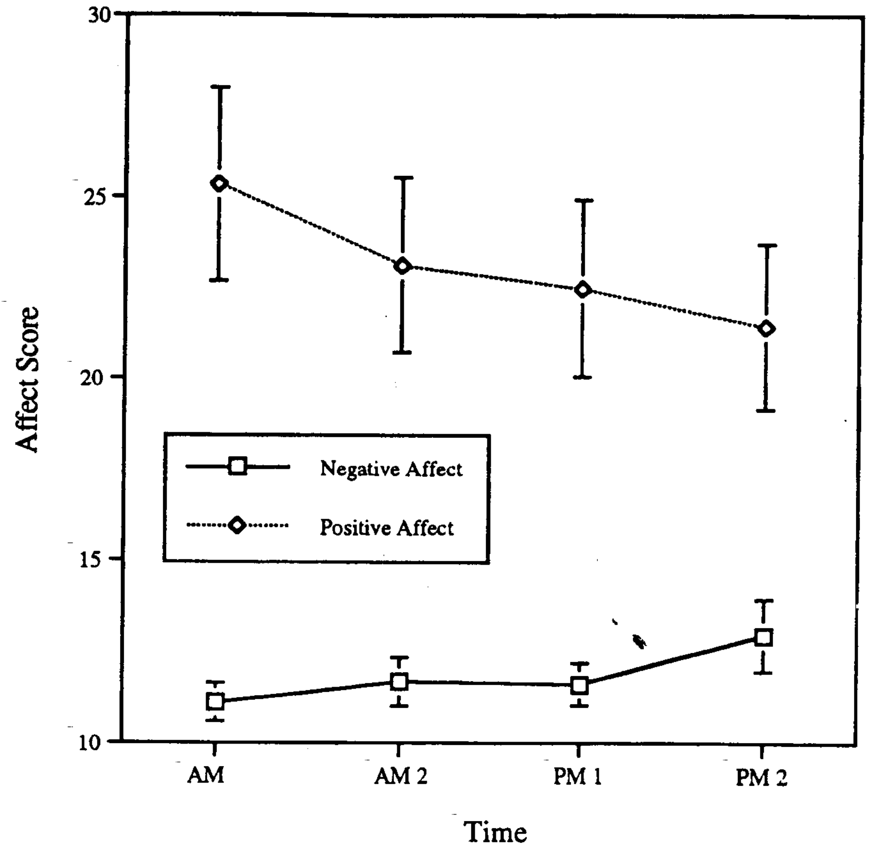


Figure 13: Mean positive and negative affect scores on the four occasions mood was measured. The error bars indicate the 95% confidence interval.

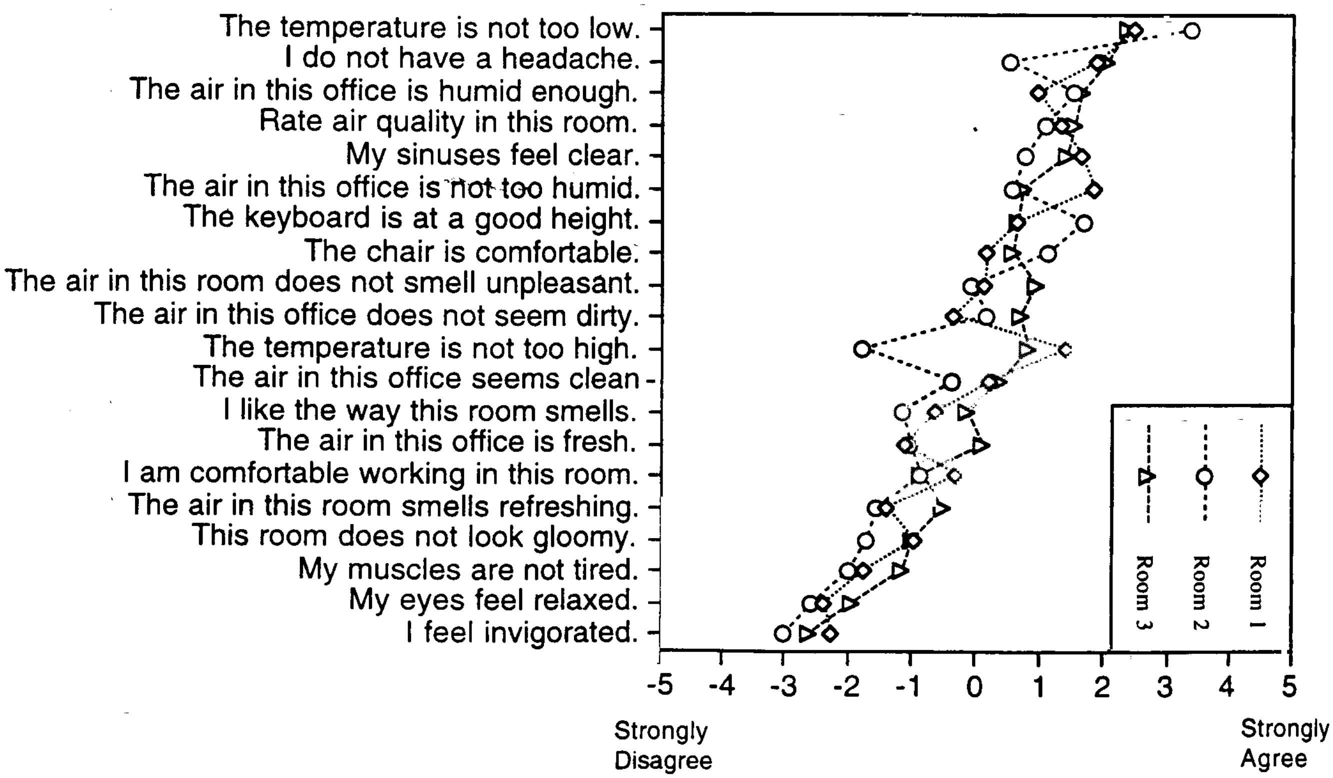


Figure 14: Mean ratings of agreement with each of the statements in the Room Assessment Questionnaire, for each room.

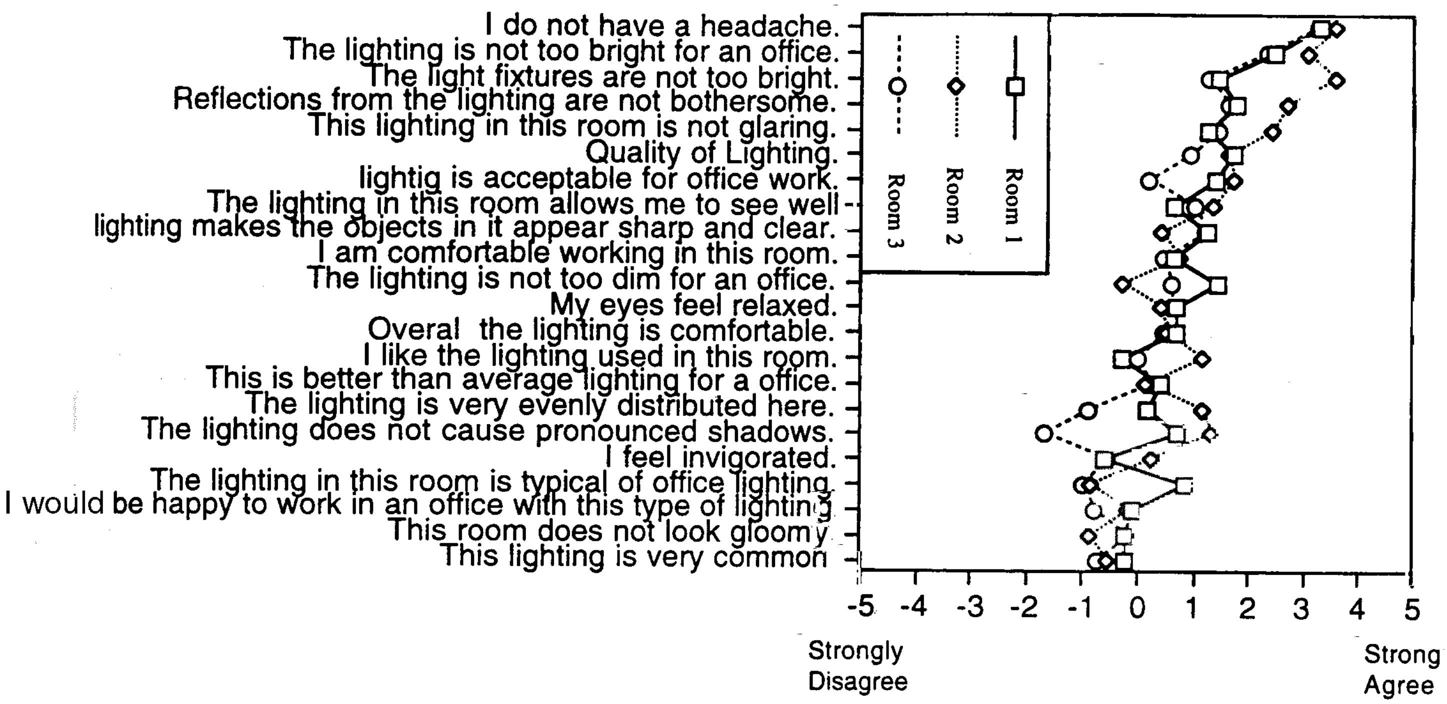


Figure 15: Mean ratings of agreement with each of the statements in the R)m Lighting Questionnair for each room.

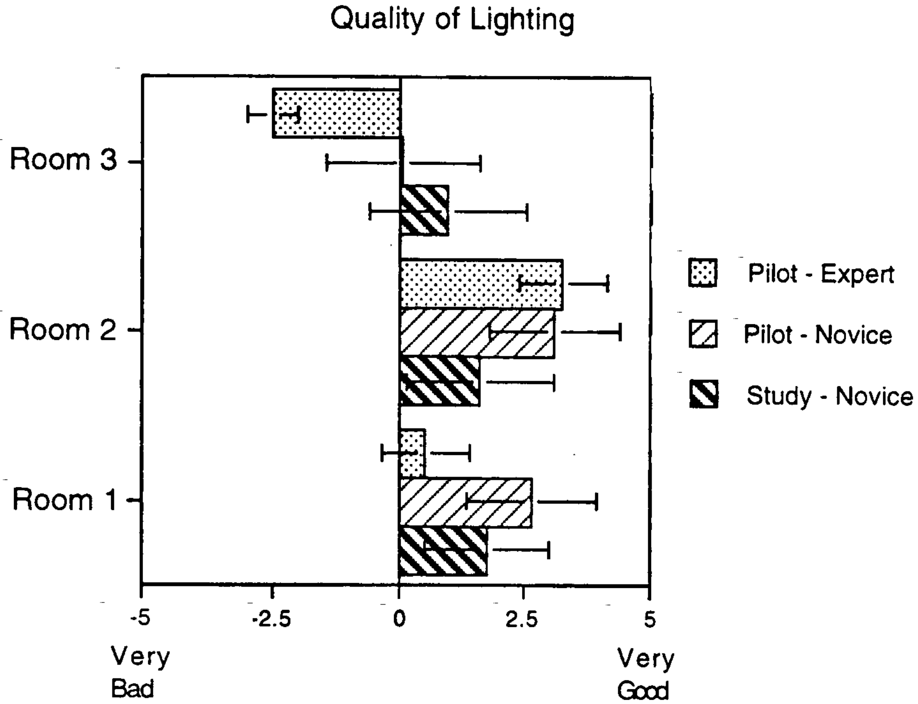


Figure 16: Mean ratings of the quality of the lighting given on a ten point scale with -5 = very bad and +5 = very good, for the three lighting installations. Means are given for subjects expert in lighting (Pilot Study 1), subjects naive in lighting (Pilot Study 1), and ten of the fifteen temporary office workers used i: this study. The error bars indicate the 95% confidence intervals.

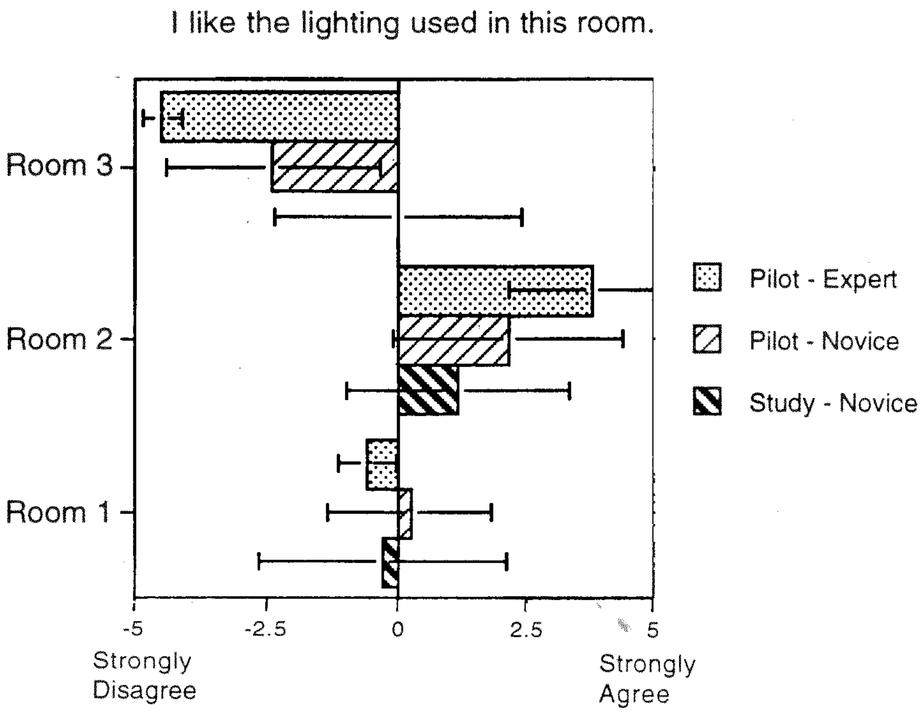


Figure 17: Mean ratings of agreement with the statement "I like the lighting used in this room" for the three lighting installations. Means are given for subjects expert in lighting (Pilot Study 1), subjects naive in lighting (Pilot Study 1), and ten of the fifteen temporary office workers used in this study. The error bars indicate the 95% confidence intervals.

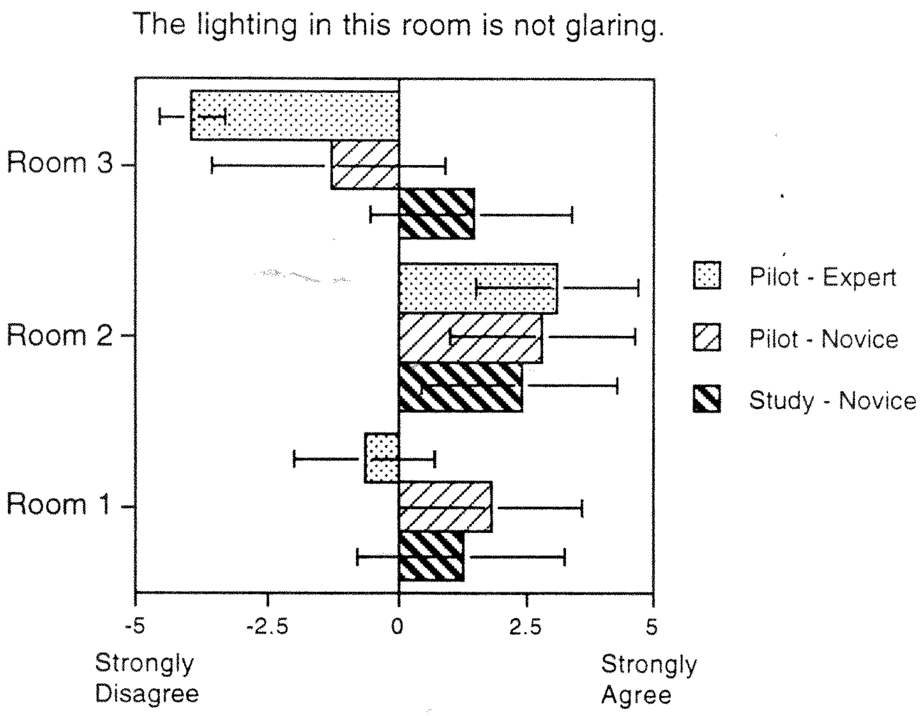


Figure 18: Mean ratings of agreement with the statement "The lighting in this room is not glaring" for the three lighting installations. Means are given for subjects expert in lighting (Pilot Study 1), subjects naive in lighting (Pilot Study 1), and ten of the fifteen temporary office workers used in this study. The error bars indicate the 95% confidence intervals.

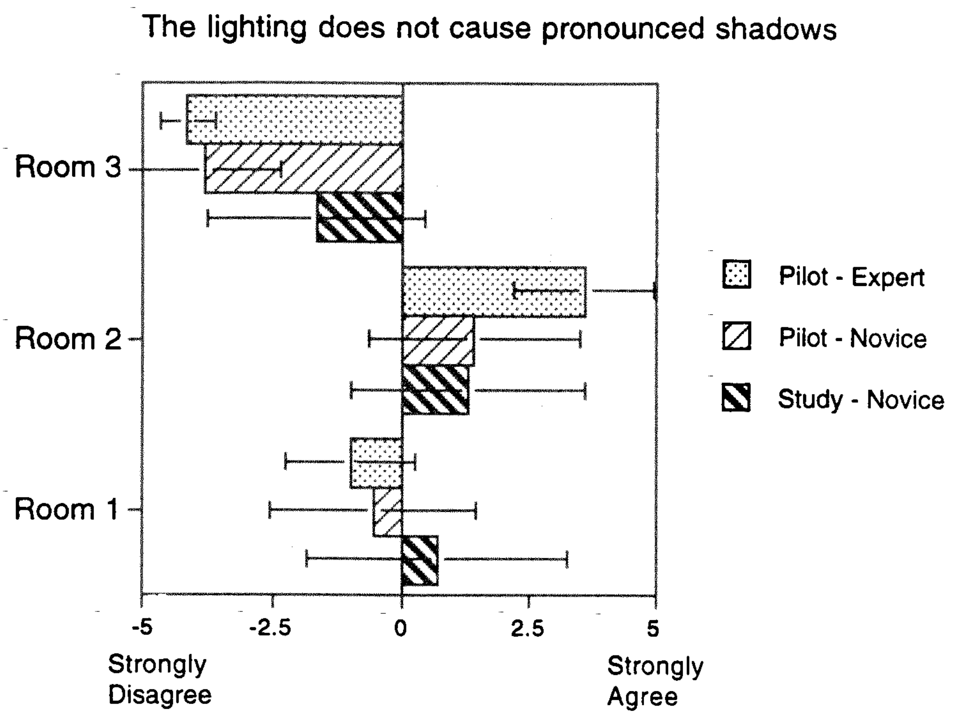


Figure 19: Mean ratings of agreement with the statement "The lighting does not cause pronounced shadows" for the three lighting installations. Means are given for subjects expert in lighting (Pilot Study 1), subjects naive in lighting (Pilot Study 1), and ten of the fifteen temporary office workers used in this study. The error bars indicate the 95% confidence intervals.

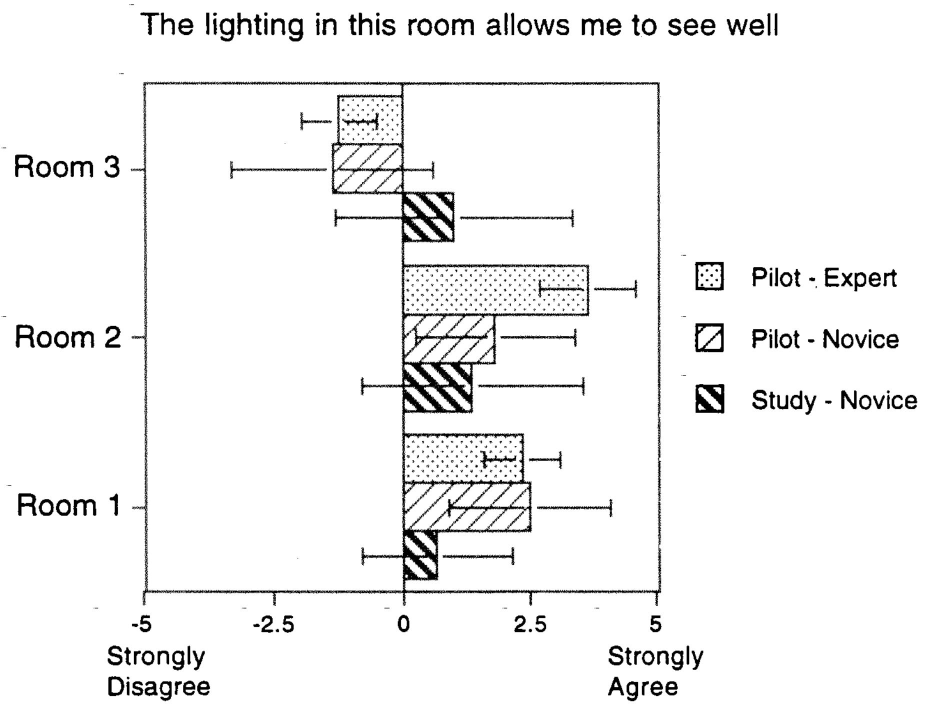


Figure 20: Mean ratings of agreement with the statement "The lighting in this room allows me to see well" for the three lighting installations. Means are given for subjects expert in lighting (Pilot Study 1), subjects naive in lighting (Pilot Study 1), and ten of the fifteen temporary office workers used in this study The error bars indicate the 95% confidence intervals.

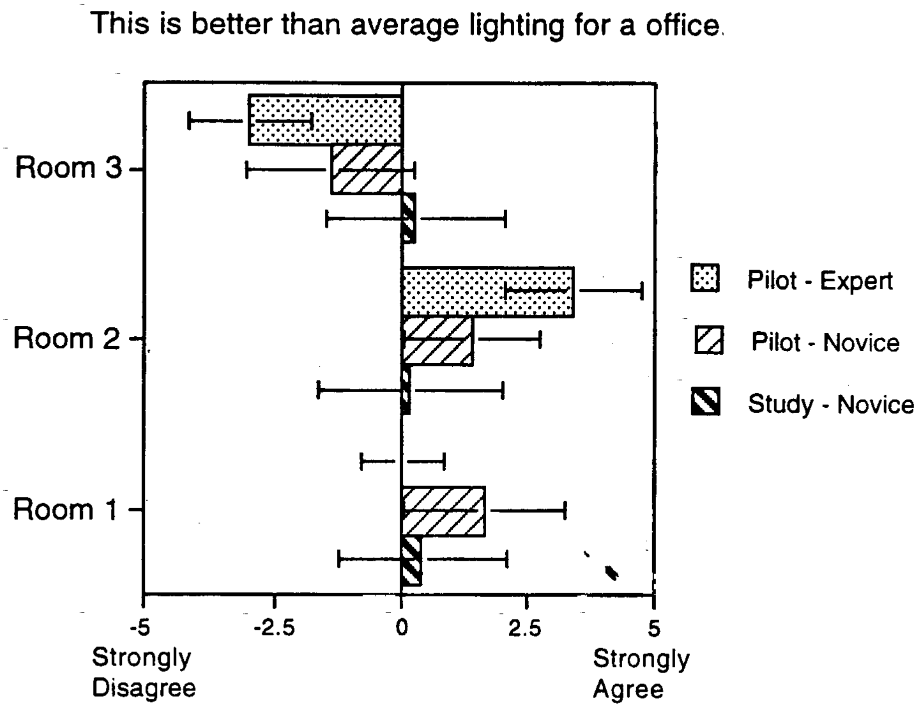


Figure 21: Mean ratings of agreement with the statement "This is better than average lighting for an office" for the three lighting installations. Means are given for subjects expert in lighting (Pilot Study 1), subjects naive in lighting (Pilot Study 1), and ten of the fifteen temporary office workers used in this study The error bars indicate the 95% confidence intervals.

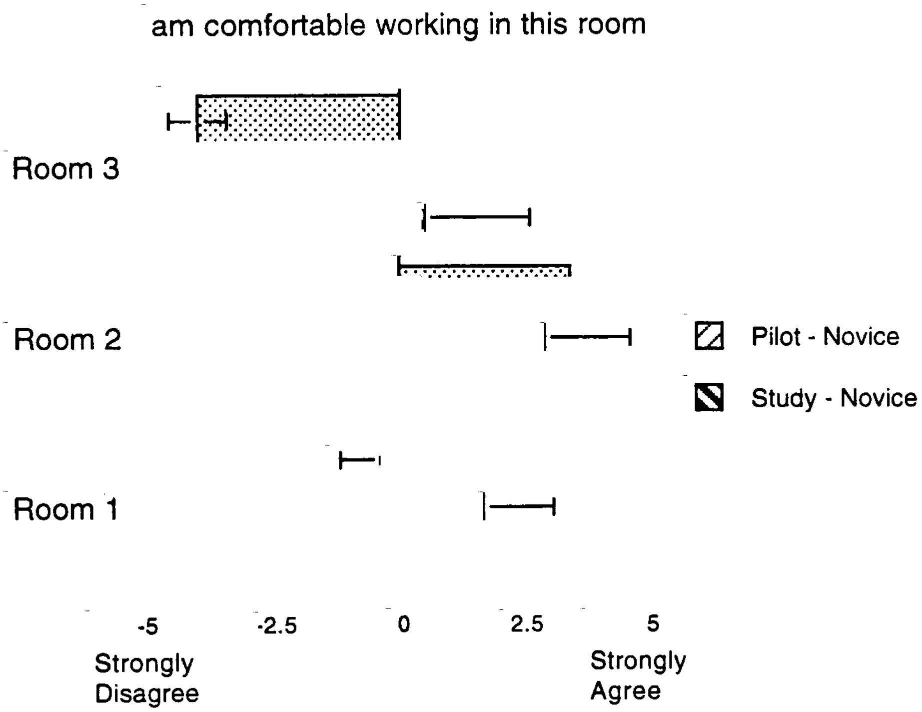


Figure 22: Mean ratings of agreement with the statement "I am comfortable working in this room" for the three lighting installations. Means are given for subjects expert in lighting (Pilot Study 1), subjects naive in lighting (Pilot Study 1), and ten of the fifteen temporary office workers used in this study. The error bars indicate the 95% confidence intervals.