SLEEP IN THE NATURAL ENVIRONMENT: PHYSIOLOGICAL AND PSYCHOLOGICAL RECORDING AND ANALYSING TECHNIQUES

K. B. Campbell and R. T. Wilkinson
MRC Applied Psychology Unit
Cambridge, England

In the past twenty years, a vast number of studies have been devoted to sleep. The large majority of these studies have been carried out in unnatural environments such as the laboratory or in the hospital. Beyond the inconvenience to certain groups of people, the actual laboratory situation itself may contribute to what has come to be known as the "first-night" effect (Agnew, Webb, & Williams, 1966) in which the first night or two of sleep have been found to be quite different from others monitored later in the week. Over the past two years, we have been engaged in a programme of adapting modern biomedical technology to the requirements of recording sleep in the natural environment of the subject or patient, as for example, in the home.

The objectives of the endeavour were established to be (1) the faithful recording of at least 4 physiological channels over an 8 hour period, (2) a considerable reduction in the size and weight of existing apparatus, (3) a complete isolation of the subject from the main power supply since he/she would not be monitored during the sleeping period, (4) for the sake of economy, the recording of at least two subjects simultaneously, and (5) the physiological data to be in a form that would allow off-line automatic computer analysis at a fast playback speed.

In the laboratory, it had been our custom to investigate the effects of the quality of the night's sleep on a number of performance indices the following day. While some of these tests could be adapted for the field with only minor modifications, others required extensive revision, yet still had to monitor essentially the identical behavioural constructs. Finally, questionnaires that asked the subjects to rate various dimensions of their sleep were, of course, readily available for field use.

The analysis of the data collected in the homes of subjects called for high-speed, off-line automatic scoring procedures. Due to the restrictions inherent in the recording system, innovative programming techniques had to be introduced.

This article summarizes each of these developments and provides examples derived from recent data.

Physiological Recording

In the mid-1970's, a small lightweight 4-channel cassette unit become available for the 24 hr monitoring of EKG activity (Cashman & Stott, 1974). The unit was modified to enable the recording of the EEG and EOG for the study of sleep (Wilkinson & Mullaney, 1976), a task simultaneously and independently undertaken as well by Ives and Woods (1975). Although the dynamic range of the instrument was limited to just over 30 dB, initial tests of sleep in the home and in such inconvenient locations as an overnight train (Wilkinson, Herbert, & Branton, 1973) suggested that the unit was capable of recording

sleep activity. There was, however, one major drawback. A sharp amplitude drop was apparent at low frequencies -- precisely where one would expect to find the presence of slow wave sleep. The problem has apparently been recently overcome. The 4-channel limit, nevertheless, was barely adequate for sleep physiology.

As a result, an entirely new system was developed (Campbell, Weller, & Wilkinson, in press). A small portable lightweight unit containing 2 EEG, 1 EOG, and 1 EKG preamplifiers as well as a fifth channel serving as an indication of signalled awakenings requiring only a switch closure was constructed. A sixth channel was linked to a sound level meter to encode the level of environmental noise in the subject's room, often a source of disturbance. The problem of storing such a large number of channels (12 channels with two subjects) was overcome by the use of pulse interval modulation (PIM) multiplexing, which is considerably less expensive than the more widely known time division multiplexing (TDM) in conjunction with pulse code modulation (PCM). In PIM, a train of fixed duration pulses is generated, one pulse for each channel, and the time interval between consecutive pulses is made to vary with signal amplitude. Each channel is allocated to a particular time slot, one of the time slots being made unique for synchronization purposes. The sampling rate is therefore proportional to signal level. Thus, with excessively large amplitude signals, the sampling rate could become very low. To prevent this, a limit has been placed on the excursions of each time slot. If a signal causes the interval to exceed upper or lower preset levels, the next channel is immediately selected and the amplitude of the previous signal is clipped. The sampling rate cannot then fall below a predetermined rate and the synchronization time slot is never confused with signal time slots. The 6 channel system uses pulse intervals of 360 μsec (180-540 μsec) compensation channel. The minumum sampling rate is therefore:

$$\frac{10^6}{(6 \times 540) + 648}$$

If only one subject is being recorded, the multiplexed signal is transmitted over the British medical band (104.6-105.0 MHz) via a single stage frequency modulated transmitter. Telemetry of course allows the subject considerable freedom of movement as well as overcoming potential fear of electrical shock. When two or more subjects have been recorded simultaneously, reception at times has been incomplete, thus, the adoption of the alternative, cable-telemetry. Even though the signal is sent via a pulse transformer along a hand-wire, the lead has not proven to be an encumbrance to the subjects. Also, as Weller (1974) has pointed out, the method is as safe and reliable as radio-telemetry. The entire system is housed in a shielded plastic case (Figure 1). The general properties of the multiplexing/telemetry package are as follows:

Channels: 6

Weight: 125 g including the power source

Full-scale deflection: 300 μV (EEG), 500 μV (EOG), 4 mV (EKG)

Bandwidth: down 3 dB at 0.2 and 43 Hz (minimum)

Figure 1. Portable sleep recording unit. In the plastic case are found 4 preamplifiers, a multiplexer and a radio-transmitter.

Signal-to-noise and crosstalk: 40 dB

Power: 6.75 V mercury battery

Current consumption: 3 mA

Battery life: 70 hrs from mercury cells of 210 mAh capacity.

As can be observed in Figure 2 the bandpass characteristic of the EEG preamplifiers are such as to enable the recording of low frequency, delta activity and high frequency spindle transients.

Unfortunately, the upper limit does not permit the recording of EMG. Whether EMG provides additional relevant information for the interpretation of the sleep profile has been a subject of debate. Some laboratories (see, for example, Smith, 1978), including the present one, believe its activity is too inconsistent to prove useful. On the other hand others (see, for example, Gaillard & Tissot, 1973) employ it for the definition of a particular period of sleep, Stage REM.

The modulated carrier signal is recorded in the direct mode on a single track Racal 4D reel-to-reel tape recorder, the triple play tape lasting 8 1/4 hours. Typically, one of the other tracks is used for recording the analogue noise (frequency range, 100-8,000 Hz) occurring during the night. Thus, if sleep is disturbed, it is relatively easy to locate and identify the environmental influence immediately prior to the disturbance.

Early in the evening, a technician arrives at the subject's home to place the silver-silver cholride electrodes (to overcome polarization during the 8 hour monitoring). EEG recording is between C_3 and M_2 with a secondary backup of $C_4 - M_1$ in case of failure of the primary location. EOG electrodes are placed on the outer canthi of each eye, one electrode being slightly inferior to the socket, the other slightly superior, enabling the complete monitoring of horizontal and partial vertical eye movement. Because of the compactness and lightness of the unit it can be placed directly on the subject's head being attached to a specially designed cap. The EEG leads can thus be of minimal length, measuring not more than 30 cm. A demultiplexer permits the reconstruction of the six original signals which may be examined on a portable oscilloscope. Once the integrity of the system is verified, the technician leaves for the night, returning in the morning to remove the electrodes and administer questionnaires and the performance battery. Schematics for the sleep environment are depicted in Figure 3a, while the details of the head-mounted unit are illustrated in Figure 3b.

Over 300 nights of sleep have now been recorded in natural environments for a variety of studies--the effects of traffic noise on sleep, variation in sleep patterns with age, the effects of shift work on sleep. The amount of data collected during an 8 hour session is indeed considerable. The task of visually analyzing them is certainly laborious and not entirely reliable. At an early stage it was, thus, decided that automatic analysis was required not only to provide an efficient means of handling the data but also as an objective method of interpreting them.

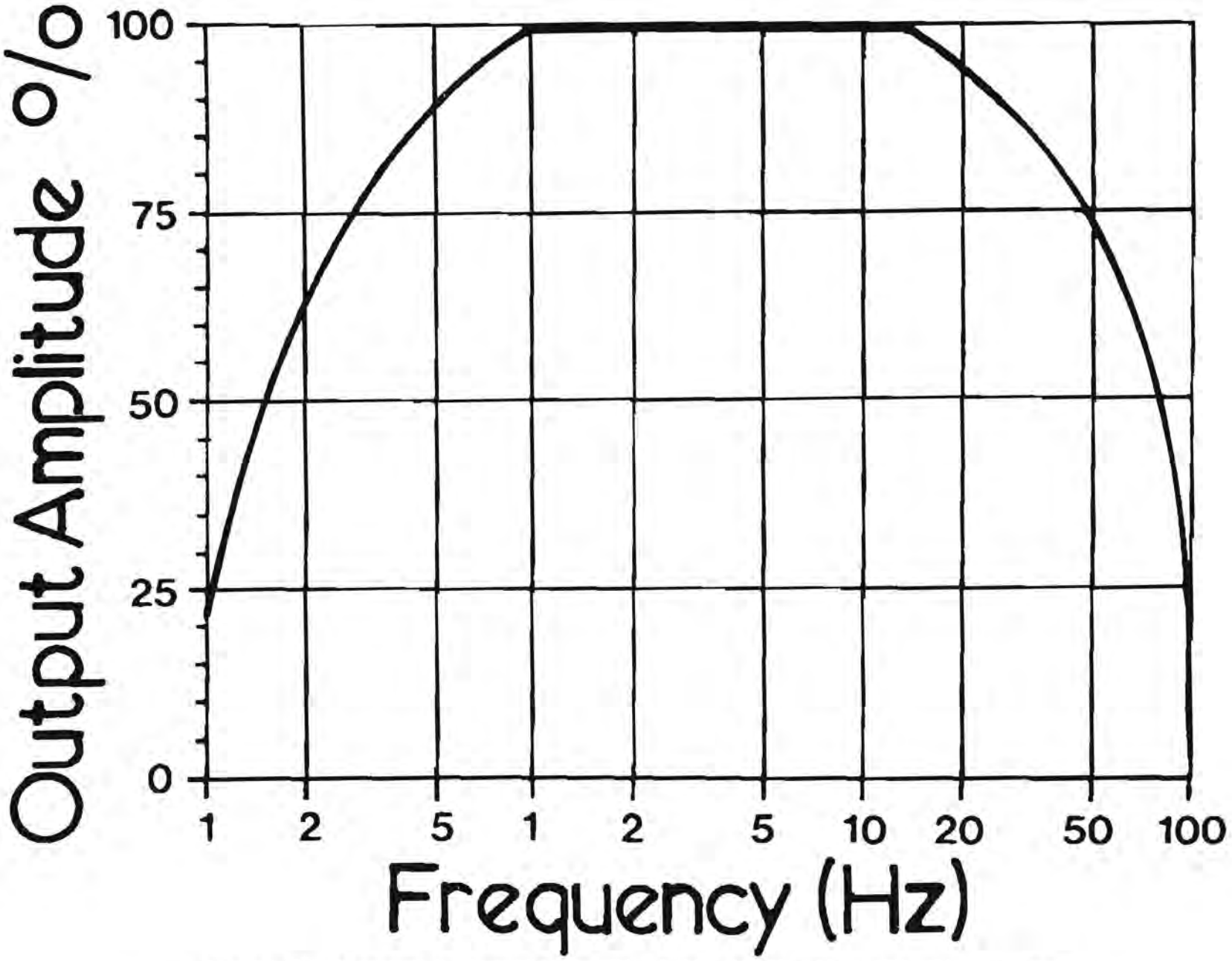

Figure 2. Frequency response curve of the sleep recording unit. Known, constant voltage sine waves of varying frequency were fed into one of the EEG inputs, amplified, multiplexed, tape recorded, played back, and demultiplexed. The output signal voltage is illustrated on the Y-axis, the input signal, the X-axis.

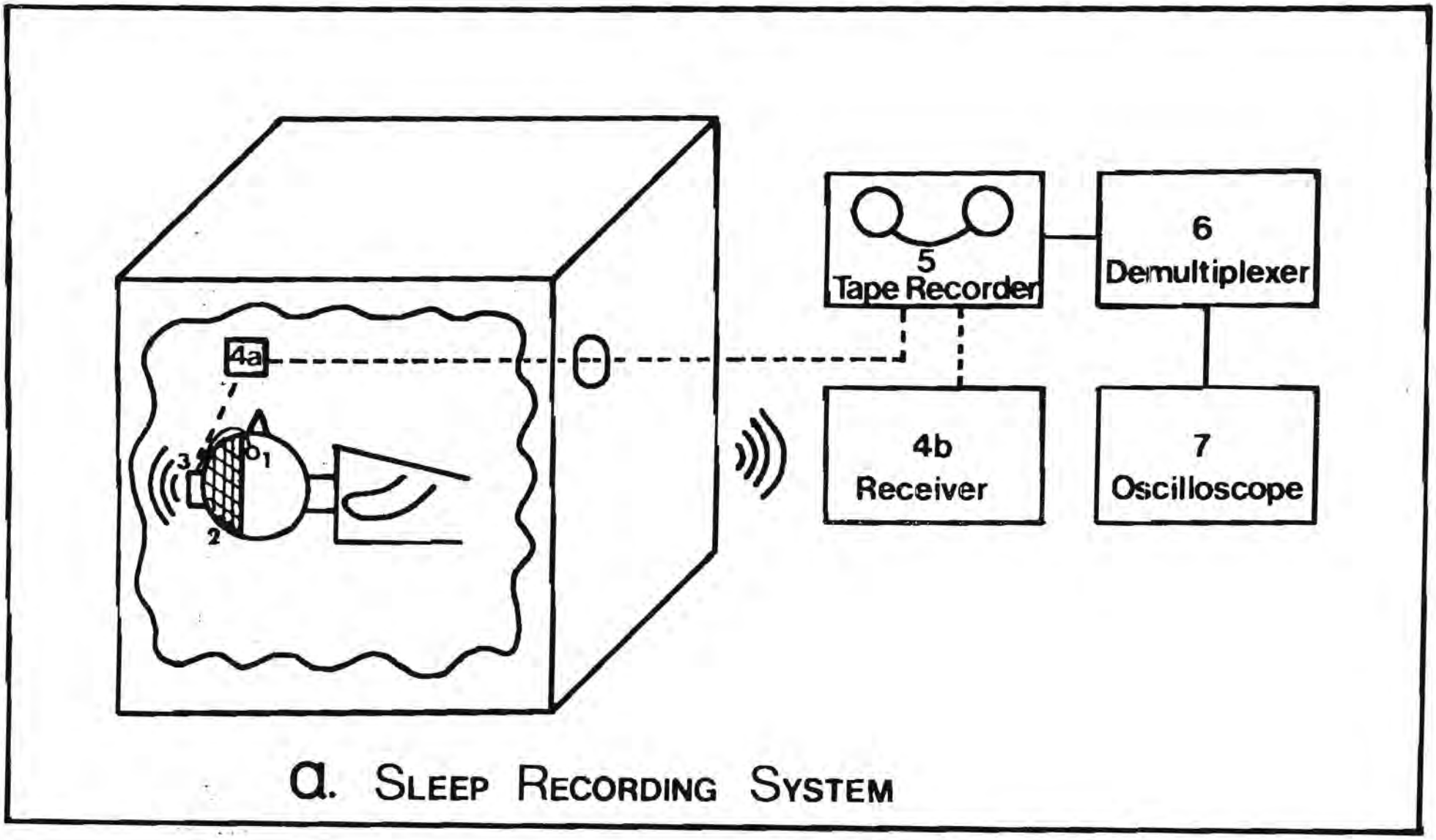

Figure 3A. Physiological activity is picked up by the electrodes (1) and amplified in the head-mounted unit (3), which is attached to a lightweight cap (2) on the subject's head. The signals are multiplexed and sent either via radio-telemetry to a receiver (4a) or via cable-telemetry to a junction box (4b). The signal from the receiver or junction box is recorded on a reel-to-reel tape recorder (5) in the direct mode. For purposes of verification of functioning of the system, the output from the tape recorder is demultiplexed (6) to reconstruct the original signals which are monitored on an oscilloscope (7).

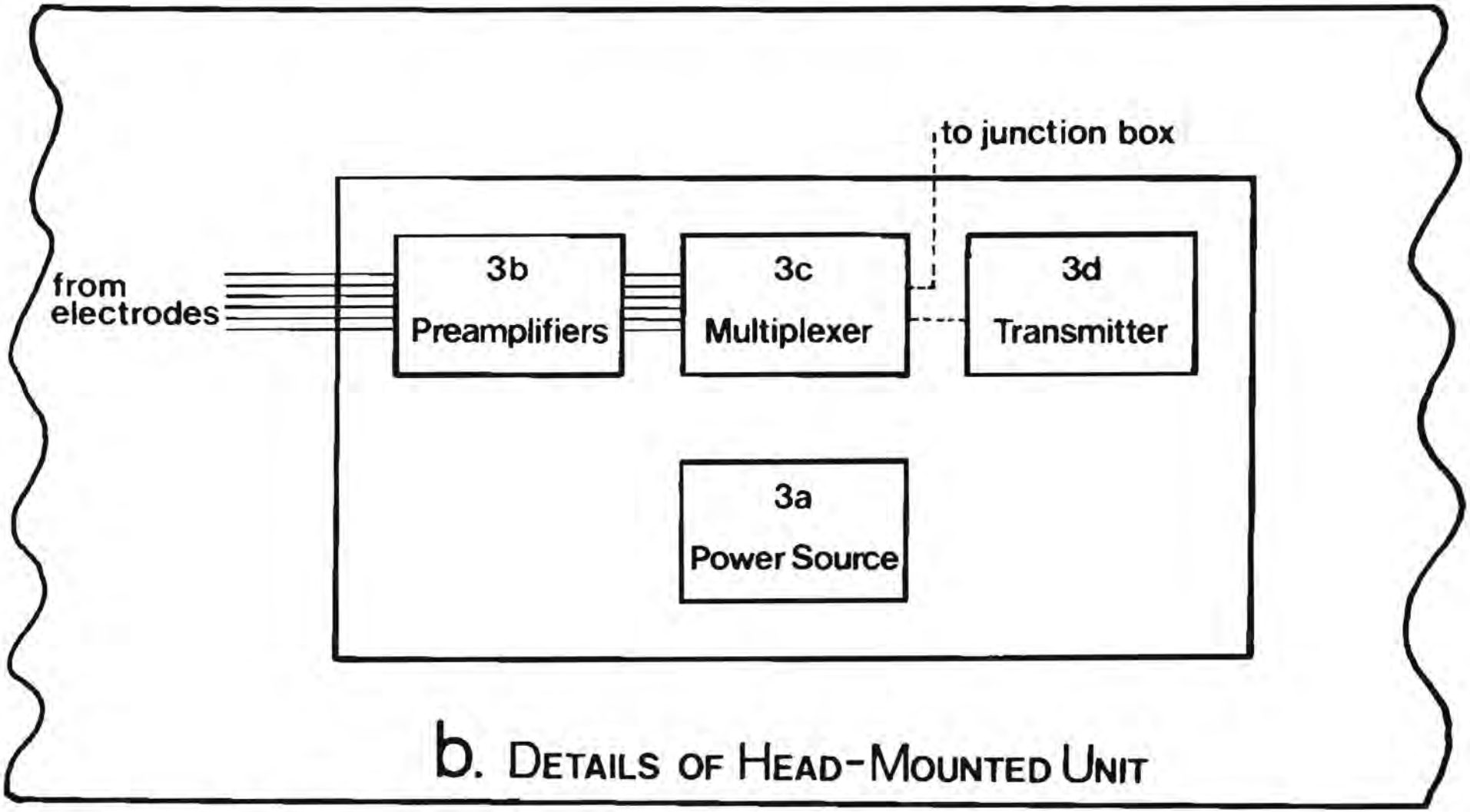

Figure 3B. Details of head-mounted unit. Two EEG, 1 EOG, 1 EKG, and 1 behavioural awakening signals (3a) are amplified and multiplexed to form a unique signal (3c) which is transmitted via radio or cable-telemetry (3d). The power source for the unit is a 6.75 V mercury battery.

Staging

Figure 4 illustrates the types (or "stages") of activity that may be present in a recording session. The waking stage, Stage W, is distinguished by the dominance of rhythmic alpha waves. There may also be muscle artifact and large amplitude eye movements. The subject first enters Stage 1 from the awake state. For purposes of classification, Stage 1 is unusual in that the EEG has no real distinguishing features. Although alpha may have largely disappeared, none of the features of other stages are present. The EOG often takes on a slow, rolling appearance which is unique to this stage. After a short duration, rhythmic 12-14 Hz transients called spindles mark the commencement of Stage 2. The first episode of Stage 2 sleep is variable, from a few minutes to almost an hour in length. Towards the end of Stage 2, the EEG comes to be characterized by the appearance of large amplitude 0.5-2.0 Hz delta waves. When this slow wave activity comes to occupy 20% of the record, it is arbitrarily classified as Stage 3. When the figure reaches 50% delta, the subject is said to be in Stage 4. After approximately 1 1/2 hours of sleep, the EEG changes to an awake or Stage 1-like pattern. What makes this activity unusual is that the EOG exhibits rapid eye movement (REM). Although the EEG is suggestive that the person is in a light stage of sleep, in reality, it is extremely difficult to awaken them, hence, the alternative name, paradoxical sleep. If awoken, the subject often reports episodes of dreaming.

The classification of more than a kilometre of paper recording into the various stages of sleep usually takes 2-3 hours to complete, often longer if there is are unusual features to interpret. While staging now appears to have become part of the standard repetoire of the sleep researcher, it is, nonetheless, accomplished at a cost of a tremendous loss of other potential data. Many other types of analyses that theoretically could be carried out, are ignored in practice due to the volume of data that can overwhem the keenest investigator. And even the objectivity of what the human can analyze has been questioned. Monroe (1969) has demonstrated that there is only approximately 75% agreement on human sleep stage classification between scorers when the same record is sent to different laboratories.

Computer analysis of sleep opened not only the feasibility of examining the minute details hidden in the data but also because of its objectivity, offered considerable scope for the sharing of these data between laboratories. Some attempt had previously been made at automation. Itil, Shapiro, Fink, and Kassebaum (1969) have used digital period analysis methods while Larsen and Walter (1970), Lubin, Johnson, and Austin (1969) and Martin, Johnson, Vigiolone, Naitoh, Joseph, and Moses (1972) have attempted software spectral analysis in conjunction with linear and nonlinear discriminant analysis. Their application was limited by the relatively slow running time (up to 3 hours) and consequent high cost. The time and cost factors also clearly limit the extent of on-line usage. Moreover, agreement between human and software scoring was unfortunately barely acceptable.

A second approach relies on both hard and software methods (Smith Karacan, 1971; Gaillard & Tissot, 1973; Kumar, 1977). These hybrid systems rely on hardware preprocessing of the physiological data, the output of the

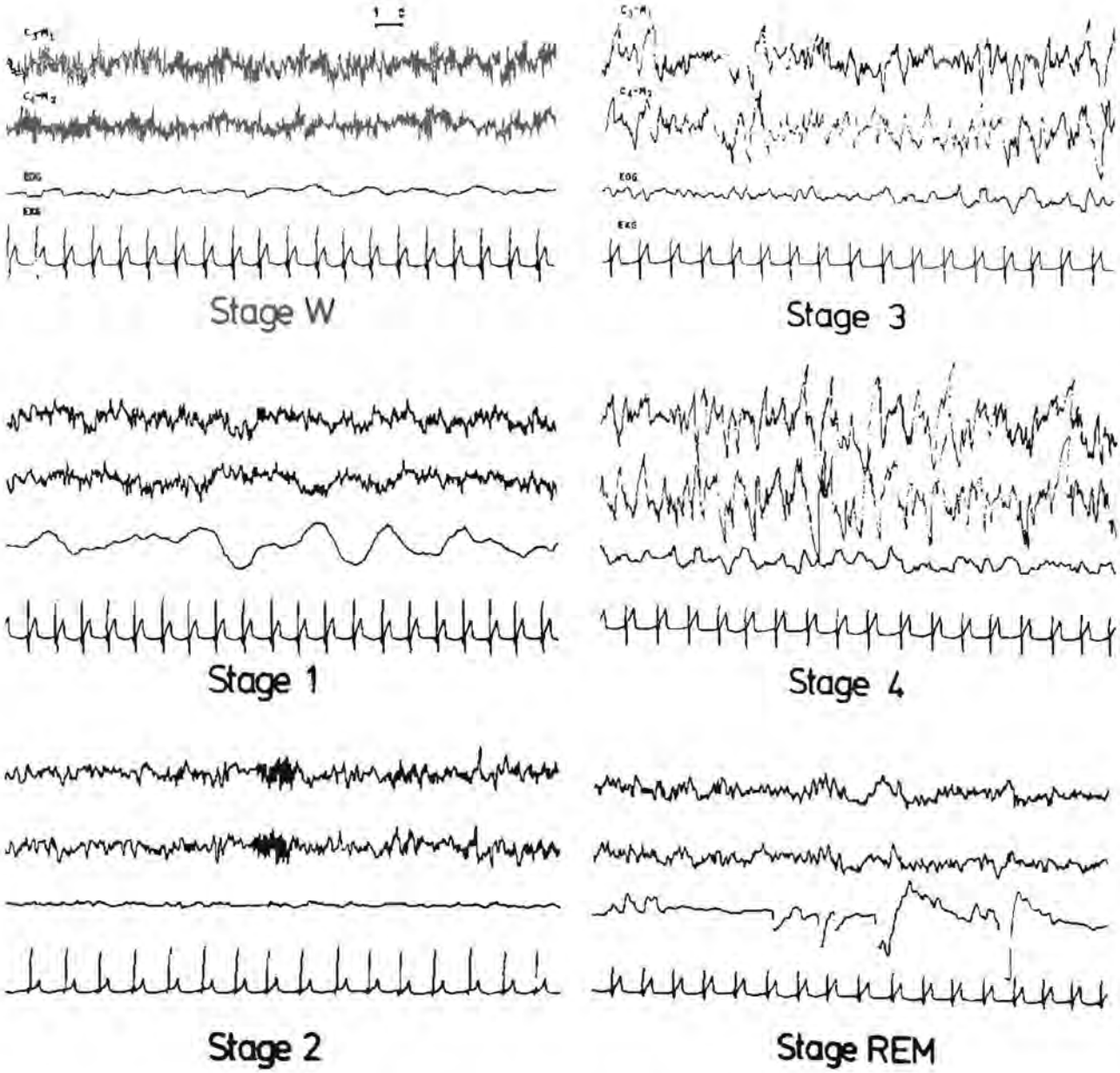

Figure 4. The stages of sleep. The awake (Stage W) EEG is typified by the presence of alpha waves. As the subject becomes drowsy and enters Stage 1, slow rolling eye movements often appear. The entrance of Stage 2 is announced by spindle transients. Stages 3 and 4 see the more and more increasing presence of slow-frequency, high amplitude delta waves. The EEG of Stage REM is similar to that of Stage 1, but is distinguished from it by the presence of rapid eye movements. The illustrations were taken from a home recording using the portable apparatus previously described. The writeout was obtained by passing the signals, unamplified and unfiltered to a Grass polygraph's paper recorder.

preprocessors being fed into a digital computer which makes the actual decisions. The hardware in the form of a bank of filters (Gaillard & Tissot,1973) or specially designed detectors (Smith & Karacan, 1971) remove a considerable burden of data processing from the digital computer, freeing it for relatively simple calculations and storage matters. The advantage of such a scheme is that data can be played back at very high speeds without loss of information. The performance of the hybrid system is impressive--approximately 85% overall agreement with human scorers.

In Canada, an analogue preprocessor has been developed (Green, 1975; Broughton, Healey, Maru, Green, & Pagurek, 1978) that closely imitates the universally employed Rechtschaffen and Kales (1968) standard scoring system. This system was modified for the purposes of our multiplexing system and linked to a minicomputer, a task that had previously not been attempted.

At present, four hardware circuits have been constructed: alpha, spindle and delta detectors for the EEG lead and a rapid eye movement detector in the EOG. The reader is referred to Broughton et al. (1968) for a description of the spindle detector and Green (1975) for the description of the alpha and delta detectors. Briefly, the spindle and alpha detectors are based on phase locked loop (PLL) circuitry, (Johns, Stear, & Hanley, 1974) except that instead of using the phase error signal of the PLL to determine the presence of spindle or alpha activity, the more reliable approach of a quadrature phase detector was substituted. Delta detection is based on zero-crossing and minimum amplitude measuring techniques similar to Smith, Funke, Yeo, and Ambuehl (1975). The REM circuit consisted of a simple filter and minimum amplitude detector. A fifth device detected abnormally high positive or negative voltage in the EEG.

The frequency ranges (at real-time) and minimum amplitudes required for binary output from each of the detectors were: for alpha, 8 to 12 Hz at 25 μV; for spindles, 11.5 to 15 Hz at 25 μV; and for delta, 0.5 to 2.5 Hz at 75 μV. The tapes recorded in the homes of the subjects were played back at 8x through a second demultiplexer. Examples of the analogue processing of various segments of one subject's sleep are illustrated in Figure 5.

Computation on and storage of the binary output of the preprocessor is carried out by a Data General Eclipse S200 minicomputer equipped with a 16 bit, 32K core memory. The data are processed in 30 second real-time epochs, a period more or less standard for most laboratories. During these 30/8 seconds (the tapes are played back at 8x), the gate of the digital interface is opened 128 times, i.e., once every 29.29 msec or at a sampling rate of 34.1 Hz. The output of each detector is digitized by one bit ("0" or "1"). For purposes other than staging, some form of storage is necessary. Once the 32K core is full, the contents are transferred to the system's 2.5 megabyte removable disc unit. To achieve a continuous stream of data through the processor, a double buffering scheme is employed (Figure 6). While one of the memory buffers is being filled from the digital input channel, the other is emptied onto the disc. The processes are switched back and forth as the operation proceeds. Once the entire disc is full (8 sleep records), it is put into permanent storage on magnetic tape, the disc then being cleared. Even with the substantial memory capacities of modern computers, it can readily be noted that a single sleep experiment can tax it to the limits. The complete data handling and storage procedures are shown in Figure 7.

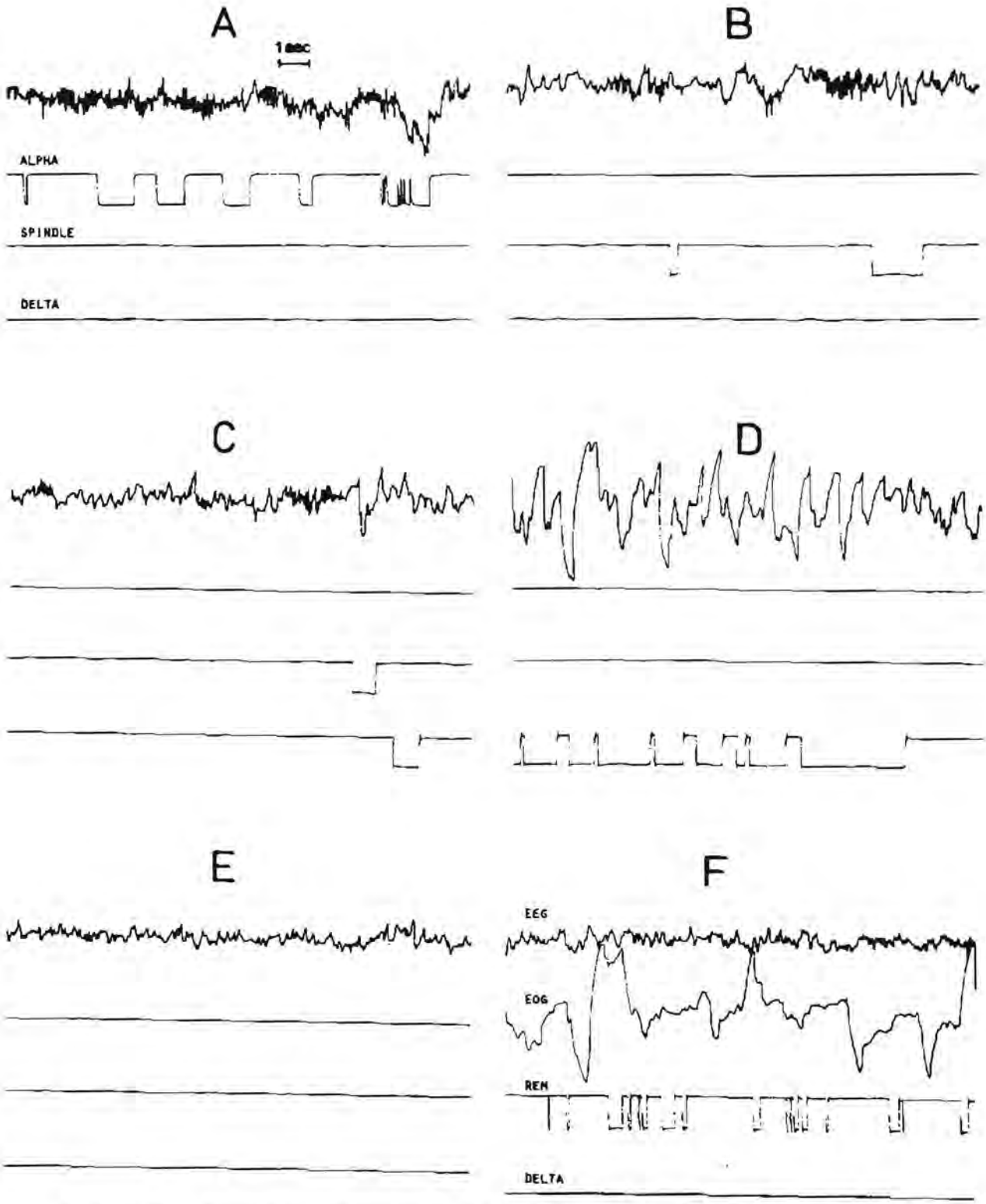

Figure 5. Analogue preprocessing of EEG and EOG activity. In (A), the dominance of alpha triggers the appropriate detector. In this and all other portions of the figure, a positive detection is noted by a downward deflection. In B, detection of poorly defined and a "classic" spindle. In C, a spindle transient followed immediately by a slow wave. In D, obvious Stage 4 activity as indicated by the delta detector. In E and F, the EEG detectors are silent. In F, EOG activity is also illustrated which quite apparently in conjunction with the "silent" EEG would be classified as Stage REM. The binary output of the preprocessor is strobed at a rate of 4.25 Hz (real time).

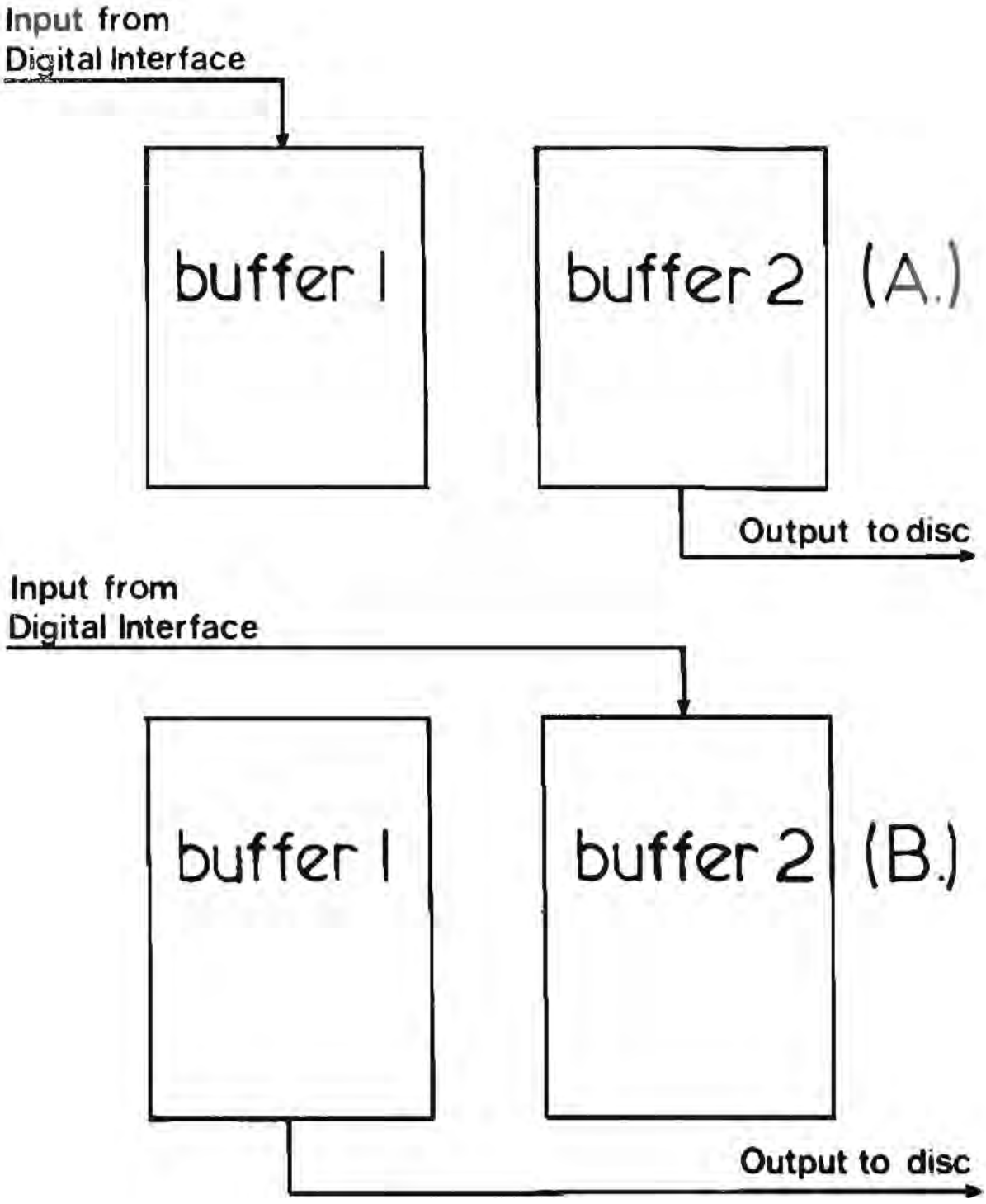

Figure 6. Double buffering technique for data streaming in the context of Figure 7.

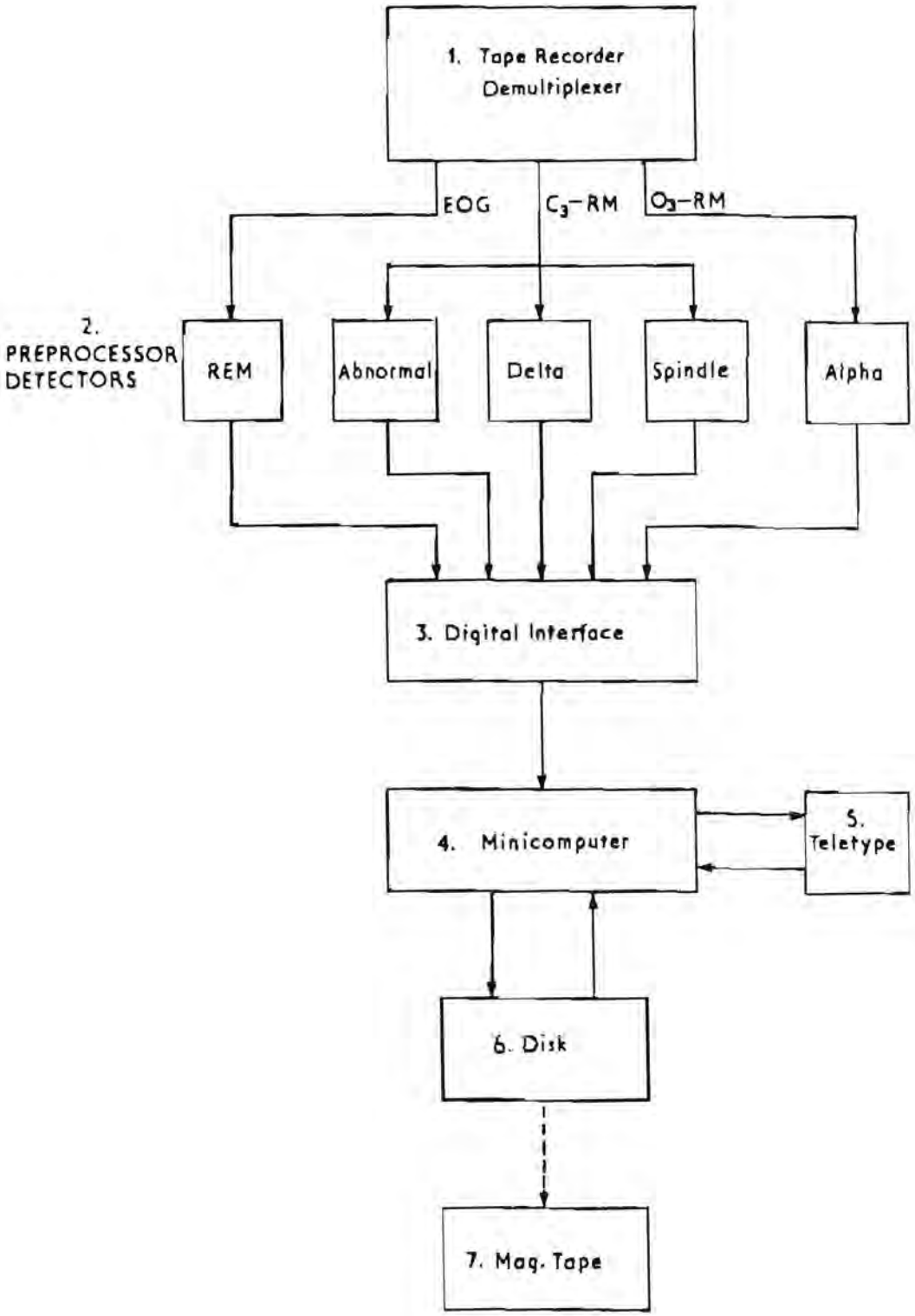

Figure 7. Schematics of the hybrid computer analysis system.

The logic of the software was designed to imitate human scoring procedures. Some changes from standard human scoring procedures were nevertheless made as these were found to provide better man/machine agreement than the direct implementation of the visual scoring criteria. A count is made of alpha, spindle, delta and REM events, a stage diagnosis being made following each 30 second real-time epoch.

Computation of the stages commences according to the following logic: if delta occupied 50% of the epoch, it was classified as Stage 4. If it occupied less than 50% but more than 25% (instead of the human criterion of 20%) it was designated to be Stage 3 sleep. If the delta count was less than 25% of the epoch but greater than 10% it was classified as Stage 2 (again, a departure from human scoring). If delta was less than 10%, then the computer had to decide between Stages W, 1 and 2, and REM. If a spindle was detected, it was considered to be Stage 2. If alpha occupied 40% of the epoch (usually 50% in the human scoring),then it was Stage W. By process of elimination, if none of the above were apparent in the EEG, then it was classified as Stage 1. Stage Rem is determined by both EEG and EOG criteria. If sufficient eye movements occur (arbitrarily set at 3) in what would be classified as Stage 1 on the basis of EEG alone, it is reclassified as stage REM. Often, human scorers label a Stage 1 epoch as stage REM even when eye movements are not apparent providing immediately prior and following epochs contain them. It was thus decided to reclassify all Stage 1 epochs following two consecutive Stage REMs as being REM provided no other change of stage was encountered. Similarly for Stage 2 scoring "if less than 3 minutes of a record which would ordinarily meet the requirements for Stage 1 intervenes between sleep spindles and/or K complexes these intervening epochs are scored as Stage 2" (Rechschaffen & Kales, 1969, p. 6).

On the basis of 8 records, the overall agreement between the automatic system and human scoring was 84% with a low of 59% agreement for Stage 1 and a high of 92% for Stage 4. The discrimination between Stages 1 and W and 1 and REM remains an area in need of further development. In defense of the present hybrid system, Stage 1 classification provides considerable difficulty for other existing automatic systems and human scorers.

Because there was a possibility of a bias in the human scoring procedure towards the classification logic of the computer, further evaluation was made with other scorers at CERN in Lyon, France employing data recorded in their laboratory. Again the comparison proved to be quite satisfactory. A further comparison of the hybrid computer's performance was made with a second, automatic system (Kumar, 1977). Although the basis of this system was entirely different from the Cambridge one, inter-system agreement ranged from 80 to 90% on 40 nights of recordings made at the TNO laboratory in Delft,the Netherlands.

Finally, because Stage 2 classification in the present format is critically dependent on precise detection of spindle activity, further verification of the functioning of the spindle circuit was deemed necessary. A second hardware detector relying on complex demodulation methods for spindle detection (Kumar, 1975) was utilized for the comparison. The two automatic devices functioned virtually identically being triggered by more than 80% of all EEG activity labelled as spindles by human judges (Campbell, Kumar, & Hofman, manuscript in preparation).

An example of the classification of an 8-hour sleep record is shown in the top part of Figure 12. Table 1 presents a breakdown of the values of classical sleep parameters for over 60 nights of sleep of young subjects. These data were recorded in the home of the subjects and analysed with the hybrid system. A comparison is made with the norms of males and females (collapsed) aged 20-29 as established by Williams, Karacan, and Hursch (1974). The values of the two laboratories are remarkably similar. Stages 3 and 4 are slightly overestimated, Stage 1 underestimated. The former may be due to individual differences between samples or to the fact that our subjects were sleeping in their natural environment. The latter discrepancy is probably a function of the logic of the automatic analysis. The larger number of stages detected by the computer as compared to the human scorers in Williams et al.'s (1974) study is almost certainly due to the fact that the latter's human scoring procedures neglect some stage transitions in favour of a more general smoothing procedure.

Temporal Organization of Sleep

Sleep scoring as typified in Table 1 furnishes only a gross overview of the total accumulation of a particular stage at the end of a night. No indication of the manner in which the respective stages have developed is provided. Recently, Gaillard (1977a, b) has developed a method for investigating the temporal organization of human sleep. Only a summary of the calculations will be made here, the reader being referred to the original sources for more complete details. The general trends in the development of the various stages of sleep can be evaluated and used to generate a theoretical model of sleep. The model has been particularly useful in the study of the effect of noise on the development of Stages 3 and 4 during the night as well as a comparison of the sleep of younger and older subjects. The same analysis technique is currently being employed to investigate the effects of shift work on sleep trends.

In making the computations, all awakenings during the night are removed from the calculations, the time base thus referring only to sleep time. All recordings are synchronized at sleep onset with each night being divided into 15 minute intervals. The number of minutes of each stage is counted in each 15 minute interval and these counts accumulated over the night (Table 2). The corresponding series of numbers over several nights for the same subject are averaged and the individual subjects' totals also averaged. The rationale behind the computation of a grand mean is that the cumulated occurrences of individual stages on a single night form an irregular "noisy" curve. The random variability disappears through the averaging process.

Figures 8, 9, and 10 present the cumulated occurrences of Stages 2, 3-4 and REM averaged across 10 nights of sleep of 5 younger (less than 32 years of age) and 5 older (+40 years) subjects. The lower portion of these figures represents the cumulation of the occurrence of a stage in minutes, while the top portion is the cumulation of a stage in terms of the proportion of the total amount of its occurrence over the night, thus providing an equal base for comparison of the different groups. It may be observed that no stage develops in a linear manner. Stage 2 (Figure 8) develops slowly early in the night but then increases in a linear fashion. Older subjects tend to have more Stage 2 sleep than younger ones. In relative terms, however, the devel-

Table 1

Comparison of Sleep Parameter Norms as Established by Williams et al. (1974) for the 20-29 Age Group with Those Determined by the Cambridge Hybrid Computer

Sleep parameter	Cambridge data	Data of Williams et al.
Total sleep time	431.4	424.3
Stage 1 (mins)	12.8	18.2
Stage 2	202.4	206.9
Stage 3	36.4	24.5
Stage 4	61.9	57.5
Stage REM	110.8	112.7
Stage 1 (% of total sleep time)	3.0	4.3
Stage 2	47.3	48.8
Stage 3	8.5	5.8
Stage 4	14.4	13.6
Stage REM	25.9	26.6
Sleep latency (mins)	12.8	13.8
Latency to Stage 3 (mins)	20.1	22.2
Latency to Stage 4 (mins)	25.3	28.6
Latency to Stage REM (mins)	71.2	94.0
Number of cycles	4.6	4.5
Mean duration of cycles (mins)	105.9	110.8
Number of stages	49.8	37.4

Table 2

Procedure for the Calculation of Cumulated Occurences of Sleep Time

	Sleep stage diagnosis	Summation of stages per 15 min Stages				Summation of successive 15 min Stages			
		1	2	3	4	1	2	3	4
	W W W								
1st 15 min interval	1 1 1 2 2 2 2 3 2 2 3 3 3 3 3	3	6	6	0	0	0	0	0
2nd 15 min interval	3 3 2 1 W W W 1 1 2 3 4 4 4 4 3 4 4	3	2	4	6	6	8	10	6

Note that all waking (W) episodes are discarded. Thus, the second interval contains 18 min of recording time but only 15 min of sleep. Although epochs are are indicated as being one minute in duration,in actuality they are 30 secs in length (after Gaillard, 1977b).

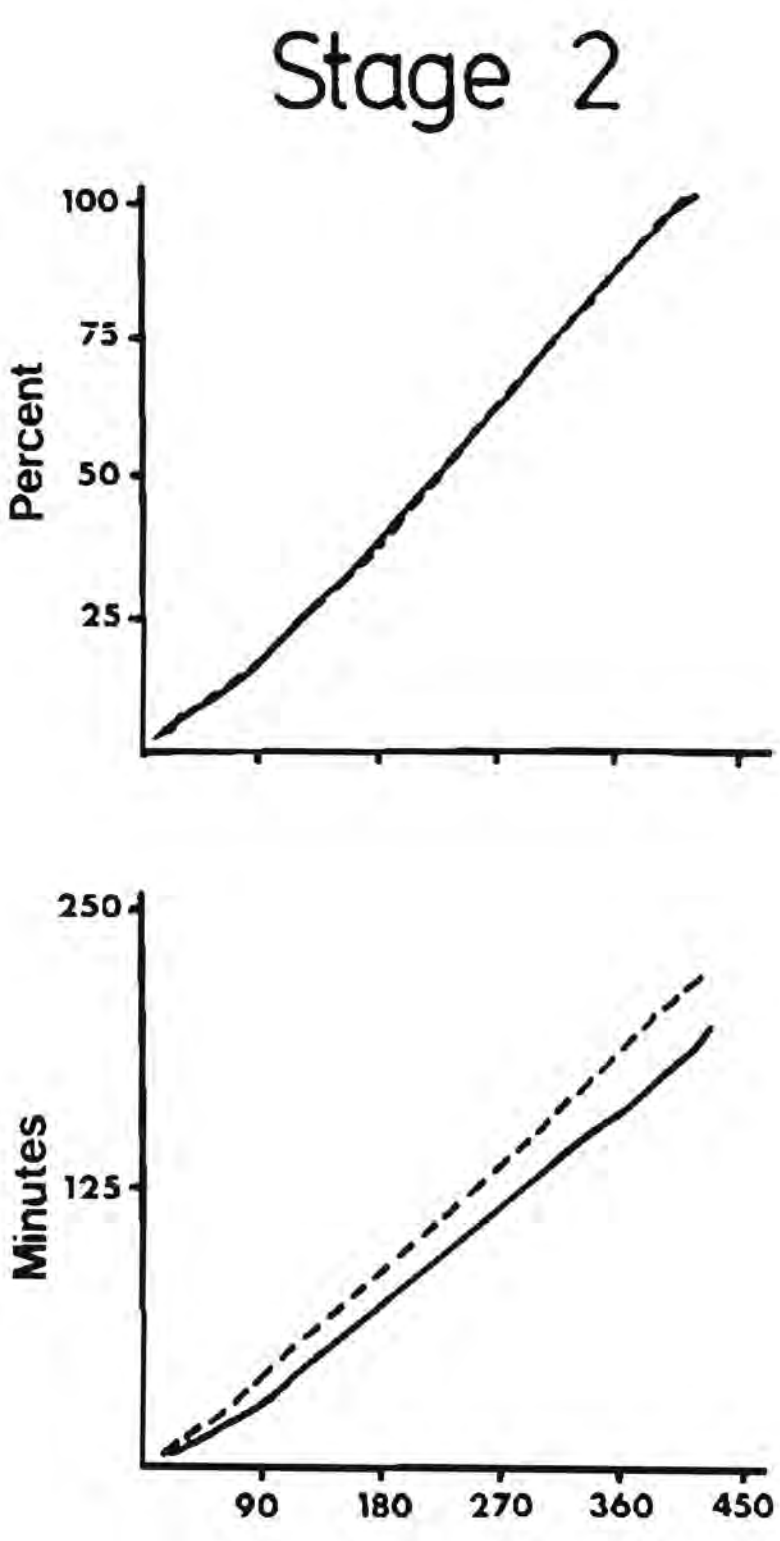

Figure 8. General trend of development of Stage 2 sleep for younger and older subjects recorded in the home.

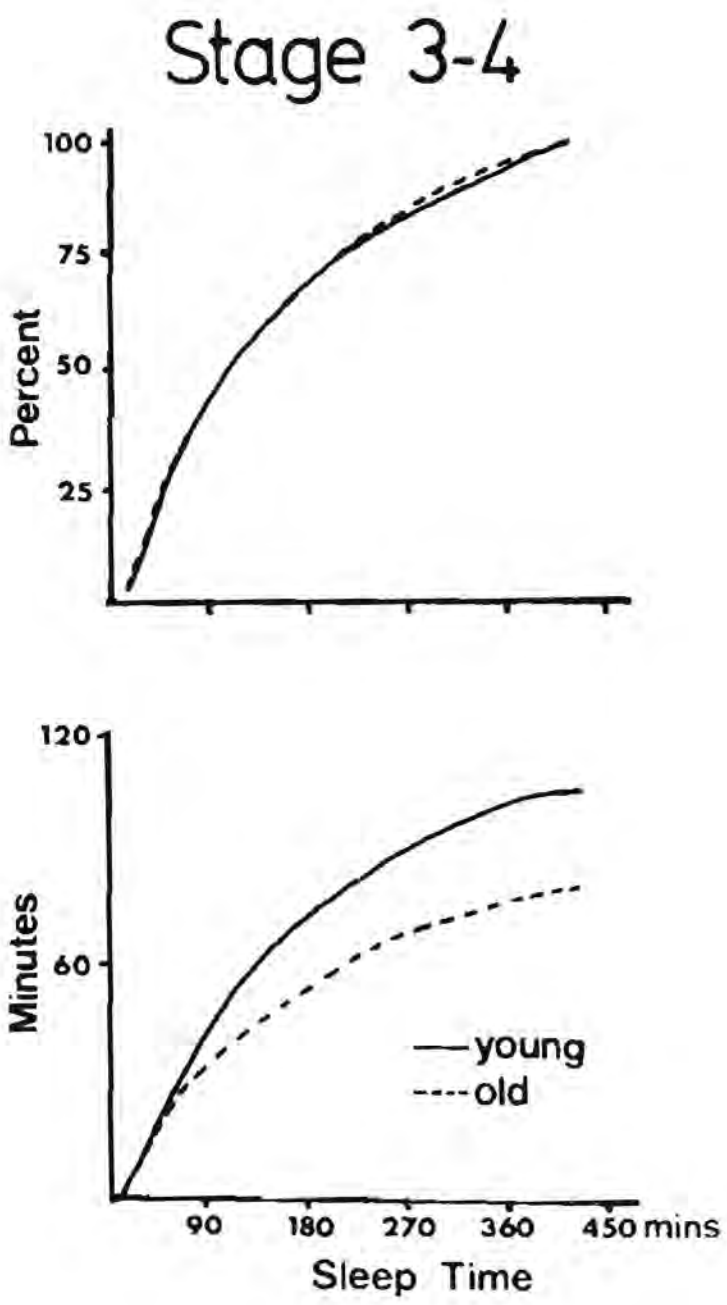

Figure 9. General trend of development of Stages 3 and 4 sleep for younger and older subjects recorded in the home.

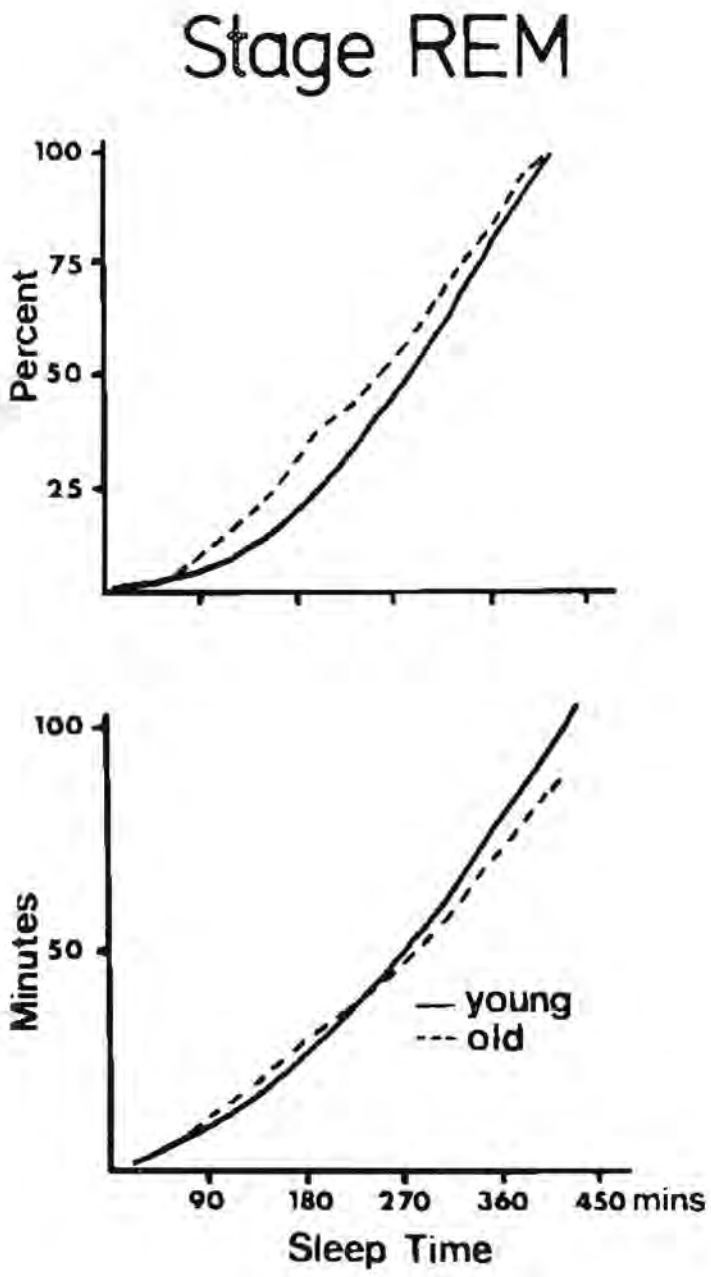

Figure 10. General trend of development of Stage REM sleep for younger and older subjects recorded in the home.

opment is essentially similar. Thus, after 60 minutes, 25% of the total amount of Stage 2 has been accumulated; after approximately 4 hours (228 minutes) 50%, and just after the 6 hour mark (330 minutes), 75% has been accumulated. There are some differences between these figures and those of Gaillard (1977a) which may be due to his subjects' longer nights of sleep. The nonlinear trend for Stage 2 is also a departure from Gaillard's findings although more recently (1979), he has described a very similar trend of evolution. There was an acceleration late in the night in his data, which of course could not be observed in our data due to the shorter sleep times.

The most interesting findings for the study of aging are those with respect to Stages 3 and 4. The well replicated finding of a decrease in slow wave sleep (Feinberg, 1974) can be observed in the bottom part of Figure 9. When the data are placed on a relative base, however, (upper portion), the development of Stages 3 and 4 manifests no age effect. For both groups, 25, 50, and 75% of the total amount of SWS occurs after 50, 120, and 240 minutes. Thus, Stages 3 and 4 develop rapidly early in the night and then decelerates so that only a further 25% of this type of sleep is gained in the last half of the evening. The very rapid emergence of SWS as opposed to Stages 2 or REM suggests an immediate and intense "need" (Tissot, 1966) for the former. Our data furthermore indicate that while the quantity of SWS is not as great in older subjects, for whatever reason, qualitatively, the general pattern of its emergence is constant throughout life.

Stage REM (Figure 10) takes on an antagonistic appearance to Stages 3 and 4. Whereas the latter emerges rapidly, REM emerges much more slowly but accelerates towards the end of the evening. Components of higher order related to ultradian variation located in the background circadian trend are only slightly visible. This probably is a function of the large sample size: cyclic variations tend to disappear due to inter- and intra-subject phase and periodicity differences.

Experimental manipulation of noise has typically pointed to a decrease in SWS as noise level increases. In the natural environment a most common source of noise is heavy traffic. Traffic levels are not constant throughout the night; early in the evening they are of moderate to heavy intensity, dropping off significantly in the early hours of the morning and reaching peak levels at around 0600 hr. From a preliminary analysis of subjects in very noisy areas of London (Figure 11), a hint of an actual shift in the development of SWS has been manifested. On noisy nights as compared to quiet ones (noise level 10 dB lower due to the installation of sound insulating windows), after about 2 hrs of sleep and continuing for another 2 1/2-3 hrs, an acceleration in the evolution of SWS can be noted. The period of acceleration corresponds to the times when traffic volume is at a minimum. Hence, if noise does indeed disturb the development of SWS, in spite of the apparent need to attain this type of sleep early in the evening, its "fulfillment" may be moved to periods when environmental conditions are more favourable. The major problem associated with such forms of trend analysis is determining when differences actually become significant. Hence while there may be an apparent dissimilarity in the development of Stages 3 and 4 over noisy and quiet nights, it is quite difficult to determine if this effect is statistically significant.

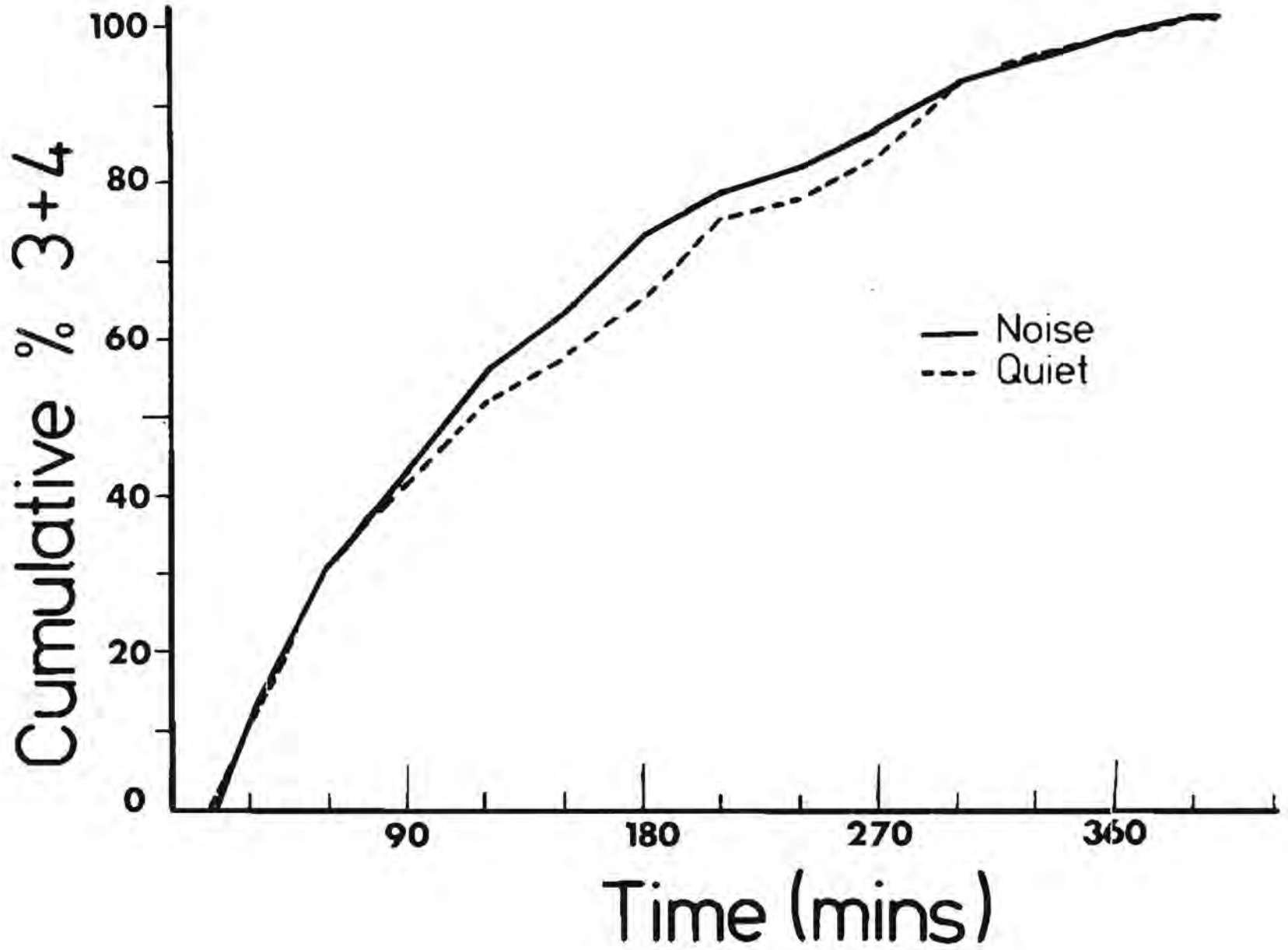

Figure 11. General trend of development of Stages 3 and 4 under noisy and quiet conditions. Note that after about 2 hours of sleep, Stage 3 and 4 evolves at a faster rate in noisy conditions than in quiet ones, then slows down later in the night, when comparatively the latter accelerates. The slowing down later in the night corresponds to a period of intense environmental noise.

Beyond Staging

It has long been recognized that the level or stage of sleep attributed to an EEG record at a given time is a classification schema based on a set of arbitrary human rules, established to put some order in what seemed to be a chaotic mass of data. With the advent of modern computers, in some sense it might have been expected that the arbitrary procedure of staging would have been abandoned. In fact, as with other fields in the neurosciences, once a standard procedure becomes rooted, it very rarely is dislodged. For purposes of inter-laboratory comparison, we continue with the traditional scoring procedures. Beyond these, however, further in-depth investigations have commenced.

Figure 12 depicts what now has become a routine analyzing format. The upper portion of the figure represents the stage profile of one night's sleep of a subject recorded in her natural environment. Because the preprocessor was strobed 128 times per 30 sec epoch (real-time), the classification summary in the form of a single stage amounts to a considerable loss of data, convenient for the human scorer who is unable to cope with more, but quite a limited usage of a computer's capabilities. It was thus decided to investigate the level of activity of each EEG event in each of these 128 samples across the entire recording session. The night was broken down into 15 minute intervals and the number of seconds of alpha, spindle, delta, and REM activity accumulated. Note that there is no concern whatsoever about arbitrary stages in this analysis. The 3 lower illustrations of Figure 12 point out intriguing trends that are entirely overlooked by simpler forms of analysis. The units (secs of activity) of course are quite variant. Thus, in this subject's record, up to 750 seconds of a 15 min interval could be occupied with delta activity, while the maximum figures for REM and spindles were 40 and 25 secs, respectively. The latter, properly speaking, are transients. Thus, only one spindle of 0.5 sec duration is enough for an EEG 30 sec epoch to be classified as Stage 2 sleep, while the same epoch would require 15 sec of delta to be classified as Stage 4. The first feature of interest is that the peaks of the respective waveforms are very predictive of particular stages of sleep. There are other features that make this form of analysis much more exciting. All three waveforms display distinctive ultradian oscillations. It is well known that Stage REM is on an approximately 90 minute cycle. Both REM and spindle activity increase in number as the night progresses, but are, however, almost exactly out of phase. When one is at a maximum the other is at a minimum and vice versa. If REM and spindle activity increase as the night progresses, some other waveform might be expected to do the opposite. This is, as suggested by Figure 9, delta. Not only does it take on a characteristic decaying oscillatory behaviour, it is, like spindles, out of phase with REM. In fact, it is the mirror image of spindle activity or from another point of view, the mirror image of the inverted REM picture.

The rhythmic nature of all three waveforms is not obvious from plots of the various stages. Gaillard (1979) thus, concludes that Stage 2 has greater rhymicity than Stages 3 or 4. The data employed in the tracings of Figure 12 are quite consistent within and between subjects and suggest that Gaillard's conclusion may not be true.

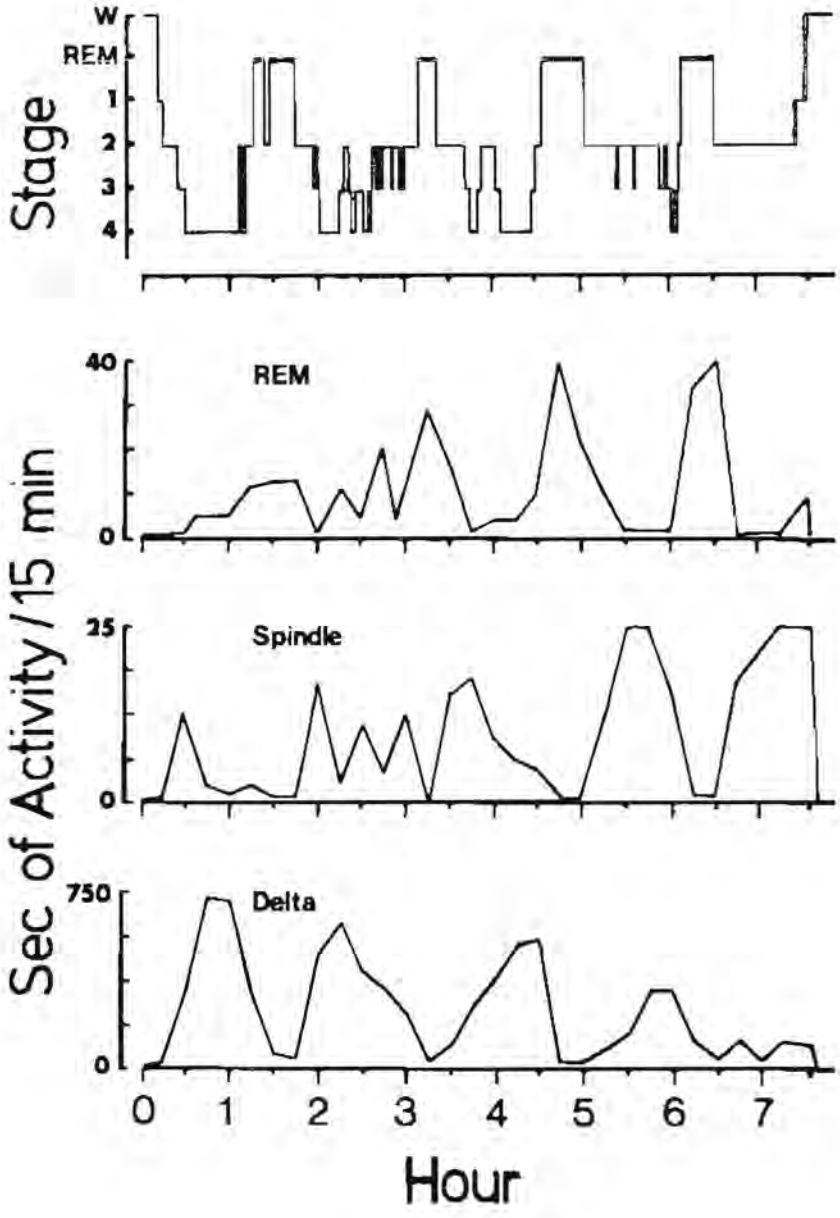

Figure 12. Computer classification of home recorded sleep into stages (upper portion) and in seconds of activity per 15 minute interval with respect to rapid eye movements, spindle transients and delta activity.

Delta activity does exhibit well-defined cyclic rhymicity, even if its quantity is limited late in the sleep period, to the extent that the threshold for Stage 3 or 4 definition is not reached. By restricting himself to stage classifications, Gaillard has overlooked slow wave activity whose total accumulation fails to reach an arbitrarily established level.

Performance Tests

Over the past two decades, it has been found that certain performance tests are sensitive indices of the effects of the quality of sleep and manipulation of the sleep-wake cycle as seen, for example, in shift work. Wilkinson (1970) in a review of such measures, has indicated that their sensitivity is dependent on a number of properties inherent in the nature of the task. Such features include: the duration of the test, level of complexity, interest value and the degree to which knowledge of results is available.

As was the case with the physiological instruments, it was quite impractical to transport the bulky, main voltage supplied psychological battery into the field setting. Two frequently used tests (Four Choice Serial Reaction Time) and the (Unprepared Simple Reaction Time) were completely redesigned while two others (Short Term Memory and the Wilkinson Vigilance Test) were modified for field situations.

Both the Four Choice Reaction Time (4CH) and the Unprepared Simple Reaction Time (USRT) tests were housed in portable battery-operated cassette recorders, enabling a significant reduction in size, weight and cost of the laboratory-based equipment. The recorders are designed to perform the triple function of housing the display and response apparatus, generating a random program of stimuli and recording the response data on the magnetic tape cassette. Back in the laboratory, the tapes are replayed, decoded and analyzed by digital computer.

Four Choice Reaction Time

The 4CH, illustrated in Figure 13 consists of a four light (light emitting diodes or LEDs) display in the form of a square. Close to these are mounted four push buttons corresponding to each of the LEDs. At the start of the test one of the LEDs is illuminated, the subject pressing the button corresponding to its position. The light is then extinguished and, after 120 msec, either the same LED or one of the other three comes on, according to a random program. The subject again makes the appropriate response and, thus, brings on the next light in the sequence and so forth. He or she continues to respond in this way as quickly and as accurately as possible for the duration of the test. Usually this is 10 minutes, although this may be optionally altered to the user's needs.

The 4CH, unlike other performance tests in our battery is self-paced--the LEDs come on at a rate corresponding to the subject's reaction time. There is of course, the possibility of error. A correct response is recorded onto the cassette as a 2 kHz tone, an error by a 4 kHz tone, each of 120 msec duration. In the laboratory, a frequency decoder is employed to reconstruct the original events and to send an appropriate TTI compatible signal to the digital interface of the computer. Because the onset of the stimulus occurs at a fixed 120

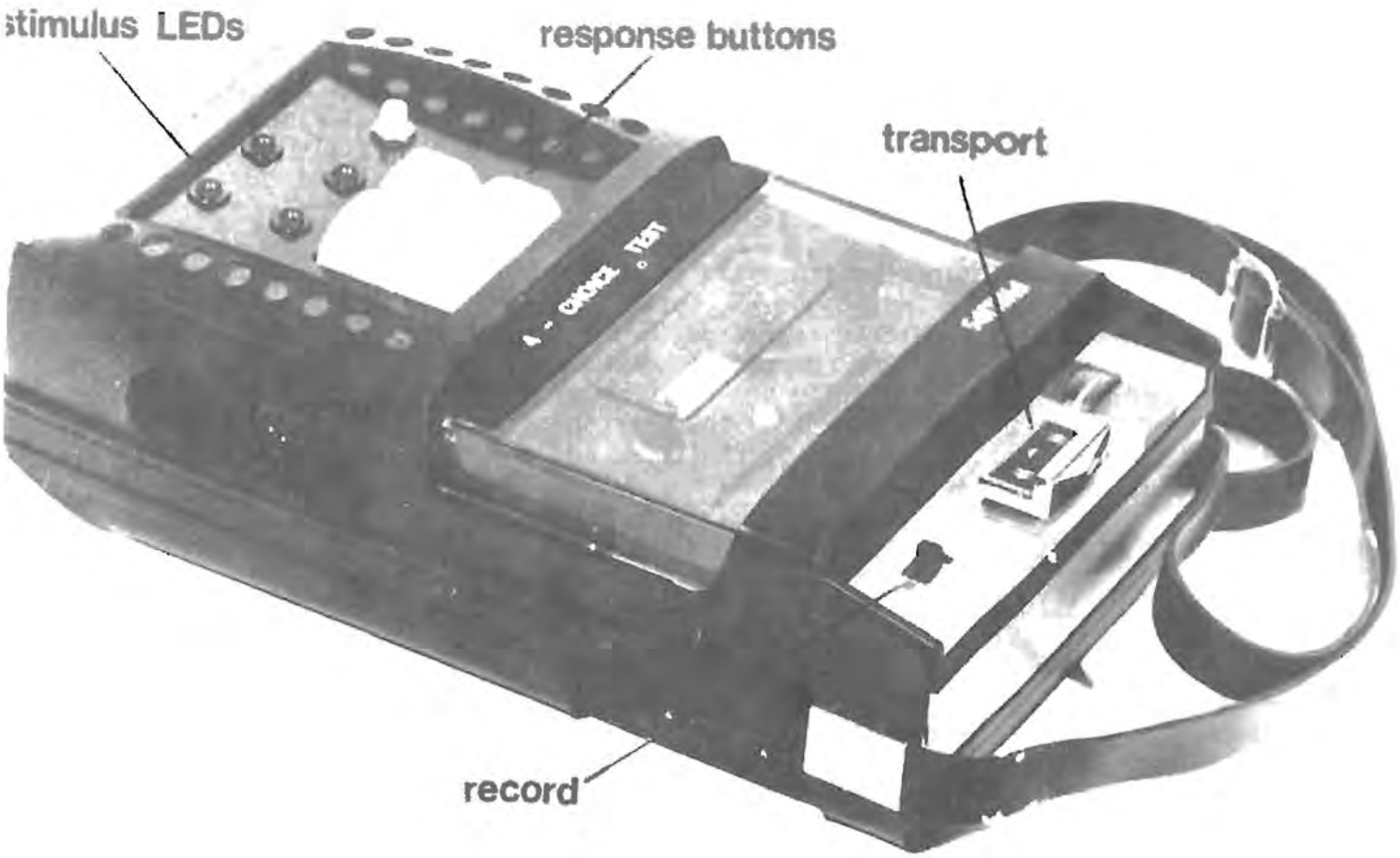

Figure 13. The portable Four Choice Reaction Time test for use in the field. One of four LEDs is illuminated, the subjects task being to respond as quickly as possible by pushing the appropriate button. Correct and error responses are coded on the cassette by 2 and 4 kHz tones.

msec after the response (corresponding to the duration of the coded tone), there is no necessity to encode it on the cassette. Thus, the reaction time for a particular correct or error response corresponds to the time between the previous response and the response in question minus 120 msec. Specific details of the circuitry and function of the test can be found in Wilkinson and Houghton (1975).

The test, because of its relative complexity, has been found to be affected by practice. The subject's performance typically improves over the first 2 or 3 sessions before leveling off. The stimulus must first be encoded, the appropriate response retrieved from memory and then the actual motor response made as quickly as possible.

Glenville and Wilkinson (1979) have described the utility of this as a measure of the effects of shift work on performance. Eleven computer operators working day, evening, and night shifts were tested over a nine week period encompassing three replications of each of the shifts. Testing took place on the first night and day shift. The subjects were examined between 0400 and 0500 hours for the night shift and 0800 and 0900 hours for the day shift. It should be noted that subjects starting the night shift had already been awake for the entire day and evening period. Thus by the end of the shift, they had been awake for approximately 24 hours. Overall, the results indicated that reaction times were significantly longer on the night than on the day shift, the effect being particularly dramatic on the second and third occurrences of each shift. While an improvement in performance was noted over the replications during the day shift, none occurred during the night shift. Moreover, performance deteriorated in the second half of the 10 minute session during the night shift as might be expected when the level of arousal declines.

With respect to specific studies on the manipulation of sleep parameters, the 4CH has been shown to be sensitive to the specific effects of 24 hour sleep deprivation (Glenville, Broughton, Wing, & Wilkinson, 1978) but less so to the minor effects of noise at night (Wilkinson et al., in press).

Unprepared Serial Reaction Time

The USRT is illustrated in Figure 14. The subject watches the window for the onset of a 000 LED display which immediately begins to count up in msec. A button is pressed as quickly as possible and the arrested display shows the reaction time. The display is then extinguished and after an interval which may randomly vary from 1 to 10 seconds, the cycle is repeated.

The reaction time is coded as a train of 1 msec pulses onto the cassette tape, the total number of which is equivalent to the reaction time (Wilkison & Houghton, in preparation). The USRT is, thus, an uncomplicated index of performance, providing knowledge of results to the subject. Because of its simplicity, it requires only a short practice session before baseline levels are attained. In most instances, a 10 minute trial is run although this can, of course, be increased or shortened depending on time available.

In the Glenville and Wilkinson (1979) study of the effects of shift work, again as in the 4CH, mean reaction time was significantly longer on the night

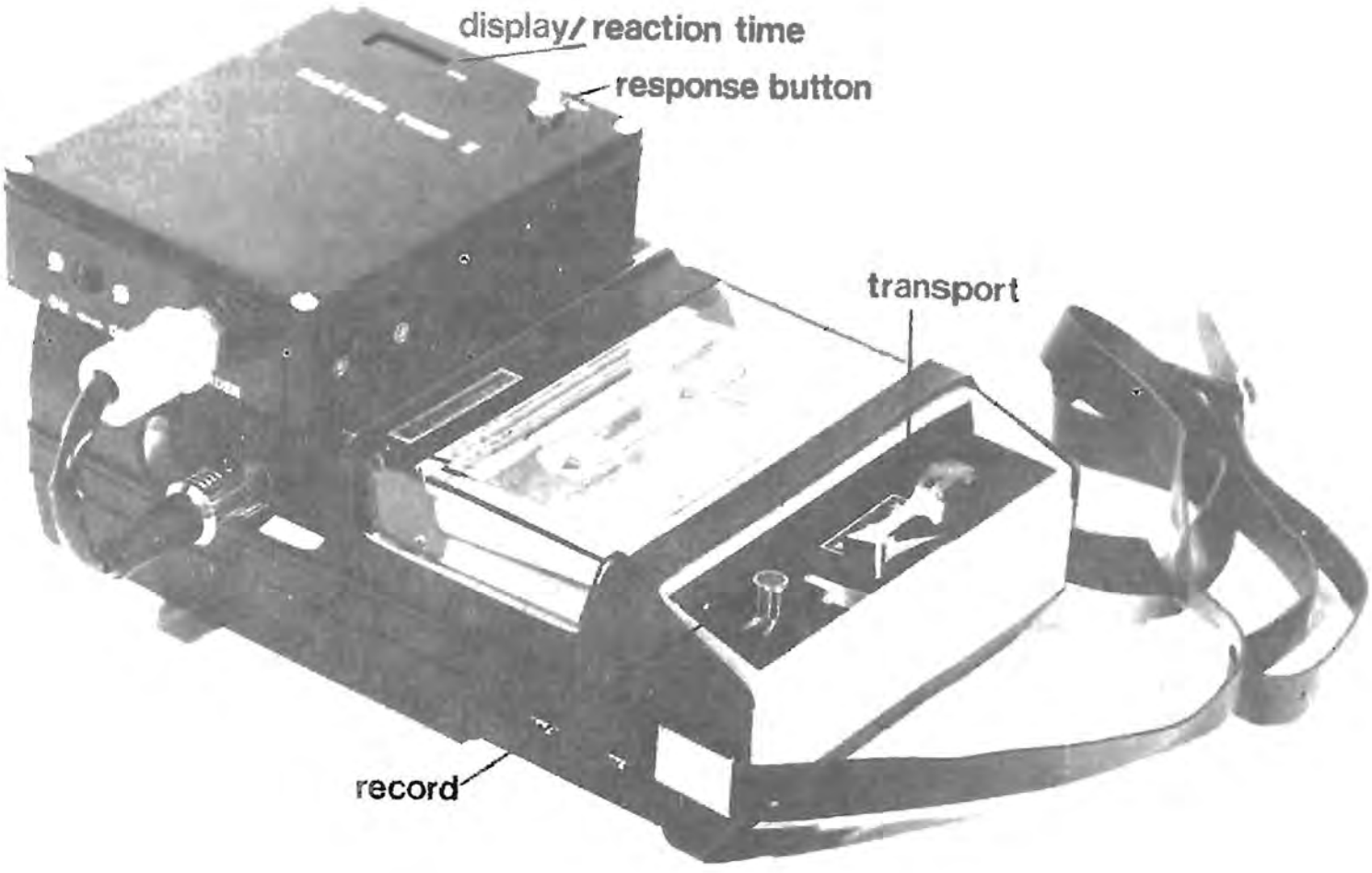

Figure 14. The portable Unprepared Simple Reaction Time test for use in the field. Subjects watch for the onset of the display and push the response button as soon as possible. The reaction time is indicated on the display and recorded as a series of pulses on the cassette tape.

than on the day shift. Moreover, during the course of the experiment the subjects become increasingly slow during the night shift but not on the day shift. Also during the former, the difference between the first and second halves of the test increased with each successive replication whereas during the latter, the difference decreased.

The results found with the portable tests are strikingly similar to the adverse effects of sleep deprivation noted by Wilkinson (1961). It was also apparent that had only an initial test been made of the effects of shift work, no significant differences would have emerged.

The USRT has also been employed in more direct investigations of sleep quality. Glenville et al. (1978) found that reduction of the number of hours of sleep resulted in a slowing of reaction times, while Wilkinson et al. (in press) noted that even relatively minor sleep disturbances, as brought about by environmental noise influences, were sufficient to alter USRT results the following day.

Short Term Memory

A process that appears to be little affected by disorders of sleep is short term memory (STM) (Hamilton et al., 1972). Indeed, some authors (see, for example, Folkard et al., 1976 for a review) claim that STM may be performed better when arousal is low than when it is high since it is complex enough to provide a high level of stimulation to the central nervous system merely as a result of carrying out the task.

In the field, the subject hears a list of digits played back from a cassette recorder at a rate of 2 digits/sec. Each list contains 8 digits randomly drawn from the set 1 through 9, but with no repetitions within the list. After each list, an interval of 6 seconds is allowed during which the subject attempts to write down the sequence he has just heard in the correct order of presentation. The test duration is again 10 minutes during which time 60 such lists are presented for recall. Errors can be one of three types: omissions, in which a digit or digits cannot be recalled; commissions, in which a digit is correctly recalled but not in its correct order of occurrence; and intrusions, in which a digit that was not in the list is falsely recalled.

The Wilkinson Vigilance Test

This test (Wilkinson, 1970) has been used quite widely in studies of sleep deprivation and other states in which arousal may vary. Of the tests mentioned, it has proven to be the most sensitive to disorders of sleep (Glenville et al., 1978). The subject listens through headphones to a cassette recording of a repetitive series of tone pips, 500 msec in duration with a regular inter-stimulus interval of 1.5 seconds, occurring in a background of "grey" noise. Occassionally and at unpredictable intervals, one of the tone pips is slightly shorter in duration than the rest (approximately 400 msec). The subject's task is to detect these signals and report them by an appropriate hand indication to the experimenter sitting out of sight of the subject. Ideally, the test continues for one hour, but can be shortened to 30 minutes depending on the time available to the subject. Performance is analysed in terms of the number of signals correctly detected (Hits) and the number of

standards incorrectly identified as signals (False Alarms). From these measures, indices of the subject's discrimination level of sensitivity (d') and willingness to respond (β) can be calculated (Tanner & Swets, 1954).

Subjective Questionnaire

While the computer profile of the subject's sleep provides an "objective" index of various parameters, a number of reports in the literature have pointed out that few of these are actually highly correlated with the subject's own impression of the quality of their sleep. Two questionnaires are therefore administered: The Stanford Sleepiness Scale (SSS) (Hoddes et al., 1973) and a locally designed scale.

The SSS is completed just prior to the subject's going to bed and immediately upon their awakening. Subjects are asked to write a number from 1 to 7 corresponding to their self-assessed level of sleepiness or fatigue.

The second questionnaire was designed to assess the subject's rating of their quality of sleep, quantity of dreams, awakenings during the night, the duration of such awakenings, and latency to the commencement of sleep. This questionnaire, which takes no more than a minute to complete, is administered in the morning.

Conclusions

The development of a sleep recording protocol in the laboratory often takes a number of years of dedicated effort before results begin to appear. The transfer of the laboratory to the natural environment has only very recently been attempted. The barriers are prodigious. Only with the advent of modern electronic trends towards miniaturization has it become feasible for progress to be made towards this end. At the same time, advances in computer technology have enabled the high-speed analysis of the vast volume of data that rapidly accumulates in the field. Automatic analysis also allows for the types of microanalyses that the human is quite incapable of carrying out.

The shift from the behavioural laboratory to the field setting may be accelerated by the methodological trends described in this article. The relatively short duration of many of the performance tests offers the possibility of an inexpensive and completely portable means of assessment for screening purposes. The research that has been conducted to date has by necessity been limited to normal samples. The application of these procedures to the patient population is an obvious "next step", but far from the only one.

Finally, the development of research tools never really ends as such. At the moment, amongst other innovations, we are investigating the feasibility of on-line analysis of sleep patterns in the home through specially constructed microprocessors.

References

Agnew,H.W., Webb,W.B., & Williams,R.L. The first night effect: An EEG study of sleep. Psychophysiology, 1966, 2, 263-266.

Broughton,R., Healey,T., Maru,J., Green,D., & Pagurek,B. A phased lock loop device for automatic detection of sleep spindles and stage 2. Electroencephalography and Clinical Neurophysiology, 1978, 44, 677-680.

Campbell,K., Kumar,A., & Hofman,W. Human and automatic validation of a phase locked loop spindle detection system. Manuscript in preparation.

Campbell,K., Weller,C., & Wilkinson,R.T. A portable pulse-interval modulation telemetry/multiplexing EEG recording system for use in the home. Electroencephalography and Clinical Neurophysiology, in press.

Cashman,P.M.M., & Stott,F.D. A semi-automatic system for the analysis of 24 hour ECG recordings from ambulatory subjects. Biomedical Engineering, 1974, 9, 54-57.

Feinberg,I. Changes in sleep cycle patterns with age. Journal of Psychiatric Research, 1974, 283-306,

Folkard,S., Knauth,P., Monk,T.H., & Rutenfranz,J. The effect of memory load on the circadian variation in performance efficiency under a rapidly rotating shift system. Ergonomics, 1976, 19, 479-488.

Gaillard,J.-M. Organisation temporelle des stades du sommeil chez l'homme. Electroencephalography and Clinical Neurophysiology, 1977a, 42, 754-760.

Gaillard,J.-M. Les tendances générales des stages du sommeil étudiées par adjustments de la polynomes orthogonaux. Electroencephalography and Clinical Neurophysiology, 1977b, 42, 847-851.

Gaillard,J.-M. Temporal organization of human sleep: General trends of sleep stages and their ultradian cyclic components. L'Encéphale, 1979, 5, 71-93.

Gaillard,J.-M., & Tissot,R. Principles of automatic analysis of sleep records with a hybrid system. Computers and Biomedical Research, 1973, 6, 1-13.

Glenville,M., Broughton,R., Wing,A.M., & Wilkinson,R.T. Effects of sleep deprivation on short duration performance measures compared to the Wilkinson Auditory Vigilance test. Sleep, 1978, 1, 169-176.

Glenville,M., & Wilkinson,R.T. Portable divices for measuring performance in the field: The effects of sleep deprivation and night shift on computer operators. Ergonomics, 1979, in press.

Green,D. A hybrid preprocessor for sleep staging using the EEG. Thesis, Carleton University, Ottawa, Canada, 1975.

Hamilton,R., Wilkinson,R.T., & Edwards,R.S. A study of four days partial sleep deprivation. In W. P. Colquhoun (Ed.), Aspects of human efficiency. London: English University Press, 1972, 101-113.

Hoddes,E., Zarcone,V., Smythe,H.R., & Dement,W.C. Quantification of sleepiness: A new approach. Psychophysiology, 1973, 10, 431-436.

Itil,T.M., Shapiro,D.M., Fink,M., & Kassebaum,D. Digital computer classitication of EEG sleep stages. Electroencephalography and Clincal Neurophysiology, 1969, 27, 76-83.

Ives,J.R., & Woods,J.F. 4-Channel 24 hour cassette recorder for long-term EEG monitoring of ambulatory patients. Electroencephalography and Clinical Neurophysiology, 1975, 39, 88-92.

Johns,M.W., Stear,E.B., & Hanley,J. Tracking the dominant frequency of the EEG by phase-lock demodulation. Electroencephalography and Clinical Neurophysiology, 1974, 37, 414-416.

Kumar,A. The complex demodulation method for detection of α waves and sleep spindles of the human EEG in real-time. Proceedings on Advanced Signal Processing Technology, Lausanne, 1975, 355-361.

Kumar,A. A real-time system for pattern recognition of human sleep stages by fuzzy system analysis. Pattern Recognition, 1977, 9, 43-46.

Larsen,L.E., & Walter,D.O. On automatic methods of sleep staging by EEG spectra. Electroencephalography and Clinical Neurophysiology, 1970, 29, 516-520.

Lubin,A., Johnson,L.C., & Austin,M. Discrimination among states of consciousness using EEG spectra. Psychophysiology, 1969, 6, 122-132.

Martin,W.B., Johnson,L.C., Vigiolone,S.S., Naitoh,P., Joseph,R.D., & Moses,J.D. Pattern recognition of EEG-EOG as a technique for all-night sleep stage scoring. Electroencephalography and Clinical Neurophysiology, 1972, 32, 417-427.

Monroe,L.J. Interrater reliability and the role of experience in scoring EEG sleep records, phase 1. Psychophysiology, 1969, 5, 376-384.

Rechtschaffen,A., & Kales,A. (Eds.), A manual of standardized terminology, techniques, and scoring system for sleep stages. BIS/BRI, UCLA, Los Angeles, 1968.

Smith,J.R. Computers in sleep research, CRC Critical Reviews in Bioengineering, 1978, 93-148.

Smith,J.R., Funke,W.F., Yeo,W.C., & Ambuehl,R.A. Detection of human sleep EEG waveform. Electroencephalography and Clinical Neurophysiology, 1975, 38, 435-437.

Smith,J.R., & Karacan,J. EEG sleep stage scoring by an automatic hybrid system. Electroencephalography and Clinical Neurophysiology, 1971, 31, 231-237.

Tanner,W.P., & Swets,J.A. A decision-making theory of visual detection. Psychological Review, 1954, 61, 401-409.

Tissot,R. Essai méthodologique d'analyse du cycle du sommeil. L'Encéphale, 1966, 2, 137-169.

Weller,C. Radio or wired telemetry for clinical applications? In P.A. Neukomn (Ed.), Biotelemetry II, Karger, Busel, 1974, 208-210.

Wilkinson,R.T. Methods for research on sleep deprivation and sleep function. In E. Hartmann (Ed.), Sleep and Dreaming, Little Brown & Boston, 1970.

Wilkinson,R.T., Campbell,K.B., & Roberts,L.D. The effects of noise at night upon performance during the day. Proceedings of the 3rd Congress on Nosie as a Public Health Problem, Freiberg, 1978, in press.

Wilkinson,R.T., Herbert,M., & Branton,P.A. A portable EEG-sleep recorder of pocket size. Sleep Research, 1973, 2, 182.

Wilkinson,R.T., & Houghton,D. Portable four-choice reaction time test with magnetic tape memory. Behavior Research Methods and Instrumentation, 1975, 7, 441-446.

Wilkinson,R.T., & Houghton,D. A simple reaction timer with magnetic tape memory. Manuscript in preparation.

Wilkinson,R.T., & Mullaney,D. Electroencephalogram recording of sleep in the home. Postgraduate Medical Journal, 1976, 52, 92-96.

Williams,R.L., Karacan,I., & Hursch,C. Electroencephalography (EEG) of human sleep. New York: John Wiley & Sons, 1974.

The research was supported by Grant 183-77-8 ENV UK from the Environmental Research Division of the Commission of the European Communities, Brussels, Belgium.

NIOSH

PROCEEDINGS

THE TWENTY-FOUR HOUR WORKDAY: Proceedings of a Symposium on Variations in Work-Sleep Schedules

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control
National Institute for Occupational Safety and Health

THE TWENTY-FOUR HOUR WORKDAY: PROCEEDINGS OF A SYMPOSIUM ON VARIATIONS IN WORK-SLEEP SCHEDULES

EDITORS

Laverne C. Johnson
Donald I. Tepas
W. P. Colquhoun
Michael J. Colligan

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control
National Institute for Occupational Safety and Health
Division of Biomedical and Behavioral Science
Cincinnati, Ohio 45226

July 1981

DHHS (NIOSH) Publication No. 81-127