Relationship Between Awareness of Sleepiness and Ability to Predict Sleep Onset

Anna Itoi
Roberto Cilveti
Mike Voth
Bezalel Dantz
Pam Hyde, M.A.
Arun Gupta
William C. Dement, M.D., Ph.D.

Stanford University Sleep Research Center
Department of Psychiatry and Behavioral Sciences
Stanford University Medical Center
Stanford, CA 94305

AAA FOUNDATION FOR TRAFFIC SAFETY
1730 M Street N.W., Suite 401
Washington, D.C. 20036

February 8, 1993

Can Drivers Avoid Falling Asleep at the Wheel?

Executive Summary

This study addresses two key questions regarding the causes of sleep-related accidents: 1) Can drivers anticipate sleep onset well enough to avoid sleep-related accidents? 2) How do drivers use physiological cues of sleepiness in making judgments regarding the riskiness of continued driving?

To explore these questions, we had 41 partially sleep-deprived subjects participate in a one-hour computerized exercise in which they predicted the likelihood of sleep. For 30 consecutive 2-minute intervals, subjects were asked to predict the likelihood of sleep onset for the coining 2-minute interval, using a scale of 0% likelihood to 100% likelihood. Subjects also reported any physiological signs of sleepiness they noticed by selecting one of six icons from the computer screen. The icons represented involuntary eye closure, involuntary head nodding, hallucinatory or wandering thoughts, yawns, and instances of sleep.

Subjects exhibited a wide variation in their ability to predict sleep onset. For all subjects, the mean prediction of sleep likelihood prior to sleep (78%) was significantly higher than the mean prediction of sleep likelihood prior to intervals in which no sleep occurred (42%). However, subjects tended to predict much lower likelihoods of sleep onset before their first sleep event (55%) than before sleep events in general. On average, the rate at which subjects reported physiological indicators of sleepiness (such as head nodding, eye closure, and wandering thoughts) was higher prior to sleep than prior to intervals in which sleep did not occur.

Subjects whose predictions seemed to ignore the frequency of physiological indicators of sleepiness tended to be poor predictors. Subjects whose physiological signs of sleepiness failed to provide a strong indication of whether or not sleep onset would occur also tended to be poor predictors. These findings suggest that people's inability to judge sleep onset, and hence their susceptibility to sleep-related accidents, may be attributable to a scarcity of meaningful physiological warning signs in some individuals and to a failure to acknowledge the importance of meaningful physiological warning signs in others.

Acknowledgments

We would like to extend our deepest thanks to Helena Kraemer for her statistical advice and to L'Ann Bingham for her efforts in making this study possible . This study was funded entirely through a grant from the AAA Foundation for Traffic Safety.

List of Figures

List of Tables

Introduction

One of the most practical questions sleep research can hope to answer is how and to what extent sleep-related auto accidents are avoidable. Researchers estimate that sleep plays a role in 1% to 10% of the 20 million automobile accidents which occur each year in the United States (Lisper et al., 1986). Several aspects of this problem have been investigated in previous studies, and a link between sleepiness and accidents has been well documented (Harris, 1977; Nelson, 1989; O'Hanlon & Kelley, 1977). Most previous studies have focused on performance degradation as a cause of sleep-related accidents. This study focuses instead on the ability of subjects to anticipate the onset of sleep as another possible factor.

It is easy to imagine many scenarios at which a prudent driver with adequate information about the consequences of his actions could make a decision to avoid an accident. What is it then that makes the sleepy driver continue to drive despite the risk? Several explanations are possible. The driver may fail to make the connection between his/her physiological state and his/her dangerously high risk of falling asleep. Current research into and development of an "alertness indicator," which detects and advises drivers of a "critical degradation in attention or alertness," (Report to the Senate Committee, 1989) is based on the assumption that lack of attention to physiological cues is the cause of sleep-related accidents. An alternative explanation is that the driver may fail to make the connection between falling asleep (perhaps only momentarily) and the resulting high risk of an accident. Since dozing while driving without an accident is a relatively commonplace occurrence -- 20% of drivers have experienced falling asleep at the wheel (Report to the Senate Committee, 1989) -- the risk of an accident may not be considered significant by a sleepy driver. Finally, the driver may have adequate information in terms of awareness of sleepiness and the consequences of his/her actions but instead may be highly motivated to continue driving so that, measuring the risks of an accident against the benefits of arriving at his destination, he decides to continue driving despite the risk. Some researchers emphasize this point in criticism of the alertness indicator. They argue that the issue is not that people do not know they are dangerously close to falling asleep, but that they continue to drive despite awareness of sleepiness (Lisper, et al., 1986).

Thus, awareness of sleepiness, awareness of the consequences of sleepiness, and motivation are all factors that can contribute to the safety of the sleepy driver. This study concentrates on the first factor: to what extent are drivers aware of their sleepiness, and what are the mechanisms behind this awareness?

Previous studies suggest that it is virtually impossible to fall asleep without any "warning" whatsoever (Lisper et al., 1986; Torsvall et al., 1989). That is, certain physiological cues invariably precede sleep, including eye straining and closure, excessive yawning, hallucinogenic and wandering thoughts, head nodding, and struggling to fight sleep (Lisper et al., 1986; Nelson 1989). Yet the question remains open whether these and other physiological and psychological cues provide sufficient information for an individual to determine accurately when sleep will occur.

In order to address these issues, we partially deprived college students of one night's sleep and on the following morning invited them to participate in a one-hour computerized exercise. This exercise required them to make repeated judgments of the likelihood of sleep onset. This report gives a description of the procedure we used, followed by a presentation of the results and a discussion of their significance.

The Research Question

Do people know when they are about to fall asleep? Two aspects of the question require clarification: 1) what it means for a person to "know" and how this should be measured; and 2) what is meant by "about to." The specific question we chose for this study was the following: How accurately can people judge the chance that they will fall asleep in the next 2 minutes on a scale of 0 to 100%?

The chosen question is intended to avoid any ambiguity in interpretation either by the subjects or by those reading the study's results. By using a set time frame, the chosen question eliminates any variability in interpretation of what it means to be "about to fall asleep." The fixed time frame also allows subjects to become familiar with the "meaning" of that particular duration of time, so that subjects' ability to estimate times is relatively unimportant to the results. The numerical probability scale allows us to capture the level of certainty in subject responses, without leaning exclusively on ambiguous probabilistic language such as "likely" and "unlikely," which has been shown by previous studies to be interpreted differently by different people and in different contexts (Kahneman et al., 1982).

We chose 2 minutes as the time threshold during which subjects make their judgments in order to make the study relevant to the experience of a sleepy driver. The threshold must be long enough to allow time for a hypothetical driver who anticipates falling asleep to take reactive measures (e.g., to pull to the side of the road). At the same time, the threshold must be short enough that subjects can have a reasonably accurate internal concept of its duration. A threshold of 2 minutes was deemed to satisfy both of these constraints.

Methods

The methods section consists of five components: the selection of the subject pool, an explanation of the testing schedule, a description of the testing rooms, a description of the prediction exercise, and a description of the detection of sleep via video and polygraphic recordings.

Subjects

Forty-two healthy Stanford students between the ages of 17 and 22 (mean age: 19.5 ± 1.5 years) participated in the study. Twenty were male and 22 were female. All students were very comfortable with using the computer and manipulating the computer's mouse input device. Subjects were recruited through flyers posted in various locations on the campus. All subjects were screened for any sleep-related complaints. Each participant received an honorarium of $75.
One female subject quit during the course of the experiment due to feelings of nausea. We used only the data on the remaining 41 subjects.

Testing Schedule

Subjects who passed the initial screening were asked to sleep at least 7 hours on the 2 nights before their scheduled arrival at the sleep laboratory and to keep a diary of their respective sleep schedules.[1] After these 2 full nights of sleep, they were asked to arrive at the sleep laboratory on the evening before the scheduled test session to undergo partial sleep deprivation. The daily testing schedule is summarized in Table 1. Time Event
23:00 Arrival, Orientation
00:00 Electrode Application
01:15 15 Minute Practice Session

01:30 Free Time
05:00 2 Hour Nap
07:30 6 Minute Practice Session
07:40 One-hour Prediction Exercise
09:00 Sleep Recovery

Table 1. Summary of the Testing Schedule

Subjects reported that they slept an average of 7.36± 1.21 hours on each of the 2 nights prior to the night of partial sleep deprivation. It is important to stress that the information from these sleep diaries ire self-reported. All subjects denied having made any drastic changes to their normal sleep schedule over the previous 2 nights.

The subjects were tested two at a time. After arriving at 23:00, the subjects signed human consent forms and were given a brief tour of the laboratory. All watches and timepieces were removed from subjects, so they had no time cues while they stayed in the laboratory. Shortly afterwards, a certified polysomnographic technologists[2] applied electrodes to each subject to be used the following morning to record any occurrences of sleep during the one-hour prediction exercise. By applying the electrodes early in the night, the subjects were able to become accustomed to them before the actual recording took place. At 01:15, the subjects participated in a fifteen-minute practice session, a shortened replica of the actual hour of testing that occurred later. The purpose of the practice session was to familiarize the subjects with all aspects of the procedure and to familiarize them with the length of a 2-minute interval.

Following the practice session, subjects were given free time but were not permitted to fall asleep. Subjects were allowed to consume non-caffeinated beverages and snacks in controlled amounts throughout the night. A researcher monitored their behavior to make sure that they did not sleep.
Since the pilot data (N=5) showed that partial sleep deprivation was sufficient to induce extreme sleepiness in most subjects, we allowed the subjects to take a 2-hour nap from 05:00 07:00. Polygraphic monitoring was carried out during the nap to note any irregular disturbances in sleep quality.[3]

After awakening from their nap, the subjects had a small snack which consisted of yogurt and a non-caffeinated beverage and then engaged in a six-minute practice session before the prediction session began at 07:40. A detailed description of the one-hour prediction exercise is given below. Subjects were allowed to recover sleep following the hour of testing. They left the laboratory when they felt sufficiently rested.

Testing Rooms

The one-hour prediction sessions took place in private, semi-dark, sound-attenuated rooms which were furnished with a video camera, a chair, a desk with a computer, a dim light (100 watts), and an intercom which was used to communicate with the subjects. The setting was designed to be as sleep-inducing as a monotonous drive or a monotonous job. Pilot tests showed that staring at a computer screen, as opposed to reading, or doing math problems, was the most sleep-inducing activity for Stanford students.
In these rooms, subjects were instructed to sit at the desk and engage in a computerized prediction exercise while being monitored by the video camera and by polygraphic recordings of their brain activity, eye movements, and submental muscle activity. [2] A "polysomnographic technologist" is defined by Thorpy & Yager in The Encyclopedia of Sleep and Sleep Disorders (1991) as "an individual trained and tested to administer the polysomnograph, the test that measures sleep activity and other physiological variables by recording brain, eye, and muscle activity in sleep."

[3] The 2-hour nap on the night before testing was recorded by the polygraph. It took an average of 5.07± 5.29 minutes for subjects to fall asleep once lights were turned off in the bedrooms (sleep-onset latency). Sleep-onset latency ranged from 0.5 minutes to 27.3 minutes.

Prediction Exercise

The computerized prediction exercise, which began at 07:40, was coordinated by a software program that repetitively asked the subject to assess the likelihood of falling asleep within the next 2 minutes, and then required the subject to report any physiological indicators of sleepiness he/she noticed. The program required these assessments and reports for 30 consecutive 2-minute intervals. The entire session lasted for one hour. Subjects were asked to try to stay awake during this hour. The structure of the session is illustrated in Figure 1. Prediction

Interval 1 p Interval 2 p Interval 30
Slept? (Y,N) 2 Slept? (Y,N) 30 Slept? (Y,N)
2 minutes

Figure 1. The Format of the One-Hour Prediction Exercise

At the beginning of each 2-rninute interval, a scale from 0% to 100% appeared on the screen as illustrated in Figure 2. The subject was instructed to ask him/herself, "How likely is it that I will fall asleep in the next 2 minutes?" The subject would then make a prediction by positioning the computer's cursor at any point on the scale using the computer's mouse, and then clicking the left mouse button. Since these students use computers almost every day, manipulating the computer's mouse proved to be no problem for them.

Once a selection was made from the prediction scale, the scale disappeared, and six icons appeared on the screen as shown in Figure 3 to enable the subject to report any physiological indicators of sleepiness. The six icons represent involuntary eye closure, head nodding, hallucinatory or wandering thoughts, yawns, sleep, and other physiological signs. While the icons were on the screen, the subject was instructed to select the appropriate icon each time he/she experienced any of these physiological indicators during the 2-minute interval. Subjects reported these events by positioning the computer's cursor over the appropriate icon and clicking the left mouse button. These icons remained on the screen until the end of the 2-minute interval when the prediction scale reappeared to prompt a prediction for the next interval. Subjects were given guidelines for icon selection on the night before the morning of testing. These guidelines are presented in Table 2.

Figure 2. Computerized Scale Used by Subjects to Record Predictions

Figure 3. Computerized Icons Used by Subjects to Record Physiological Indicators of Sleepiness

THOUGHTS * Hallucinations, unintentionally wandering thoughts, scattered and unconnected visions, etc. **YAWN** * Self-explanatory. **OTHER** * Any physiological sign of sleepiness other than the five identified by the icons. **EYES** * Involuntary partial or full eye closure for greater or equal to 1 second. Blinking, though also involuntary, does not count. • Eyes unintentionally out of focus. **HEAD** * Involuntary head nodding. Head nodding is characterized by either a gently swaying or a spasmodic jerk of the head. **SLEEP** * Sleep can be very momentary. Even 3 seconds should be counted as sleep.

Table 2. Criteria for Icon Selection

Video and Polygraphic Recording of Sleep

We defined sleep to include attacks as short as 3 seconds ("microsleep" events).[4] Sleep was identified by the polygraph[5] and confirmed by the video recordings[6] of subjects during the one-hour prediction exercise. Although there is no set criterion for "microsleep" duration, the 3-second criterion we chose is relevant to sleep-related accidents, as "it takes an unresponsive driver only three seconds to traverse a four-meter-wide highway shoulder at a speed of 60 m.p.h. and an angle of departure of three degrees" (O'Hanlon & Kelley, 1977).

If a subject was sleeping at the moment the prediction scale appeared, he/she was awakened with a single beep over the intercom. The subject was also awakened if he/she seemed to be in danger of falling to the floor. Any occurrence of sleep throughout the hour was otherwise left undisturbed.

[4] We scored "microsleeps" according to the guidelines delineated by Guilleminault et al. (1975) who defined a "microsleep" as a "short-lasting burst of typical stage 1, slow wave sleep as described by Rechtshaffen and Kales (1968) and/or a short burst of `synchronous theta activity' recorded in central monopolar derivations (C3/A2, C4/A1,02/Al)." The Encyclopedia of Sleep and Sleep Disorders offers a more general definition, defining a "microsleep" as "an episode of sleep lasting only a few seconds that occurs during wakefulness" (Thorpy & Yager, 1991).

[5] We can detect sleep by using the polygraph to record changes in the subjects' brain waves (EEG), eye movements (EOG), muscle tone (EMG), and breathing.

[6] The video recordings were reviewed both during and after the hour of testing for visually identifiable sleep episodes. We defined "visually identified sleep" as greater than or equal to 3 seconds of eye closure, head nodding, and muscle relaxation. Instances in which subjects were moving the mouse to record either predictions or physiological indicators of sleepiness were not considered instances of sleep. Time synchronization of the polygraph and video camera enabled us to confirm "visually identified sleep" with "polygraphically identified sleep."

Subjects made predictions about the likelihood that they would sleep in each 2-minute interval rather than about how often they would sleep. Thus the primary measure of sleep in the study is a simple binary indication of whether or not sleep occurred in a given 2-minute interval. We refer to any of the 30 consecutive 2-minute intervals in which the subjects experienced at least one sleep episode according to the video and polygraphic recordings as a "sleep interval" or simply a "sleep." Likewise, any interval in which sleep did not occur is referred to as a "no-sleep interval" or simply a "no-sleep." Therefore, a subject could have fallen asleep much more than 30 times, but the maximum possible number of sleep intervals would never exceed 30.

Results

The results of the study are divided topically into five sections. The first section summarizes the general results of the prediction exercise. The second section describes the ability of subjects to predict sleep onset -- that is, the relationship between subject predictions and actual sleep onset. The third section examines the extent to which physiological indicators of sleepiness warn of sleep onset. The fourth section probes the extent to which subjects considered physiological indicators of sleepiness in making their predictions. The fifth section and the discussion that follows it explores interactions among several categories of data.

General Results

The research protocol required 41 subjects to make 30 predictions each for a total of 1,230 predictions. However, subjects occasionally slept during times when the prediction scale appeared on the computer screen and consequently missed making their predictions. The 48 missed predictions leave 1,182 predictions recorded over the course of the study. One subject slept at least once in all 30 minute intervals, while 9 subjects never slept during the one-hour prediction exercise.

Generally, more subjects slept in each successive interval during the course of the study. The mean prediction of sleep likelihood also followed a general upward trend over time, as did the mean complaint rate (including eyes, head, thoughts, yawn, prior sleep events, and other complaints).

Ability of Subjects to Predict Sleep Onset

In general, subjects had some ability to discriminate sleep from no-sleep with their predictions. However, subjects were frequently surprised by their first sleep event. That is, predictions prior to the first sleep event were much lower than predictions prior to sleep events in general. In addition, first sleep predictions were not significantly different from predictions prior to intervals when no sleep occurred. In addition, there was a wide variation among individual subjects' ability to predict sleep onset, with a significant number of subjects exhibiting a limited ability to predict sleep.

For all subjects, the mean prediction of sleep likelihood prior to sleep was 78%, and the mean prediction of sleep likelihood prior to no-sleep was 42%. Note that if subjects were perfect predictors of sleep, they would always predict 100% prior to sleep and 0% prior to no-sleep. The mean prediction prior to subjects' first sleep was 55%. Table 3 summarizes these results. Though the mean prediction prior to sleep is significantly higher than the mean prediction prior to no-sleep, there is no significant difference between the mean prediction prior to the first sleep and the mean prediction prior to no-sleep:

Statistic:	Predictions Prior to No-Sleep	Predictions Prior to Sleep	Predictions Prior to First Sleep
Mean	42%	78%	55%
SD	25%	22%	22%
Different from No-sleep ($P<0.05$)[7]	--	Yes	No

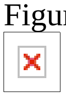
Table 3. Comparison of Predictions Prior to No-Sleep/Sleep/First Sleep

Figure 5 illustrates the relationship between subject predictions and subsequent sleep across all subjects and all intervals. The upward trend in the data indicates that higher predictions correspond to a higher frequency of sleep occurrence. The data points in Figure 5 give the percentage of times sleep occurred when predictions were in various ranges. For example, the first data point shows that when predictions were between 0 to 10%, sleep occurred in the subsequent interval approximately 5% of the time. The last data point shows that when predictions were between 90% to 100%, sleep followed approximately 90% of the time.

We also examined the relationship between predictions and sleep on a subject-by-subject basis. We evaluated the degree of difference between subjects' mean predictions prior to periods in which sleep occurred and their mean predictions prior to periods in which sleep did not occur. This difference can be taken as a measure of subjects' ability to discriminate sleep from no-sleep when considering their prior predictions. The larger the difference, the better the subject is able to discriminate between sleep and no-sleep. For example, a subject who always predicts a 100% likelihood of sleep before sleeping and a 0% likelihood before not sleeping is considered

[7] Paired t-tests were performed based on individual subject means.

Figure 5. Relationship Between Predictions and Sleep Onset



maximally able to discriminate sleep from no-sleep (100-0 = 100). Table 4 summarizes the results, tabulating subjects by the difference between their mean predictions prior to sleep and their mean predictions prior to no-sleep. For 2 subjects, the mean gap was between 50% and 60%, suggesting a fairly high degree of accuracy in discriminating between sleep and no-sleep intervals. However, for 19 of the 31 subjects, the gap between their mean prediction prior to sleep and prior to no-sleep was less than 30%, indicating a relatively poor degree of accuracy. (Appendix I lists the complete results of the subject-by-subject analysis summarized in Table 4).

$Mean_s - Mean_{NS}$	Number of Subjects
60-50%	2
50-40%	3
40-30%	7
30-20%	9
20-10%	6
10-0%	4
TOTAL	**31**

Table 4. Difference Between Mean Predictions Prior to Sleep and No-Sleep

Relationship Between Physiological Indicators and Sleep Onset

The rates at which subjects reported feeling signs of sleepiness were in general significantly higher before sleep than before no-sleep. A notable exception was that complaints of yawning were significantly more frequent prior to intervals in which sleep did not occur, indicating that yawning is perhaps not a good indicator of imminent sleep. Subject by subject analysis reveals that subjects differ widely in the relationship between physiological indicators and sleep onset. In particular, some subjects showed very little difference in complaint rate between sleeps and no-sleeps.

Subjects clicked on computer icons during the course of the study to record physiological indicators of sleepiness, including eye complaints, head-nodding complaints, prior sleeps, complaints or yawning, reports of wandering and hallucinatory thoughts, and any other indicators of sleepiness they felt (hereafter referred to as "other" complaints). For each subject, each interval, and each complaint type, we calculated a waking complaint rate (hereafter referred to simply as a it complaint rate") which measures the number of complaints recorded per minute the subject was awake during the specified interval. We calculated a total complaint rate for each subject for each interval as the sum of the individual complaint rates for that subject and interval. To illustrate how complaint rates for the various physiological indicators measured in the study relate to sleep onset, we compared complaint rates prior to intervals in which sleep occurred to those prior to intervals in which no sleep occurred.

We found that the mean total complaint rate was significantly higher prior to sleep (3.50 complaints/minute) than prior to no-sleep (1.82 complaints/minute). The rates at which subjects selected the "sleep," "head," "eyes," and "thought" icons were significantly higher before sleep than before no-sleep ($p<.05$). There was no significant difference, however, in the mean rate at which subjects selected the "other" icon between sleep and no-sleep intervals, and the rate at which subjects complained of yawning was actually significantly higher before no-sleep than before sleep ($p<.05$). Table 5 summarizes the comparisons of complaint rates prior to sleep (S) and no-sleep (NS) in terms of means and standard deviations.

Click here for table 5

Figure 6 illustrates the relationship between total complaint rate and subsequent sleep across all subjects and intervals. The upward trend in the data indicates that higher complaint rates correspond to a higher frequency of sleep occurrence. The data points in Figure 6 show the percent of time subjects slept when the total complaint rate was in various ranges.

[8] Paired t-tests were performed based on individual subject means.

Figure 6. Relationship Between Physiological Indicators and Sleep Onset

We also compared individual subjects' complaint rates prior to intervals in which sleep occurred to those prior to intervals in which no sleep occurred. Table 6 below summarizes the results, tabulating subjects by the difference between the mean total complaint rates prior to sleep and the mean total complaint rates prior to no-sleep. This table excludes the 9 subjects who never slept and the I subject who slept for all 30 intervals. Two of the 31 subjects analyzed reported on average at least 3 more complaints of sleepiness before sleep occurred than before sleep did not occur. Fourteen subjects exhibited a mean complaint rate difference of less than 1 complaint/minute, 3 of whom actually had higher complaint rates before no-sleep than before sleep. (Appendix 2 lists the complete results of the subject-by-subject analysis summarized in Table 6).

Mean Total Complaint Rates - Mean Total Complaint RateNS (Complaints/Minute)	Number of Subjects
>3	2
3-2	8
2-1	7
1-0	11
<0	3
TOTAL	**31**

Table 6. Relationship Between Physiological Indicators and Sleep Onset - Individual Subjects

Relationship Between Physiological Indicators and Subject Predictions

In general, subjects tended to give high predictions of sleep likelihood after experiencing high complaint rates and low predictions after low complaint rates. Thus, subjects seemed to take into account physiological indicators of sleepiness in making their predictions. However, subject-by-subject analysis reveals broad variations: some individuals in this study exhibited no clear association between their predictions and the rate at which they reported physiological indicators of sleepiness.

In order to investigate to what extent subjects' predictions responded to physiological cues, we grouped predictions into "high predictions" (greater than or equal to 50%) and "low predictions" (less than 50%) and then compared complaint rates prior to low predictions with those prior to high predictions.

We found that the mean total complaint rate was significantly higher prior to high predictions (3.10 complaints/minute) than prior to low predictions (1.82 complaints/minute). Table 7 summarizes the comparisons of complaint rates prior to high predictions (H) and low predictions (L) in terms of means and standard deviations.

The rates at which subjects selected the "sleep," "head," "eyes," and "thought" icons were significantly higher before high predictions than before low predictions ($p<.05$). There was no significant difference, however, in the mean rate at which subjects selected the "other" icon between high and low predictions, and the rates at which subjects complained of yawning were actually significantly higher before low predictions than before high predictions ($p<.05$).

Click here for table 7

Figure 7 illustrates the relationship between total complaint rate and subsequent sleep across all subjects and all intervals. There is a general upward trend in the data points. The upward trend indicates that subjects were more likely to predict high values after intervals in which their complaint rates were high.
We also compared individual subjects' mean complaint rates prior to intervals in which they predicted above 50% to those in which they predicted below 50%. Table 8 summarizes the results of this analysis, excluding the 5 subjects who never predicted above 50% and the 4 subjects who never predicted below 50%. Table 8 tabulates subjects by the

extent of the difference between total mean complaint rates before high predictions (predictions higher than 50%) and the total mean complaint rates before low predictions (predictions lower than 50%).

[9] Paired t-tests were performed based on individual subject means.

Figure 7. Relationship Between Physiological Indicators and Predictions

As shown in Table 8, we found that the mean difference between complaint rates before high predictions and low predictions varied widely among subjects. Two subjects reported, on average, at least 3 more complaints per minute before high predictions than before low predictions. However, 9 subjects reported less than I more complaint per minute before high predictions than before low predictions. Of these 9 subjects, 2 actually reported more complaints before low predictions than before high predictions. (Appendix 3 lists the complete results of the subject-by-subject analysis summarized in Table 8).

Mean Total Complaint Rates (H) - Mean Total Complaint Rate (L)	Number of Subjects
>3	2
3-2	10
2-1	11
1-0	7
<0	2
TOTAL	**32**

Table 8. Relationship Between Physiological Indicators and Prediction-Making - Individual Subjects

Relationships Among Physiological Indicators, Sleep Onset, and Predictions

In this final section, we combine the three areas of results presented above (relation of predictions to subsequent sleep onset, relation of complaints to subsequent sleep onset, and relation of complaints to subsequent predictions) and look at whether or not subjects who showed a strong relationship in one area (i.e., predictions were correlated with complaint rate) also showed a strong relationship in another (i.e., predictions were correlated with subsequent sleep). This allows us to gain additional insight into the reasons why subjects' ability to predict sleep onset is limited.

The first question we consider is whether subjects whose complaint rates are especially predictive of sleep onset (that is, their total complaint rate is higher prior to sleep than prior to no-sleep) also tend to predict well (defined for this analysis as those subjects whose predictions are much higher before sleep intervals than before no-sleep intervals). Figure 8 plots the difference between mean total complaint rates prior to sleep and prior to no-sleep against the difference between mean predictions prior to sleep and prior to no-sleep for each subject. The general upward trend indicates that subjects whose physiological signs provided a good indicator of whether sleep would occur tended to do a good job of predicting sleep onset, while subjects whose physiological signs were less informative tended to predict more poorly. The suggestion here is that one reason subjects predicted poorly is simply because they did not have (or did not notice) adequate physiological warning signs.

Figure 8. Ability to Judge Sleep vs. Predictiveness of Complaints

The second question we consider is whether subjects whose complaint rates are strongly related to their predictions (that is, their total complaint rate is higher prior to high predictions - predictions greater than or equal to 50% - than it is prior to low predictions - predictions less than 50%) also tend to predict well (defined for this analysis as those subjects whose predictions are much higher before sleep intervals than before no-sleep intervals). Figure 9 plots the difference between mean total complaint rates prior to high and prior to low predictions against the difference between mean predictions prior to sleep and prior to no-sleep for each subject. The general upward trend indicates that subjects who seemed to consider their physiological signs in making predictions also tended to do a good job of predicting sleep onset, while subjects who apparently failed to consider physiological signs tended to predict more poorly. This suggests a second reason why some subjects tended to predict poorly: they failed to adjust their judgments of sleep likelihood according to the physiological signs that they experienced.

Figure 9. Ability to Judge Sleep vs. Responsiveness of Predictions to Complaints

Discussion

The results of the study establish that people do have a limited ability to predict the onset of sleep. What is of particular interest to our research question is the extent of that ability. The wide variation in predictions both before sleep and before no-sleep establishes that subjects certainly do fall asleep at times when they think sleep is highly unlikely and fail to fall asleep at times when they think sleep is highly likely. The mean prediction before subjects' first sleep was low (55%), suggesting that subjects in general judged sleep to be fairly unlikely when, in fact, it first occurred. There was no significant difference between predictions prior to subjects' first sleep (55%) and predictions prior to no-sleep (47%). We found, however, that the mean prediction prior to sleep in general (78%) was significantly higher than the mean prediction prior to no-sleep (47%).

Since there is a significant chance that a sleepy driver will not be alive to experience his or her second, third, or fourth sleep attack, it is important to note that predictions made in this study prior to the first sleep event are of particular importance. There was broad variation in subjects' ability to predict their first sleep onset. Predictions prior to the first sleep event ranged from as low as 4.9% to as high as 93.3%. Interestingly, the subject who predicted 4.9% reported feeling quite confident that he could accurately predict sleep before the study began. After the study, he admitted to being surprised by his inaccuracy. This instance was one of 4 instances in which subjects predicted a less than 10% chance of falling asleep before a sleep event actually occurred. The possibility that subjects can predict a very low likelihood of sleep and then actually fall asleep easily translates into a potential to make misjudgments on the road during a prolonged state of sleepiness.

There are many reasons why subjects' ability to predict sleep onset may be limited. A number of possible reasons are highlighted by these results. First, the information people receive in terms of physiological signs of sleepiness is imperfect. There is some positive relationship between physiological signs measured in this study (eyes, head, thoughts, and prior sleep) and imminent sleep onset (see Table 5), but this relationship is imperfect in the sense that there are times, particularly during extended periods of struggling to stay awake, when sleep can occur suddenly and without clear physiological warning signs. In addition, there is individual variation in the strength of the relationship between indicators of sleepiness and imminent sleep onset. Figure 8 indicates that the subjects whose physiological complaints were not strongly correlated with sleep onset tended to predict sleep onset rather poorly. Thus, poor predictions can be viewed as being caused by poor quality of information (i.e., physiological signs) available upon which to base predictions. The quality of information issue is a very difficult but probably very important problem to solve. The results of this study suggest that work on a device such as a sleepiness indicator, which can somehow provide better information on imminent sleep onset than is available through one's normal awareness, may be worthwhile in reducing judgment errors attributable to poor quality of information.

A second reason why subjects' ability to predict sleep onset may be limited is that people may not adequately consider the physiological indicators of sleepiness that they can, in fact, notice and record. That is, though they have information that could provide a good indicator that sleep will occur, they may ignore this information in judging the likelihood of sleep onset. Again, the results of this study show that people can use physiological indicators of sleepiness such as eyes straining, closing, or wanting to close, head-nodding, and hallucinogenic or wandering thoughts as a warning sign that sleep is imminent.

However, if subjects ignore this information, they will fare no better than those who perceive no signals at all. Figure 10 shows that subjects who paid little heed to physiological indicators in making their predictions tended to be poor predictors. Presumably this lack of attention to warning signs in prediction-making can be readily linked to poor 'judgments in critical contexts such as driving. The problem of inattention to the warning signs of sleep onset is also a difficult one. It is particularly difficult in that it can confound solutions to the aforementioned problem of inadequate warning signs.

For example, some researchers have suggested that the idea of an alertness indicator is worthless because of this second problem of ignoring sleepiness warning signs. However, it is our contention that both problems need to be solved, and that the best solution to the second problem is education. People need to understand that their ability to judge sleep onset is, in fact, limited, and that they may not experience a drastic change in their physiological state just prior to falling asleep. The alertness indicator is a good idea given the low quality of information normally available for sleep-related decision making, but it is only worthwhile to the extent that people can be educated as to their inability to judge sleep onset accurately without it.

Closely related to the problem of limited information is the problem of overconfidence regarding the ability to anticipate sleep onset. Though this study does not specifically investigate this phenomenon, some attempts have been made to fill in the gaps. After the study, a questionnaire was administered to a separate group of subjects participating in a different experiment. The protocol of our current experiment was explained to these subjects and they were asked whether they thought they could accurately judge sleep onset in this context. Most subjects indicated that they thought they would be able to predict above 90% before every sleep and below 10% before every no-sleep. In our study, no subjects were able to predict sleep onset with anywhere near this degree of accuracy. The fact that people expect to be able to judge sleep onset very accurately seems to make them more willing to test the limits of their ability to stay awake, even in risky situations such as driving. As suggested above, the solution to this problem is education. People need to be made aware that they often do not have adequate information through signs of sleepiness to avoid sleep-related accidents.

Even if people do have the ability to predict the likelihood of sleep onset in a 2-minute period with great accuracy, it is still unclear whether they would take appropriate action to avoid sleep-related accidents. We frequently hear of people who continue to drive despite feeling extremely tired because they are trying to drive until the last second, and they are sure that they will have some clear physiological warning moments before sleep occurs. However, a study examining accounts of incidents in which drivers had fallen asleep at the wheel found that drivers do not appear to have any last-second warning. The researcher observes that "fatigue ... develops insidiously, and even those who know they are tired may actually lose consciousness suddenly and without trying to decelerate" (Parsons, 1986). The fact that we are able to control our sleep-impulse to some extent and in some situations can easily be mistaken as an ability to control sleep indefinitely and in all situations. As sleep-deprived individuals are subjected to prolonged, uninterrupted, and monotonous tasks, the chance of failing asleep increases with time as suggested in our study. It is our contention that sleepy drivers should not count on any last-second warning signs or on their ability to fight sleep; if a driver feels that sleep is even somewhat likely within the next few minutes, that driver should immediately stop driving.

Finally, the subjective probability scale methodology employed in this study provides a useful short term subjective measure of sleepiness that promises a high degree of consistency in meaning across subjects and also a meaningful and direct interpretation in terms of the physical reality of future sleep onset. Future studies comparing the ability of different groups (e.g., true insomniacs versus psychological insomniacs, children versus the elderly, drivers who have been in accidents versus drivers who have not) to predict sleep onset would be useful in establishing differences among these groups with greater cross-subject precision than is possible using traditional verbal subjective scales. Most importantly, it would be useful to conduct studies to examine whether people can improve their abilities to predict sleep onset through practice and/or training.

References

Department of Transportation: Report to the Senate Committee on Appropriations and the House Committee on Appropriations. (1989). Transportation-Related Sleep Research.

Guilleminault, C., Billiard, M., Montplasir, J., & Dement W. (1975). Altered states of consciousness in disorders of daytime sleepiness. Journal of Neurological Sciences, 26, 377-393. Harris, W. (1977). Fatigue, circadian rhythm, and truck accidents. In R.R. Mackie (Ed.), Vi2ilance, New York: Plenum Press, 133-146.

Kahneman, D., Slovic, P., & Tversky A. (1982). Judgment Under Uncertainty: Heuristics and Biases, New York: Cambridge University Press.

Lisper, H.-O., Laurell, H., & Van Loon, J. (1986). Relation between time to falling asleep behind the wheel on a closed track and changes in subsidiary reaction time during prolonged driving on a motorway. Ergonomics, 29, 445-453.

Nelson, T. (1989). Subjective factors related to fatigue. Alcohol, D 5 (3), 193214.

O'Hanlon, J., & Kelley G. (1977). Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation. In R.R., Mackie (Ed.), Vigilance, New York: Plenum Press, 87-1 10.

Parsons, M. (1986). Fits and other causes of loss of consciousness while driving. Quarterly Journal of Medicine, 58 227), 295-303.

Rechtschaffen, A., & Kales, A. (1968). A Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages of Human Subjects, Los Angeles: University of California, Brain Information Service/Brain Research Institute.

Thorpy, M.J., &Yager, J. (1991). The Encyclopedia of Sleep and Sleep Disorders, New York: Facts on File.

Torsvall, L., & Akerstedt, T. (1987). Sleepiness on the job: continuously measured EEG changes in train drivers. Electroencephalography and Clinical Neurophysiology, 66, 502-51 1.

Torsvall, L., Akerstedt, T., Gillander, K., & Knutsson, A. (1989). Sleep on the night shift: 24hour EEG monitoring of spontaneous sleep/wake behavior. Psychophysiology, 26, 352358.

Appendix I

Click here for Appendix I

Appendix 2

Relationship Between Physiological Indicators and Sleep Onset

Subjects' mean total complaint rates prior to sleep and prior to no-sleep are presented in order of decreasing difference between the two total complaint rate means.

	Total Complaints					Total Complaints		

S#	NS	S	Diff.		S#	NS	S	Diff.
24	1.58	6.62	5.04		46	0.94	1.61	0.67
27	2.25	5.50	3.25		17	1.28	1.92	0.64
36	2.78	5.29	2.50		10	4.50	5.05	0.55
35	1.73	4.23	2.49		22	1.69	2.24	0.55
5	5.08	7.58	2.49		45	0.61	0.77	0.16
34	2.40	4.81	2.40		7	3.18	3.30	0.12
30	2.61	4.96	2.35		29	1.95	1.99	0.05
20	1.92	4.25	2.34		13	0.75	0.50	-0.25
28	3.06	5.33	2.26		39	0.61	0.25	-0.36
26	2.27	4.45	2.17		44	3.06	1.86	-1.20
25	2.48	4.37	1.89		23	0.03	-	-
38	1.32	3.07	1.75		16	0.19	-	-
8	5.37	6.88	1.51		40	0.48	-	-
33	0.71	2.08	1.37		21	0.59	-	-
6	2.17	3.48	1.31		37	0.90	-	-
14	1.32	2.33	1.01		9	1.59	-	-
31	0.00	1.01	1.01		18	2.83	-	-
15	0.92	1.90	0.98		43	5.22	-	-
11	1.78	2.74	0.95		12	-	4.41	-
41	2.94	3.87	0.93		19	-	0.17	-
42	2.30	3.18	0.88					

Appendix 3

Relationship Between Physiological Indicators and Predictions

Subjects' mean complaint rates prior to high predictions (P>50%) and prior to low predictions (P<50%) are presented in order of decreasing difference between the two total complaint rate means.

	Total Complaints					Total Complaints		
S#	P< 50%	P> 50%	Diff.		S#	P< 50%	P> 50%	Diff.
24	0.00	5.81	5.81		17	0.92	1.96	1.04
27	2.25	5.50	3.25		18	2.33	3.36	1.02
25	1.06	3.97	2.91		7	2.74	3.61	0.86
35	1.50	4.18	2.68		31	0.18	0.95	0.77

26	1.63	4.29	2.66		46	0.83	1.43	0.60
34	1.00	3.66	2.66		40	0.08	0.59	0.50
28	3.36	6.00	2.64		37	0.74	1.20	0.46
11	0.63	2.88	2.26		41	3.28	3.39	0.11
38	0.42	2.62	2.21		13	0.67	0.74	0.07
30	2.85	4.96	2.11		22	1.92	1.79	-0.14
36	3.32	5.35	2.04		19	0.20	0.00	-0.20
8	4.92	6.94	2.02		23	0.03	-	-
12	2.50	4.48	1.98		16	0.19	-	-
9	0.63	2.26	1.64		21	0.57	-	-
5	5.47	7.10	1.63		39	0.59	-	-
6	1.97	3.58	1.61		43	5.22	-	-
10	3.50	5.07	1.57		42	-	3.06	-
44	0.50	1.95	1.45		20	-	2.88	-
15	0.25	1.70	1.45		29	-	1.96	-
14	1.00	2.26	1.26		45	-	0.68	-