THE RELIABILITY OF A DEPTH OF SLEEP MEASURE AND THE EFFECTS OF FLURAZEPAM, PENTOBARBITAL, AND CAFFEINE ON DEPTH OF SLEEP

By

MICHAEL H. BONNET

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL
OF THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1977

ACKNOWLEDGMENTS

The author wishes to thank the members of his committee, Dr. Keith Berg, Dr. C. Michael Levy, Dr. William Yost and Dr. Steven Zornetzer for their time, interest, and guidance not only in this project but throughout his graduate career. Also the author must thank Dr. Doris Chernik and Hoffman La Roche Company for their academic and monetary interest in funding this project.

It is a humbling experience to look at the effort others have invested in oneself. And the small return is often shameful. A very long list of people appears to whom the author owes gratitude for showing him what an educator, a psychologist, and a friend could be. To parents and friends and, perhaps most of all, to Dr. Wilse B. Webb, the present work is dedicated as a very small interest payment on a very large debt. As the son is the father to the man, so the gifts many have given are expressed herein and so shall they be expressed in any work of value which this author produces.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the
Graduate Council of the University of Florida
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

THE RELIABILITY OF A DEPTH OF SLEEP MEASURE AND THE EFFECTS OF FLURAZEPAM, PENTOBARBITAL, AND CAFFEINE ON DEPTH OF SLEEP

By

Michael H. Bonnet

June 1977

Chairman: Dr. Wilse B. Webb
Major Department: Psychology

In the initial experiment a reliability estimate of the presently used depth of sleep measure was ascertained, both for the raw threshold data and for data which resulted from three transformations to add control. For six trials (an average night) in six subjects, the reliability estimate for the raw data was .90 and for the controlled data was .95.

The sensitivity of the threshold measure was then tested in a study involving flurazepam, pentobarbital, caffeine, and placebo conditions. Thresholds were measured from a standardly defined segment of Stage 2 sleep for six subjects for eight nights. Flurazepam and pentobarbital were seen to increase thresholds, both in a pure sleep measure and in awake thresholds, in a time course fashion

across the night. Trend analyses indicated that flurazepam was both a faster-acting and a longer-lasting drug than pentobarbital, but the main effects of both drugs on sleep thresholds were seen in the first 3 1/2 hours of sleep. Caffeine administration resulted in lower thresholds, as a concomitant largely of lower awake thresholds, than the placebo or the other drugs. Such evidence favors the contention that depth of sleep is, at least in some aspects, a different measure of sleep behavior from the EEG because depth of sleep has been shown to vary within EEG sleep stage. However, hypotheses designed to show the operation of a dual arousal system which could explain both sleep depth and EEG findings by showing differential arousal characteristics for the barbiturate versus the benzodiazepine were not upheld. A central mechanism for sleep depth which was independent but related to EEG controlling mechanisms was implied by the findings but not identified.

Concurrent threshold results included evidence that the awakening threshold is inversely related to the time it took subjects to fall asleep again, and thus, that the latency measure could also serve as an index of drug activity. Both the latency information and the threshold time course information were upheld to some extent by subjective measures collected the following morning. Substantial agreement among the three data bases was considered a validation of all three measures of the sleep process.

INTRODUCTION

> Where no part of the soul remained behind, concealed in the limbs, as fire remains concealed when buried under much ash, whence would sense be suddenly rekindled through the limbs, as flame can spring up from hidden fire?
>
> Lucretius

The threshold of response is basic to all of perception, discrimination, and higher psychological functions. As such, threshold maintains a predominant role in all of behavior. It is perhaps fitting that the earliest studies of sleep sought to determine awakening thresholds throughout the night as their definition of the sleep process. Elegant threshold curves were produced by De Sanctis and Neyroz (1902), and others, and their data closely approximated later electroencephalographic (EEG) pictures of the time course of the sleep process. Threshold measures of sleep were common for about 70 years, although very few studies were done until the use of EEG became commonplace. The appeal of the threshold studies was simply that behavioral responsivity as a function of signal intensity is an intuitively obvious way to approach sleep and one that seems fairly simple.

There were some good reasons for the switch to EEG as a measure, though. There are many serious problems involved in sleep threshold experiments. Results from different sensory modalities differ and may not be directly comparable. Different types of required response result in different data. Effects of instruction, expectancy, and subject motivation can drastically alter results. The manner of signal presentation affects measurement. The number and order (Sokolov, 1963) of stimulus presentation can affect threshold both in waking and in sleeping subjects. Some of these problems as they apply to threshold measurements during sleep have been briefly noted by Webb and Agnew (1969), but it must be stressed that these problems apply not only during sleep but also in waking subjects. The further problem of an interaction of sleep/waking with any of the above factors must also be of interest.

As a part of the problem of threshold measures in general, "the depth of sleep, like its quality, is an elusive characteristic" (Kleitman in the 1939 edition of his book). Several investigators have vainly sought to untangle the complexities of the problem of sleep depth. Four major problems have obscured a clear understanding of depth of sleep. First is the disassociation of autonomic measures during sleep. Heart rate, respiration, body temperature, and the whole spectrum of autonomic responses do not simply relate to sleep depth (Kleitman,

1939; Rechtschaffen, Hauri, & Zeitlin, 1966; Williams, 1967), nor do they relate to each other during sleep. The implication is that if any one psychological measure is used as the indicant of sleep depth, it will not agree with either awakening thresholds or other physiological measures. The second problem involves subject responses. Threshold studies are based upon a subject response. If the presentation of a tone is followed by a button press response in two subjects and one of those subjects wakes up and remembers the whole incident the next morning while the other shows no other response or memory for the event, what does one conclude about the two subjects? They appear to have equal and yet unequal thresholds at the same time. Both of the first two problems are "what are you interested in" problems. If one is interested in a general physiological index, the answer does not exist, but if one asks a specific question such as, "How loud a sound does it take for a person to say, 'I am awake,'" such problems dissolve except perhaps at a philosophical level. A third problem is perhaps a blessing in disguise. It is the discovery and popularity of the EEG as a measure. Since the more complete delineation of the EEG sleep cycle in 1953, there has been a tendency to use EEG as the criterion of depth of sleep rather than in investigating depth of sleep. For this reason few systematic studies of sleep depth have been done (Williams, 1967). The fourth problem is really a function of EEG measures and depth and involves the lack of

agreement between human and animal studies of sleep depth. This has limited the exploration of sleep depth.

The use of the EEG as a measure of sleep avoids several of the rather spectacular problems associated with sleep depth measures, but the EEG has several problems of its own as a measure. The major problems of EEG deal with its relation to actual, observable behavior. At one level, for example, Johnson (1973), in a review of the relation of sleep stages to subsequent performance behavior, said the relation "remains a mystery" (p. 337). At another level, a whole series of research has observed that the relation of ongoing EEG to behavior is not causal but only correlational. The result is that an experimental manipulation can separate the two. As early as 1938, Davis, Davis, and Thompson reported that subjects breathing gas mixtures containing half as much oxygen as normal showed delta activity while still being responsive to commands to open or close their eyes or to write. Drugs produced similar effects. For example, atropine produced slow waves similar to delta in animals that were neither drowsy or unresponsive (Bradley, 1958; Wikler, 1952, cited in Lacey, 1967). As such, if drugs disassociate EEG from behavior, EEG would seem a very poor measure to use in describing the effects of drugs because any effects seen might have little relation to behavior. Feldman and Waller (1962) found that animals with lesions in the reticular formation showed slow wave cortical activity while behavioral activity was

present. Conversely, lesions in the posterior hypothalamus resulted in animals behaviorally comatose who still showed fast wave, "waking" brain activity in response to reticular stimulation. Lacey in a 1967 review concludes that "electrocortical arousal, autonomic arousal, and behavioral arousal may be considered to be different forms of arousal, each complex in itself" (p. 15). In short, EEG is never a perfect correlate of behavior and occasionally is randomly or even inversely related to behavior. Further, the disassociation may be rather common. It may even be seen in sleep deprivation in normal human subjects (Blake, Gerard, & Kleitman, 1939).

There is one more disturbing problem with the use of EEG as a measure. The Feldman and Waller study (1962) hypothesized that while the reticular system mediated cortical tone, the posterior hypothalmic region had a functionally separate role in maintaining waking. If that hypothesis were correct, and the data did support it, it would mean that the simple use of EEG measures would overlook an entire dimension of behavior--specifically that mediated by the hypothalamic system.

Threshold measures, which form the basis for the label "sleep depth" herein, traditionally have been oriented toward the realm of physical behavior (here defined as requiring a motor response). As such, the problem of the degree of correlation with motor behavior is nonexistent. Because of this the hypothalamic components proposed by Feldman and Waller could perhaps best be measured by a

threshold type study. The problem is that a whole gamut of learning effects and all the shortcomings of threshold measures are also introduced. The potential experimenter is faced with a mass of variables and interrelationships. This is both good and bad. Because of this, researchers in the area call for more studies and bemoan the fact that depth of sleep is a useless variable. In 1967, Williams rhetorically asked if the term "depth of sleep" should be given up. He answered "no," because no sequence of experiments controlling all the relevant variables had been done. In a following discussion Snyder agreed with Williams in these words (Note 5):

> It has become kind of a cliche among workers in this field that the concept of depth of sleep has little meaning, yet I think that we all have a great deal of conviction that in terms of common sense experience is a very important idea. Perhaps if our science cannot provide an operational definition of depth of sleep, then this is the short-coming of our science. (p. 314)

To complete the circuit of perhaps the three most knowledgeable researchers in the area, Rechtschaffen et al. (1966) said,

> Awakening threshold has probably not received the interest it deserves in its own right. . . . Surely such an adaptive mechanism represents the product of "careful" natural selection and deserves attention as a fundamental biological phenomenon. (p. 937)

A recent review (Bonnet, Note 1) cites about 150 references in the area of depth of sleep. A complete review is clearly inappropriate to the present work with three broad and important exceptions--the reliability of

the measure, the present use and usefulness of the measure, and the application of the measure in new areas.

Central to the experimental use of any measure are questions concerning the reliability and the validity of that measure. While several sources have commented on large between- and within-subject differences in threshold during sleep, only one source (Note 10) has commented on the reliability of such measures. In that study, as the result of a complex awakening schema, six subjects were classified as either light, medium, or deep sleepers. Using the same schema on a second night, three were rated the same and three shifted by one group. On these results, reliability was deemed "sufficient." The lack of measurement reliability data in a measure known to exhibit wide variability is unfortunate because it limits conclusions on the validity of findings. A major reason for the paucity of data in this area is probably the fact that threshold measures, in general, have been taken to present considerable face and content validity. Validity, of course, is not reliability, and a truly valid waking measure is not necessarily a reliable or valid measure of sleep.

Even with the assumption that depth of sleep measures are both reliable and valid, one must question their value. Are they responsive to parameters that sleep as a behavior should be responsive to? Can information not already obtained from EEG studies be obtained by the use of such measures? Although systematic studies have not been done

to answer these questions, studies relevant to both have been done. A large body of information has grown concerning information processing during sleep (see, for example, Williams, 1973). A number of studies have been reported spun off from the classic Oswald, Taylor, and Treisman (1960) study, in which both EEG and arousal probabilities were greater to a subject's name than to another name or the name played backwards. Sleep depth obviously varies with signal relevance. Either as a part or independently of signal relevance, subject motivation, in terms of payment for hearing a signal versus nonpayment, has also been shown to have a large effect on responsivity during sleep (Wilson & Zung, 1966; Zung & Wilson, 1961). On the other hand, a study designed to find the effects of subject motivation, again in terms of payment, on the EEG sleep stage distribution across the night (Bonnet & Webb, 1976) found only marginal changes in EEG patterns.

Another variable which sensibly should affect the sleep process is the length of prior wakefulness. A large volume of work (see, for example, Agnew & Webb, 1971; Webb & Agnew, 1971) has documented the EEG effects of prior wakefulness. One study (Williams, Hammack, Daly, Dement, & Lubin, 1964) has examined the effects of prior wakefulness on behavioral threshold. Very briefly, it was found that 64 hours of sleep deprivation, in addition to changing the EEG sleep distribution, also lowered the number of responses to tones within sleep stage. While the response

percentage in Stage 2 sleep was 24% before deprivation, it was only 4% after deprivation. As a comparison, the response level in Stage 4 before deprivation (9%) was a good deal greater than responsiveness in Stage 2 after deprivation. This study has the interesting conceptual addition that sleep depth can vary within sleep stage and could therefore represent the operation of different systems in the sleep response.

The characteristic EEG distribution across the night was documented in detail as early as 1964 (Williams, Agnew, & Webb). Five sources (Goodenough, Lewis, Shapiro, & Sleser, 1965; Keefe, Johnson, & Hunter, 1971; Rechtschaffen et al., 1966; Shapiro, Goodenough, & Gryler, 1963; Watson & Rechtschaffen, 1969; Zimmerman, 1970) point to a characteristic pattern of depth of sleep across the night which operates not only with sleep stage but also within sleep stage. The major characteristic of that pattern is a lightening of sleep within sleep stage as the night progresses.

Circadian effects on the sleep process as measured by EEG have been reported (see, for example, Agnew & Webb, 1973; Webb & Agnew, 1971; Webb, Agnew & Williams, 1971). One study, reported by abstract only (Note 3), has examined the effects of sleep placement on depth of sleep. Corvalan and Hayden (Note 3) determined awakening thresholds for subjects in sleep periods beginning at 10:00 p.m., 8:15 a.m., 9:15 a.m., 10:15 a.m., and

noon. A wide range of threshold values was reported in each stage of sleep, and it appeared that the values from REM and Stage 4 depended upon real time. The highest values were recorded between 10:00 p.m. and 8:00 a.m. Corvalan (1969) reports reaction time to a constant level stimulus during sleep across day and night. Only reaction time during REM appeared time locked, and it was seen to decrease with absolute time across the 24 hours. From the brief information, there appear to be many problems with these experiments. There was unequal distribution of time of sleep period with none occurring in the evening, and there appeared to be no attempt to control for effects of prior wakefulness. In regard to Stage 4, results are inconsistent--one method found time dependency and one did not. The results on REM agree with those studies examining sleep depth during the normal sleep time. However, evidence does indicate the presence of some circadian effects.

Two other variables, drugs and age, have been found to have profound effects on EEG patterns. However, neither of the variables has been examined for sleep depth characteristics. There is a need for further information in both areas, not only because such information will establish further the covariation of EEG and threshold measures, but also because there is a tremendous potential for both applied and theoretical data in both areas. In the present studies, the area of drug effects on depth of sleep will be examined more closely.

There have been two common approaches in the study of the effectiveness of drugs in human sleep. The earliest very simply involved administering the drug and asking the subject in the morning how he had slept. Such procedures, of course, allowed several types of bias, and the relationship between sleep depth and such ratings is unknown. Jick (1969) reviewed this data and found it difficult to differentiate hypnotics used within this methodology.

The second approach, which is at present the common approach, was to use the EEG as a measure. Several recent reviews are readily available including Oswald (1968, 1969, 1973, 1974), Hartman (1969, 1973), and Winters (1969). The approach of all the reviews and of the studies reviewed has been similar. Following drug administration, EEG measures of sleep latency, sleep stages, sleep stage shifts, and sleep length were recorded and compared. A very recent and interesting application of the EEG in the examination of ongoing brain activity to chart drug effects has been the development of what Itil has called "sleep prints" (1976). Such prints chart occurrence of various wave frequencies in the EEG more completely than is seen in sleep staging. Flurazepam, for example, results in a sharp decrease (over placebo) in wave frequency seen about two hours after sleep onset followed by an increase (above placebo) in frequencies, peaking at about four hours after sleep onset and staying above placebo for the remainder of the night. The meaning of these changes, however, is not clear. Although

shifts in stage distributions have been associated with "deeper" or "lighter" sleep as a result of the general relation between sleep stage and sleep depth, the variable of sleep depth within sleep stage was never mentioned. A rather thorough review of the literature revealed a rather alarming paucity of studies (Bonnet, Note 1).

The single most applicable drug study involving sleep depth in the published literature was performed by Lindsley in 1957. He was interested in response rate (a button push) to reduce tone intensity as a result of several sleep conditions including sleep deprivation and a dose of seconal. Seconal is a short-lasting barbituate which usually induces sleep in 15-20 minutes. Lindsley's data showed normal response rate during about a 15-minute sleep latency (slightly shorter than the two subjects' normal latency). Sleep onset was followed by approximately 4 hours of virtual nonresponse mirroring almost exactly the nonresponse characteristic of 38 hours of sleep deprivation. At this point responding began and increased. With responses during sleep latency subtracted, about 600 responses were made during the control night, about 440 on the drug night, and about 300 on the night following sleep deprivation. It is obvious that the experimental conditions had an effect, but, of course, it cannot be determined if those effects were due to increased amounts of Stage 4 or increased difficulty of responding within stages or both.

Sedatives have been found not to alter brief latency evoked potentials but to modify the threshold and waveform of longer latency responses (Davis, 1973; Price & Goldstein, 1966). However, with the exception of comments by Monninghof and Presbergen in 1883 that consumption of a large amount of alcohol made subjects less sensitive to stimulation in earlier parts of the night, sleep depth during the night has simply not been examined except circuitously. Mullin, Kleitman, and Cooperman (1933) examined the effects of alcohol and caffeine on motility during sleep. The results indicated that alcohol consumption caused a distinct reduction in body movement during the first half of the night and a possible reduction in movement for the entire night and that caffeine produced a marked increase in motility during sleep. Mullin, Kleitman, and Cooperman (1937) found that sleep depth was related to the length of time following a body movement. Such would imply that alcohol deepened sleep while caffeine lightened it. Still these findings could have been the result of shifts in sleep stage distribution as easily as a threshold shift within sleep stages.

In a pair of more recent studies (Itil, Saletu, Marasa, & Mucciardi, 1972; Itil, Saletu, & Marasa, 1974) subjects were awakened at 5:00 a.m. by a tone series (1000 Hz, 1 second in duration) beginning at 40 dB (reference not given) and increased in 10 dB steps to 80 dB. The 5:00 a.m. awakening was not controlled for sleep stage and occurred

seven hours after drug administration and attempted sleep onset time. Flurazepam and U-31,889, which is a triazolo-benzodiazepine derivative, were tested against placebo in the 1972 study. At 5:00 a.m., thresholds were significantly higher to the highest dose of U-31,889 (2 mg) than to any other drug or placebo condition. The major problem with the study is that both flurazepam and U-31,889 are classed as short-acting drugs with time courses usually estimated at less than 7 hours. As such one would predict the null hypothesis in all conditions at a test 7 hours after administration. The finding with the highest dose of U-31,889 would, however, additionally raise the question that not only depth of sleep but also drug time course might hypothetically vary with drug dosage. However, the potential results of this study are further confounded by the fact that no control was exercised on awakening parameters (with the exception of time). It is a fact that most drugs alter sleep stage characteristics in complex fashions so that the probability of sleep stage occurrence in the reported experiment differed with the drug given. There is, therefore, a nonrandom probability of occurrence of any sleep stage at 5:00 a.m. This could account for the reported differences (or cover more extreme differences) even though threshold differences between Stage 1-REM and Stage 2 appear minimal (see Note 1). Further, it is conceivable that some condition might predispose subjects to have increased awakenings or to be awake naturally around

5:00 a.m. (i.e., early morning). Because thresholds are lower shortly after an awakening (Mullin, Kleitman, & Cooperman, 1937), this could also confound results. The 1974 study contains the same methodology and faults. In it, methaqualone, triazolam, and flurazepam were tested against placebo, and thresholds were significantly greater than placebo with triazolam and flurazepam at 5:00 a.m. In a third study reported in Williams and Karacan (1976) no difference from placebo was found at the 5:00 a.m. awakening for diazepam or clorazepate dipotassium.

Some information on drug effects is also available from other sources. As a part of some methaqualone studies done in the Florida Sleep Laboratories (Note 9), Rorer provided some brief results of studies they had done on their product (Quaalude). They reported that a dose of 10 mg/kg usually increased voltage of sciatic nerve stimulation necessary for EEG arousal in the cat by 100%. A dose of 5 mg/kg elevated threshold by 70%. Sleep stage tested was not specified or cited as controlled, and the nonrandom variations earlier discussed probably did exist. Natural within-sleep threshold shifts may or may not have played a factor. However, if those variables could be considered random, the study provided evidence for a drug threshold shift in the cat.

Other animal studies, recording from various sites ranging from the limbic system to the reticular formation, have found increases of threshold to various arousal

responses (Gogolak & Pillat, 1965; Randall, Schallek, Scheckel, Stefko, Banziger, Pool & Moe, 1969; Schallek & Kuehn, 1965; among others). The usual stimulus has been stimulation in the reticular formation. Selective threshold increase at limbic versus reticular sites in these studies has been taken as evidence of site of drug action, a point which will be discussed later. It should be mentioned that these findings are not universal. Lanoir and Killam (1968) found that their cats did not even go to sleep after administration of valium or mogadon. However, for present purposes, an in-depth review of this data beyond the statement that threshold increases have been in general found is perhaps not strictly relevant for the following reasons. In all of these studies the point of stimulation and the point of measurement have usually been deep within the brain, and the relation to "real world" stimuli or behavioral responses is uncertain. Also, the initial criticisms of the results provided by Rorer hold generally. The largest problem, however, is that it has been well established that sleep depth in animals does not correspond well with sleep depth in humans (Jouvet, 1961; Williams, 1967; or see Bonnet, Note 1, for review) as it relates to EEG sleep stage. The strong implication is that the sleep system in man is different from that of the rat or cat. Human data, conceivably, relates only tangentially to the animal data, and the present concern is primarily with the human.

The state of the area of human research on thresholds after drug administration is summarized by Itil in a chapter published in late 1976:

> Despite the fact that awakening threshold seems to be one of the most important parameters for determination of the quality of sleep, we were unable to find any recent literature regarding that measurement. (p. 236)

As a part of a pilot study in 1975, two human subjects have been run in our laboratory. Subjects were given a standard 300 mg tablet of Quaalude 15 minutes before retiring on the drug night. Subjects had been previously adapted to the laboratory. Five-second segments of auditory stimuli were presented in an ascending method of limits design which terminated when a button push response was made. Stimuli were presented only in Stage 2 sleep, and 5-7 arousals were made per night. Data from each subject were plotted with data points connected by straight lines. Interpolated threshold at eight times across the night was averaged between subjects. The results were clear. (a) Threshold drops on both the drug night and control night were seen over the sleep period within Stage 2. The finding replicates Rechtschaffen et al. (1966), Watson and Rechtschaffen (1969), Zimmerman (1970), and Keefe, Johnson, and Hunter (1971) among others. (b) The Quaalude appeared to stably increase Stage 2 threshold by about 8 dB throughout most of the night. Because neither of these results could have been predicted from simple monitoring of the EEG, it seems that a more

encompassing view of sleep as a process is necessary. One possible approach is the dual arousal system approach advanced by Routtenberg (1966, 1968) and modified to apply specifically to studies of responsivity during sleep by Bonnet (Note 1). Very briefly, the models proposed that two arousal systems mutually interact and that responsivity is a function of activity in both. Arousal System I is the classic reticular system, which greatly influences the cortical EEG. Arousal System II is a limbic midbrain system involved primarily with incentive and less completely represented in cortical EEG (Feldman & Waller, 1962). This implies that threshold shifts mediated by Arousal System II activity may not be adequately represented in ongoing EEG. Such is a conceivable explanation of why some insomniacs complain of getting little or no sleep while displaying normal EEG sleep activity. This proposition (and others) could be tested only by use of sleep depth measures.

In terms of drug effects in the two hypothesized systems, there are a large number of animal studies at two levels of examination. At the more gross level, which can be taken to mean results by experimenters recording brain rhythm activity from various brain structures after drug administration, evidence suggests that (1) barbiturates act at the level of the reticular formation (Kido & Yamamoto, 1962; Schallek & Kuehn, 1965; Schallek, Kuehn, & Jew, 1962); and (2) that minor tranquilizers work by inhibiting limbic structures (Schallek & Kuehn, 1965; Schallek et al.,

1962; Schallek & Thomas, 1971; Steiner & Hummel, 1968). Olds & Olds (1969) concluded,

> There now appears to be general agreement that there are drug receptors in the mesencephalic reticular formation which are directly acted upon by barbiturates before action on other structures occurs, and that such action is largely responsible for the behavioral and EEG effects noted with these compounds.

On the other hand, the benzodiazepines appear to have a "selective action" on hippocampal neurons (p. 100).

However, Fuxe and his coworkers, working at the neurotransmitter level, have identified several transmitter pathways in the rat brain but have had little luck in differentiating barbiturate and benzodiazepine action in at least eight studies in the past eleven years (see Fuxe, Hokfelt, & Ungerstedt, 1970; or Lidbrink, Corrodi, & Fuxe, 1974).

There is at present no easy way to combine these two bases of data and additionally no promise that neurophysiology in the rat is similar to neurophysiology in the human. It is the human case, of course, in which one cannot implant electrodes or lesion, that is of primary interest. Gross effects are all that one could hope to find.

Routtenberg (1968) reviewed a finding by Kornetsky and Bain (1965) that pentobarbital and chlorpromazine administration altered waking responses to a choice task in animals. Pentobarbital increased errors of commission while chlorpromazine increased errors of omission. This

was suggestive to Routtenberg that the locus of the action of the drugs might be different and the chlorpromazine might act to depress Arousal System II. Such would be a logical place of action for a tranquilizer. Because cortical synchrony is primarily controlled by Arousal System I (Routtenberg, 1966), barbiturate effects would easily be seen in the EEG (Olds & Olds, 1969). Arousal System II depression would be seen only to the extent that it affected Arousal System I. If overall behavioral responsiveness (i.e., depth of sleep) were modified by levels in both systems, drugs acting on Arousal System II might increase sleep depth without significantly altering the recorded EEG. Further, it is conceivable that arousals themselves differ in EEG terms depending upon the locus of action of the drug involved.

Obviously, the arousal mechanism is extremely complex. Chances of being able to differentiate drug activity by EEG arousal characteristics are possibly slim but conceivable within the model proposed and on the basis of the site of action data.

In addition to these theoretical points it is held that the measurement of thresholds during the sleep process would give several sorts of useful applied data. Perhaps the single most important contribution is the ability of threshold measures to track a behavioral time course of drug effects. As such drugs may be differentiated by different time course attributes and classed according to

behavioral speed of action, length of action, and amount of impairment. Amount of behavioral impairment might also vary with dose. The practical question here is what happens to a patient treated with a specific drug if a fire alarm goes off during the night (etc.). Can he be awakened? Can he stay awake? If awakened, can he go back to sleep if he wants? How quickly can he go back to sleep? Or, what should be given to a patient with noisy neighbors? With sufficient time course information drug and dose could be matched to result in maximum depth at time of maximum noise. The point is simple. There is only one way to determine the effects on behavior of any drug as it interacts with the sleep process. That way is to sample behavior at appropriate times during the drug activity. The first step must explicate the effect of drugs on sensory mediated responses.

The present state of knowledge in threshold studies and drug studies indicates that a dual experiment is necessary to begin to behaviorally document drug activity during the sleep period. As a first step and initial experiment the reliability of a threshold measure must be established. Given measurement reliability, an attempt to control the threshold variable with various drugs can be made.

Finally, it must be recognized that an arousal from sleep is dependent upon two factors. An awakening threshold is dependent upon the threshold of the subject when

awake plus the amount of threshold impairment caused by the sleep process. The behavior of an awake subject will always be referred to as waking behavior or waking threshold. The sum score will be referred to as awakening threshold or arousal threshold. Awakening threshold minus waking threshold equals "pure" sleep depth. The waking threshold and pure depth of sleep measures may operate independently.

EXPERIMENT 1

Hypotheses

In terms of the reliability determinations, it was hypothesized that there would be substantial measurement reliability but that that reliability would be dependent upon several controlled variables. (1) Because auditory thresholds are routinely taken on many people, it was hypothesized that waking threshold measures would be extremely reliable and would form a ceiling that would not be surpassed by the sleep measures. Differences between the waking and sleeping reliability figures might be attributable to the sleep process itself. (2) In the sleep data four graded sets of data were examined with the hypothesis that each "grade" or extra step of control would increase measurement reliability. Step 1 was raw awakening threshold data. Step 2 was Step 1 data with the corresponding waking threshold subtracted ("pure" sleep depth data). This controlled for movement of an earphone in the ear and circadian effects as well as removing any other effects of waking threshold. Step 3 was a transformation of the Step 2 data. All data were plotted as a function of time of night when the threshold determination was made, and the points were connected by straight lines. The night was divided into seven equal time periods by six points.

Six points were used because an average of more than six but less than seven data points could be collected per subject per night. The interpolated value for each subject was read from the graph at each point. This step was taken so that an adjustment could be made to account for time of night effects as well as individual patterns in sleep depth across the night. The Step 4 analysis was to eliminate data from the first laboratory night, which is often uncharacteristic.

A final aspect of reliability is time. Subject nights in Experiment 1 were three to five nights apart. The effects of a longer measurement interval (two months) will be reported as a part of Experiment 2.

Method

Six male subjects aged 21 to 23 were selected from responses to an ad in the student newspaper for sleep subjects. They were screened with the Florida Sleep Inventory to have normal sleep habits including an approximate 11:30 p.m. bedtime and eight-hour sleep length. Subjects were paid $15 per night for their services. Each subject was given a series of tests including the Spielberger Trait Anxiety Scale, the depression scale of the MMPI, the Adjective Check List, and a selected medical history. In addition, subjects were chosen who produced alpha and reported minimal drug use.

Thresholds were determined by a modified Tracor RA 214 Rudmose screening audiometer. The audiometer was built to

ANSI 1969 specifications and calibrated by an audio specialist at the ear insert output before experiments were begun. Readings reported in this study are SPL (reference .0002 dynes/cm^2) levels. Modifications of the audiometer included the addition of a resistor such that signal intensities measured at hearing-aid ear inserts could be measured between -10 and 102 dB SPL at 1000 Hz in 2 or 3 dB steps. Signals were presented in a three-seconds on, three-seconds off sequence with interpolated catch trials. Normal, waking thresholds generally are 7 ± 15 dB (personal communication from W. Yost), and all of the subjects were in that normal range. All thresholds were obtained from stepwise procedures, which with the exception of catch trials, changed direction after a subject response change. All subjects had half an hour of practice previous to any experimental nights during which their "normal" thresholds were determined.

Each subject spent five nights in the lab. Nights were three to five days apart. Subjects reported to the lab at 10:00 p.m., had electrodes attached so that recordings could be made from F_1-F_7, P_1-T_5, O_3-O_zP_z (O_zP_z is halfway between O_z and P_z) and an eye channel. A navel body temperature probe was also attached. Subjects completed the Spielberger State Anxiety Scale and a day events questionnaire before entering the sleep rooms at 11:00 p.m. Subjects were allowed to insert the earphone into their ear, and a response button was taped into their preferred hand. Subjects were instructed to press the button

whenever they heard the tone and their responses, as well as a signal marking stimulus onset and stimulus offset, were fed through the EEG machine and written out as they occurred. Any response which occurred during a stimulus presentation was counted as correct. Responses during catch trials and in between stimulus presentations (false alarms) were minimal. On the rare instances of responding on catch trials subjects were innocuously warned that there might be catch trials so that they should try to listen for the tone. From 11:00 to 11:30 p.m. subjects were allowed to read or study while equipment was tested. During this time, each subject's waking threshold was determined once more. Lights were turned out at 11:30 p.m.

Subjects had practiced the threshold task and were aware that their threshold would be tested between five and eight times during the night. Each waking and awakening threshold determination began with an ascending series and ended when stable ascending values were obtained. Subjects were instructed that when awakened during the night, they were to say "I am awake" the first time they heard the tone in addition to pushing the button in their hand each time they heard the tone. When both responses were made (and a waking EEG had appeared), a small night light was turned on in the room to define the waking threshold determination period, a period lasting about two minutes except when subjects fell asleep during it. Waking threshold was

defined as "stable" when the value for three ascending series was the same determined from a stepwise procedure.

On one of the five nights (night 3 for 2 subjects, night 4 for 2 subjects and night 5 for 2 subjects), subjects remained awake so that data on threshold variability across the night without sleep could be compared with the data on arousal threshold and waking threshold on sleep nights. On the other four nights, the first threshold determination was made five minutes into the first Stage 2 period. Thereafter, the following criteria were imposed before a threshold determination was made: at least 5 minutes into Stage 2; at least 30 minutes since the last natural or experimentally produced awakening; at least 10 minutes since the last transitory body movement or muscle artifact greater than 6 seconds; well-defined Stage 2. Subjects were not awakened more than eight times on any night.

Subjects were awakened finally at 7:30 a.m., and apparatus was removed. Subjects filled out the Post-sleep Inventory (Webb, Bonnet, & Blume, 1976) and were allowed to leave the lab.

Results and Discussion

The major statistical tool used in estimating the reliability of the threshold data was the method of intraclass correlation (Guilford, 1954), an ANOVA procedure. Between four and eight observations were collected on each

subject on each night. To keep equal numbers in each ANOVA cell, ANOVA's for Step 1 and Step 2 comparisons were based on the first four observations for each subject. Step 3 and Step 4 ANOVA's were based on the average number of data points collected per night per subject (six). ANOVA's had variance terms for trials across a night, nights, subjects, error and interactions. Because degrees of freedom for all interactions were greater than six and interaction F-values were less than 2.00, the interaction sums of squares were pooled with the error sums of squares to form one error term (Hays, 1973). Degrees of freedom were similarly pooled. Results of the Step 1 through Step 4 ANOVA's can be seen in Table 1. The ANOVA for the waking threshold equivalent to Step 4 (without waking threshold subtracted) can be seen in Table 2. The $\bar{r}_1$ values, expressing the average strength of trial to trial relation of threshold (approximating Pearson r) for a single trial, most graphically illustrated the effects of adding controls. Moving from Step 1 through Step 4 respectively, the $\bar{r}_1$ values were .61, .65, .80, and .86, and the error term was reduced on each step. The waking value was .92. In terms of variance accounted for in the sleep data, the Step 4 data accounts for twice as much as the Step 1 data (74% vs. 37%) but is still smaller than the 85% accounted for in the waking data.

Simultaneous with the present experiment, Johnson (personal communication) collected some threshold data in various stages of sleep as a part of a larger experiment.

Table 1. Step 1, 2, 3, and 4 Threshold Reliability ANOVA's for Arousal Threshold

Step 1 ANOVA (see text)

source	df	SS	MSE	F - for Ss Effect	Ss Intraclass Correlation
TR	3	208	69	F = 25.6	$\bar{r}_{16}$ = .96
Nights	3	1736	579		$\bar{r}_{4}$ = .86
S	5	17311	3462		$\bar{r}_{1}$ = .61
pooled error	84	11320	135		

Step 2 ANOVA

source	df	SS	MSE	F - for Ss Effect	Ss Intraclass Correlation
TR	3	138	46	F = 30.3	$\bar{r}_{16}$ = .97
Nights	3	1165	388		$\bar{r}_{4}$ = .88
S	5	20479	4096		$\bar{r}_{1}$ = .65
pooled error	84	11318	135		

Step 3 ANOVA

source	df	SS	MSE	F - for Ss Effect	Ss Intraclass Correlation
TR	5	609	122	F = 94.6	$\bar{r}_{24}$ = .99
Nights	3	1443	481		$\bar{r}_{6}$ = .96
S	5	33662	6732		$\bar{r}_{1}$ = .80
pooled error	130	9261	71		

Step 4 ANOVA

source	df	SS	MSE	F - for Ss Effect	Ss Intraclass Correlation
TR	5	272	54	F = 113	$\bar{r}_{18}$ = .99
Nights	2	558	279		$\bar{r}_{6}$ = .97
S	5	26678	5336		$\bar{r}_{1}$ = .86
pooled error	95	4489	47		

S = subjects (These symbols are used on all tables.)
R = replication
DRUG = drug condition
TR = time trial across night

Table 2. Step 4 Waking Threshold Reliability ANOVA

Step 4 ANOVA

source	df	SS	MSE	F - for	Ss
TR	5	53	10	Ss Effect	Intraclass
Nights	2	240	120	F = 209	Correlation
S	5	10049	2010		$\bar{r}_{1} = .99$
pooled error	95	911	10		$\bar{r}_{18} = .99$
					$\bar{r}_{6} = .92$
					$\bar{r}_{1}$

In what would correspond to a presented $\bar{r}_1$ of .61 for a single trial unadjusted arousal threshold in the present data, his average correlation in Stage 2 was .36. Several large methodological differences as well as data from two small samples could account for the rather large discrepancy.

In the present results, it was reported that reliability was increased (to .65 for a single trial) by adjusting the sleep threshold scores for the corresponding threshold when awake. This effect, which accounted for movement of the earphone in the ear canal and circadian threshold shifts, appears to be real--i.e., not a function of the simple effect of measure combination. Waking threshold was not significantly correlated to sleep threshold between or within subjects.

The second reported increase in reliability (from .65 to .80) presented is probably an artifact in the present data. The interpolation method involved using a weighted average of observations such that two observations were used for each value. The increase in reliability found closely approximates that predicted by the Spearman-Brown formula if dual observations are used. This lack of effect for interpolation is interpreted to be a result of the fact that there was no trials effect in the data. As such MSE for trials was very small and accounted for very little of the total variance. The magnitude of the effect of the interpolation on the trials variance can be seen by

comparing the trials variance in Step 2 and Step 3. Interpolation increased trials variance by more than 2 1/2 times. In a case where time course across the night was important, such an increase in resolution could be of extreme importance. In the present data interpolation by itself added only about .03 to the reliability estimate.

The final control step was to eliminate effects from the first night in the laboratory. The step added about .06 to the single trial reliability.

The examination of the adjustment process, then, has left the conclusion that removal of the waking threshold and elimination of first night effects has added real increases in reliability and that the smoothing effects of interpolation increased reliability primarily through the "artifact" of increased observation. If the Spearman-Brown formula is used to correct the final Step 4 reliability figure to that of a true single observation, that value is .75. It is proposed that the base reliability of .61 was improved .04 (to .65) by inclusion of waking threshold, .06 by exclusion of night 1 (to .71) and .04 (to .75) by the interpolation. This figure would then predict a measurement reliability of .95 for an average night of six trials with the given adjustments.

Visual representations of depth of sleep results can be seen in Figures 1 and 2. Figure 1 presents threshold data as a function of time of night (Step 4) for two subjects--the lightest and deepest sleepers of the six

Figure 1. Within and Between Subject Variance in Arousal Threshold of Two Subjects over Three Baseline Nights

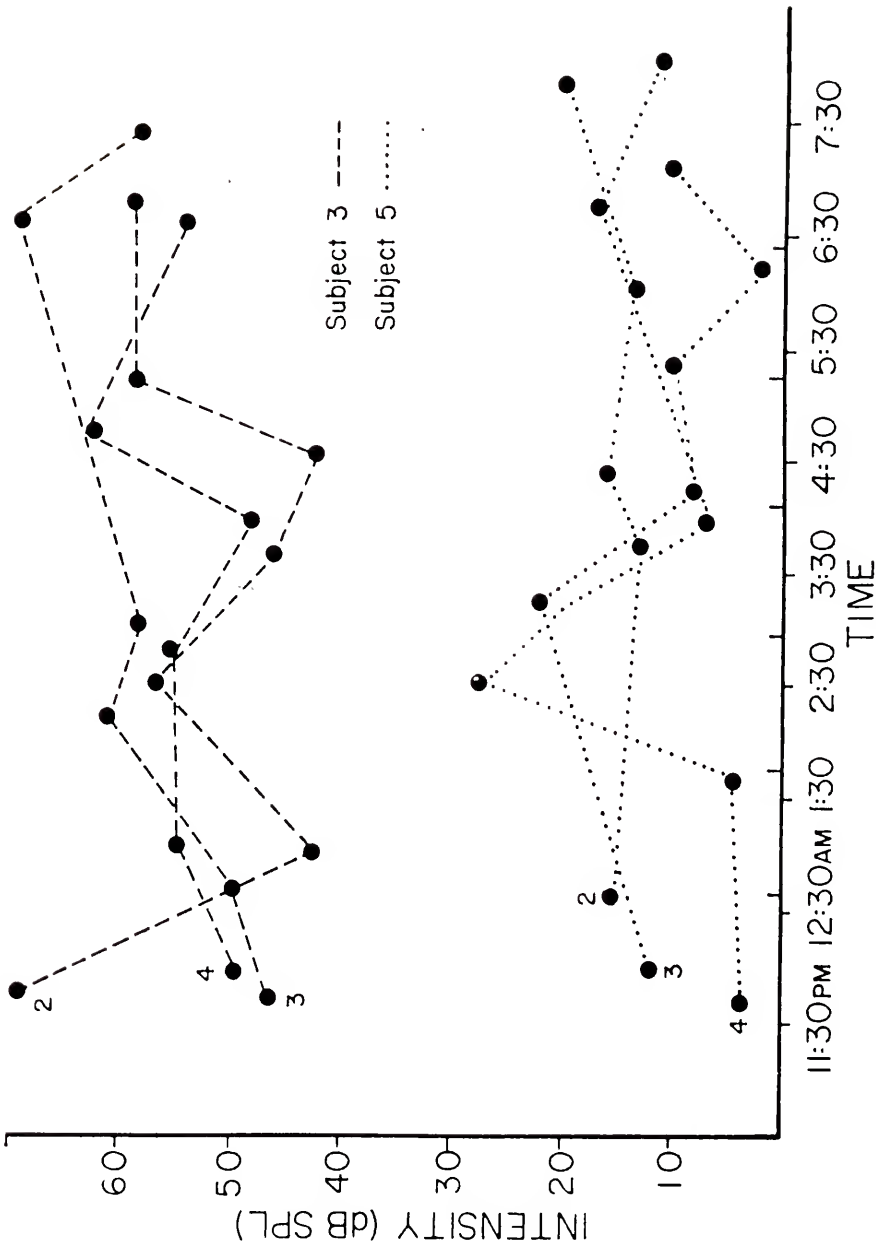
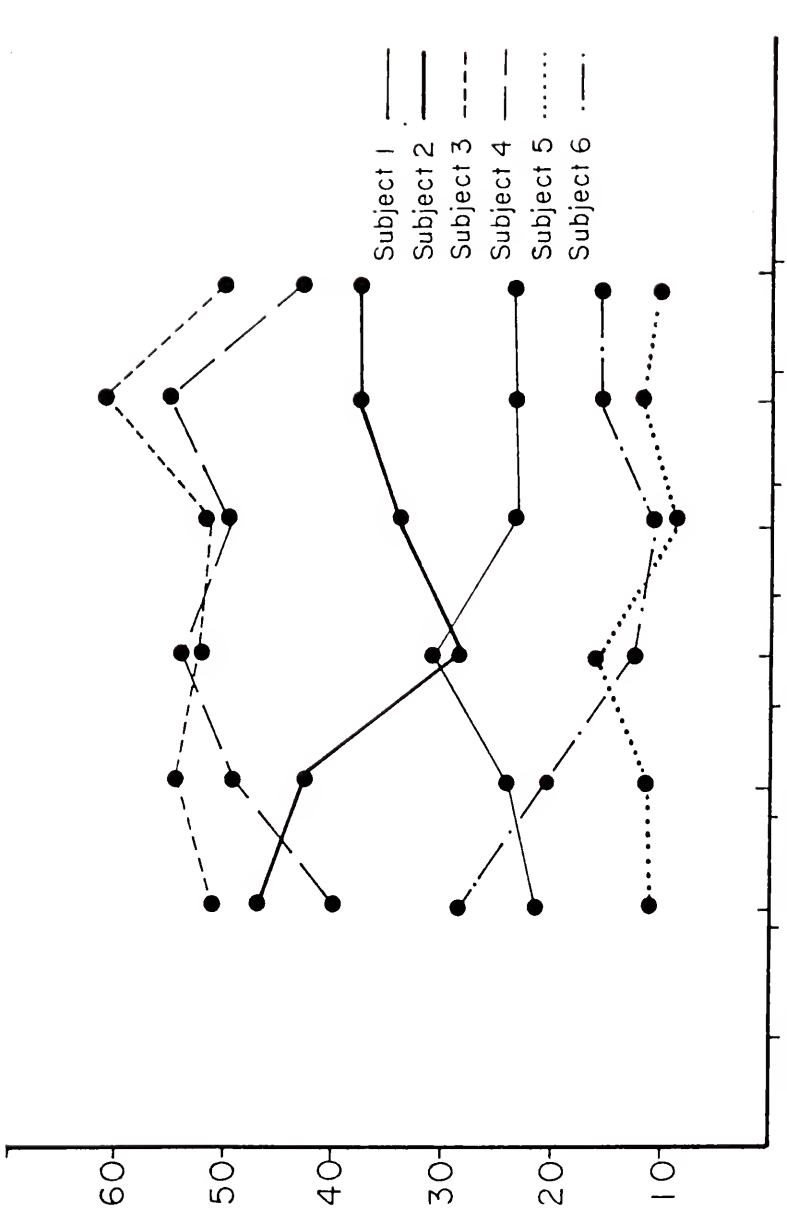

INTENSITY (dB SPL)
60
50
40
30
20
10
11:30PM
12:30AM
1:30
2:30
3:30
4:30
5:30
6:30
7:30
TIME
Subject 3
Subject 5
2
4
3
2
3
4

Figure 2. Three-night Average of Arousal Threshold (Corrected for Waking Threshold) for the Six Subjects of Experiment 1


INTENSITY (dB SPL)
60
50
40
30
20
10
11:30PM
12:30AM
1:30
2:30
3:30
4:30
5:30
6:30
TIME
Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Subject 6

subjects for their final three baseline nights. Each data point represents arousal threshold minus waking threshold at that point in time. The marks on the horizontal axis represent the arbitrary points at which the night was sectioned. Figure 2 is a plot for all six subjects of their three-night average of data as taken from the time points represented for two subjects in Figure 1. It can be interpreted as the average depth of sleep curve for the subjects across their night of Stage 2 sleep. Measurement reliability in Figure 1 is evidenced by the closeness of observations across nights within the subjects and the spread between subjects.

In comparing the waking threshold data to the sleep data (Tables 1 and 2), it can be seen that the increased reliability in the waking data is more a function of decreased within-subject difference than that of the also decreased between-subject differences. There are two possible reasons for this difference. It is possible that "attention" is simply more variable in sleeping than in waking subjects. It is also possible that the experimental situation played a role. Arousals from sleep were distributed rather randomly across the night in what might be described as a vigilance experiment with about eight "signals." However, the waking threshold determinations were time locked--they occurred immediately after arousal. The night without sleep was designed to attack this point. Data from the awake night, which actually lasted from

11:30 p.m. to 4:30 a.m., were transformed in a Step 3 analysis. An initial ascending value (waking) and a stable waking value resulted. The variance for each of these measures was calculated and compared to the variance of the arousal threshold and waking value recorded on the fourth night of sleep during the same time period. The variance of the arousal from sleep measures was greater than the variance in both stable waking measures in all subjects, and that difference was significant with an F-test for variances ($F_{.05,4,4} > 6.39$) in five of the six subjects. The variance of arousal from sleep was greater in five subjects than the variance for the initial ascending series on the awake night, and that difference was significant ($F_{.05,4,4} > 6.39$) in four of the subjects. As no other variance comparisons were significant, it must be concluded that some process attributable to sleep tends to increase not only absolute threshold but also the variability of the measure when compared to waking values.

Several additional types of data were collected at each arousal in the belief that they might relate to the threshold. Those variables were body temperature, amount of delta in the record, the length of the awakening and the latency to sleep onset after the awakening. Within-subject correlations for each item with the arousal threshold value were done for random subjects. Patterns of significant correlations were not found except in isolated subjects with the exception of the relation between arousal

threshold and the time it took to fall asleep again after the threshold procedure. These latency data are examined in Appendix A.

EXPERIMENT 2

Hypotheses

Given a reliable measurement procedure (Experiment 1) it was hypothesized that arousal in terms of both a motor response (button push) and vocalization would be modified after a single drug administration in Stage 2 sleep as follows:

1) Flurazepam (30 mg) and pentobarbital (100 mg) would elevate response thresholds over placebo depending upon dosage and length of action.
2) Caffeine (400 mg) would lower response thresholds depending upon dosage and length of action.
3) The sleeping medications were given in a standard therapeutic dose; attempts to equate dosage on any other dimension were not made; and any dose differences reflect only that point. It was hypothesized, however, that time course of activity would be the major variable in identifying the different compounds. Flurazepam was hypothesized to have a short onset to peak activity from studies of its time course effects on average frequency EEG patterns. Caffeine, a short-acting stimulant given at the equivalent of four cups of coffee, was also

hypothesized to show a short initial threshold shift with a subsequent return to baseline values. Pentobarbital, classed as acting for three to six hours, might show a slightly later peak in activity than flurazepam.

4) Pentobarbital, a barbiturate, and flurazepam, a benzodiazepine, were chosen as a test of the model developed by Bonnet (Note 1) and based on earlier work by Routtenberg (1966, 1968). Pentobarbital, working at the level of the reticular formation, was hypothesized to decrease EEG responsivity up to the point of waking threshold. Flurazepam, on the other hand, was hypothesized to have less effect on the ongoing EEG with the result that EEG desynchronization would not be as rapidly followed by a response on flurazepam nights as on pentobarbital nights. These EEG/behavior disassociations would also be a function of drug time course.

Method

The methodology of Experiment 2 completely paralleled that of Experiment 1 with the exceptions to be noted here.

Experiment 2 was an eight-night design with six subjects. Five of the subjects had participated in the first experiment. The sixth subject had two adaptation nights preceding this eight-night sequence. Experiment 2 took

place two months after Experiment 1. All laboratory conditions were similar.

Of the eight nights, the first night was designated as laboratory readaptation. Methods were exactly the same as Experiment 1. Lab nights were three to five nights apart (subjects coming on Monday night returned on Friday night). On the last seven nights subjects received a pill at 11:15 p.m. when already in bed. Caffeine was given on either night two or night eight and was only single blind. On the other six nights subjects received a numbered, uniform pink capsule under double-blind conditions. The order of administration was randomized within-subject such that each subject had flurazepam, pentobarbital, and placebo in one random order on the first three nights and in another random order on the second three nights.

The same subjective report measures as used in Experiment 1 were used with the exception that in Experiment 2 the Nowlis Mood Scale and the Spielberger State Anxiety Scale were completed both in the evening and in the morning.

Three threshold measures were chosen to be examined. The first corresponded to the Step 4 measure reported in Experiment 1. It was the point of consistent correct button push responses on an initial ascending series minus a following measure of ascending threshold when awake interpolated at six equally spaced points across the night (12:39, 1:48, 2:57, 4:06, 5:15, and 6:24 a.m.). The measure was planned to be a measure of possible shift in

sleep depth independent of any shift seen in waking threshold due to conditions. The second measure was the stabilized waking threshold also reported in Experiment 1. The third measure was designed to be a measure of "total effect" on behavioral response. This threshold was calculated by using the stabilized ascending waking threshold found at 11:15 p.m. each night. The 11:15 value was subtracted from the intensity at which each subject verbally said he was awake on all trials during the night. Thus waking and sleeping effects attributable to drugs were added in this measure, which is called the combined measure.

The code of drug conditions was devised by the Hoffman La Roche Company and kept in the University Infirmary in case of adverse reaction. It was returned unopened to the Hoffman La Roche Company at the end of the experiment. Upon receipt of the data reported here, Hoffman La Roche supplied the conditions code.

Results

An analysis of variance was done on the data from each of the three threshold measures. Effects for drug condition (DRUG), trial across the night (TR), replication of the experiment (R), subject (S), and all possible interactions were found. Because only the placebo, flurazepam, and pentobarbital were replicated, only those three conditions were included in the initial ANOVA to keep all cells of the analysis filled. The ANOVA's may be seen in Tables 3, 4,

and 5. Two striking results were found (unless otherwise specified all tests and confidence intervals were chosen at the .05 level of error probability). First, in terms of pure sleep depth, waking threshold, and the combination, there were significant main effects for drug condition. Both drugs resulted in higher thresholds than placebo with the exception of the effect of pentobarbital on waking threshold, which missed being significantly different from placebo by .14 dB. Second, in both measures involving thresholds during sleep, there was an unexpected trials by replications interaction. As no experimental conditions varied from the first to the second half of the study and a subject-by-subject examination of the data revealed no obvious explanation for a replications effect, such an effect was taken to represent a complex carryover effect. Because effects of carryover or repeated use were not the major topic of the present study and because significant drug condition effects did exist, it was decided that a further examination of the first administration of each drug was most proper.

Tables 6, 7, and 8 report the first administration analyses of variance for placebo, caffeine, flurazepam, and pentobarbital. A trial-by-trial plot of the three types of threshold drug may be seen in Figures 3, 4, and 5. In the waking condition, again, main effects for drug condition were found. The waking threshold after caffeine (7.58 dB) was less than that after placebo (10.17 dB), and the

Table 3. Replication by Drug Condition by Trial by Subject ANOVA for Flurazepam, Pentobarbital and Placebo from "Pure" Sleep Data

Source	df	MS
S	5	10164
R	1	16
S×R	5	494
DRUG	2	1318
S×DRUG	10	214
R×DRUG	2	440
S×R×DRUG	10	542
TR	5	764
S×TR	25	140
R×TR	5	403
S×R×TR	25	89
DRUG×TR	10	77
S×DRUG×TR	50	98
R×DRUG×TR	10	151
S×R×DRUG×TR	50	136
TOTAL	215	418

			Significance
1) 3-way interaction	$F_{10,50} = \frac{R\times DRUG\times TR}{S\times R\times DRUG\times TR}$	= 1.100	NS
2) 2-way interaction	$F_{2,10} = \frac{R\times DRUG}{S\times R\times DRUG}$	= .812	NS
	$F_{10,25} = \frac{R\times TR}{S\times R\times TR}$	= 4.540	.005
	$F_{10,50} = \frac{DRUG\times TR}{S\times DRUG\times TR}$	= .783	NS
3) Main effect	$F_{2,10} = \frac{DRUG}{S\times DRUG}$	= 6.170	.02

Drug Condition	Means	
Placebo	42.06 dB	.05 Confidence Interval = 6.57 dB
Flurazepam	50.10 #	
Pentobarbital	48.63 #	# Greater than placebo using t confidence intervals

Table 4. Replication by Drug Condition by Trial by Subject ANOVA for Flurazepam, Pentobarbital, and Placebo from Waking Threshold Data

Source	df	MS
S	5	2430
R	1	36
S×R	5	79
DRUG	2	221
S×DRUG	10	43
R×DRUG	2	9
S×R×DRUG	10	31
TR	5	20
S×TR	25	8
R×TR	5	4
S×R×TR	25	6
DRUG×TR	10	11
S×DRUG×TR	50	10
R×DRUG×TR	10	12
S×R×DRUG×TR	50	11
TOTAL	215	72

			Significance
1) 3-way interaction	$F_{10,50} = \frac{R\times DRUG\times TR}{S\times R\times DRUG\times TR}$	= 1.04	NS
2) 2-way interaction	$F_{2,10} = \frac{R\times DRUG}{S\times R\times DRUG}$	= .29	NS
	$F_{5,25} = \frac{R\times TR}{S\times R\times TR}$	= .72	NS
	$F_{10,50} = \frac{DRUG\times TR}{S\times DRUG\times TR}$	= 1.08	NS
3) Main effect	$F_{1,5} = \frac{R}{S\times R}$	= .45	NS
	$F_{2,10} = \frac{DRUG}{S\times DRUG}$	= 5.18	.03
	$F_{5,25} = \frac{TR}{S\times TR}$	= 2.46	NS

Drug Condition	Means
Placebo	10.95 dB
Flurazepam	14.40 #
Pentobarbital	13.24

.05 Confidence Interval = 2.43 dB

\# Greater than placebo (.05) using a $\underline{t}$ confidence interval

Table 5. Replication by Drug Condition by Trial by Subject ANOVA for Flurazepam, Pentobarbital and Placebo from Combined Waking and Sleeping Data

Source	df	MS
S	5	17244
R	1	1005
S×R	5	543
DRUG	2	3319
S×DRUG	10	402
R×DRUG	2	233
S×R×DRUG	10	281
TR	5	953
S×TR	25	138
R×TR	5	179
S×R×TR	25	58
DRUG×TR	10	67
S×DRUG×TR	50	133
R×DRUG×TR	10	88
S×R×DRUG×TR	50	101
TOTAL	215	594

			Significance
1) 3-way interaction	$F_{10,50} = \frac{R\times DRUG\times TR}{S\times R\times DRUG\times TR}$	= .872	NS
2) 2-way interaction	$F_{2,10} = \frac{R\times DRUG}{S\times R\times DRUG}$	= .830	NS
	$F_{5,25} = \frac{R\times TR}{S\times R\times TR}$	= 3.060	.03
	$F_{10,50} = \frac{DRUG\times TR}{S\times DRUG\times TR}$	= .510	NS
3) Main effect	$F_{2,10} = \frac{DRUG}{S\times DRUG}$	= 8.260	.01

Drug Condition	Means	
Placebo	46.50 dB	.05 Confidence Interval = 7.44 dB
Flurazepam	59.68 #	
Pentobarbital	55.92 #	# Greater than Placebo (.05) using a t confidence interval

Table 6. Drug Condition by Trial by Subject ANOVA for Flurazepam, Pentobarbital, Caffeine, and Placebo from "Pure" Sleep Data

Source	df	MSE
S	5	5179
DRUG	3	2275
S×DRUG	15	456
TR	5	476
S×TR	25	89
DRUG×TR	15	263
S×DRUG×TR	75	107
TOTAL	143	393

Significance

1) 2-way interaction $F_{15,75} = \frac{DRUG \times TR}{S \times DRUG \times TR} = 2.46$ .006

Condition by Trial Means

	Trial 1	2	3	4	5	6	T.D.
Placebo	37.92 dB	42.50	43.50	41.17	38.1	34.33	
Flurazepam	68.08	62.17	52.58	45.17	42.75	43.76	1>3=4=5=6 2>4=5=6
Pentobarbital	57.42	60.25	51.75	48.08	39.75	40.5	1>4=5=6 2>5=6 3>5
Caffeine	31.25	36.25	32.5	34.6	38.75	41.1	
Condition Differences	C=P<N=F	C=P<N=F	C<N=F	C<N			

Table 6 - continued

.05 Confidence Interval = 11.9 dB

P = Placebo

F = Flurazepam

N = Pentobarbital

C = Caffeine

T.D. = Trial Differences

Table 7. Drug Condition by Trial by Subject ANOVA for Flurazepam, Pentobarbital, Caffeine and Placebo from Waking Threshold Data

Source	df	MS
S	5	1374
DRUG	3	320
S×DRUG	15	19
TR	5	13
S×TR	25	11
DRUG×TR	15	9
S×DRUG×TR	75	7
TOTAL	143	64

			Significance
1) 2-way interaction	$F_{15,75} = \frac{DRUG \times TR}{S \times DRUG \times TR}$	= 1.29	NS
2) Main effect	$F_{3,15} = \frac{DRUG}{S \times DRUG}$	= 17.2	.0001
	$F_{5,25} = \frac{TR}{S \times TR}$	= 1.15	NS

Drug Condition	Means
Placebo	10.17 dB
Flurazepam	14.31
Pentobarbital	12.89
Caffeine	7.58

C<P<N = F*

*.05 level difference using a t confidence interval
Confidence Interval = 2.20 dB

Table 8. Drug Condition by Trial by Subject ANOVA for Flurazepam, Pentobarbital, Caffeine, and Placebo from Combined Waking to Sleep Data

Source	df	MS
S	5	6884
DRUG	3	6438
S×DRUG	15	534
TR	5	638
S×TR	25	80
DRUG×TR	15	197
S×DRUG×TR	75	106
TOTAL	143	544

Significance

1) 2-way interaction $F_{15,75} = \frac{DRUG \times TR}{S \times DRUG \times TR} = 1.85$ .04

Drug Condition by Trial Means

	Trial 1	2	3	4	5	6	T.D.
Placebo	41.50 dB	47.92	53.83	52.25	45.83	39.75	6<3,4
Flurazepam	70.17	71.75	69.17	61.92	53.83	54.92	5=6<1,2,3
Pentobarbital	58.10	70.67	62.75	59.17	50.25	45.08	1<2 5<2,3 6<1,2,3,4
Caffeine	29.77	33.50	32.25	35.83	41.83	27.01	1=6<5
Condition Differences	C<P<N<F	C<P<N=F	C<P<F	C<P=N=F	C<F	C<P<F	

.05 Confidence Interval = 11.9 dB

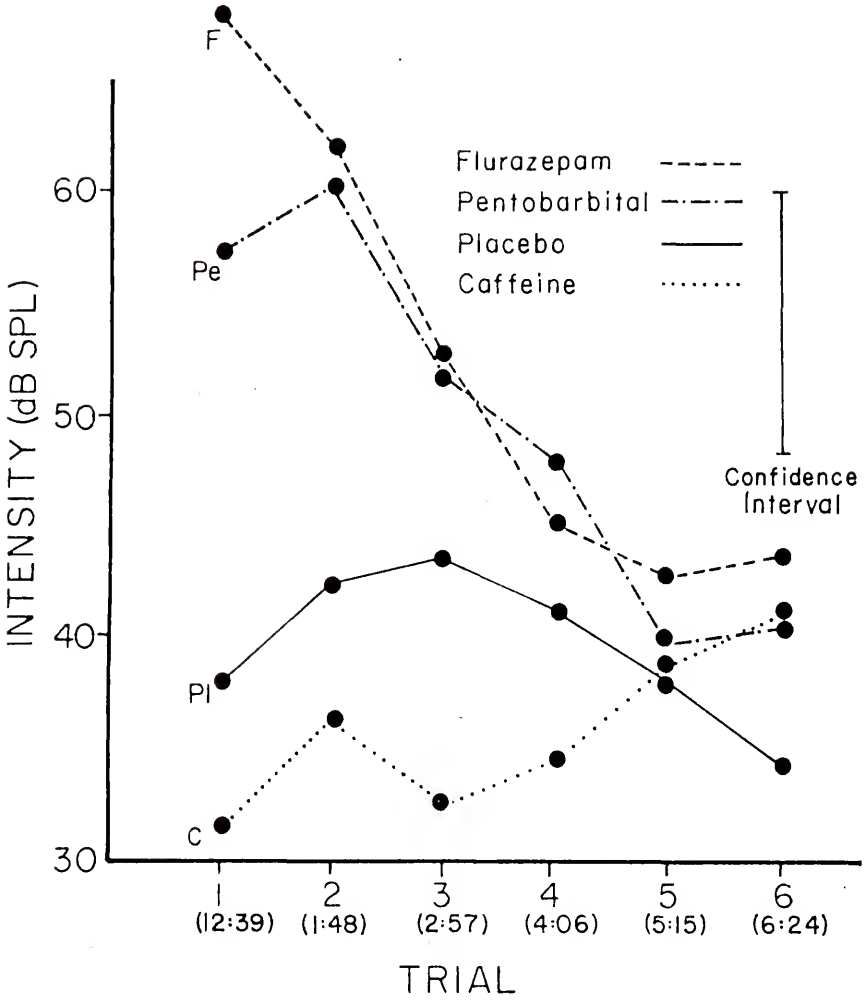


Figure 3. Average Depth of "Pure" Sleep Across the Night after Administration of Flurazepam, Pentobarbital, Caffeine, or Placebo

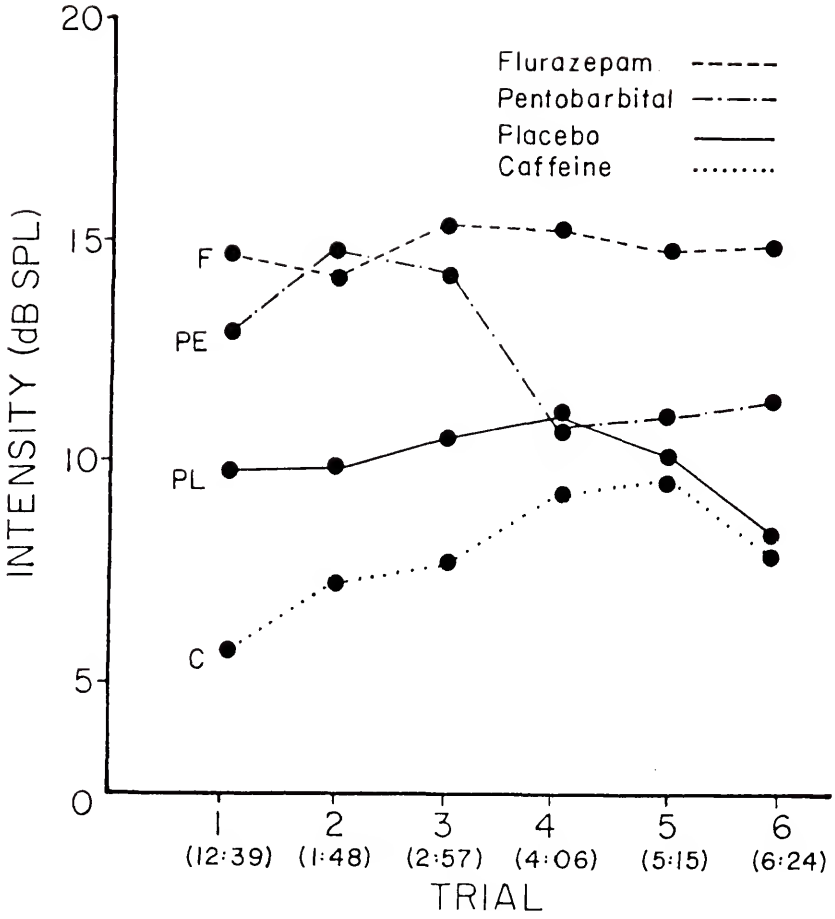


Figure 4. Average Threshold of Awake Subjects Across the Night after Administration of Flurazepam, Pentobarbital, Caffeine, or Placebo

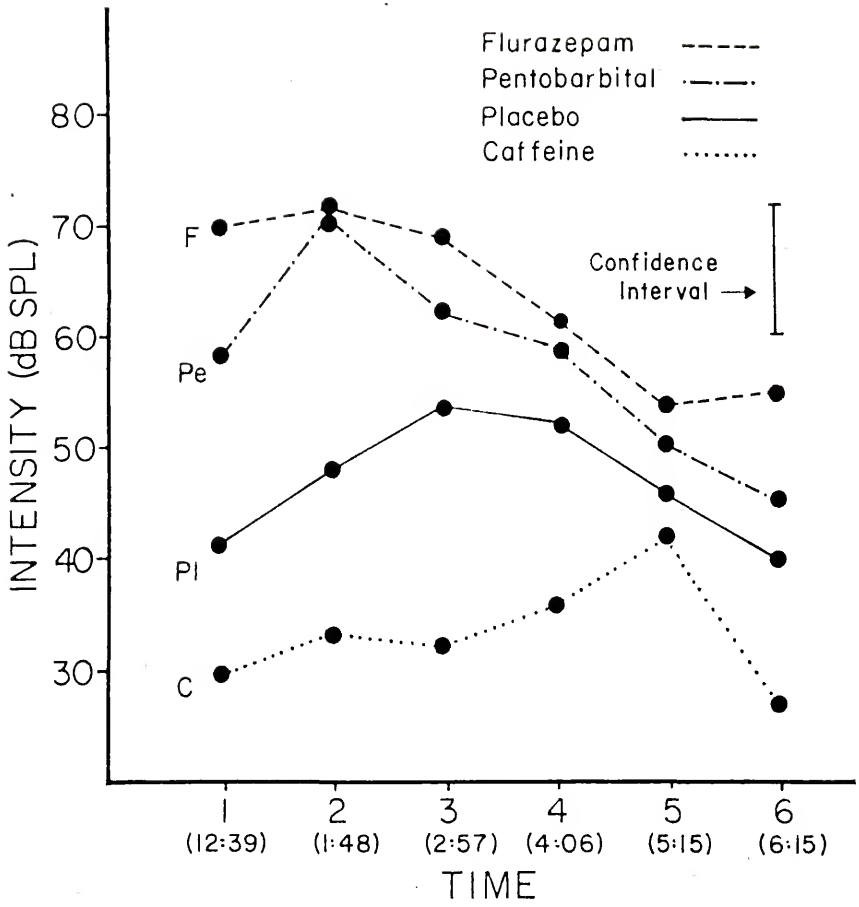


Figure 5. Average Combined (Waking and Sleep) Threshold Across the Night after Administration of Flurazepam, Pentobarbital, Caffeine, or Placebo

threshold after placebo was less than that after either pentobarbital (12.89 dB) or flurazepam (14.31 dB). There were no trial effects.

There was a trial by drug condition interaction in both of the sleep measures. Confidence intervals were computed, and trial-by-trial comparisons were made both across trials and across conditions. Those results are noted in Tables 6, 7, and 8 at the foot of each column and the end of each row. Briefly, condition differences in the "pure" sleep measure were seen only for the first four trials. In the first two trials administration of flurazepam and pentobarbital resulted in higher thresholds than the administration of caffeine or placebo, which did not differ. Thresholds after caffeine administration remained lower than those after the other two drugs in Trial 3 but were lower than only pentobarbital by Trial 4. Trial effects were seen only within flurazepam and pentobarbital. An initial peak was seen in both which differentiated the first trial from the last three or four trials and the second trial from the last two or three trials.

The results for the combined threshold measure, as would be expected, contain components found in both the waking and "pure" sleep measures and are therefore more extensive than either of the others. Thresholds after caffeine were lower than those after flurazepam for all six trials and less than placebo and pentobarbital for all but Trial 5. Placebo thresholds were lower than those for

flurazepam for the first three trials and Trial 6 and were lower than those for pentobarbital on the first two trials. Of all the threshold comparisons, flurazepam and pentobarbital differed significantly only on the first trial in the total data, where thresholds were higher with flurazepam. In the combined measure there were trial effects (time of night) for all conditions including placebo. Placebo was marked by one low threshold trial (Trial 6) and caffeine by one high threshold trial (Trial 5). The picture for flurazepam and pentobarbital was very similar to that seen in the "pure" sleep data--higher thresholds on the first two or three trials than on the last two (see Table 8 for the complete rundown).

To more fully address Hypothesis 3 (differentiating drugs by time course effects), trend analyses (Note 4) were done for the three threshold measures. The intent was to see if additional differences in the form of drug activity in trials across the night could be found. Illustration of all results should be referred to Figures 3, 4, and 5, which graphically display the various trends in the three threshold measures. The trend analysis yields an F statistic with 3 and 5 degrees of freedom if the strict criteria of Box, which allow only n - 1 ("n" being the number of different subjects) degrees of freedom for error, are followed. A more lax criterion, commonly used in analysis of variance, is to allow (G - 1)(n - 1) degrees of freedom for error (where "G" refers to the number of

groups) even if the same subjects appear in all groups. Using the strict criterion an F of 5.41 is required for significance (.05 level) in the present data, and only the linear differences (slope) between both flurazepam and pentobarbital as compared to the linear trend in the caffeine condition in the "pure" sleep data (Figure 3) meet this criterion (F = 10.39 and F = 7.39 respectively). Using the less stringent criterion ($F_{.05,3,15} = 3.29$), flurazepam has both linear characteristics which differ from the linear characteristics in the placebo and also curvilinear characteristics which differ from those seen in placebo. Specifically some quadratic curvilinearity can be seen throughout the placebo condition and a reverse quadratic is seen in the final flurazepam trials.

In the waking data (see Figure 4) all significant trend effects involved the pentobarbital data. The linear trend in pentobarbital differed from both caffeine and flurazepam conditions. In addition the quadratic curve in the placebo condition allowed it to be differentiated from the reverse quadratic effect seen in the last five trials of pentobarbital. Finally a quartic component ($\cap\cup\cap$ versus $\cup\cap\cup$) was found between pentobarbital and flurazepam indicating an early, middle, and late peak for flurazepam (Trials 1, 3 and 6) versus middle peaks for pentobarbital (Trial 2).

In the data combining sleep and waking effects (Figure 5), the linear effects of both flurazepam and

pentobarbital differed from caffeine and the pentobarbital also differed from placebo. Both flurazepam and pentobarbital, by virtue of early peaks, differed from caffeine, which had a late peak, in the cubic trend.

Tests of Hypothesis 4 were made by visual examination of several events in the EEG records for evidence of disassociations. Stimulus intensities were recorded at the first point of breakup of EEG sleep patterns, at the point of the beginning of button push responses and at the point of verbalization. Alpha production was examined around the button push and verbalization points as well. Previous work has not indicated time course effects in any of these variables, but because the drugs might institute time course effects a simplified procedure to control for such possible effects was deemed advisable. Each night was split into five parts as a function of time, and all values falling within that time period were averaged within each subject within each condition. The groupings were as follows: Group 1 $\leq$ 80 minutes, Group 2 $>$ 80 minutes and $\leq$ 180 minutes, Group 3 $>$ 180 minutes and $\leq$ 280 minutes, Group 4 $>$ 280 minutes and $\leq$ 380 minutes, and Group 5 $>$ 380 minutes. An initial subjects by replications by drug conditions by group ANOVA was done. The only significant effects were found in the measure of the difference between initial breakup of EEG patterns and the point at which subjects verbally said they were awake. Because all interaction terms containing replications and the replications

main effect were nonsignificant, the data for the two replications on flurazepam, pentobarbital and placebo were combined within condition so that they could be compared with the caffeine condition in an analysis with minimal empty cells. The results of this subject by drug condition by group ANOVA can be seen in Table 9. An increasing tendency for initial body movement to be disassociated from reported awakening can be seen as one moves from caffeine to placebo to flurazepam to pentobarbital. The difference between caffeine and the other drugs is significant, and the difference between placebo and pentobarbital just misses significance.

Table 9. Drug Condition by Trial by Subject ANOVA for the Disassociation of EEG Breakup and Verbalization as a Function of Tone Intensity

Source	df	MS
S	5	1395
DRUG	3	842
S×DRUG	15	195
GROUP	4	251
S×GROUP	20	63
DRUG×GROUP	12	88
S×DRUG×GROUP	57	49
TOTAL	116	

			Significance
1) Interaction	$F_{12,57} = \frac{\text{DRUG×GROUP}}{\text{S×DRUG×GROUP}}$	= 1.77	NS
2) Main Effects	$F_{3,15} = \frac{\text{DRUG}}{\text{S×DRUG}}$	= 4.32	.02
	$F_{4,20} = \frac{\text{GROUP}}{\text{S×GROUP}}$	= 4.01	.02

Drug Condition Means

Placebo	8.72 dB
Flurazepam	14.56*
Pentobarbital	16.21*
Caffeine	4.40*

*Caffeine differs at .05 level

Group (Time across the Night) Means

1	6.24** dB
2	13.35**
3	14.51**
4	11.69**
5	9.90

**Group 1 differs at .05 level

DISCUSSION

The two major purposes of the present work were first to estimate the reliability of a threshold measurement of the sleep process and second to examine the effects of common drugs on threshold. The reliability of the measurement process could only be classed as little short of remarkable given the single, small subject class. With appropriate controls the data indicate that from a single night of measurement (six trials) in a laboratory-adapted subject measurement reliability is estimated to be $\bar{r}_6 = .95$.

In one final reliability analysis using data (Step 4 transformation) from the five subjects participating in both experiments, an average threshold value was found for each subject from the last two baseline nights in Experiment 1 and from the two placebo nights of Experiment 2. The Pearson r was .804. The figure implied a degree of reliability in a two-night sample over a two-month period, but a reliability estimate for a single trial would have been lower than for closer nights.

There is also little question that flurazepam, pentobarbital and caffeine affected responsiveness during sleep when compared with placebo in the present experiment.

Further, the ability to make a statistical statement with six subjects is indicative of a gross effect as can be seen by the size of the confidence intervals employed in the sleep threshold data (approximately 12 dB). The more telling and perhaps more important findings most probably lie in two other interrelated dimensions. Those dimensions are the interplay of drug effects on "pure" sleep as opposed to waking thresholds and the time course effect of drug action.

The results to be discussed with noted exceptions are those which achieved a degree of statistical significance. In a study with a very small number of subjects, statistical significance is often indicative of either a very gross effect or chance. In a single study it is impossible, of course, to certainly separate the two. In a later part of this section the attempt to replicate the present results will be discussed, and from both sets of data, the overall effect for drugs appears to be a gross effect. However, the individual time course effects and trends are succeedingly less likely to be found to be gross effects and are more likely to be idiosyncratic to this experiment. A part of this effect comes from the unquestionably inflated alpha intervals from the multiple comparisons described here. Such an avenue was chosen because it was considered more important to trace possible variables which may be borne out by replication than to ignore such variables and risk a similar fate for them in future replications. The

present discussion, then, is certainly descriptive and hopefully predictive for replications in several important areas.

As judged from significance levels in the ANOVA tables, the major effect of caffeine (as compared to placebo) is on waking threshold. When awakened after caffeine administration subjects had lower thresholds than when awakened in any other condition, but the "pure" sleep measure did not differ from placebo. The present data does not allow speculation as to the subjects being "more awake" in the caffeine awakenings. It should be stressed that the same criterion of a stabilized threshold and waking EEG was always used but that the EEG criteria were difficult to control because a subject with his eyes open and/or concentrating on a threshold task may produce little alpha although awake. Also, in some conditions subjects had a difficult time staying awake for the three seconds between tones. The final question is a difficult one never completely resolved at a practical level--was the subject really awake? Obviously, spindle and K-complex activity was never included in a "waking" threshold. But latencies to spindles were in some conditions incredibly fast (see Appendix A); threshold could jump 70 dB in three seconds, and the data on waking thresholds must be viewed in that realistic context. Let it suffice to say that it occasionally took several awakenings to get subjects to both wake

up and stay awake for the few requisite seconds that form the basis for what is called waking threshold.

The more fine-grained trend analyses found the slightly positive linear slope of the caffeine nights to be different from the larger negative linear slope of the other drugs in the sleep data. Both effects, of course, were expected as indicators of lessening drug effects over time. Still the trend analyses could not differentiate caffeine from placebo.

Flurazepam and pentobarbital present a picture quite different from that of caffeine. Because the analyses of variance could separate these two drugs at only one point (higher thresholds for flurazepam in the first trial of the combined data), their ANOVA effects can be explained together. Both were characterized by high peaks at the first trial point (1 hour and 24 minutes after drug ingestion) in the pure sleep and combined data, and those peaks remained significantly higher than placebo for the first two or the first three trials respectively. In addition, both drugs resulted in higher waking thresholds than did placebo.

These drugs additionally displayed clear time of night effects in both sets of sleep data. Thresholds were always higher on the first two or three trials than they were on the last two or three trials. This may be contrasted with placebo and caffeine conditions, which both had only one

extreme point that may have occurred by chance. The placebo condition will be discussed further in a later section.

The trend analysis picture of flurazepam versus pentobarbital is an intriguing one with implications for differential effects on basic sensory processing and with ties to other data bases. An unstated hypothesis apparent in the present treatment of results is that waking threshold and "pure" depth of sleep are to some extent independent measures. There is, for example, an approximate zero correlation between the two measures in baseline data. To the extent that an independence exists, the measures could be selectively influenced by many variables. In the present case all drugs modified waking thresholds. The effects of caffeine and flurazepam were fairly linear across the eight hours of sleep, but the effects of pentobarbital (as gleaned from trend analysis) began to dissipate about the fourth trial. This trial was about five hours after drug ingestion, and the predicted duration of pentobarbital action (from the pentobarbital package insert) is 3-6 hours. The trend analyses also uncovered a quartic effect between flurazepam and pentobarbital in the waking data. If quartic curves are examined, it is seen that the maximum difference is found in a set of curves which display both a steep onset and a steep ending in opposite directions. In this case flurazepam has been fit with a curve $\smile\!\!\smile$ and pentobarbital with a curve $\frown\!\!\smile$. In terms of drug time

course, this says that flurazepam has a faster onset than pentobarbital (as is borne out by the significant difference in those two drugs in the first trial in the combined data) and continues to act for a longer time. Both of these results are borne out by aspects of sleep latency (Appendix A) and questionnaire data (Appendix B).

A few studies have examined performance on a battery of tasks in the morning 8-10 hours after drug administration. In two studies, Bixler, Leo, Mitsky, Pollini, and Kales (1974 abstract) first reported no performance effects (tasks not specified) in the morning eight hours after ingestion of either flurazepam (30 mg) or secobarbital (100 mg). In a later study with about twice as many subjects, Bixler, Leo, Mitsky and Kales (1976 abstract) found performance decrements with flurazepam (30 mg), secobarbital (100 mg) and phenobarbital (100 mg) as compared to placebo. The different results were explained as a result of allowing subjects to eat breakfast before testing in the first study. Roth, Kramer, and Lutz (Note 7) reported the effects of flurazepam, triazolam, secobarbital and placebo on performance. Ten hours after drug administration performance on a battery of tests including arithmetic, digit symbol substitution and card sorting was significantly impaired by only flurazepam (30 mg) as compared to placebo. The weight of evidence from these studies directly supports the present evidence for continued action of flurazepam until the end of the sleep period (and

longer) and offers minimal support that the barbiturate effects may be shorter lasting. Further, the present results indicate that the reported performance decrements are a direct result of sensory mediated deficits.

In the sets of sleep data, no trend analysis was able to split the effects of flurazepam and pentobarbital. However, in the pure sleep data, a late quadratic effect was seen with flurazepam as compared to placebo. This levelling effect, most apparent in the last three trials, seems to indicate a levelling effect of flurazepam on sleep. However, this result is confounded by the fact that the final trial threshold with flurazepam is almost 10 dB higher (nonsignificant) than the last placebo observation. If the trend find is appropriate, it would indicate differential effects of flurazepam on sleep and waking thresholds. The waking effect continues through Trial 6 while the sleep effect has dissipated by Trial 4. Regardless, a strong trials effect was seen in the sleep data and virtually no trials effect was seen in the waking data. In the pentobarbital, on the other hand, the trials effect on the pure sleep data does appear to be (nonsignificantly from ANOVA) reflected by a downward shift in waking threshold on Trial 4. Because waking thresholds have not been previously examined with rigor in sleep studies, the present evidence for independence in sleep and waking measures is the only evidence either for or against such a contention. A final answer must rest in replication. On

the basis of the less ambiguous waking results, the present conclusion would be that flurazepam is both faster acting and longer lasting than pentobarbital.

Several previous studies have reported a time of night effect for sleep depth in both REM and Stage 2 sleep under baseline conditions. Most of those studies have dichotomized data (early/late) in their comparison and found thresholds from the early part of the night to be higher than those from the later part. In data from Experiment 1 and the baseline and placebo nights in Experiment 2, a quadratic effect was always seen in the data, although with one exception a trials effect was never present in the analyses of variance. If the Experiment 1 data are dichotomized and the values of each subject are averaged for the first half and the last half of the four baseline nights, the group means show thresholds higher in the first half of the night in the combined data (51.1 dB versus 48.4 dB) and in the pure sleep data (36.1 dB versus 32.2 dB) but not in the waking data (15.0 versus 16.2 dB). None of these differences was significant, and none was significant even if the first trial of the night was excluded from the "early" data. The differences reported in the literature have been small, and the lack of significance here could be a result of the small sample or in the selection of awakening times. Thirdly, a methodological question could be raised. Evidence dating from the 1930's (Mullin et al., 1933; Mullin et al., 1937) has shown that depth of sleep

varies for up to 25 minutes after an arousal and also that body movements increase strikingly across the night. If sleep depth also varied after body movements and studies were not controlled for signal presentations shortly after body movements (a 10-minute criterion was set in the present study), more extreme low values might be found in the second half of the night corresponding to an increase in the probability of a body movement.

One of the reasons for the choice of flurazepam and pentobarbital in the present study was to examine the effects of two classes of drugs on EEG arousal characteristics. While there was some tendency for both drugs to disassociate the verbal response from the first breakup in the EEG as compared to placebo and more certainly caffeine, the two drugs could not be differentiated by any EEG arousal characteristics. There are two possible reasons. Most simply it could be that differential effects on behavior just do not exist although this does not agree with evidence from Kornetsky and Bain (1965). An equally possible and more plausible explanation is that the data analysis procedure, which consisted primarily of visual alpha counts before and after button push and verbal responses, was grossly insensitive. This last possibility could be tested in future studies by the use of computerized frequency analysis, but the meaningfulness of such an analysis would be limited because it is already known that these two drugs differentially affect some aspects of the EEG. Any finding

on differing arousal characteristics would be confounded by that knowledge.

The failure to support the dual arousal system hypothesis coupled with the present and other evidence of within-subjects shifts in depth of sleep under highly controlled EEG specifications leaves the present results in a theoretical quandary. The evidence showing EEG/behavior disassociation is extensive. The present results indicate that that disassociation cannot be easily identified in terms of EEG arousal characteristics except for the EEG breakup/verbal response dichotomy. Both the present results and the Kornetsky and Bain finding could be reconciled in a framework which posited that information from both (or all) arousal systems contribute EEG patterns in an averaged fashion so that inhibitory effects in either of two systems could have a similar EEG effect. But these rules might not apply to actual behavior as tested by Kornetsky and Bain (1965). A test of this hypothesis would require a more clearly-defined signal detection task in addition to the awakening threshold collected in the present study. As a rough estimator, there was a trend for more button push responses to be made before a verbalization in the present study after pentobarbital than after flurazepam (means 3.08 versus 2.04 responses, $\underline{t} = 1.66$, NS) in agreement with the similar findings of Kornetsky and Bain. Perhaps the most important conclusion, however, must be made a step further away from the data. The present

results indicate an EEG/behavior discrepancy after drug use. This is exactly predicted by a large amount of drug work dating to 1938 which has shown behavior/EEG disassociation after drug use. The importance of the many EEG/behavior disassociations is not the mechanisms by which they are caused but rather their mere existence and proclivity. Their existence raises doubts about the validity of EEG as a measure under important conditions. The EEG studies of flurazepam and the benzodiazepines in general, for example, have indicated a tendency for suppression of slow wave sleep and little effect on REM. These results, indeed, were found in a study by Itil et al. (1974). In that study several doses of benzodiazepines and placebo were given and a significant but seemingly paradoxical positive correlation was found between subjective "lightness" of sleep and the amount of Stage 4. This correlation is understandable if the possibility that the benzodiazepines both decrease Stage 4 and increase behavioral threshold irrespective of sleep stage is considered. In short it is possible to increase subjective "depth" of sleep and behavioral threshold while decreasing Stage 4, which normally has been associated with deep sleep in humans. The implication is that EEG as a single measure of sleep may be deceptive at best and inappropriate at worst when it is confounded by the experimentally-produced artifact of drug use or other trauma.

The present picture of drug activity differs considerably from standard EEG results. That difference is a difference of focus and methodology and, of course, is a difference of question asked.

Experiment 2 was designed to have a replication of flurazepam, pentobarbital and placebo effects within itself. This replication failed on several dimensions. In the initial analyses of variance, two significant effects were found. One, for drug conditions, has been discussed at length. The other, a trials by replications interaction, has not been examined. A replications interaction means that somehow the first half of the experiment was different from the second half. When confidence intervals were constructed in the "pure" sleep threshold data, it was found that thresholds on the first trial of the second replication were lower than thresholds on the first trial of the first replication. A similar finding, extending for three trials, was found in the combined data. Both EEG and subjective report data bases uphold the contention that the second replication was different from the first. In terms of EEG effects, initial latencies to Stage 1 sleep onset were ranked within each subject for the eight experimental nights. When the ranks were compared, it was found in both the placebo and pentobarbital conditions that the rank latency for all subjects was greater in the second replication than in the first replication (there was one rank tie for placebo, but the sign test for n = 5 was still

significant), an effect significant at the .02 level in a sign test. There was only one reversal in the flurazepam group, but that made the results nonsignificant. On the Post-sleep Inventory, 6 of the 30 items differed from replication one to replication two. Those items included reports that subjects had more thoughts at the time of going to bed and during the night, thought the room temperature and the bed were less comfortable, and thought that they tossed and turned more in replication two.

There are several possible reasons for a failure of replication in the present study. Five will be briefly discussed here although it is certainly possible to find others and no clear answer will emerge. Perhaps the most well-founded explanation is a tolerance explanation. While previous studies have used a similar three-night minimum washout between drug conditions (Itil et al., 1972, 1974) and a three-night examination of recovery effects from extended use is a standard Kales' procedure, it is possible that effects continue longer than Kales has measured and that the Itil studies could not or did not examine tolerance effects. A tolerance explanation would predict longer latencies on the second drug night and is considered a conceivable explanation to Roth (Note 6) from work he has done with flurazepam. Another explanation is a "time of the quarter" explanation. All subjects were students and might have been under increasing academic pressure as the quarter progressed. The final night on which flurazepam,

pentobarbital or placebo was given was on the Monday of dead week. Such an explanation would posit subjects had more thoughts and longer latencies because they were more worried about school in the second replication. Obviously it cannot be shown if more thoughts caused longer latencies or longer latencies caused more thoughts. A third explanation is an "end of experiment" explanation. Subjects spent 13 nights in the lab over a three-month period. For some subjects the last three nights were the second drug replication. As such, their dissatisfaction with nonoptimal laboratory accommodations (uncomfortable bed and temperature) could have increased sufficiently to interfere with their sleep process. This explanation is probably doubtful because adaptation trends usually go the other way and do not reverse. Roth (Note 6) has commented that end of experiment effects have occasionally been encountered in his work but that they are usually seen after multiple consecutive nights, which did not occur in the present study. Fourthly, the threshold data could be explained by state dependent learning within drug states. However, this argument would not predict a latency or a subjective sleep change and could probably be discounted for those reasons. Finally, the early trial threshold shifts could be explained by the increased sleep latencies. In a time sense, peak threshold obviously depends upon latency and the data was not corrected for that factor. The thresholds of the second replication, which came, on the average, more

closely to sleep onset, could be lower for that reason, but this explanation fails in that it cannot explain why the latencies were longer.

In short, no single, satisfactory explanation of replication effects has presented itself. An examination of individual data is no more enlightening. However, this discussion has brought forward, as an assumption, a very important fact. Three relatively independent measures of the sleep process have been discussed. It is indeed very relevant that when an effect (hypothesized or not) was found in the threshold measure, it was directly supported by findings in both EEG and subjective measures. Such a finding not only documents the existence of an underlying factor, but also validates the measures used.

SUMMARY

The present study has demonstrated high reliability in a threshold measure of the sleep process. It has further documented the effects of flurazepam and pentobarbital in increasing auditory thresholds of sleeping and waking subjects in a time course fashion as compared to placebo. Effects of caffeine in decreasing auditory thresholds were also seen. Trend analyses indicated that flurazepam was both a faster and a longer-acting drug than pentobarbital but those results (and really all results in a six-subject study) must be considered tentative. Flurazepam and pentobarbital could not be separated by their EEG arousal characteristics. The finding was considered in opposition to a dual-arousal system hypothesis. Finally, failure to achieve a complete replication of all time course effects was taken to represent possible tolerance or time of the academic quarter effects and also as validating evidence of three separate measures of the sleep process--EEG, subjective report, and behavioral threshold.

APPENDIX A

THE EFFECT OF FLURAZEPAM, PENTOBARBITAL AND CAFFEINE ON LATENCY TO SLEEP ONSET ACROSS THE NIGHT

The latency to sleep onset at the beginning of a sleep period is reported in virtually every EEG study of sleep. An estimate of how much time subjects spend awake over a period of sleep (Stage 0) is also routinely reported. Both types of data are also routinely collected on almost all sleep questionnaires.

Students of insomnia have used both latency and Stage 0 time as criteria for assessing the degree of an insomnia problem. They have also used a third measure, usually referred to as early awakening and inability to fall asleep again. While at least one source is available (Webb & Agnew, 1975) that shows a circadian effect on initial latency to sleep onset, a systematic study on the speed with which subjects fall asleep after an awakening during the night has not been done. It was proposed that such sleep latency measures across the night might be an important variable in determining drug time course and effectiveness.

Method

As a part of the earlier reported experiments, subjects were awakened from Stage 2 sleep several times during the night. When each subject was awake, the experimenter turned on a small night light in the subject room and measured the subject's waking threshold. Subjects had been instructed that their waking threshold would be tested at each awakening, and that they were to remain awake and work on the threshold task as long as the night light was on. The turning off of the night light was the signal that the threshold determination was complete and that the subject should fall asleep again. Length of awakening (i.e., length of threshold determination) was not controlled. The criterion for awakening was rather a stabilized waking threshold where stabilized was defined as the same or close (5 dB maximum) value on three successive ascending series in the threshold determination. With rare exceptions this criterion was met.

As sleep stage of awakening was controlled, the absolute time of night of awakening could not be controlled. Because Webb and Agnew (1975) had displayed circadian effects in sleep onset, each night for each subject was arbitrarily split into five parts by time (0-80 minutes, 81-180 minutes, 181-280 minutes, 281-380 minutes, 381+ minutes after an 11:30 p.m. bedtime), and all observations for each subject for each condition were averaged within

each time block to get a single subject value for each time block for each condition. The few empty cells were filled with the average of blocks on each side. An empty cell in the first or last time block was left unfilled.

Latency was measured as the time from when the night light was turned off to the appearance of the first spindle or K-complex in the record. Latency values are invariably skewed. As a result all within sleep latency values for each subject were rank ordered across all conditions, and it is those ranks which will be reported as data.

Results

As in previous data analyses an initial ANOVA was run primarily to test for a replications effect in the flurazepam, pentobarbital and placebo data. No significant interactions were found, but there were significant main effects both for drug condition and time of night. Because there was no replications effect, data for the placebo nights and corresponding drug nights were averaged and entered into a second analysis of variance also containing data from the single caffeine night. That analysis is presented in Table 10. Three things should be noted about the ANOVA. First, the mean square for subjects does not equal zero as it should for ranks because of averaging over conditions, because the laboratory adaption night was ranked but not entered into the analysis and because cells were averages of ranks. Second, there were eight empty

Table 10. Drug Condition by Time of Night by Subject ANOVA for Sleep Latency Rank after an Awakening

Source	df	MS
S	5	52
DRUG	3	1850
S×DRUG	15	78
TIME	4	434
S×TIME	20	77
DRUG×TIME	12	122
S×DRUG×TIME	53	42
TOTAL	112	124

Significance

1) 2-way interaction $F_{12,53} = \frac{122}{41.9} = 2.91$ .005

Drug Condition by Time Means

	1	2	3	4	5	T.D.
Placebo	34.8	16.1	21.0	25.5	33.3	1>2,3,4 5>2,3,4 2<4
Pentobarbital	22.5	16.8	16.9	29.3	31.4	4>2,3 5>1,2,3
Flurazepam	22.5	14.9	14.5	16.4	19.8	1>2,3
Caffeine	44.7	41.8	32.4	32.4	35.5	1>3,4,5 2>3,4

Table 10 - continued

Condition Differences	F=N<P<C	F=N=P<C	F=N=P<C	F<N=P=C	F<N=P=C

Confidence Interval is 7.55 at the .05 level

cells. Third, there was a significant drug condition by time of night interaction ($F = 2.90$). Individual time by condition comparisons were made. The latency rank by time plot can be seen in Figure 6 for the four drug conditions and differences are noted in Table 10. Briefly, latencies were very short with flurazepam throughout the night with a relatively small time-of-night effect.

At the second and third time points placebo, flurazepam, and pentobarbital did not differ, but with flurazepam latencies remained short throughout the sleep period. Pentobarbital latencies differed from placebo only in the first time period. Latencies after caffeine administration were longer than all other conditions for the first three trials and longer than flurazepam for all trials. Definite time of night effects were seen in all conditions except flurazepam. Latencies in the early and late parts of the night were high, and with placebo and pentobarbital latencies in the second and third periods were lowest.

Initial sleep onset latency has been briefly discussed. It was not included with latency during sleep results because stimulus conditions differed, values were generally higher, and there was a replications effect in that variable. The median rank for initial sleep onset latency corresponded to an average latency of about 6.7 minutes. During the sleep period a similar median rank was about 1.5 minutes. Only 6 percent of the sleep period latencies were longer than 6.7 minutes, which indicated

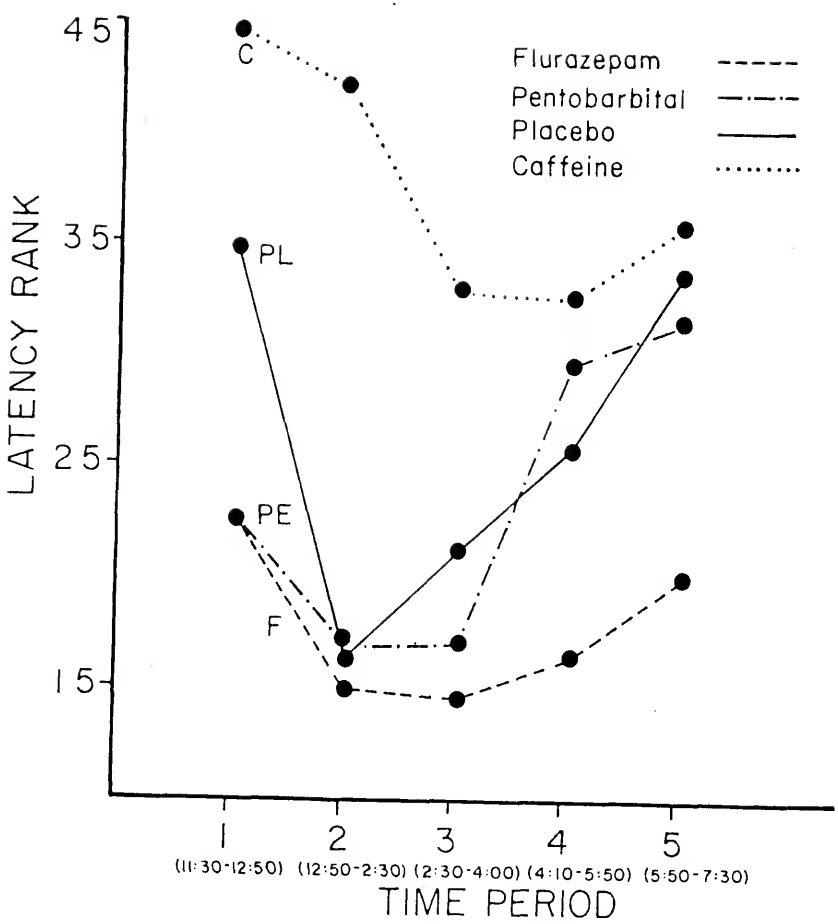


Figure 6. Sleep Latency after an Awakening Across the Night in Flurazepam, Pentobarbital, Caffeine, and Placebo Conditions

little overlap of distributions. However, there were condition effects in the initial sleep latency data. In a night-to-night comparison a sign test requires no reversals for a statistical level of significance to be reached with six subjects. Latencies to initial sleep onset were significantly longer after caffeine than after flurazepam or after the first pentobarbital administration. There was one reversal in the comparison of caffeine to the first placebo night (caffeine night longer) and one reversal in comparison of the first flurazepam night to the first placebo night (placebo night longer).

Discussion

Circadian and drug effects on sleep onset are obvious and profound. The time of night differences may be explained to be a dual function of circadian and sleep process effects. For this reason, the present results are not directly comparable to the Webb and Agnew (1975) data, which examined only circadian effects on initial latency. The Webb and Agnew data suggested a circadian trough in sleep latency somewhere around 7 a.m. The present data suggest an earlier trough (around 1:30 a.m.), which may be the result of the interaction of circadian and sleep process effects--i.e., it is very hard to fall asleep after just having slept for 10 hours. Long latencies in early and late parts of the night would then reflect initiation and termination of the sleep process effects and the curvilinearity of the latency function would correspond to the

curvilinearity in threshold. This would underscore the biorhythmic aspects of the sleep process without necessarily supporting any theoretical position. Further work is obviously needed to examine sleep process effects on latency in sleep placed at different circadian times to test the relative contributions of both factors.

The relatively small circadian influence seen after flurazepam administration and the lack of condition differences (between placebo, flurazepam, and pentobarbital) at the second and third time points is probably indicative of a basement effect. The average latency in the six subjects for a rank of 15 is slightly under a minute. Since latency was measured as the time to the production of a spindle or K-complex and spindles or K-complexes may occur only two or three times during a minute of normal Stage 2 sleep, an average shorter than about 30 seconds is probably impossible. Latencies are approaching this value in the 12:45 to 2:30 a.m. range of the data except after caffeine administration.

The latency data, as presented in Figure 6, offer interesting corroborative information concerning drug time course of action. In the combined threshold data (sleep plus waking thresholds) of Experiment 2, caffeine differed from flurazepam on every trial just as caffeine differs from flurazepam on every trial in the latency data. Caffeine differs from placebo in the combined threshold data for the first four trials only (until 4:06 a.m.). In the latency data caffeine differs from placebo for the

first three trials (until 4:10 a.m.). In the threshold data, flurazepam differed from placebo on the first three trials and the last trial. In the latency data it differed from placebo on the first trial and on the last two. Only the comparison of flurazepam with pentobarbital does not follow the earlier script exactly. From the conclusions of Experiment 2 it would be predicted that flurazepam would differ from pentobarbital in that latencies after its use would be shorter than after pentobarbital in the first and last time period because flurazepam was predicted to have faster initial activity and to last longer than pentobarbital. As predicted, latencies after flurazepam were shorter than after pentobarbital at the last time period. However, the drugs did not differ at the first time period, when it was predicted that they should.

From the sheer magnitude of the similarities here presented between threshold effects and sleep latency effects, it could be predicted that latency and threshold would be highly correlated within subject. If those correlations are done in the present data set, all except one are significant within-subject (the one nonsignificant correlation was at the .07 level). If those within-subject correlations are transformed to z-scores, averaged, and transformed back to a correlation, that average correlation is -.39, which for the average of 46 observations per subject, is significant at the .01 level. It does appear that when a subject is deeply asleep and is awakened, he

can fall asleep more quickly (sleep stage controlled) than when he is not so deeply asleep. This, of course, does not prove a causative mechanism but does suggest the possibility that drugs (and various other sources) might allow subjects to fall asleep more quickly through a common central mechanism which also controls how deeply a person sleeps. Such a view is consonant within most any theory of arousal.

Espoused similarity to the earlier set of results is important at another level. The bulk of results presented before this appendix concerned only data from the first replication. The decision rule had been to average data to compare it with the caffeine data in all cases, but that rule was limited by occasional significant replication effects. The result was that the data base for the threshold results was different from that of the sleep latency results. It could be argued that the sleep latency data, although without a replications effect, also suffer from whatever caused the threshold effect. This would imply that only data from the first replication should be used in the latency analyses. But, as might be guessed from the lack of replication effect, if such an analysis is done, the results are essentially the same as using the entire data set. It might also be added that a continued problem in many studies of arousal is lack of agreement among measures and concurrent definitional problems.

A correlation averaging technique was also done with latency and sleep depth in the data from Experiment 1. The

average correlation was -.23, which was not significant. Besides chance, this difference might be due to fewer observations with less variance on the baseline than on the drug nights. The possibility that the correlation in Experiment 2 data is caused, not magnified, by drugs is mediated against by other analyses to be reported more fully in Appendix C. Briefly in an analysis comparing placebo and baseline nights of deep sleep to placebo and baseline nights of light sleep within-subject, there were significant differences in both sleep latency questions on the Post-sleep Inventory. Even with seven subjects, both initial latency to sleep onset and the length of awakenings during the night were subjectively described as significantly shorter ($p < .05$) on nights of deep sleep than on nights of light sleep. A test of actual initial sleep latency on these nights found that six of the seven subjects had shorter initial latencies on the high threshold night, which is nonsignificant with a sign test ($p = .0625$) although close. The single reversal was by two minutes.

APPENDIX B

THE EFFECTS OF DRUGS ON TWO SUBJECTIVE REPORT MEASURES

There is a fairly large literature on the subjective effects of various drugs on the sleep process. It is reviewed by Jick (1969). The present study does not seek to deal with that topic in general but rather with two subtopics directly related to other areas of the present paper.

Two primary subjective measures were used in the present study. They were the Post-sleep Inventory and the Nowlis Mood Scale. The Post-sleep Inventory is in developmental stages (Webb, Bonnet, & Blume, 1976; Note 2, Note 8). Sensitivity to drug conditions would help validate that scale. Conversely, items on that scale such as evening sleepiness and ease of awakening in the morning could help bolster conclusions concerning the differentiation of drug time course effects by threshold determination.

Method and Results

On each morning of Experiment 2 subjects were finally awakened between 7:30 and 8:00 a.m. During the next 30 minutes they had electrodes removed and filled out the Post-sleep Inventory and the Nowlis Mood Scale.

Responses on 29 individual Post-sleep Inventory items, the six factors derived from that scale and six subscales (aggression, anxiety, elation, fatigue, vigor, and deactivation) from the Nowlis Mood Scale were analyzed in individual subject by condition (placebo, flurazepam, and pentobarbital) by replication analyses of variance. While there were main effects for condition and replication in several items, no interaction effects were found.

Therefore the data for the replicated drug conditions were averaged and an analysis of variance with the caffeine data was run. The analysis had effects for subjects, drug condition (placebo, flurazepam, pentobarbital and caffeine) and error.

Drug conditions were differentiated significantly by eight individual Post-sleep Inventory items and three composite factors. Of these 11 items the caffeine night played the primary divergent role in six. The caffeine condition was accompanied by reports of longer latency, more evening thoughts, a more uncomfortable room temperature and lighter sleep than in any other condition. Additionally subjects felt less exhausted in the evening and awoke more easily in the morning after caffeine than after the other drugs.

Flurazepam administration led to reports that subjects were significantly more sleepy at bedtime than in any other condition. Subjects also reported higher sleepy values on the "sleepy PM" factor, which includes exhaustion and sleep

latency, on flurazepam nights than on caffeine or placebo nights. Also the total scale score, indicating "good" sleep was higher for flurazepam than for caffeine. Both flurazepam and pentobarbital nights were accompanied by higher sleep factor scores than placebo. This composite includes latency, awakening length, body movement and depth of sleep.

No differences for any condition were found on the scales from the Nowlis Mood Scale.

Discussion

In the Post-sleep Inventory data, it was the caffeine night which was in general the clearly divergent night. This finding is consonant with the fact that subjects were chosen who were good sleepers. The law of initial values would state that it would be difficult to improve the sleep of good sleepers over baseline on any dimension but that it would be relatively easier to disrupt it.

Clearly flurazepam and pentobarbital were found to differ from placebo on few measures. It was claimed by subjects that awakenings were shorter than placebo with both drugs (corroborated in Appendix A with latency measures) and that both drugs improved sleep process factors such as latency, body movement and sleep depth. In one item, evening sleepiness, flurazepam was differentiated from all other conditions including pentobarbital. This last finding serves as subjective evidence that onset of

flurazepam activity was faster than that of pentobarbital. Similar results were found in the evening sleepiness factor.

The second conclusion about flurazepam from Experiment 2 was that flurazepam continued to act longer in the morning than did pentobarbital. Of the two specific questions to assess that issue, one (woke up extremely tired) did not differentiate conditions. The other (had a very hard time awakening) differentiated only caffeine from the other drug conditions. The report was that flurazepam made it (nonsignificantly) harder to wake up in the morning than pentobarbital or placebo. The difference from placebo narrowly missed significance. Still, it cannot be concluded from this data that the time course of flurazepam continued beyond the sleep period. However, at least one other study (Itil et al., 1974) has found flurazepam to induce more difficulty in becoming fully alert in the morning than did placebo.

Of the four recent studies which have examined the effects of drugs on mood in normal subjects, only a non-sleep study looking at mood an hour after drug administration (Sambrooks, MacCulloch, & Rooney, 1975) has found significant mood shifts as a result of nitrazepam and flurazepam administration. Of studies testing morning mood one (Schwartz, Roth, Kramer, & Hlasny, 1974) found no effects after triazolam administration and one using averages from three consecutive drug nights (Kales,

Malmstrom, Kee, Kales, & Tan, 1969) found largely only trend effects from REM suppressants and little effect on mood with chloral hydrate, which did not affect REM. In light of these results, it was not surprising that no clear trends for mood shift existed in the present, one-night data.

APPENDIX C

THE RELATION OF DEPTH OF SLEEP TO SUBJECTIVELY REPORTED SLEEP PARAMETERS

Zimmerman (1968) reported on the relationship of depth of sleep to the EEG and physiological patterns. He found remarkably few differences between his light sleepers and his deep sleepers. One possible explanation is that depth of sleep, being a measure of threshold, is not related to many within-sleep variables directly just as average visual acuity would not be expected to be related to a range of sleep variables. But just as some theories of REM sleep state that differing amounts of eye activity within an individual may affect his visual processing abilities when awake (Berger, 1969), it is possible that differences in an individual's depth of sleep may be reflected in other sleep variables. Unfortunately, the measurement of sleep depth interferes with the measurement of EEG parameters of sleep, and therefore, those parameters will not be reported in this paper. Instead, the relation of depth of sleep to subjective reports of the sleep process will be reported.

Method and Results

The data base and methodology for the present appendix were the same as in the previously reported results. Only

baseline and placebo nights were used in the present analysis. Initial laboratory nights were excluded. From the maximum of six laboratory nights per subject the night of highest threshold and the night of lowest threshold were chosen for each subject. Two criteria were used in choosing threshold. In one analysis mean nightly threshold values were used and in the second analysis the last threshold observation of the night was used. Because there was one subject who participated in only Experiment 1 and one subject who participated in only Experiment 2, data from seven subjects were available and were used.

Nights of high threshold and nights of low threshold were compared using paired t-tests for each of the 29 items on the Post-sleep Inventory and for six derived factors. In the analysis of overall threshold means it was found that when thresholds were high, subjects reported a shorter initial sleep latency and shorter awakenings during the night. In the analysis using the last threshold measurement of the night, both sleep latency measures remained significantly shorter in nights of deep sleep. Subjects also reported (a) being more sleepy in the evening ("sleepy PM" factor); (b) better sleep on the sleep process factor item, which includes latency ratings along with noise, body movement, and depth of sleep; and (c) a borderline increase ($p < .06$) in the feeling good in the morning factor, which includes adequate sleep, tiredness, mood, and physical feeling in the morning on the nights of deep sleep than on

the nights of light sleep. Results on the individual depth of sleep item were nonsignificant in both analyses, although the t-value (2.49) in the second analysis was significant at the .055 level.

Discussion

The results suggest that variation in depth of sleep is indeed related to shifts in at least some parameters of subjective sleep quality. Obviously, sweeping generalities cannot be made from seven subjects regardless of significance values. Two statements perhaps are warranted. Large between-subject differences in threshold and Zimmerman's work with threshold between groups suggest that within-subjects designs may be superior in testing threshold effects. Zimmerman's negative findings for relations of between-subject differences in examinations of between-subject differences in depth of sleep and other variables were upheld in the present study. Briefly, a correlation was done between a final sleep depth value of an average of Post-sleep Inventory items on the placebo nights. No correlations were significant. No significant correlations were found in attempt to relate preexperimental general sleep Post-sleep Inventory estimates to between-subject differences in sleep depth. Further, scales of the Adjective Check List, the Spielberger Trial Anxiety Scale and the depression scale of the MMPI failed to correlate between subjects with depth of sleep.

Secondly, subjective report has a bad name in sleep research. The problem stems from the fact that subjective reports just do not agree with EEG results. The present author has worked on a paper (unpublished) trying to relate EEG measures to the Post-sleep Inventory. Results were largely uninterpretable. At present it is less conceivable that threshold measures can answer this particular data relationship question than it is that a within-subjects approach can answer such questions. In the simplest sense, it gives the subject something to base his responses on.

There is no reason to conclude that subjects use depth of sleep information in judgments of sleep quality. From the Post-sleep Inventory data it is not certain that subjects are aware of how deeply they are sleeping. It does appear (again) the depth of sleep is related to latency, which subjects are conscious of and probably do use in making sleep quality statements. Replication with a larger number of subjects is obviously needed before any firm conclusion can be reached.

Summary

The three appendices have attempted to draw together aspects of three measures of the sleep process as they are commonly affected by drugs. In the first appendix the apparently strong relationship between EEG-derived latency to sleep onset and awakening threshold was documented. Both variables had a similar time course (high thresholds

as compared to short latencies) and were similarly responsive to the drug conditions. Aspects of both measures varied from the first to the second replication of the experiment.

In the second appendix, the effects of drugs on subjective sleep parameters were examined. While caffeine administration was still clearly disruptive, differential effects of flurazepam, pentobarbital, and placebo were more difficult to find. Corroborative evidence for earlier results implicating faster onset of activity for flurazepam was found, but similar significant support for evidence for longer effects of flurazepam was not found.

Finally, the relation of within-subject shifts in threshold across baseline and placebo nights was examined in terms of subjective responses. The primary differentiating items were sleep latency items, and this finding led to the conclusion that the primary relation of sleep depth to subjectively reported sleep parameters might lay in the relation of sleep depth to sleep latency.

REFERENCE NOTES

1. Bonnet, M. H. The depth of sleep research. Unpublished area paper, University of Florida, 1976.

2. Bonnet, M. H., & Webb, W. B. Long and short sleepers revisited. Manuscript in submission, 1976.

3. Corvalan, J. C., & Hayden, M. P. Depth of sleep and auditory thresholds during 24-hour recording. Unpublished abstract.

4. Graham, F. BIGANOVA trend analysis program. University of Wisconsin.

5. Kleitman, N., Snyder, F., Hebb, D. O., Williams, H. L., & Dement, W. C. Discussion. In S. S. Kety, E. V. Evarts, & H. C. Williams (Eds.), Sleep and altered states of consciousness. Baltimore: Williams & Wilkins, 1967.

6. Roth, T. Personal communication, February 1977.

7. Roth, T., Kramer, M., & Lutz, T. The effects of hypnotics on sleep, performance, and subjective state. Unpublished paper, University of Cincinnati, 1977.

8. Webb, W. B., & Bonnet, M. H. The sleep of "morning" and "evening" types. Manuscript in preparation, University of Florida, 1977.

9. Williams, R. L., & Agnew, H. W., Jr. The effects of Quaalude and Doriden on human sleep. Unpublished study, 1967.

10. Zimmerman, W. B. Psychological and physiological differences between "light" and "deep" sleepers. Unpublished dissertation, University of Chicago, 1967.

REFERENCES

Agnew, H. W., Jr., & Webb, W. B. Sleep latencies in human subjects: Age, prior wakefulness, and reliability. Psychonomic Science, 1971, 24, 253-254.

Berger, R. J. Oculomotor control: A possible function of REM sleep. Psychological Review, 1969, 76, 144-164.

Bixler, E. O., Leo, L. A., Mitsky, D., & Kales, A. Performance following hypnotic drug administration: Evaluations of flurazepam (Dalmane), phenobarbital, and secobarbital (Seconal). Sleep Research, 1976, 5, 100. (Abstract)

Bixler, E. O., Leo, L. A., Mitsky, D., Pollini, S., & Kales, A. Effects of alcohol, hypnotic drugs, and marijuana on performance. Sleep Research, 1974, 3, 46. (Abstract)

Blake, H., Gerard, R. W., & Kleitman, N. Factors influencing brain potentials during sleep. Journal of Neurophysiology, 1939, 2, 48-60.

Bonnet, M. H., & Webb, W. B. Effect of two experimental sets on sleep structure. Perceptual and Motor Skills, 1976, 42, 343-350.

Bradley, P. B. The central action of certain drugs in relation to the reticular formation of the brain. In H. H. Jasper, L. D. Proctor, R. S. Knighton, W. C. Noshay, & R. T. Costello (Eds.), Reticular formation of the brain. Boston: Little, Brown, 1958.

Davis, H. Classes of auditory evoked responses. Audiology, 1973, 12, 464-469.

Davis, P. A., Davis, H., & Thompson, J. W. Progressive changes in the human EEG under low oxygen tension. American Journal of Physiology, 1938, 123, 51-52.

De Sanctis, S., & Neyroz, U. Experimental investigations concerning the depth of sleep. Physiological Review, 1902, 9, 254-282.

Feldman, S. M., & Waller, H. J. Dissociation of electrocortical activation and behavioral arousal. Nature, 1962, 196, 1320-1321.

Fuxe, K., Hokfelt, T., & Ungerstedt, U. Morphological and functional aspects of central monoamine neurons. International Review of Neurobiology, 1970, 13, 93-126.

Gogolak, G., & Pillat, B. Effect of mogadon on the arousal reaction in rabbits. Progress in Brain Research, 1965, 18, 229-230.

Goodenough, D. R., Lewis, H. B., Shapiro, A., & Sleser, I. Some correlates of dream reporting following laboratory awakenings. Journal of Nervous and Mental Disease, 1965, 140, 365-373.

Guilford, J. Psychometric methods. New York: McGraw-Hill, 1954.

Hartman, E. Sleep and dependence on amphetamine and other drugs. In A. Kales (Ed.), Sleep physiology and pathology. Philadelphia: J. B. Lippincott, 1969.

Hartman, E. The functions of sleep. New Haven: Yale University Press, 1973.

Hays, W. L. Statistics for the social sciences. New York: Holt, Rinehart, and Winston, 1973.

Itil, T. M. Discrimination between some hypnotic and anxiolytic drugs by computer-analyzed sleep. In Williams, R. L., & Karacan, I. (Eds.), Pharmacology of sleep. New York: John Wiley & Sons, 1976.

Itil, T. M., Saletu, B., & Marasa, J. Determination of drug induced changes in sleep quality based on digital computer "sleep prints." Pharmakopsychiatric Neuro-Psychopharmakologie, 1974, 7, 265-280.

Itil, T. M., Saletu, B., Marasa, J., & Mucciardi, A. N. Digital computer analyzed awake and sleep EEG (sleep prints) in predicting the effects of a triazolobenzodiazepine (U-31, 889). Pharmakopsychiatric Neuro-Psychopharmakologie, 1972, 5, 225-240.

Jick, H. Clinical evaluation of hypnotics. In A. Kales (Ed.), Sleep physiology and pathology. Philadelphia: J. B. Lippincott, 1969.

Johnson, L. C. Are stages of sleep related to waking behavior? American Scientist, 1973, 61, 326-338.

Jones, B. E., Bobillier, P., Pin, C., & Jouvet, M. The effect of lesions of catecholamine-containing neurons upon monoamine content of the brain and EEG and behavioral waking in the cat. Brain Research, 1973, 58, 157-177.

Jouvet, M. Telencephalic and rhombencephalic sleep in the cat. In G. E. W. Wolstenholme & M. O'Connor (Eds.), The nature of sleep. Boston: Little, Brown, 1961.

Kales, A., Malmstrom, E. J., Kee, H. K., Kales, J. D., & Tan, T. Effects of hypnotics on sleep patterns, dreaming, and mood state: Laboratory and home studies. Biological Psychiatry, 1969, 1, 235-241.

Keefe, F. B., Johnson, L. C., & Hunter, E. J. Electroencephalographic and autonomic response patterns during waking and sleep stages. Psychophysiology, 1971, 8, 198-212.

Kido, R., & Yamamoto, K. Analysis of tranquilizers in chronically electrode implanted cat. International Journal of Neuropharmacology, 1962, 1, 49-53.

Kleitman, N. Sleep and wakefulness. Chicago: University of Chicago Press, 1939.

Kornetsky, C., & Bain, G. The effects of chlorpromazine and pentobarbital on sustained attention in the rat. Psychopharmacologia, 1965, 8, 277-284.

Lacey, J. T. Somatic response patterning and stress: Some revisions of activation theory. In M. H. Appley & R. Trumball (Eds.), Psychological stress. New York: Meredith Publishing Co., 1967.

Lanoir, J., & Killam, E. K. Alterations in sleep wakefulness patterns by benzodiazepines in the cat. Electroencephalography and Clinical Neurophysiology, 1968, 25, 530-542.

Lidbrink, P., Corrodi, H., & Fuxe, K. Benzodiazepines and barbiturates: Turnover changes in central 5-hydrorytryptamine pathways. European Journal of Pharmacology, 1974, 26, 35-40.

Lindsley, O. R. Operant behavior during sleep: A measure of depth of sleep. Science, 1957, 126, 1290-1291.

Mullin, F. J., Kleitman, K., & Cooperman, N. R. The effect of alcohol and caffeine on motility and body temperature during sleep. American Journal of Physiology, 1933, 106, 478-487.

Mullin, F. J., Kleitman, N., & Cooperman, N. R. Studies on the physiology of sleep changes in irritability to auditory stimuli during sleep. Journal of Experimental Psychology, 1937, 21, 88-96.

Olds, M. E., & Olds, J. Effects of anxiety relieving drugs on unit discharges in hippocampus, reticular midbrain, and pre-optic area in the freely moving rat. International Journal of Neuropharmacology, 1969, 8, 87-103.

Oswald, I. Drugs and sleep. Pharmacological Reviews, 1968, 20, 272-303.

Oswald, I. Psychophysiological and biochemical changes following use and withdrawal of hypnotics. In A. Kales (Ed.), Sleep physiology and pathology. Philadelphia: J. B. Lippincott, 1969.

Oswald, I. Drug research and human sleep. Annual Review of Pharmacology, 1973, 13, 243-252.

Oswald, I. Pharmacology of sleep. In O. Petre-Quadens & J. D. Schalg (Eds.), Basic sleep mechanisms. New York: Academic Press, 1974.

Oswald, I., Taylor, A. M., Treisman, M. Discriminative responses to stimulation during human sleep. Brain, 1960, 83, 440-453.

Price, L. L., & Goldstein, R. Averaged evoked responses for measuring auditory sensitivity in children. Journal of Speech and Hearing Disorders, 1966, 31, 248-255.

Randall, L. O., Schallek, W., Scheckel, C. L., Stefko, P. L., Banziger, R. F., Pool, W., Moe, R. A. Pharmacological studies on flurazepam hydrochloride (RO 5-6901), a new psychotropic agent of the benzodiazepine class. Archives Internationales de Pharmacodynamic et de therapie, 1969, 178, 216-241.

Rechtschaffen, A., Hauri, P., & Zeitlin, M. Auditory awakening thresholds in REM and NREM sleep stages. Perceptual and Motor Skills, 1966, 22, 927-942.

Routtenberg, A. Neural mechanisms of sleep: Changing view of reticular formation function. Psychological Review, 1966, 73, 481-499.

Routtenberg, A. The two-arousal hypothesis: Reticular formation and limbic system. Psychological Review, 1968, 75, 51-80.

Sambrooks, J. E., MacCulloch, M. J., & Rooney, J. F. The automated assessment of the effect of flurazepam and nitrazepam on mood state. Acta Psychiatrica Scandinavica, 1975, 51, 201-209.

Schallek, W., & Kuehn, A. Effects of benzodiazepines on spontaneous EEG and arousal responses of cats. Progress in Brain Research, 1965, 18, 231-238.

Schallek, W., Kuehn, A., & Jew, N. Effects of chlordiazepoxide (Librium) and other psychotropic agents on the limbic system of the brain. Annals of the New York Academy of Science, 1962, 96, 303-312.

Schallek, W., & Thomas, T. Effects of benzodiazepines on spontaneous electrical activity of subcortical areas in the brain of the cat. Archives of International Pharmacodynamics and Therapeutics, 1971, 192, 321-337.

Schwartz, J., Roth, T., Kramer, M., & Hlasny, P. A new benzodiazepine hypnotic and its effect on mood. Current Therapeutic Research, 1974, 16, 964-970.

Shapiro, A., Goodenough, D. R., & Gryler, R. B. Dream recall as a function of method of awakening. Psychosomatic Medicine, 1963, 25, 174-180.

Sokolov, E. N. Perception and the conditioned reflex. New York: Macmillan, 1963.

Steiner, F. A., & Hummel, P. Effects of nitrazepam and phenobarbital on hippocampal and lateral geniculate neurons in the cat. International Journal of Neuropharmacology, 1968, 7, 61-69.

Watson, R., & Rechtschaffen, A. Auditory awakening thresholds and dream recall in NREM sleep. Perceptual and Motor Skills, 1969, 29, 635-644.

Webb, W. B., & Agnew, H. W., Jr. Measurement and characteristics of nocturnal sleep. In L. E. Abt & B. F. Riess (Eds.), Progress in clinical psychology (Vol. 8). New York: Greene & Stratton, 1969.

Webb, W. B., & Agnew, H. W., Jr. Stage 4 sleep: Influence of time course variables. Science, 1971, 174, 1354-1356.

Webb, W. B., & Agnew, H. W., Jr. Sleep efficiency for sleep-wake cycles of varied length. Psychophysiology, 1975, 12, 637-641.

Webb, W. B., Agnew, H. W., Jr., & Williams, R. L. Effect on sleep of a sleep period time displacement. Aerospace Medicine, 1971, 42, 152-155.

Webb, W. B., Bonnet, M., & Blume, G. A post-sleep inventory. Perceptual and Motor Skills, 1976, 43, 987-993.

Williams, H. L. The problem of defining the depth of sleep. In S. S. Kety, E. V. Evarts, & H. L. Williams (Eds.), Sleep and altered states of consciousness. Baltimore: Williams & Wilkins, 1967.

Williams, H. L. Information processing during sleep. In W. P. Koella & P. Levin (Eds.), Sleep. Basel: Karger, 1973.

Williams, H. L., Hammack, J. T., Daly, R. L., Dement, W. C., & Lubin, A. L. Responses to auditory stimulation, sleep loss, and the EEG stages of sleep. Electroencephalography and Clinical Neurophysiology, 1964, 16, 269-279.

Williams, R. L., Agnew, H. W., Jr., & Webb, W. B. Sleep patterns in young adults: An EEG study. Electroencephalography and Clinical Neurophysiology, 1964, 17, 376-381.

Williams, R. L., & Karacan, I. Pharmacology of sleep. New York: John Wiley & Sons, 1976.

Wilson, W. P., & Zung, W. K. Attention, discrimination, and arousal during sleep. Archives of General Psychiatry, 1966, 15, 523-528.

Winters, W. D. Antidepressants and sleep: Clinical and theoretical implications. In A. Kales (Ed.), Sleep physiology and pathology. Philadelphia: J. B. Lippincott, 1969.

Zimmerman, W. B. Psychological and physiological differences between "light" and "deep" sleepers. Psychophysiology, 1968, 4, 387. (Abstract)

Zimmerman, W. B. Sleep mentation and auditory awakening thresholds. Psychophysiology, 1970, 6, 540-549.

Zung, W. K., & Wilson, W. P. Response to auditory stimulation during sleep. Archives of General Psychiatry, 1961, 4, 548-552.

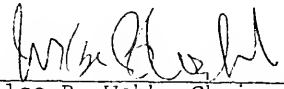
BIOGRAPHICAL SKETCH

There was a child who tried to add one to infinity and got nowhere. Now he tends the fields of orange clover. He has a knapsack that he carries each day. It holds cheese and bread and wine and the scent of clover. The fields are of hills, of clover that waves as it is tended, of the child. The child is of the hills, of the clover that shines as if it were the sun set upon the fields, of bread and wine and the scent of clover.

The child walks through the fields and his feet are covered with the orange of clover, his hands with the work of it. His songs are wordless. His movement is ruddy, gentle, and ceaseless as is that of the eyes of clover.

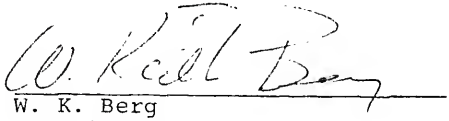
In the sea there is a world of clover. It is the alley-streets of London and the texture of Simon's "Zoo." And more than can be written can be read within the fields of orange clover.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



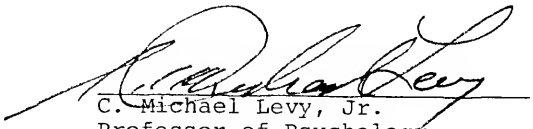
Wilse B. Webb, Chairman
Graduate Research Professor
of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



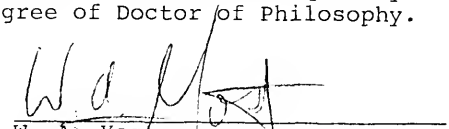
W. K. Berg
Associate Professor of
Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



C. Michael Levy, Jr.
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



W. A. Yost
Associate Professor of
Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

S. F. Zornetzer
Associate Professor of
Neuroscience

This dissertation was submitted to the Graduate Faculty of the Department of Psychology in the College of Arts and Sciences and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

June 1977

Dean, Graduate School