



Performance and alertness on 8 h and 12 h rotating shifts at a natural gas utility

ROGER R. ROSA & MICHAEL H. BONNET

To cite this article: ROGER R. ROSA & MICHAEL H. BONNET (1993) Performance and alertness on 8 h and 12 h rotating shifts at a natural gas utility, Ergonomics, 36:10, 1177-1193, DOI: [10.1080/00140139308967987](https://doi.org/10.1080/00140139308967987)

To link to this article: <https://doi.org/10.1080/00140139308967987>



Published online: 31 May 2007.



Submit your article to this journal [↗](#)



Article views: 83



View related articles [↗](#)



Citing articles: 44 View citing articles [↗](#)

Performance and alertness on 8 h and 12 h rotating shifts at a natural gas utility

ROGER R. ROSA*

National Institute for Occupational Safety and Health, Division of Biomedical and Behavioral Science, Cincinnati, OH 45226, USA

and MICHAEL H. BONNET

Sleep Laboratory, Veterans Administration Hospital, Dayton, OH, USA

Keywords: Extended workdays; Work scheduling; Shiftwork; Human performance; Sleep.

An 8 h/5-7 day shift schedule was compared with a newly instituted 12 h/2-4 day schedule in this, our second worksite study of extended workshifts. Workers completed a performance/alertness test battery, and a questionnaire on sleep patterns and other personal habits, 2-4 times a week on all shifts. After 10 months adaptation to the 12 h shift schedule, there were decrements in performance/alertness attributable to the extra 4 h on the extended shift. There were also reductions in sleep across the workweek which were most apparent on 12 h night shifts. The results are consistent with our first worksite study of 12 h shifts and indicate extra caution should be exercised when scheduling critical activities for extended workshifts, especially extended night shifts.

1. Introduction

Several investigations have demonstrated the disruptive effects of night and shiftwork on various physiological, psychological, and social circadian rhythms (e.g., Johnson *et al.* 1981, Rutenfranz *et al.* 1977). These concerns, along with evolving industrial needs and demands for larger blocks of leisure time, have stimulated interest in alternatives to the common five-day workweek of 8 h per day (Colligan and Tepas 1986). Many alternative work schedules 'compress' the workweek into fewer days by adding more work hours per day (usually 10-12 h). For the shiftworker, the extra days off afforded by these schedules presumably allow larger blocks of time for recuperation from night work, for family and friends, and for other leisure pursuits. A disadvantage of such schedules, however, is the potential for excessive fatigue from long workshifts leading to safety or health hazards (Colligan and Tepas 1986, Hodge and Tellier 1975). In Singapore, for example, 12 h shift systems in several industries were abandoned after a brief trial because they were judged to have an adverse impact on worker health and social life (Kogi *et al.* 1989). The reported health factors associated with fatigue included insufficient sleep and weight loss by workers on 12 h shifts. Other factors potentially associated with fatigue included lower productivity, high turnover, and more frequent part-time work on off-duty days. Lees and Laundry (1989), on the other hand, reported no change in some health indices, and improvement in other indices (headaches, gastric upsets, and alcohol-related complaints), after a switch from an 8 h to a 12 h shift

*Requests for reprints should be sent to Roger R. Rosa, National Institute for Occupational Safety and Health, 4676 Columbia Pkwy, Cincinnati, OH 45226, USA.

system at a yarn manufacturing plant. Analyses of factors more directly associated with fatigue, however, were not reported.

Other researchers have attempted more direct assessments of fatigue associated with 12 h shifts through analysis of job factors, or administration of laboratory-type performance tests and self-report scales. Kelly and Schneider (1982) conducted a complex reliability assessment of nuclear power plant operators and concluded that 12 h shifts would produce a 70% increase in the probability of a public accident. With a standard nursing-care job analysis, Mill *et al.* (1983) concluded that 12 h shifts had no adverse impact. Todd *et al.* (1989), on the other hand, concluded that *overall* nursing care was adversely affected by 12 h shifts, while direct physical care was not affected. In his analysis of accidents in truck drivers, Hamelin (1987) reported that accident risk was particularly high after 11 h of work, especially if work occurred at night.

Assessment of extended workshifts with standard laboratory performance tasks and self-report scales has indicated some improvements and some decrements associated with these shifts. Volle *et al.* (1979) reported decreased grip strength and decreased critical flicker fusion frequency (CFF) in a factory on 10 h shifts compared with a similar factory on 8 h shifts. Peacock *et al.* (1983) reported improved subjective alertness, sleep, and cardiovascular fitness in police officers after a switch from 8 h to 12 h shifts, and no effect on CFF or grammatical reasoning performance. Mills *et al.* (1983) reported increased subjective fatigue across a 12 h shift, and more frequent errors in a grammatical reasoning task and a medical record reviewing task. No comparisons, however, were made to 8 h shifts. Daniel and Potasova (1989) concluded that, overall, workers on 12 h shifts performed more poorly on several cognitive and psychomotor tasks when compared with 8 h shift workers. They suggested, however, that differences in the capabilities of the workers selected for each shift system contributed to these results.

In our own worksite studies of 12 h shifts, we elected to assess worker fatigue with a battery of standard performance tests and self-report scales. (See Rosa and Colligan 1988, and Rosa *et al.* 1985, for descriptions of the test battery and laboratory demonstrations of its sensitivity.) The availability of portable microcomputers has allowed us to place this test battery at the worksite for several weeks to track worker alertness and fatigue on all shifts.

Two worksite evaluations of extended workshifts now have been completed. In our first worksite study (Rosa *et al.* 1989), the battery was used to compare an 8 h/3-shift rotation with a 12 h/2-shift rotation in a group of control-room operators at a continuous processing plant. After seven months adaptation, decreased test battery performance and increased subjective fatigue were observed on 12 h shifts as compared to 8 h shifts. Daily sleep logs kept by each operator indicated a mild sleep debt by the end of the 12 h/3-4 day workweek. Test battery scores, however, did not deteriorate across the workweek, indicating that the shorter workweek compensated somewhat for the longer workday. After 3-5 years on the 12 h shift schedule, declines in alertness with time on-shift and reductions in total sleep time still were apparent, and few improvements were observed relative to the 7 month test phase (Rosa 1991).

The present report documents our second worksite investigation at a large natural gas utility. In a trial attempt to make shiftwork more tolerable, certain worksites within the utility system were temporarily switched from a 5-7 day workweek with three 8 h rotating shifts to a 2-4 day workweek with two 12 h rotating shifts. After 6-10 months of adaptation to the new shift schedule, we compared 8 h to 12 h

shifts with test battery measures of performance and subjective alertness/fatigue, and with a daily questionnaire quantifying subjective reports of sleep patterns and other personal habits potentially affected by shiftwork.

2. Method

2.1. Subjects

All workers were male, ranging in age from 25 to 59 years, with a mean age of 37 years. Testing occurred at Gas Control and two gas storage-transmission stations (Stations 1 and 2). Workers at Gas Control were tested on both 8 h and 12 h shifts. Workers at Station 1 were tested only on 8 h shifts, and workers at Station 2 were tested only on 12 h shifts.

Gas Control workers monitored control panels and computer terminals to make decisions about gas supply and allocation. Physical demands were light. Workers at Stations 1 and 2 operated, repaired, and maintained equipment such as gas engines, turbines, compressors, and pumps. According to their position descriptions, the gas company rated the physical demands at Stations 1 and 2 as 'very heavy'. They stated further that these jobs required frequent lifting and carrying of objects weighing 22 kg or more (1/3 to 2/3 of the time), and occasional lifting of objects in excess of 50 kg (up to 1/3 of the time).

2.2. Work schedules

The 8 h and 12 h rotating shift schedules are illustrated in figure 1. Under the 8 h schedule, individuals worked seven days in a row. Under the 12 h schedule, they typically worked 2 or 3 (and occasionally 4) days in a row. Shift times for Gas Control and Station 1 on the 8 h schedule were: 07:00 to 15:00 (day), 15:00 to 23:00 (evening), and 23:00 to 07:00 (night). Shift times for Gas Control on the 12 h schedule were: 09:30 to 21:30 (day) and 21:30 to 09:30 (night). Shift times for Station 2 on the 12 h schedule were: 08:00 to 20:00 (day) and 20:00 to 08:00 (night).

8-HOUR SCHEDULE

```
DAY  M T W T F S S M T W T F S S M T W T F S S
SHIFT D D D D . . . E E E E E E . . N N N N N N . . D D D
```

12-HOUR SCHEDULE

```
DAY  M T W T F S S M T W T F S S M T W T F S S
SHIFT D D . . N N N . . D D . . N N . . D D D . . N N . . .
```

D=DAY E=EVENING N=NIGHT .=REST

Figure 1. 8 h and 12 h rotating shift schedules for a typical crew.

2.3. Test schedule

Figure 2 shows the timing of the test periods at each of the worksites. As shown in the figure, five weeks (more than one shift cycle) of testing on the 8 h schedule began in October at Gas Control and Station 1. Ten months later, workers were tested for five weeks on the 12 h schedule at Gas Control, and re-tested on the 8 h schedule at Station 1. Workers at Station 2 were tested for 5–10 weeks after both six and ten

months of experience on the 12 h schedule. (Logistical problems precluded testing workers at Station 2 on the 8 h schedule before their change to the 12 h schedule.)

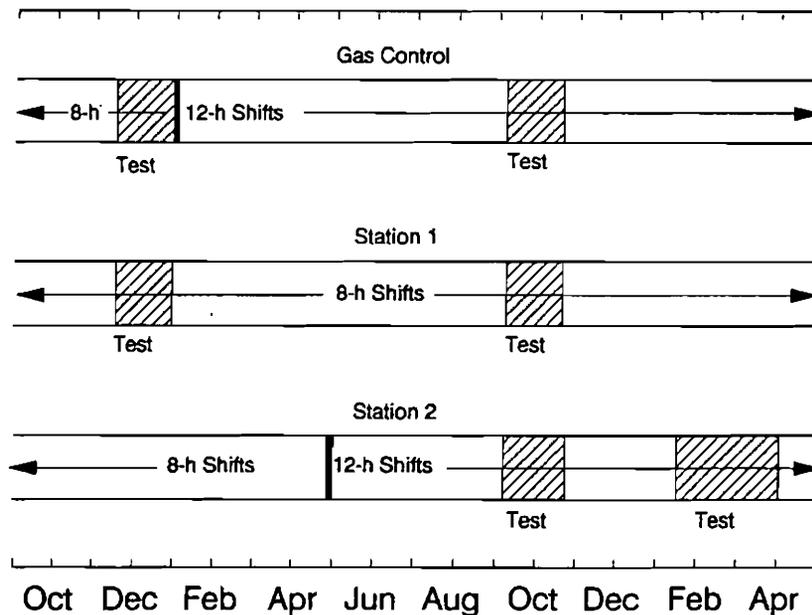


Figure 2. Test periods at the three worksites.

2.4. Measurements

2.4.1. *Performance tests and self-report scales:* The computerized test battery is a flexible system designed to evaluate a range of psychological functions including cognitive, perceptual-motor, and motor skills, as well as subjective alertness or fatigue. The tasks are brief in order to minimize interference with the worker's regular job. Any number of tests can be selected as general performance measures or as analogues to specific job demands. Details about the components of the battery are reported in Rosa and Colligan (1986, 1988) and Rosa *et al.* (1985).

As in our previous worksite study (Rosa *et al.* 1989, Rosa 1991), management would permit access to the worksite only if testing minimally interfered with the workers' job responsibilities. For this reason, test battery sessions were limited to a maximum of 20 min. To permit comparison with our first worksite study, we used the same set of measures.

Cognitive abilities were evaluated with a 3 min test of mental arithmetic (a digit addition task adapted from Williams and Lubin 1967), and a 4 min dual task consisting of grammatical reasoning combined with auditory reaction time. The reasoning component of the dual task was based on Baddeley's (1968) test except that response time was measured for each conditional statement rather than for the entire test. Four minutes (60 trials) of simple auditory reaction time (adapted from Lisper and Kjellberg 1972), and a 2 min hand steadiness task were included as more general indices of alertness and fatigue. Self-reported measures of sleepiness (Hoddes *et al.* 1973), positive mood (mean of the attentiveness, energy, stress, boredom, and irritability scales of Thackray *et al.* 1977), gastro-intestinal distress, and workload

also were taken. Subjective workload was rated on a numeric scale for the 3 h period preceding administration of the test battery.

2.4.2. Daily sleep and habits questionnaire: A questionnaire containing items about sleep and other personal factors potentially affected by shiftwork (see Johnson *et al.* 1981, Rutenfranz *et al.* 1977, Scott 1990, for reviews) was given at the beginning of every workshift. The workers were queried on their times of retiring and arising (including nap times), sleep latency, number of awakenings, time awake during the sleep period, depth of sleep, quality of sleep, and feelings of drowsiness upon awakening for the 24 h period immediately preceding the workshift. The workers also reported use of sleep aids (including alcohol, and prescription and non-prescription medication), and noted personal schedule adjustments attributable to shiftwork, including adjustments of mealtimes, exercise time, and absenteeism.

2.5. Participation

2.5.1. Test battery: Five workers at Gas Control and one worker who transferred from Station 1 to Station 2 were tested on 8 h shifts and subsequently on 12 h shifts. Nine workers at Station 1 were tested only on 8 h shifts. Twelve workers at Station 2 were tested only on 12 h shifts.

Workers completed the test battery at various times on all shifts under both the 8 h and 12 h schedules. Our original test schedule was designed so that testing would occur at the beginning, middle, and end of each workshift or workweek. In practice, however, participation rates varied because of job demands and worker preferences. Frequency of participation over the entire test period ranged from 5 to 37 test battery sessions per person (mean = 13, SD = 8).

2.5.2. Daily sleep and habits questionnaire: Nine workers from Gas Control and one worker who transferred from Station 1 to Station 2 completed daily questionnaires on both 8 h and 12 h shifts. Nine workers from Station 1 completed questionnaires only on 8 h shifts, and twelve workers from Station 2 completed questionnaires only on 12 h shifts. The questionnaires were given at the beginning of the workshift. Frequency of participation over the entire test period ranged from 5 to 30 questionnaires completed person (mean = 12, SD = 7).

Inspection of the number of test battery sessions and questionnaires completed for each 8 h and 12 h shift, and each day of the workweek, indicated that participation was more frequent on the first two 12 h workdays. Across time of day, test battery participation was more frequent in the evening and at night compared to late morning and early afternoon. These rates, however, were similar for both the 8 h and 12 h schedules.

2.6. Procedure

Two Kaypro II portable computers were distributed in accessible locations at each worksite. The computers controlled the instructions for, and presentation of, all test battery performance tasks and self-report scales (except the daily sleep and habits questionnaire). A brief introduction to the purpose of the study, the computers, and the test battery was the only experimenter intervention before testing began. Except for a brief training session, the experimenters were not present during data collection. To initiate the test, the subject sat at the computer console and typed a simple command.

2.7. Test battery data analysis

2.7.1. *Data transformations:* Before analyses, several dependent variables in the test battery were transformed to approximate the normal distribution (see Myers 1979: 72–73). Grammatical reasoning response time, dual reaction time, and simple reaction time were transformed to their inverses. Proportion scores, such as grammatical reasoning errors, dual and simple reaction time misses, and hand steadiness time off-target, were transformed to the arcsine of their square roots.

2.7.2. *Analyses of variance:* Average scores for 8 h and 12 h shifts, and average scores across the workweek, first were compared with analyses of variance (ANOVAs). Separate sets of repeated-measures and between-groups ANOVAs were computed, respectively, for subjects tested during both the 8 h and 12 h shift schedules (Gas Control), and subjects tested only during one of these schedules (Stations 1 and 2).

The first repeated-measures ANOVA compared scores averaged for each subject within each shift (8 h day, evening, and night, or 12 h day and night). A one-way ANOVA (across five shifts) was performed on these scores. *Post hoc* comparisons among significant sets of means were computed using the Newman–Keuls procedure (Myers 1979). Two other repeated-measures ANOVAs tested differences across the workweek and shifts *within* a schedule. To maximize the cell frequencies in these ANOVAs, the scores were first divided into early (8 h days 1–2, 12 h day 1), middle (8 h days 3–5, 12 h day 2) and late (8 h days 6–7, 12 h days 3–4) segments of the workweek. The scores for each subject were then averaged for each shift within these segments. For the 8 h schedule, a 3 (shifts) by 3 (week segments) ANOVA was performed on these scores. For the 12 h schedule, a 2 (shifts) by 3 (week segments) ANOVA was performed on the scores.

The first between-groups ANOVA also compared scores averaged for each subject within each shift (8 h day, and night, or 12 h day and night). In the between-groups analyses, however, different subjects contributed scores to the 8 h and 12 h shifts. A one-way ANOVA (across the five shifts) was performed on these scores using the interaction of subjects by shifts nested within groups as the error term. *Post hoc* comparisons among significant sets of means were computed using the Newmann–Keuls procedure (Myers 1979). Two other repeated-measures ANOVAs were performed on these subjects' scores to test differences across the workweek and shifts *within* a schedule. The procedures described above were used to divide the workweek, and analogous 3 (8 h shifts) by 3 (week segments), and 2 (12 h shifts) by 3 (week segments) ANOVAs were calculated.

Because of unequal cell frequencies, least-squares regression solutions were computed for all ANOVAs (SAS General Linear Models Procedure, SAS Institute 1985). Consequently, the *F*-ratios from these analyses are not exact and the degrees of freedom do not conform to a balanced design. An alpha level of $p < 0.05$ was adopted as the criterion for statistical significance in all analyses.

2.7.3. *Stepwise regression analysis:* Multiple regression analyses were performed to explore parameters in the data which were obscured in the average scores of the ANOVAs. In particular, the regression approach allowed a simultaneous, fine-grained analysis of the number of hours worked, time of day or circadian rhythm, and consecutive days worked. Since these are confounded variables, the regression approach was the only way to test their concerted effects while simultaneously controlling the degree of correlation among them. Multiple regression also allowed

the partitioning of variance attributable to practice. Statistical control of practice effects was necessary because there was no opportunity for extensive practice on the performance tasks before the first test phase. Consequently, the subjects were not at an asymptotic level of performance before critical testing began.

A final reason for the regression approach was variable rates of participation, which resulted in unequal cell frequencies in the ANOVAs. These inequalities can drastically reduce the power of the ANOVAs and increase the probability of Type II error (see Myers 1979). By combining all subjects' observations and using an omnibus error term in the regression analysis, we maximized power and reduced the probability of Type II error. This procedure, however, violated the statistical assumption of independence of observations by mixing both 'between-subjects' and 'within-subjects' observations. Non-independence can introduce positive bias in the significance tests, and might have increased the probability of Type I error (see Myers 1979). We feel that the information gained from this approach, however, outweighs the increased probability of Type I error.

A separate regression equation was calculated for each dependent variable in the test battery. An alpha level of $p < 0.05$ was adopted as the criterion for statistical significance.

The following variables, assumed to affect performance or alertness, were entered stepwise into each regression equation:

- (1) work schedule (dummy codes for 8 h and 12 h schedules);
- (2) number of hours worked or time on-shift;
- (3) consecutive days worked;
- (4) circadian rhythm; (This variable was determined by the 'cosinor' curve fitting method of Naitoh *et al.* [1985]. With this method, a sinusoidal 24 h rhythm is calculated from a weighted combination of the sine and cosine of time.)
- (5) the two-way interactions among the above variables; (separate terms were entered for interactions with the sine and cosine of time.)
- (6) subjective workload;
- (7) practice; (To partition variance attributable to practice, both linear and curvilinear practice effects were entered in the regression equation. When significant practice effects were obtained, this procedure de-trended the data and clarified effects from other variables. Two practice curves were entered in the regression analyses to maximize the possibility of obtaining a "best-fit" practice effect. The stepwise regression approach determined statistically the practice effect [or effects] which partitioned the greatest amount of variance. The linear effect entered in the regression equation was the cumulative number of times the test battery was taken at the time of a particular observation [i.e., number of repetitions]. The curvilinear effect was the natural logarithm of this value.)
- (8) dummy codes for subjects also were included in the model to partition variance attributable to individual differences. (These codes were entered [i.e., 'forced'] into the model before stepwise analysis of the other effects.)

A Total of 332–383 observations (depending on missing values) from three worksites were entered in the regression equations to maximize statistical power.

2.8. Questionnaire data analysis

The same ANOVA procedures applied to the test battery were used to compare

average questionnaire scores across 8 h and 12 h shifts, and average scores across the workweek. Separate sets of repeated-measures and between-groups ANOVAs were computed, respectively, for subjects tested during both the 8 h and 12 h shift schedules (Gas Control), and subjects tested only during one of the schedules (Stations 1 and 2).

3. Results

3.1. Test battery

3.1.1. *Analyses of variance across shifts:* Means for each statistically significant test battery score across 8 h and 12 h shifts are shown in Table 1 for the repeated measures analyses and the between-groups analyses. Corresponding *F*-ratios and *post hoc* comparisons among means also are shown.

The repeated-measures ANOVAs of average shift scores indicated significantly faster grammatical reasoning response times, and more frequent simple reaction time misses, during 12 h shifts compared to 8 h shifts. The workers also reported feeling most sleepy on 8 h shifts. No significant differences across shifts were observed for grammatical reasoning errors, digit addition total score or errors, dual reaction time RT or misses, simple reaction time RT, hand tremor, or positive mood.

The between-groups ANOVAs of average shift scores indicated significantly more frequent simple reaction time misses, and more hand tremor, during 12 h shifts compared to 8 h shifts. The workers reported the highest subjective sleepiness on 8 h night shift, and higher subjective workloads on 12 h shifts compared to 8 h shifts.

In comparing the results of the repeated-measures and the between-groups analyses, the most consistent findings in table 1 were more reaction time signals missed on 12 h shifts, and higher subjective sleepiness on 8 h night shift.

3.1.2. *Analyses of variance across workdays:* Although some statistically significant results were obtained, few consistent findings were observed when the results from subjects tested during both the 8 h and 12 h work schedules were compared with subjects tested only during one of the work schedules.

3.1.3. *Multiple regression analyses:* Table 2 lists significant effects from the regression analyses for each dependent variable. A brief description of each effect also is listed in the table, along with a change or difference score derived from the associated beta weight (or regression parameter). The difference score characterizes the effect when the variance from all other significant effects is held constant (i.e., partitioned).

As shown in Table 2, the regression analyses revealed circadian variations in most of the test battery measures. Practice effects (not shown) were observed in the grammatical reasoning, digit addition, and dual reaction time tasks. Partitioning of variance attributable to practice revealed several effects associated with the 12 h shift schedule. Grammatical reasoning errors, for example, were more frequent on the 12 h schedule. Digit addition performance was best on 8 h evening shift, but addition performance was better on 12 h night shift compared to 8 h night shift. Consistent with the ANOVAs, simple reaction time signals were missed more frequently on 12 h shifts.

The interaction of time on-shift with the sine of time was significant for simple reaction time performance, self-reported sleepiness, and self-reported positive mood. These interactions suggest that ascending circadian arousal during the day

Table 1. Repeated measures and between-groups means and ANOVA results for test battery scores across 8 h and 12 h shifts.

		Shift					F (p)	Post hoc tests
		Day 1	8 h Eve 2	Night 3	12 h Day 4	Night 5		
REPEATED MEASURES (df=4,15)								
<i>Dependent variable</i>								
<i>Performance</i>								
Grammatical reasoning RT (sec)	M	4.63	4.06	4.46	3.77	4.01	3.05	1 > 4
	SD	0.31	0.69	0.50	0.49	0.45	= 0.05	1+2+3 > 4+5
Simple RT misses (%)	M	0.2	0.7	1.2	4.6	2.7	4.46	4 > 1,2,3
	SD	0.2	0.4	1.3	3.6	2.2	< 0.01	4+5 > 1+2+3
<i>Self-report</i>								
Sleepiness (1=least, 7=most)	M	3.34	2.22	3.48	2.33	3.01	7.30	3 > 2,4
	SD	1.16	0.52	0.87	0.68	0.86	< 0.01	1 > 2,4
BETWEEN GROUPS (df=4,47)								
<i>Dependent variable</i>								
<i>Performance</i>								
Simple RT misses (%)	M	0.8	0.6	0.5	2.6	1.3	4.16	4 > 1,2,3
	SD	0.6	0.5	0.8	2.2	1.1	< 0.01	4+5 > 1+2+3
Hand tremor (% Time Off-Target)	M	0.4	2.0	3.5	11.3	7.4	6.28	4 > 1,2,3
	SD	7.6	3.1	6.3	9.1	5.9	< 0.01	4+5 > 1+2+3
<i>Self-report</i>								
Sleepiness (1=least, 7=most)	M	2.23	2.22	3.32	2.10	2.86	3.67	3 > 1,2,4,5
	SD	0.81	0.59	0.97	0.83	1.04	< 0.02	
Workload (1=low, 9=high)	M	3.31	4.17	3.99	5.24	4.70	3.64	5,4 > 1
	SD	1.39	0.99	1.66	1.22	0.85	< 0.02	5+4 > 1+2+3

Table 2. Significant regression effects from the fatigue test battery. The intercept is shown in parentheses beside each variable name. The significance level of each regression parameter is shown in parentheses beside each parameter name.

<i>Dependent variable Significant effects</i>	<i>Nature of the effect</i>
<i>Grammatical reasoning RT (5.29 sec) Workday ($p < 0.02$) Sine of time ($p < 0.001$)</i>	0.08 s faster per day. 0.39 s peak-to-peak amplitude, slowest at 06:00. (<i>Note:</i> This effect is modulated by the significant cosine interaction for this variable.)
<i>Grammatical reasoning errors (32%) 12 h schedule ($p < 0.01$)</i>	9% more errors on 12 h shifts.
<i>Digit addition total (45.38 correct) Time on-shift by cosine of time ($p < 0.02$)</i>	Most correct on 8 h evening shift, more correct on 12 h night than 8 h night shift.
<i>Digit addition errors (4.55%) 12 h schedule by cosine of time ($p < 0.01$)</i>	Circadian variation on 12 h shifts (3% peak-to-peak amplitude), fewest errors at 24:00 h.
<i>Dual reaction time (417 msec) Cosine of time ($p < 0.03$)</i>	29 ms peak-to-peak amplitude, fastest at 24:00 h.
<i>Dual reaction time misses (0.24%) None</i>	
<i>Simple auditory reaction time (123 ms) Time on-shift by sine of time ($p < 0.001$)</i>	Slowest reaction time at the end of 12 h night shift.
<i>Subjective workload ($p < 0.01$)</i>	Slower at higher workloads.
<i>Simple reaction time misses (0.02%) 12 h schedule ($p < 0.001$)</i>	1.1% more misses on 12 h shifts.
<i>Hand steadiness (6.24%) Workday ($p < 0.02$) 12 h schedule by cosine of time ($p < 0.04$)</i>	0.5% more steady per day. Circadian variation on 12 h shifts (3.7% peak-to-peak), most steady at 24:00.
<i>Subjective sleepiness (2.46; 1 = least sleepy, 7 = most sleepy) Time on-shift ($p < 0.01$) Cosine of time ($p < 0.001$)</i>	0.04 units more sleepy per hour. 0.70 units peak-to-peak amplitude, most sleepy at 24:00. (<i>Note:</i> This effect is modulated by the significant sine-of-time interaction for this variable.)
<i>Time on-shift by sine of time ($p < 0.001$)</i>	Increased sleepiness with time on-shift, most sleepy on 12 h night shift. (<i>Note:</i> This effect is modulated by the significant effect for cosine of time for this variable.)
<i>Positive mood (6.35; 1 = least, 9 = most) Time on-shift ($p < 0.02$) Cosine of time ($p < 0.002$)</i>	0.03 units less positive per hour. 0.49 units peak-to-peak amplitude, least positive at 24:00. (<i>Note:</i> This effect is modulated by the significant sine-of-time interaction for this variable.)

Table 2.—continued.

<i>Dependent variable</i>	<i>Nature of the effect</i>
<i>Significant effects</i>	
12 h schedule by workday	0.23 units less positive per day on 12 h shifts.
Time on-shift by sine of time ($p < 0.001$)	Less positive with time on-shift, least positive on 12 h night shift. (<i>Note:</i> This effect is modulated by the significant effect for cosine of time for this variable.)
<i>Gastro-intestinal distress</i> (1.43; 1=least, 4=most)	
Time on-shift by cosine of time ($p < 0.03$)	increased distress with time on-shift, most distress at 24:00, more distress on 12 h night shift than 8 h night shift.

counteracts some of the fatigue resulting from hours worked, and descending circadian arousal at night potentiates fatigue resulting from hours worked. (For the self-report measures, the phase and amplitude of this effect are modulated by a significant effect for cosine of time.)

To illustrate the nighttime effects, figure 3 shows predictive functions for simple auditory reaction time and subjective sleepiness for the 8 h and 12 h night shifts. Values in this figure were calculated from the multiple regression equations by inserting the appropriate number of hours worked for parameters involving time on-shift, and the appropriate transformation of clock time for parameters involving the sine or cosine of time. (Subjective workload was held constant in the reaction time equation.) As shown in figure 3, workers would be expected to report feeling sleepier, and their reaction time would be expected to be slower, as the night shift progressed. Because of more hours worked, these decrements would be expected to be greater during the 12 h night shift than during the 8 h night shift. The same pattern of decrements would be expected in the positive mood measure (not shown in figure 3).

Across the workweek, the regression analysis indicated improvements in performance in the grammatical reasoning and hand steadiness tasks, and a decline in performance in the reaction time task. These changes were not specific to either work schedule. Subjective state (sleepiness and positive mood) generally became worse across the workweek during the 12 h schedule.

In summary, the regression analyses revealed circadian variations in performance and subjective state which, in most measures, were associated with decline performance at night. In addition, the regression analysis indicated that, except for digit addition, the worst performance occurred during the 12 h shift schedule.

3.2. Daily sleep and habits questionnaire

3.2.1. *Analyses of variance across shifts:* Means for statistically significant questionnaire scores across 8 h and 12 h shifts are shown in table 3 for the repeated measures analyses and the between-groups analyses. Corresponding *F*-ratios and *post hoc* comparisons among means also are shown. (*Note:* Workers reported no absenteeism or use of sleep aids. Consequently, these variables were eliminated from further analyses.)

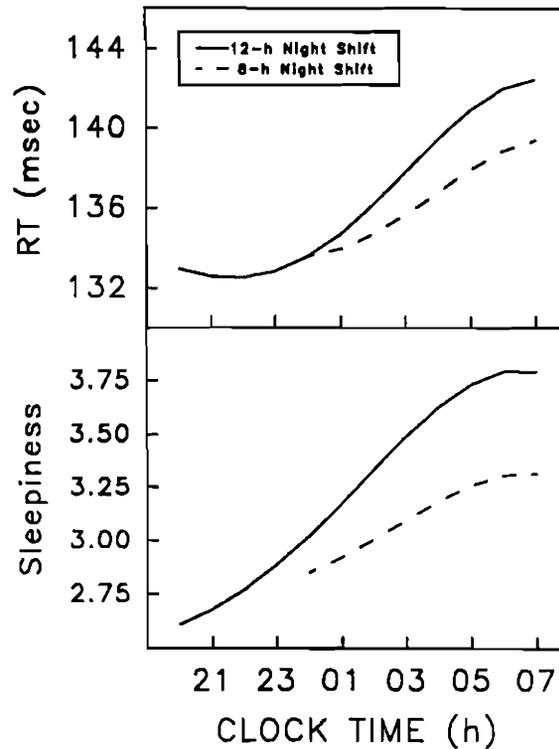


Figure 3. Multiple regression predictive functions for simple auditory reaction time and subjective sleepiness for 8 h and 12 h night shifts.

The repeated-measures ANOVAs of average shift scores indicated more total sleep on 8 h evening compared to 8 h night shift, the most frequent napping on 12 h night shift, and deeper sleep ratings on 8 h shifts. No significant differences across shifts were observed for sleep latency, number of awakenings, time awake during the sleep period, sleep quality rating, rating of feeling on awakening, adjustments of personal schedules, or minutes of daily exercise.

The between-groups ANOVAs of average shift scores indicated the most frequent napping on 8 h night shift, and more frequent mealtime adjustments on night shift. No significant differences across shifts were observed for sleep latency, total sleep time, number of awakenings, time awake during the sleep period, sleep depth and quality ratings, rating of feeling on awakening, missed social events, or minutes of daily exercise.

In summary, there were few similarities between the repeated-measures and between-groups ANOVAs of average shift scores.

3.2.2. Analyses of variance across workdays: In comparing the ANOVAs from subjects tested on both schedules to ANOVAs from subjects tested only on one schedule, the only statistically significant finding common to both sets of analyses was a decline in total sleep time across the workweek of 12 h shifts. Means for 12 h day and night shifts, from early, middle, and late in the workweek, are shown for each group in figure 4. In Group 1, tested on both work schedules, total sleep time decreased across the workweek primarily on night shift. This result was supported by

Table 3. Repeated measures and between-groups means and ANOVA results ($df=4,20$) for questionnaire scores across 8 h and 12 h shifts.

		Shift					<i>F</i> (<i>p</i>)	Post hoc Tests
		Day 1	8 h Eve 2	Night 3	12 h Day 4 Night 5			
REPEATED MEASURES								
<i>Dependent variable</i>								
Total sleep time (min)	M	440	444	353	426	441	4.21	2>3
	SD	45	44	34	37	47	<0.02	
Percent naps	M	05	07	15	02	53	11.34	5>1,2,3,4
	SD	10	15	20	05	15	<0.001	4+5>1,2,3
Sleep depth rating (1=Light, 9=Deep)	M	8.02	7.74	7.30	6.85	5.88	4.39	1,2,3>5
	SD	0.87	0.96	0.98	0.58	0.98	<0.02	1+2+3>4+5
BETWEEN GROUPS								
<i>Dependent variable</i>								
Percent naps	M	02	09	31	02	17	3.48	3>1,2,4
	SD	07	25	34	06	22	<0.02	
Adjusted meal times (%)	M	29	50	84	48	78	4.57	3,5>1
	SD	31	46	33	32	27	<0.01	

a significant interaction of shift with week segment ($F(2,2)=37.0, p<0.02$). In Group 2, tested only during the 12 h shift schedule, overall sleep time was shorter on night shift, and sleep time decreased across the workweek on both day and night shifts. These results were supported by a significant main effect for shift ($F(1,11)=11.56, p<0.01$), and a significant main effect for week segment ($F(2,40)=3.85, p<0.03$), from the ANOVAs conducted solely on data from Station 2 (see Sections 2.7.2 and 2.8).

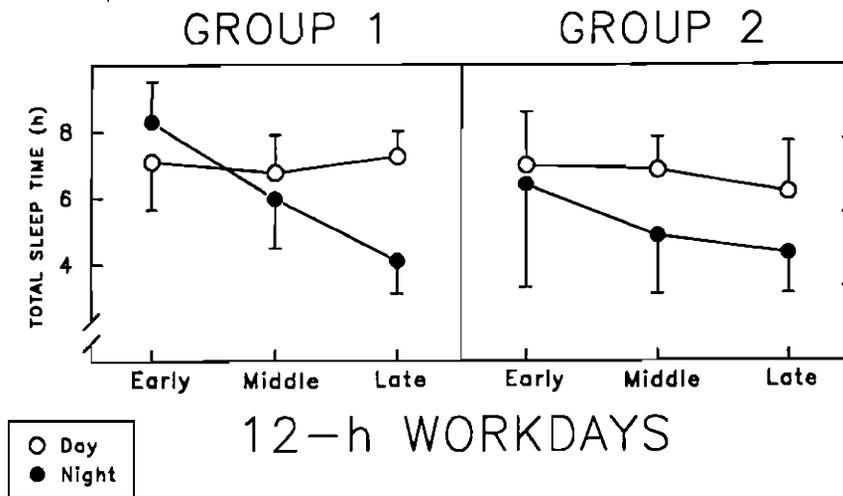


Figure 4. Total sleep time on 12 h day and night shifts, from early, middle, and late in the workweek, for the group tested during both the 8 h and 12 h shift schedules (Group 1), and for the group tested only during the 12 h shift schedule (Group 2). Error bars indicate standard deviations.

4. Discussion

4.1. 8 h vs 12 h shift schedules

4.1.1. *Test battery:* Our analyses of the test battery measures suggested lower alertness after 10 months of working on 12 h shifts. The most consistent support for this suggestion was more frequent missed signals in the simple reaction time task during 12 h shifts. This result was observed in the repeated-measures and between-groups ANOVAs, and in the multiple regression analysis. After partitioning variance attributable to practice, the regression analysis also suggested more frequent errors in the grammatical reasoning task on the 12 h shift schedule compared with the 8 h shift schedule.

More frequent reaction time misses reflect periodic lapses in alertness during which no attention is given to external stimuli (Bills 1931). More frequent reasoning errors might also be an indication of lapses in alertness, or they might suggest greater difficulty performing logical operations that place a demand on working memory (see Baddeley 1976).

The regression analysis also indicated some parallel effects between simple reaction time performance and subjective sleepiness which were obscured by the ANOVA tests of average shift scores. The parallelism between reaction time and subjective sleepiness would be especially apparent at night when increased fatigue with hours of work would be exacerbated by decreased arousal attributable to the

circadian rhythm. Thus, one might expect performance and alertness on the 12 h night shift to be the worst of any shift tested because of the extra work hours and declining circadian arousal. Similar night-time effects on performance and subjective alertness were associated with 12 h shifts at the first worksite we studied (Rosa *et al.* 1989, Rosa 1991). Replication in separate worksite studies increases confidence in these results.

4.1.2. *Daily sleep and habits questionnaire:* Of particular concern from the questionnaire results is reduced total sleep time across the workweek, which was most consistent on 12 h shifts. Similar reductions in sleep time were observed in our first worksite study. Other researchers (Mills *et al.* 1983, Peacock *et al.* 1983), however, have reported improvements in sleep duration after introduction of a 12 h shift schedule. In the laboratory, the levels of sleep loss observed in our studies have been associated with increased sleepiness and decrements in performance (Carskadon and Dement 1981, Webb and Agnew 1974). At the worksite, however, this association was less clear. In the present investigation, the regression analysis indicated decreases in self-reported positive mood. These changes paralleled reductions in sleep time across the compressed workweek. Performance loss, however, was more difficult to demonstrate.

Despite unclear relationships between partial sleep loss and worksite alertness measures, enough laboratory research has been conducted to suggest that sleep deprivation is a potential risk for workers on a 12 h shift schedule (see Johnson 1982, for review of sleep deprivation). The reduced sleep time observed in both worksite studies indicates a need for workers to make a strong effort to maximize sleep during the workweek, even at the expense of other activities.

4.2. *Final remarks*

Our worksite studies have suggested declines in alertness after increasing shift length by 50%. These results are consistent with some reports in the literature (Daniel and Potasova 1989, Hamelin 1987, Volle *et al.* 1979), but not with others (Peacock *et al.* 1983). The declines in alertness observed in our studies were most apparent at night when lowered circadian arousal potentiates fatigue resulting from hours of work (also see Hamelin 1987). These results emphasize a need to schedule critical or hazardous activities early in the shift whenever possible.

Some of the cost of working 12 h days may be partially offset by the shortened workweek. Loss of sleep during the compressed workweek, however, is a concern at both the present and previous worksites (Rosa *et al.* 1989, Rosa 1991).

Despite the increased fatigue associated with 12 h shifts, it is our impression that the compressed workweeks are quite popular at both organizations we studied. At the first worksite, more than 80% of the employees voted to retain the new work schedule. At the second worksite, other groups have requested that a compressed workweek be scheduled. Apparently, the extra off-duty days are proving valuable for personal activities.

Because of the popularity of the compressed workweek, personnel at the first worksite have expressed their willingness to tolerate the increased fatigue that results from 12 h workdays. Management is in favour of the compressed workweek because they expect a popular work schedule to reduce attrition. Retention of experienced workers is considered an advantage because of reduced training costs, and because experienced workers are assumed to work more safely and efficiently than

inexperienced workers. Management has agreed to a permanent schedule change based on their analysis showing continued safe and efficient plant performance under the 12 h schedule.

Two cautions, however, should be considered relative to conditions at the first worksite. First, many workers in the present study have more physically demanding jobs than workers in the previous study. The combination of high physical demands and 12 h shifts could produce extremely high levels of fatigue. We planned to test this possibility in the present study, but the number of participants doing light or heavy physical work was not sufficient for statistical analysis.

The second caution is that fatigue is an issue both on and off the job. For example, workers in the dense urban location of the present study must travel to their jobs in heavy traffic. Excessive fatigue may be a greater concern in that location than in the semi-rural location of the previous study. These conditions point to the need for long-term monitoring of work-related and non-work-related accidents and injuries as a function of extended workshifts.

References

- BADDELEY, A. 1968, A 3 min reasoning task based on grammatical transformation, *Psychonomic Science*, **10**, 341–342.
- BADDELEY, A. D. 1976, *The Psychology of Memory* (Basic, New York).
- BILLS, A. C. 1931, Blocking: a new principle in mental fatigue, *American Journal of Psychology*, **43**, 230–245.
- CARSKADON, M. A. and DEMENT, W. C. 1981, Cumulative effects of sleep restriction on daytime sleepiness, *Psychophysiology*, **18**, 107–113.
- COLLIGAN, M. J. and TEPAS, D. I. 1986, The stress of hours of work, *American Industrial Hygiene Association Journal*, **47**, 686–695.
- DANIEL, J. and POTASOVA, A. 1989, Oral temperature and performance in 8 h and 12 h shifts, *Ergonomics*, **32**, 689–696.
- HAMELIN, P. 1987, Lorry driver's time habits in work and their involvement in traffic accidents, *Ergonomics*, **30**, 1323–1333.
- HODDES, E., ZARCONI, V., SMYTHE, H., PHILLIPS, R. and DEMENT, W. C. 1973, Quantification of sleepiness: a new approach, *Psychophysiology*, **10**, 431–436.
- HODGE, B. J. and TELLIER, R. D. 1975, Employee reactions to the four-day week, *California Management Review*, **18**, 25–30.
- JOHNSON, L. C. 1982, Sleep deprivation and performance, in W. B. Webb (ed.) *Biological Rhythms, Sleep and Performance* (John Wiley, New York).
- JOHNSON, L. C., TEPAS, D. I., COLQUHOUN, W. P. and COLLIGAN, M. J. (eds) 1981, *The Twenty-Four Hour Workday: Proceedings of a Symposium on Variations in Work-Sleep Schedules* (DHHS [NIOSH] Publication No. 81-127, US Government Printing Office, Washington, DC).
- KELLY, R. J. and SCHNEIDER, M. F. 1982, The 12 h shift revisited: recent trends in the electric power industry, *Journal of Human Ergology*, **11** (suppl.), 369–384.
- KOGI, K., ONG, C. N. and CABANTOG, C. 1989, Some social aspects of shift work in Asian developing countries, *International Journal of Industrial Ergonomics*, **4**, 151–159.
- LEES, R. E. M. and LAUNDRY, B. R., Comparison of reported workplace morbidity in 8 h and 12 h shifts in one plant, *Journal of the Society of Occupational Medicine*, **39**, 81–84.
- LISPER, H.-O. and KJELLBERG, A. 1972, Effects of 24 h sleep deprivation on rate of decrement in a 10 min auditory reaction time task, *Journal of Experimental Psychology*, **96**, 287–290.
- MILLS, M. E., ARNOLD, B. and WOOD, C. M. 1983, Core 12: a controlled study of the impact of 12 h scheduling, *Nursing Research*, **32**, 356–361.
- NAITOH, P., ENGLUND, C. E. and RYMAN, D. H. 1985, Circadian rhythms determined by cosine curve fitting: analysis of continuous work and sleep-loss data, *Behavior Research Methods, Instruments, and Computers*, **17**, 630–641.
- PEACOCK, B., GLUBE, R., MILLER, M. and CLUNE, P. 1983, Police officers' responses to 8 h and 12 h shift Schedules, *Ergonomics*, **26**, 479–493.

- ROSA, R. R. 1991, Performance, alertness, and sleep after 3-5 years of 12 h shifts: a follow-up study, *Work and Stress*, **5**, 107-116.
- ROSA, R. R. and COLLIGAN, M. J. 1986, The NIOSH fatigue test battery: laboratory validation of a portable system for field study of extended workdays and work scheduling, NIOSH Internal Report available from the National Institute for Occupational Safety and Health, Division of Biomedical and Behavioral Science, 4676 Columbia Parkway, Cincinnati, OH 45226, USA.
- ROSA, R. R. and COLLIGAN, M. J. 1988, Long workdays versus restdays: assessing fatigue and alertness with a portable performance battery, *Human Factors*, **30**, 305-317.
- ROSA, R. R., COLLIGAN, M. J. and LEWIS, P. 1989, Extended workdays: effects of 8 h and 12 h rotating shift schedules on performance, subjective alertness, sleep patterns, and psychosocial variables, *Work and Stress*, **3**, 21-32.
- ROSA, R. R., WHEELER, D. D., WARM, J. S. and COLLIGAN, M. J. 1985, Extended workdays: effects on performance and ratings of fatigue and alertness, *Behavior Research Methods, Instruments and Computers*, **17**, 6-15.
- RUTENFRANZ, J., COLQUHOUN, W. P., KNAUTH, P. and GHATA, J. N. 1977, Biomedical and psychosocial aspects of shift work, *Scandinavian Journal of Work Environment and Health*, **3**, 165-182.
- SAS INSTITUTE INC. 1985, *SAS User's Guide: Statistics, Version 5 Edition* (SAS Institute, Inc., Cary, NC).
- SCOTT A. J. (ed.) 1990, *Occupational Medicine State of the Art Reviews vol. 5: Shiftwork* (Hanley and Belfus, Philadelphia).
- THACKRAY, R. I., BAILEY, J. P. and TOUCHSTONE, R. M. 1977, Physiological, subjective, and performance correlates of reported boredom and monotony while performing a simulated radar task, in R. R. Mackie (ed.) *Vigilance: Theory, Operational Performance, and Physiological Correlates* (Plenum, New York), 203-211.
- TODD, C., REID, N. and ROBINSON, G. 1989, The quality of nursing care on wards working 8 h and 12 h shifts: a repeated measures study using the MONITOR index of quality care, *International Journal of Nursing Studies*, **26**, 359-368.
- VOLLE, M., BRISSON, G. R., PERUSSE, N., TANAKA, M. and DOYON, Y. 1979, Compressed work-week: psychophysiological and physiological repercussions, *Ergonomics*, **22**, 1001-1010.
- WEBB, W. B. and AGNEW, H. W. 1974, The effects of a chronic limitation of sleep length, *Psychophysiology*, **11**, 165-174.
- WEVER, R. A. 1981, On varying work-sleep schedules: the biological rhythm perspective, in L. C. Johnson, D. I. Tepas, W. P. Colquhoun and M. J. Colligan (eds) *The Twenty-four Hour workday: Proceedings of a Symposium on Variations in Work-Sleep schedules* (DHHS [NIOSH] Publication No. 81-127, US Government Printing Office, Washington, DC), 51-86.
- WILLIAMS, H. L. and LUBIN, A. 1967, Speeded addition and sleep loss, *Journal of Experimental Psychology*, **73**, 313-317.