The Effects of Total Sleep Deprivation on Bayesian Updating*

David L. Dickinson Dept. of Economics Appalachian State University	Sean P.A. Drummond Dept. of Psychiatry Laboratory for Sleep and Behavioral Neuroscience University of California—San Diego Psychology Service Veterans Affairs San Diego Healthcare System

ABSTRACT

Recent evidence suggests that nearly 25% of U.S. adults (47 million) suffer from some level of sleep deprivation. The impact of this sleep deprivation on the U.S. economy includes direct medical expenses related to sleep deprivation and related disorders, the cost of accidents, and the cost of reduced worker productivity. Sleep research has examined the effects of sleep deprivation on a number of performance measures, but the effects of sleep deprivation on decision-making under uncertainty are largely unknown. In this article, subjects perform a decision task (Grether, 1980) in both a well-rested and experimentally sleep-deprived state. The experimental task allows us to explore the extent to which subjects weight prior odds versus new evidence (i.e., information) when forming subjective (posterior) beliefs of a particular event. Well-rested subjects display a tendency to overweight the evidence in forming subjective posterior probability estimates, which is inconsistent with Bayes rule but possibly consistent with use of a 'representativeness' heuristic. In his original Bayes rule experiment, Grether (1980) also found that typical student-subjects overweighted the evidence relative to the prior odds in making posterior assessments. Ironically, behavior following sleep-deprivation is more consistent with the use of Bayes rule, because this treatment significantly reduces the (over)weight that subjects place on the new evidence. Because choice accuracy is *not* significantly affected by sleep deprivation, the significant difference in estimated decision-model parameters may indicate that the brain compensates under adversity in certain risky choice decision environments.

JEL Key Words: Bayes Rule, Uncertainty, Information, Experiments, Sleep.
JEL Codes: D81, D83, C91

*The authors wish to thank the U.S. National Academy of Sciences and the Japanese Society for the Promotion of Sciences for sponsoring the Japanese-American Beckman Frontiers of Science Symposium, which lead to this research collaboration. Helpful comments were also provided by seminar participants at Appalachian State University. The authors are particularly grateful to Henry J. Orff and Benjamin S. McKenna for their work on experiment administration and data management. This study was funded in part by the US Department of the Army award #DAMD17-02-1-0201 and NIH M01 RR00827. The US Army Medical Research Acquisition Activity is the awarding and administering acquisition office. The content herein does not necessarily reflect the position or policy of the Government, and no official endorsement should be inferred.

A large volume of evidence suggests that individuals in industrialized nations are becoming increasingly sleep-deprived. According to a recent poll conducted by the *National Sleep Foundation*, the average American adult slept less than 7 hours per night in 2005. The nightly average was 7.5 hours in 1975 and 9 hours per night in 1910. This trend towards less and less sleep has significant implications given the known effects of sleep deprivation: decreased motor and cognitive performance, reduced vigilance and reaction time, worsened mood, and reduced ability to think flexibly (Pilcher and Huffcutt, 1996; Harrison and Horne, 1999; Harrison and Horne, 2000). Indeed, even 7 hours of sleep per night leads to significantly diminished cognitive performance relative to 8 or 9 hours (Van Dongen, et al, 2003; Belenky, et al., 2003). Nearly 50 million Americans, close to 25% of all adults, are estimated to suffer from some level of sleep deprivation. Sixty percent of adults surveyed reported driving while drowsy, while 37% reported falling asleep or nodding off at some point while driving.[1] Estimates of the cost of lost U.S. worker productivity caused by sleep deprivation vary, but a conservative estimate—based on a 4% reduction in productivity for sleep-deprived working adults—is over $40 billion dollars annually (Stoller, 1997).

Many professions that give rise to more significant sleep deprivation as a matter of routine—emergency personnel, medical residents, military personnel, long-haul truck drivers—are also those where impaired functioning can put lives at risk. A study of long-haul truck drivers (Mitler, et al. 1997) in Canada and the U.S. found that they averaged only about 5 hours of sleep per night. A recent study of first- and second-year medical residency students found that two-thirds reported sleeping an average of six or less hours per night (Baldwin, et al. 2004). A smaller fraction (20%) averaged five hours of sleep a

[1] This data is reported by the *National Sleep Foundation*, and can be accessed at www.sleepfoundation.org.

night, and such residents were more likely to report, among other things, having made significant medical errors. Weinger and Ancoli-Israel (2002) concluded that sleep deprivation significantly impairs doctors' performance, thereby impacting patient safety, in part due to poor decisions made by sleep deprived physicians. Also, sleep deprivation has been considered at least partially responsible for several major historical disasters, including the Space Shuttle Challenger explosion, the Exxon Valdez oil spill, and the Chernobyl Nuclear plant explosion (Coren, 1996). In sum, the impact of sleep deprivation on society as a whole, while difficult to measure precisely, is massive.

This paper reports results from a laboratory study that examines the information processing abilities of subjects in a well-rested versus an experimentally sleep-deprived state. Much of the existing sleep deprivation research examines subject performance on sustained attention, mathematical and/or verbal tasks, such as simple reaction time tasks, arithmetic processing, grammatical reasoning, or verbal learning. Examinations of flexible thinking, strategy updating, and risk assessment are relatively new to sleep research (see references in Harrison and Horne, 2000). There is some evidence that complex or interesting tasks may be less likely to show deficits under TSD (e.g., Horne, 1988), but this remains a controversial proposition (Harrison and Horne, 1999; Pilcher and Huffcutt, 1992; Wimmer, et al, 1992).

Our present focus on Bayesian updating as a particular decision model is meant to examine the fundamentals of how information is processed by decision-makers. We examine differences in subjects' propensity to incorporate new information as they update prior probabilities to form posterior (subjective) probability estimates. A Bayes rule experiment is administered to subjects both well-rested and after 22-24 hours of total sleep deprivation (TSD). For comparison to existing economics research, we utilize the Bayes rule experiment presented in Grether (1980). His results indicate that subjects tend to overweight new evidence relative to prior odds when forming subjective beliefs. The result, further confirmed in Grether (1992), is largely due to subjects' tendencies to utilize a 'representativeness' heuristic in cases where new sample information looks representative of one population versus another (see, e.g., Kahneman and Tversky, 1972).[2]

The results from our pooled sample (well-rested and sleep-deprived data) are quite similar to those in Grether (1980)—subjects tend to weight new evidence more heavily than the prior odds, which is inconsistent with the Bayes rule prediction. This consistency is largely due to the well-rested subsample. Results from the TSD subsample reveal that subjects do *not* overweight the evidence relative to the prior odds of event occurrence. In fact, TSD subjects place less weight on both prior odds and evidence (and roughly equal weight on each), relative to when well-rested, in forming subjective posterior estimates. This change is not statistically significant, though, for the prior odds variable. Interestingly, the significant reduction in the decision weight placed on new evidence following TSD does not necessarily result in less accurate probability estimates. There is, however, some evidence that decision model error terms have higher variance in the TSD subsample, which implies somewhat less consistent behavior under TSD. Because information updating is a fundamental component of decision making, this unique examination of Bayesian updating following sleep deprivation is relevant to a large variety of behavioral applications.

[2] Grether (1992) indicates that the representativeness heuristic is used when available, and it is available a high proportion of the time in his earlier (1980) design. When not available as often in the (1992) experimental design, overweighting of the evidence is not borne out as a more general result.

2. Background

An examination of Bayesian updating under sleep deprivation contributes significantly to both the literatures in economics and sleep. A limited amount of sleep research indirectly points towards failed information assimilation under sleep deprivation (e.g., increased hesitance and reduced focus among sleep-deprived junior doctors in Goldman et al, 1972). However, direct evidence on decision making under uncertainty and information updating is needed, and Harrison and Horne (2000) recognize the lack of sleep deprivation research on specific decision models. Bayes rule is a fundamental decision model of belief revision and decision-making under uncertainty, and it has application to a variety of contexts. The relevance of this research to economists stems from our desire to understand decision-making behavior, and the evidence indicates that a good portion of decision-makers are, in their typical state, sleep-deprived to some degree. So, any identified differences in behavioral responses relevant to economic decision-making models highlight the significance of an individual's sleep-deprived state in behavioral analysis. Such differences also identify a previously uncontrolled confound in experimental data sets (for example, student subject experiments during exam periods may include relatively more sleep-deprived subjects).

Only a small amount of economics research has examined sleep. Biddle and Hammermesh (1990) incorporate labor productivity effects of sleep in a theoretical model of the allocation of time. Their empirical results from a variety of sources lead them to conclude that increased labor market time reduces sleep, as opposed to leisure activities. Their estimates from a system of demand equations indicate that higher wages reduce sleep—more so for men than women. This is consistent with the aggregate evidence on sleep reduction in many industrialized countries, and it implies that sleep deprivation will be an inevitable byproduct of wage growth in a society. Kamstra et al. (2000) examine the effects of daylight saving time changes on financial market returns. Interestingly, stock market returns drop both after losing an hour (Spring) *and* gaining an hour (Fall) of sleep. Reduced performance after gaining an hour of sleep can be attributed to what sleep researchers call 'desynchrony', or being out-of-sync with one's internal (biological) circadian rhythm. As a whole, sleep is largely unexplored by economists, and we believe that this research is a significant first step towards examining potentially important decision-making effects caused by reduced sleep.

Sleep deprivation can be either partial or total, where total sleep-deprivation implies no sleep at all during a given day(s) (i.e., one or more 24 hour periods). Intuition might suggest that total sleep deprivation impairs functioning more than partial sleep deprivation. If this were true then one might not feel as concerned about the average *partially* sleep deprived adult—a college student studying all night for an exam would be the exception. However, existing research indicates that there are just as many reasons to be concerned about the effects of partial sleep deprivation. Van Dongen et al. (2003) found that chronic *partial* sleep deprivation of 4 or 6 hours per night for as few as six consecutive nights resulted in significant deficits on cognitive performance. In fact, the deficits were *equivalent* to those from up to two nights of *total* sleep deprivation experienced by a separate treatment group. In other words, chronic partial sleep deprivation can cause performance deficits equivalent to those from 1-2 nights of zero sleep. And yet, partially sleep deprived subjects did not (subjectively) report feeling as sleepy as TSD subjects. Pilcher and Huffcutt (1996) also find that the average partial sleep deprivation study included in their meta-analysis reported evidence of significant performance and mood effects, and they note that these partial sleep deprivation effects have perhaps been underestimated in some narrative reviews of the sleep literature.[3]

Even when sleep deprivation might not affect some behavioral outcome measures, there is still much to understand about how underlying decision processes might be altered. Drummond et al. (2000) is an intriguing study that shows how versatile the brain can be under adversity. In their study, recognition memory on a verbal learning task showed *no* significant change as a result of a TSD treatment, though there was evidence that additional brain regions became activated following sleep deprivation. The subjects' parietal lobes, especially in the left hemisphere, came 'on-line' after total sleep deprivation. Because the parietal lobes are related to performance, their activation after TSD compensated for any decreased performance resulting from deficits in other brain regions. Others have reported similar increases in brain activation and resultant intact performance during TSD on a variety of tasks (Drummond et al., 2001, 2004, 2005; Portas et al, 1998; Chee and Choo, 2004). These results indicate that our understanding of decision-making under sleep deprivation is incomplete at best, and more exploration is needed even in cases where individuals apparently retain functional ability (i.e., cognitive function, information processing, etc). For the present paper we only examine behavioral outcomes (not neural outcomes). The evidence we find in support of distinct decision-

[3] In this paper, we focus on the effects of TSD on information processing, which are relatively understudied compared to cognitive tasks of various sorts. There is another family of effects of sleep deprivation, which includes decreased glucose metabolism, increased risk of obesity, and decreased release of growth hormone, among others.

weights across sleep-states may, however, be a clue indicating neural activation differences in such information-updating environments.

3) The Experiments

As noted, the experiments replicate the Grether (1980) design for a hand-run Bayes rule decision task, which we administer to one or two subjects at a time. Two bingo cages are each filled with six colored balls: Cage A is filled with four green and two red balls, and Cage B is filled with three red and three green balls. Six draws, with replacement, are to be made from one of the cages. Each subject was informed of a 'prior' probability of using Cage A in terms of a die roll. For example, a 1/3 prior odds of Cage A was implemented by informing the subject that Cage A would be used if the die roll was 1-2 (3-6 implied use of Cage B). Subjects did not see the actual die roll, but its result tells the experimenter from which Cage to make the six draws from behind an opaque divider. The subject was shown each draw from the bingo cage, and after six draws was asked to indicate whether the balls were drawn from Cage A or B. A correct cage response resulted in payment of $12, whereas an incorrect response paid $2.

The procedure—choose the cage, draw a sample of six balls, subject indicates cage used—was repeated six times, and subjects were informed that only one of these times would count for payment as determined by a random draw at the end of the experiment. The design was balanced across prior A odds of 1/3, 1/2, and 2/3.[4] This creates a reasonably high proportion of samples, on average, that yield three or four green balls out of six, which are samples 'representative' of Cage B or A, respectively. This allows for an examination of subjects' propensity to utilize a 'representativeness'

[4] One implementation accidentally utilized one instance each of the prior odds of 1/6 and 5/6.

heuristic, or rule-of-thumb, decision process in making their cage decision. Note that because compensation is higher when correctly indicating the cage, it is incentive compatible to indicate Cage A only if subjects perceive the (posterior) probability of Cage A to be greater than 50%.

If subjects use Bayes rule in their cage choice, they will form a posterior probability that Cage A is used from the particular sample of green/red balls drawn--

$$P(\text{Cage A} \,|\, \#\text{ green balls}) = \frac{\mathrm{P_A} \cdot P(\#\text{ green balls} \,|\, \text{Cage A})}{\mathrm{P_A} \cdot P(\#\text{ green balls} \,|\, \text{Cage A}) + \mathrm{P_B} \cdot P(\#\text{ green balls} \,|\, \text{Cage B})},$$

where $\mathrm{P_i}$ is the prior odds of the cage *i* being used. That is, the new sample information is used to update the prior probability of Cage A. Table 1 shows the Bayesian updated posterior odds of Cage A in this Grether (1980) design. If subjects use a representativeness heuristic to make their choice of which cage is used, then a sample draw of three or four green balls out of six will induce a choice of Cage B or A, respectively, simply because the sample drawn looks like the population of one of the cages. One can see from Table 1 that the use of the representativeness heuristic can lead to an incorrect cage choice. For example, the posterior probabilities indicate that Cage A is more likely when $\mathrm{P_A}$=2/3 and three green balls are drawn, but this sample looks like the Cage B population. Similarly, when $\mathrm{P_A}$=1/3 and four green balls are drawn, Bayesian updating would lead one to indicate that Cage B was used—the posterior probability of Cage A is less than 1/2. The design is balanced so that, on average, the proportion of representative samples should not bias accuracy in favor of (or against) Bayesian choices.

A total of 24 subjects were administered the Bayes rules experiment. These subjects participated in a total sleep deprivation study, which involved 6 consecutive nights and days in the Laboratory for Sleep and Chronobiology at the University of

California-San Diego. Subjects were compensated several hundred dollars for the entire stay at the lab, but it was made clear to the subjects that these experiments afforded the opportunity to earn additional payments that were unrelated to their fixed compensation. Lab staff generally indicated that the subjects were more engaged in these Bayes rule experiments than in other cognitive task experiments in which they participated during their lab stay, and so the extra compensation appeared salient to the subjects. Subjects were tested on various cognitive dimensions during their entire lab stay, with testing occurring approximately every two hours. This basic Bayes rule experiment was performed twice by each subject (so, they had the opportunity to earn $12 twice); once in a well-rested state, and once after 22-24 hours of total sleep deprivation. Each administration of the Bayes rule experiment lasted approximately thirty minutes.

Screening criteria for this study only allowed subjects who were right-handed, healthy, and considered 'normal' sleepers—those who had a consistent sleep-wake schedules that included 7-9 hours in bed each night. Subjects are indirectly monitored for one week prior to reporting to the sleep lab by keeping a sleep journal and wearing an actigraph.[5] Because we motivated the relevance of this research by indicating how common it is to *not* be a normal sleeper, one may question the external validity of using only normal sleepers. As experimentalists, however, we face the usual trade-off of internal control versus external validity in conducting a sleep deprivation study. Only by using otherwise normal sleepers can we be confident of having removed other confounds that may limit our ability to attribute treatment affects to sleep deprivation itself. During

[5] The actigraph measures wrist movement as a proxy of gross motor activity. This movement, in turn, is used to determine sleep and wake. These data verify that subjects are engaged in normal sleep patterns prior to their lab stay and are not partially sleep deprived at the beginning of the experiment. The complete list of experimental inclusion/exclusion criteria is fairly standard for sleep deprivation research, and they are available on request.

sleep deprivation, subjects were not allowed any sleep, not allowed stimulants of any sort, and they were under constant supervision by lab staff to ensure no sleep during this time. Figure 1 describes the basic timeline of the subjects' lab stay relative to their participation in these decision experiments.

In a more recent paper, Grether (1992) notes that there are limits to what can be gleaned from the data using his simpler 1980 design. Because the design favors generating samples that are representative of Cage A or B, we are somewhat limited in our ability to generalize towards instances in which new information is not necessarily representative. On the other hand, we chose the more simple design in order to present subjects the most straightforward decision task that involved prior and new-sample information. As stated above, this design also provides an efficient evaluation of the use of a representative heuristic compared to a Bayes rule in subjects' decision making. The dichotomous choice of Cage A or B does not allow us to infer strength of belief (i.e., 55% versus 95% certain that the balls came from Cage A), as does Grether (1992) in a modified design. However, given the known debilitating effects of sleep deprivation on vigilance, we felt this was a reasonable trade-off in design choice in order to be more assured that subjects understood the decision task, even after total sleep deprivation.

The particular placement of our Bayes rule task during the subjects' lab stay implies that all subjects complete their second Bayes rule decision task in their sleep deprived state.[6] As such, one might be concerned that subject learning may be generating some of the data. To explore this possibility, the Bayes rule experiment was also

[6] Due to an un-planned deviation from the sleep lab protocol for these experiments, one subject was administered the Bayes rule task under the TSD treatment first, in which case the coding of the TSD dummy variable distinguishes this one subject from the others. Ideally, the ordering of the TSD and well-rested administration would be counter-balanced but, as described above, the surrounding evidence does not indicate that subject learning is generating the TSD treatment effects.

administered to an additional eight control subjects (N=96 observations). These control subjects performed the Bayes rule decision task twice, at the same 22-24 hour interval, but the control subjects were well-rested in both instances. Decision model estimates for these control subjects (see Appendix) find no significant difference in the weight placed on the evidence during the second Bayes rule experiment—contrary to the main finding in the TSD data.

In other words, we find no evidence that the differences in decision-making we report in the next section are due to subject learning across the two administrations of the experiment. In addition, subject learning would imply that choice accuracy should be higher the second Bayes rule experiment, but it is not. Or, learning might imply that a particular empirical model should better fit the data as choices converge to a particular set of model parameters—Grether (1980) finds this among experienced subjects, for example. Our results also show that this is not the case. We are therefore confident in attributing the second trial effects to the sleep deprivation treatment.

4) Results

The data are from 24 subjects who ranged from 18 and 39 years of age, and each submitted voluntary consent for the total sleep deprivation study. Because each Bayes rule experiment involves 6 trials of the cage choice task, the total number of observations is N=144 in the well-rested state and N=144 in the sleep deprived state. The econometric estimations reported in this section account for the potential non-independence of decisions of a given subject across different trials as a subject-specific random effect.

Table 1 shows the posterior probabilities of Cage A, which imply posterior odds of either Cage A or B being more likely. For example, the posterior probability of Cage A of .584 indicates a posterior odds of Cage A of approximately 1.40:1. Certain prior odds and sample draws imply a relatively easier choice for the subject in the sense that the posterior odds of the more likely cage are quite high (e.g., if P_A=1/3 and only one green ball is drawn, the posterior odds of the more likely cage (Cage B in this case) are about 11:1. The bold cells in Table 1 highlight the sample possibilities for Cage A that lead to the most difficult choices among all possibilities. These highlighted cells represent all instances when posterior odds of the more likely choice are about 1.40:1 (some in favor of Cage A, some in favor of Cage B). Grether (1980) initially restricts attention to this subsample of data in order to compare choices of equal difficulty that include cases where the representativeness heuristic favors the right choice, cases where it favors the wrong choice, and cases where it provides no direction on cage choice.

Table 2 shows the summary data for this subsample of cases of relatively difficult subject choices. Interestingly, a breakdown of the TSD versus well-rested data indicates that, after TSD, subjects get a significantly higher proportion of responses correct when the representativeness heuristic favors the Bayesian updated cage choice (p=.10). When well-rested, a higher proportion of difficult choices are correct when representativeness is not available or at odds with the Bayesian updated choice, although these differences are not significant.[7] This subset of the data also shows evidence consistent with some of the sleep deprivation literature, which has found that performance (i.e., accuracy) does not

[7] We use a binomial test of the null hypothesis that the proportion of correct choices in TSD subsample is equal to the proportion of correct choices in the WR subsample. We avoid testing this hypothesis in case C, when the representativeness heuristic is unavailable, due to extremely low number of cases (N=2) in the WR data.

necessarily decline under TSD when the task is interesting and/or financially motivated (see Harrison and Horne, 2000). However, a simple look at the percentage of correct choices in Table 2 examines only a subsample of less than half of the total data. Furthermore, a model of the posterior probability estimates is necessary in order to identify any general difference in the decision model used by the subjects. Such a difference is implied if subjects apply compensatory effort following TSD.

A more complete analysis of subject choice is shown in Table 3. Here, following Grether (1980) for comparison, we estimate the following decision model:

$$Y_{it}^* = \alpha + \beta_1 \ln LR(A)_t + \beta_2 \ln\left(\frac{P_A}{1-P_A}\right)_t + \mu_i + \varepsilon_{it} \tag{1}$$

where Y_{it}^* is the subject i's subjective log odds in favor of Cage A in trial t, $LR(A)_t$ is the likelihood ratio for Cage A, and $\left(\frac{P_A}{1-P_A}\right)_t$ is the prior odds ratio for Cage A. The dichotomous variable Y_{it} is observed equal to 1 if $Y_{it}^* \geq 0$, and so we estimate (1) using a random effects probit estimation. Grether (1980) estimates logit results for this model, and does not account for subject-specific random effects, and so our econometric specifications are similar but not identical. A Bayes rule hypothesis amounts to testing jointly whether $\alpha=0$, $\beta_1=\beta_2>0$, while the representativeness heuristic would be supported if $\beta_1>\beta_2\geq 0$. In other words, a Bayesian subject will weight the evidence and the prior odds equally, while a subject who uses the representativeness heuristic would place more weight on the evidence than the prior odds of cage A. The basic findings of Grether (1980), who estimates a version of (1) as a logit model, support the representativeness heuristic hypothesis. That is, $\beta_1>\beta_2\geq 0$ for most of his subject groups, indicating that subjects overweight the evidence (i.e., the likelihood ratio) relative to the prior odds.

Table 3 shows our random effects probit estimation of model (1) for our subjects.[8] A test for structural change is performed on the data to test whether or not the same model parameters (α, β_1, and β_2) apply to the well-rested and TSD data. Using the likelihood ratio test on the restricted model of pooled data and the unrestricted models of the separate TSD=1 and TSD=0 subsamples, we reject the null hypothesis that a single set of model parameters applies to both sets of data (the chi-squared statistic=13.36—significant at the p=.01 level for the test of three restrictions). Thus, the results indicate a structural change in the parameter estimates following TSD, and so we next turn our focus to the model estimates for the separate well-rested and sleep-deprived subsamples. As noted earlier, results from additional control subjects do *not* support the hypothesis that the differences in the well-rested and TSD data are due to subject learning. The supporting evidence from these control subjects is given in detail in the Appendix.

One difference that stands out in Table 3 is that well-rested subjects place more weight on the evidence than the prior odds. This difference is statistically significant using the chi-squared test for the restriction that $\beta_1 = \beta_2$ (p=.06). When subjects are well-rested, the estimated decision model replicates a key result from Grether (1980) using the same basic experimental design. When sleep-deprived, however, there is no significant

[8] For comparison, we perform a *logit* estimation of the Grether (1980) model that is similar to (1) above, but does not include a random effects error-term specification. The pooled results that Grether reports for his financially motivated subjects yield the estimated model $Y_{it}= -.11+2.25*\ln LR(A)_{it}+1.82*P_A/(1-P_A)_{it}$, where α, β_1, and β_2 are statistically significant. In estimating the same logit model for our pooled data, the results are $Y_{it}= .04+2.26*\ln LR(A)_{it}+1.95*P_A/(1-P_A)_{it}$, with β_1 and β_2 being statistically significant (p=.00). So, our results are quite comparable to those reported in Grether (1980), and logit estimations of any of the models in this section are consistent with the results we find in the probit estimations that we report (logit estimation results available from the authors on request).

difference in the weight the subjects place on the prior odds versus the sample evidence (p=.91). Sleep deprivation reduces the weight the decision-maker places on the evidence relative to the prior odds. Ironically, the decision model under sleep deprivation is consistent with the Bayes rule hypothesis, because sleep deprivation apparently eliminates the overweighting that well-rested subjects tend to place on the evidence. In all cases, the models do a reasonably good job of predicting the Cage A and Cage B choices of the subjects, correctly predicting their choice between 83% and 85% of the time.[9]

The marginal effects are shown in Table 3 in addition to the coefficient estimates for interpretability. Consider the marginal effect on the log odds term, $\ln\left(\frac{P_A}{1-P_A}\right)_t$. With our particular experimental parameterization, this term increases by about one when comparing P_A=1/3 to P_A=2/3. So, the marginal effects of .54 and .34 for the well-rested and TSD data, respectively, imply that this increase in prior odds makes subjects 54% more likely to choose Cage A when well-rested, but only 34% more likely to choose Cage A when sleep-deprived. This difference between marginal effects on the log odds terms may not be statistically significant, however. Consider an alternative formulation for the pooled data set with a dummy variable for TSD=1, along with interaction terms

[9] An alternative model that Grether (1980) estimates includes dummy terms for samples that are representative of either Cage A or B. Our key results appear to hold under this alternative empirical model, although the model failed to converge properly for the well-rested subsample of data. Nevertheless, relative to the pooled data, the TSD sample estimates for weight placed on the prior odds and the evidence are both *less* that those estimated for the pooled data, and significant in both cases. Some evidence for use of the representativeness heuristic is found more specifically in this alternative estimation, though it is only significant for the case when the sample looks like Cage B—subject are then significantly *less* likely to choose Cage A.

$$Y_{it}^{*} = \alpha + \beta_1 \ln LR(A)_t + \beta_2 \ln\left(\frac{P_A}{1-P_A}\right)_t$$

(2)

$$+ \beta_3 * TSD_{it} + \beta_4(\ln LR(A)_t * TSD_{it}) + \beta_5\left(\ln\left(\frac{P_A}{1-P_A}\right)_t * TSD_{it}\right) + \mu_i + \varepsilon_{it}$$

We estimate this random effects probit specification to allow a more direct parameter estimate comparisons. The results are:

Parameter	α	β_1	β_2	β_3	β_4	β_5
***Marginal* effect**	.07	**.91**	**.58**	-.09	**-.56**	-.24
p-value (two-tailed test)	.64	.00***	.00***	.61	.00***	.23

These estimates are consistent with Table 3 results in showing that the tendency to significantly overweight the evidence ($\beta_1 > \beta_2$) is mitigated when the subject is sleep deprived ($\beta_4 < 0$). The coefficient on $\beta_5 < 0$ is in the direction indicating that sleep deprivation significantly reduces the weight one places on the prior odds, but the estimate is not statistically significant.[10]

The marginal effect on the evidence, ln*LR(A)*, term in Table 3 represents the marginal change to the probability of choosing cage A for a one-unit change in the log-likelihood ratio for Cage A. For our design, a sample of two green and four red balls, for example, generates a likelihood ratio of -1.05, while a sample of three green and three red balls generates ln*LR(A)*= -.353, which is an increase in ln*LR(A)* of about .70. The estimated marginal effect for well-rested subjects implies that this change in ln*LR(A)*

[10] For a similarly estimated logit model, β_5 significance is at p=.12. We also examine the relative difficulty of the different choices subjects would make, as proxied by the Bayesian posterior-odds of the more likely choice—higher odds represent an easier choice. As expected, we find that more difficult choices reduce the likelihood that subjects pick the correct cage. However, the TSD treatment does not significantly affect subject choice-accuracy, neither in general—noted earlier—nor for varying difficulty levels of choice, relative to when subjects are well-rested. These results are available from the authors on request.

from drawing one additional green ball would make subjects 57% more likely to choose Cage A. For sleep-deprived subjects the comparative marginal effect is only about 25%, and so the magnitude of the estimated effect is also meaningful. Of course, this does not take into account the fact that 'representative' samples may affect decisions independent of their effect on the likelihood ratio, but it is clear that these effects are behaviorally, as well as statistically, significant. The different sample draws in our experiment created a range of likelihood ratios from $\ln LRA(A)$= -2.50 for the case where six red balls were drawn, to $\ln LRA(A)$= 1.73 for the case where six green balls were drawn, though the extreme draws were rare.

These differences in the parameter estimates for the decision models when comparing subjects well-rested versus sleep-deprived are significant given that they indicate that TSD causes subjects to place a decreased decision-weight on new evidence and on prior odds. The estimated effect is significant in the case of the likelihood ratio (i.e., the evidence), and the effect is robust to model specification (compare sub-sample estimates in Table 3 with estimates of model (2)). We also estimate that TSD reduces the decision-weight that subjects place on the prior odds, though the effect is not as large in magnitude and did not reach statistical significance.

It is intriguing, however, that the accuracy of the subjects' choices is no worse when sleep-deprived than when well-rested, on average. For all N=144 observations of both well-rested and TSD data, subjects indicated the correct cage 67-68% of the time. To the extent that well-rested subjects have a tendency to *over*weight the evidence in the Bayes rule experiment, which would contribute to incorrect choices, sleep deprivation mitigates this tendency.[11] Sleep deprivation may also reduce the decision-weight placed on the prior odds, resulting in a net effect of unchanged choice accuracy. For the control subjects, choice accuracy drops for the *second* administration of the task, and so maintained choice accuracy following TSD may indicate compensatory effort of some sort. Research on sleep deprivation has found that performance deficits may *not* occur when the task is interesting or complex, or financially motivated (Harrison and Horne, 2000), although the underlying cognitive processes may be quite different (Drummond et al., 2000).

An examination of the residuals from estimating (1) indicate that the TSD sample yields somewhat higher-variance residuals, though the difference is not statistically significant (two-sample F-test for variance, p=.20). This may suggest that choices following TSD are not as convergent upon the decision model in (1) as when well-rested. Though our residuals-variance result is statistically insignificant, it is similar to Grether's (1980) finding of less consistent behavior for inexperienced subjects. Our lack of significance may be due to our limited sample size, but the result is consistent with results in the sleep literature that indicate increased variability and statistical variance under TSD.

It is also worth noting that our result of unaffected choice accuracy in the Bayes rule experiment following sleep deprivation only implies that subjects are equally accurate in assessing the likelihood of being in state A versus state B. This does not imply that a TSD subject is as adept at dealing with any further ramifications of being in

[11] Choices and accuracy are not consistent with random decisions. In the well-rested subsample, the actual Cage A frequency is 54.2%, and subjects chose Cage A 52.8% of the time (actual accuracy was 68.1%). In the TSD subsample, Cage A frequency was 43.8%, and Cage A choice occurred 46.5% of the time (67.4% accuracy).

one state versus the other. This latter consideration will also be a function of TSD effects on factors like vigilance and reaction time, e.g., that have well-established cognitive function effects on individuals. Furthermore, because the Bayes rule experiment does not allow subjects to sort themselves out of the uncertain choice environment, it is important to complement these research findings with an examination of preferences for risk. Such an examination is the topic of some of our related research.

5. Conclusions

The topic of sleep deprivation is virtually unexplored in research on economic decision models. Because of the evidence indicating that, as a society, we are more sleep deprived at present than in any previous generation, the implications this has on decision-making under uncertainty across many environments are worth exploring. Not only are the monetary costs of sleep-deprivation significant to an economy (e.g., lost worker productivity), but the implications of sleep-deprivation take on increased significance when one considers the public health/safety ramification of sleep-deprivation in certain susceptible professions (e.g., medical residency, long-haul truck driving, the military). Because recent sleep research indicates that performance of selected tasks may be just as affected under chronic partial sleep deprivation as under total sleep deprivation (Van Dongen et al., 2003), the effects of sleep deprivation on decision-making are not likely to be limited to only the short-term totally sleep deprived individual.

This paper examines the effects of sleep deprivation on a particular type of decision-making that is of interest to decision scientists, in general, and is unexplored by sleep researchers. We administer a Bayes rule decision experiment to twenty-four subjects in experimentally controlled well-rested and sleep-deprived states. Because the general population does not exactly fit either of these experimentally induced states, the results can be viewed as indicative of the decision processes of a given individual when approaching either the well-rested or TSD state. This decision experiment provides a fundamental look at how subjects process and filter information in uncertain choice environments. That is, a Bayesian subject is assumed to update a prior belief with new information on a situation in order to form a posterior belief of event occurrence. So, the experiment examines a basic decision model that may serve as a building block for many more complicated decision environments.

Our main result is that, following sleep deprivation, subjects decrease the decision-weights placed on new information and prior odds (i.e., prior information). This result is significant and robust to model specification for the estimated effect of TSD on the weight subjects place on new sample evidence. The result that subjects reduce the decision-weight placed on the prior odds following TSD is less conclusive. Nonetheless, the net effect is that sleep-deprived subjects behave more in accordance with Bayes rule, whereas well-rested subjects behave more non-Bayesian by *over*weighting the evidence relative to prior odds—Grether (1980) finds this overweighting of the evidence among a typical sample of student subjects in his design that we replicate. Using a simple outcome measure, we find that the choice accuracy following TSD is not significantly different than when well-rested, though there is some indication that decisions may be somewhat less consistent (i.e., a higher error term variance in the TSD sample estimates, though statistically insignificant at p=.20).

This experiment involves an unavoidable risky decision environment. The authors' related research shows evidence that sleep-deprived subjects are less risk-averse for gambles over monetary gains (Drummond, et al., 2006). As such, the present experiments do not capture a potential interesting secondary effect of sleep deprivation. Namely, a TSD individual may be less likely to *avoid* a risky decision environment, when the opportunity to sort oneself out of the decision exists. Sleep deprivation may therefore lead individuals to choose more risky decision environments, on average, where the cost of error is significantly higher. This may have interesting implications for, among others, military personnel choosing to engage or not engage in a riskier outcome scenario, or a physician choosing between two courses of surgical action.

Finally, because we find that subject choices are, on average, no less accurate in the Bayes rule experiment under TSD, the significant difference in the estimated parameters of the decision model merit further exploration. These results may indicate an important difference in how the brain processes information in different states, or they may indicate compensatory neural activation that is not captured by the parameters of our econometric model.[12] Drummond et al. (2000), for example, uses neuroimaging to document compensatory brain activity following TSD for subjects engaged in a verbal learning task, and other research further supports the compensatory activation hypothesis (e.g., Drummond et al., 2001, 2004, 2005; Portas et al, 1998; Chee and Choo, 2004). So, although behavioral outcomes, such as choice accuracy, may be unaffected following TSD, neural activation differences may provide important clues as to how the brain functions under adverse conditions. Our finding of significant differences in the parameter estimates of a simple decision model applied to subjects both well-rested and

[12] The hypothesis of compensatory activation following TSD is further supported by our finding that choice accuracy in our well-rested control subjects is actually lower during the second administration of the experiment, though subjects are still well-rested.

following TSD may be an initial indication of compensatory neural activity that we intend to explore further.

Table 1: Posterior probabilities of Cage A

Prior probability of Cage A	Number of Green Balls Drawn						
	0	**1**	**2**	**3**	**4**	**5**	**6**
2/3	.149	.260	**.413**	**.584**	.737	.849	.918
1/2	.081	.149	.260	**.413**	**.584**	.737	.849
1/3	.042	.081	.149	.260	**.413**	**.584**	.737

Table replicated from Grether (1980) Table 1. Bold cells represent approximately equal posterior odds of the more likely Cage (i.e., choices of approximately equal difficulty for subjects)

Table 2
Proportion correct by sample type when posterior odds are approximately 1.40:1 (subsample of data: $N_{well\text{-}rested}$=60, N_{TSD}=62)

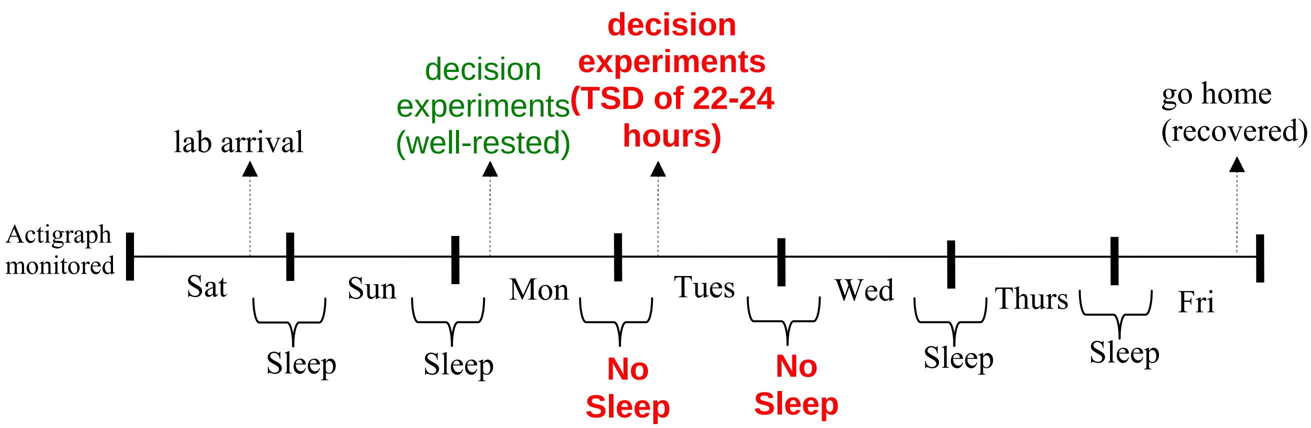
Well-rested			Sleep Deprived (TSD)		
A[a]	**B[b]**	**C[c]**	**A[a]**	**B[b]**	**C[c]**
.52 (N=27)	.58 (N=31)	1.00 (N=2)	.67 (N=24)	.52 (N=25)	.46 (N=13)
Weighted average=.57			Weighted average=.57		

a. Representativeness heuristic favors Bayesian updated cage choice
b. Representativeness heuristic does not favor the Bayesian updated cage choice
c. Representativeness heuristic not available.

Table 3: Probit estimates of Y_{it}^{*} model
(random effects specification. p-values given in parenthesis)

Variable	Pooled (N=288) Coeff.	Pooled (N=288) marg. effect	Well-rested (N=144) Coeff.	Well-rested (N=144) marg. effect	Sleep-deprived (N=144) Coeff.	Sleep-deprived (N=144) marg. effect
Constant	.03 (.83)	.01 (.83)	.13 (.42)	.05 (.42)	-.08 (.68)	-.03 (.68)
ln*LR(A)*	1.27 (.00)***	.48 (.00)***	2.20 (.00)***	.81 (.00)***	1.01 (.00)***	.36 (.00)***
$\ln\left(\frac{\boldsymbol{P_A}}{1-\boldsymbol{P_A}}\right)$	1.10 (.00)***	.42 (.00)***	1.46 (.00)***	.54 (.00)***	.97 (.00)***	.34 (.00)***
% correctly predicted	84.38%		85.42%		83.33%	

FIGURE 1
A week in the sleep lab: time-line



Note: Some subjects stayed in the lab one less day and participated in a one-night of TSD study. Our examination of TSD effects after one night of TSD allows us to combined subjects from different sleep studies, whether or not they participated in a one or two night TSD lab stay.

REFERENCES

Baldwin, DeWitt C. Jr., and Steven, R. Daugherty. 2004. "Sleep deprivation and fatigue in residency training: Results of a national survey of first- and second-year residents." *Sleep*, 27(2): 217-23.

Biddle, Jeff E., and Daniel S. Hamermesh. 1990. "Sleep and the allocation of time." *Journal of Political Economy*, 98(5, part 1): 922-43.

Chee, M. W., and W. C. Choo. 2004. "Functional imaging of working memory after 24 hr of total sleep deprivation." *Journal of Neuroscience*, 24: 4560-4567.

Coren, Stanley. 1996. *Sleep Thieves*. New York: Free Press.

Drummond, Sean P.A., Gregory G. Brown, J. Christian Gillin, John L. Stricker, Eric C. Wong, and Richard B. Buxton. 2000. "Altered brain response to verbal learning following sleep deprivation." *Nature*, 403(Feb 10): 655-57.

Drummond, Sean P.A., Gregory G. Brown, J. S. Salamat, and J. Christian Gillin. 2004. "Increasing task difficulty facilitates the cerebral compensatory response to total sleep deprivation." *Sleep*, 27: 445-451.

Drummond, Sean P.A., M. J. Meloy, M. A. Yanagi, Henry J. Orff, and Gregory G. Brown. 2005. "Compensatory recruitment after sleep deprivation and the relationship with performance." *Psychiatry Research Neuroimaging*.

Drummond, Sean P. A., J. Christian Gillin, and Gregory G. Brown. 2001. "Increased cerebral response during a divided attention task following sleep deprivation." *Journal of Sleep Research*, 10: 85-92.

Goldman, L.T., McDonough, M.T., and G.P. Rosemond. 1972. "Stresses affecting surgical performance and learning." *Journal of Surgical Research*, 12: 83-86.

Grether, David M. 1980. "Bayes Rule as a descriptive model: The representativeness heuristic." *Quarterly Journal of Economics*, 95(3): 537-557.

Grether, David M. 1992. "Testing Bayes rule and the representativeness heuristic: Some experimental evidence." *Journal of Economic Behavior and Organization*, 17: 31-57.

Harrison, Yvonne, and James A. Horne. 2000. "The impact of sleep deprivation on decision making: A review." *Journal of Experimental Psychology: Applied*, 6(3): 236-49.

Harrison, Yvonne, and James A. Horne. 1999. "One night of sleep loss impairs innovative thinking and flexible decision making." *Organizational Behavior and Human Decision Processes*, 78(2): 128-45.

Horne, James A. 1988. "Sleep deprivation and divergent thinking ability." *Sleep*, 11: 528-36.

Kahneman, Daniel., and Amos Tversky. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica*, 47(2): 263-91.

Kahneman, Daniel., and Amos Tversky. 1972. "Subjective probability: A judgment of representativeness." *Cognitive Psychology*, 3: 430-54.

Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi. 2000. "Losing Sleep at the Market: The daylight saving anomaly.' *American Economic Review*, 90(4): 1005-11.

Mitler, Merrill M., James C. Miller, Jeffrey J. Lipsitz, James K. Walsh, and C. Dennis Wylie. 1997. "The sleep of long-haul truck drivers." *New England Journal of Medicine*, 337(11): 755-61.

Pilcher, June J. and Allen I. Huffcutt. 1996. "Effects of sleep deprivation on performance: A meta-analysis." *Sleep*, 19(4): 318-26.

Portas, C. M., G. Rees, A. M. Howseman, O. Josephs, R. Turner, and C. D. Frith. 1998. "A specific role for the thalamus in mediating the interaction of attention and arousal in humans." *Journal of Neuroscience*, 18: 8979-8989.

Stoller, M.K. 1997. "The socio-economics of insomnia: The materials and methods." *European Psychiatry*, 12(Suppl 1): 41s-48s.

Tversky, Amos., and Daniel Kahneman. 1991. "Loss aversion in riskless choice: A reference-dependent model." *Quarterly Journal of Economics*, 106(4): 1039-61.

Van Dongen, Hans P.A., Greg Maislin, Janet M. Mullington, and David F. Dinges. 2003. "The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation." *Sleep*, 26(2): 117-26.

Weinger, M.B., and S. Ancoli-Israel. 2002. "Sleep deprivation and clinical performance." *Journal of the American Medical Association*, **287**(8): 955-7.

Wimmer, F., R. F. Hoffmann, R. A. Bonato, and A. R. Moffitt. 1992. "The effects of sleep deprivation on divergent thinking and attention processes." *Journal of Sleep Research*, 1: 223-230.

Appendix: Control Subject Data

The experimental protocol was administered to an additional eight subjects, who were well-rested for both the first and second administration of the Bayes rule experiment—well-rested was verified using similar measures as for the sleep deprivation subjects. The results of estimation equation (1) from the text for the sample of N=96 Bayes rule decisions are shown in Table A1 below.

Table A1: Probit estimates of Y_{it}^* model for CONTROL SUBJECTS (random effects specification. p-values given in parenthesis)

Variable	Pooled (N=96) Coeff.	marg. effect
Constant	.10 (.61)	.03 (.61)
ln*LR(A)*	1.75 (.00)***	.61 (.00)**
$\ln\left(\frac{\boldsymbol{P_A}}{1-\boldsymbol{P_A}}\right)$	1.97 (.00)***	.69 (.00)***
% correctly predicted	87.50%	

As can be seen, results are similar to those from the main data set, except that the estimated weights on evidence and prior odds are somewhat higher. Estimation up to an unknown scale parameter, however, prohibits a direct comparison across models. A test for structural change in the data across the first- and second-administration would require estimation of two small subsamples of just N=48 observations. Rather, we estimate a model similar to (2) in the text. That is, the pooled control-subject data is analyzed with dummy variables for second-administration of the experiment, with interaction terms that allow for the second-administration effects to potentially differ with respect to decision weights on evidence and prior odds. Thus, we estimate

$$\begin{aligned} Y_{it}^* = \alpha &+ \beta_1 \ln LR(A)_t + \beta_2 \ln\left(\frac{P_A}{1-P_A}\right)_t + \beta_3 * 2\text{ndAdmin}_{it} + \\ &\beta_4(\ln LR(A)_t * 2\text{ndAdmin}_{it}) + \beta_5\left(\ln\left(\frac{P_A}{1-P_A}\right)_t * 2\text{ndAdmin}_{it}\right) + \mu_i + \varepsilon_{it} \end{aligned} \quad (1A)$$

The results are

Parameter	α	β_1	β_2	β_3	β_4	β_5
Marginal **effect**	-.001	**.46**	**.64**	-.64	.64	-.25
p-value (two-tailed test)	.99	.00***	.00***	.23	.17	.48

So, the only significant variables are the prior odds and the evidence (we fail to reject the null hypothesis that $\beta_1=\beta_2$, p=.24). The second administration of the Bayes rule task does not significantly affect the likelihood of choosing Cage A. If anything, one could say that β_4 approaches the *opposite* significance compared to the TSD subjects. Though statistically insignificant, $\beta_4>0$ would indicate that control subjects place *more* weight on the evidence in the second trial administration, which would only strengthen our TSD findings.

Alternatively, one could examine the entire pooled data set, regular and control subjects, and include distinct dummy and interactions terms for the *Control* subjects and second administration (2*ndAdmin*) effects in order to examine if their Y_{it}^* decision model differs from that of the regular subjects. For the pooled N=384 observations, we estimate

$$Y_{it}^* = \alpha \; + \beta_1 \ln LR(A)_t + \beta_2 \ln\left(\frac{P_A}{1-P_A}\right)_t + \beta_3 * TSD_{it} + \beta_4 (\ln LR(A)_t * TSD_{it}) +$$

(2A) $$\beta_5\left(\ln\left(\frac{P_A}{1-P_A}\right)_t * TSD_{it}\right) + \beta_6 * Control_{it} + \beta_7 * \text{2ndAdmin}_{it} +$$

$$\beta_8 (\ln LR(A)_t * \text{2ndAdmin}_{it}) + \beta_9 \left(\ln\left(\frac{P_A}{1-P_A}\right)_t * \text{2ndAdmin}_{it}\right)\mu_i + \varepsilon_{it}$$

The results are

Parameter	α	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
***Marginal* effect**	.05	**.69**	**.62**	-.07	**-.33**	-.27	-.05	-.59	.60	.42
p-value (two-tailed test)	.61	.00***	.00***	.51	.01***	.17	.81	.59	.59	.54

These results indicate that there are no significant differences in the control subjects, except that their decision weights for Y_{it}^* are no different when performing the Bayes task the first or the second time (compare with results from estimating (2) in the text). Additionally, the control subjects were correct in their cage choice 75% of the time during the first administration of the experiment, but only 56% of the time in the second administration. In sum, the evidence is in support of our conclusions that the sleep deprivation treatment, not learning, is generating the behavioral differences we estimate in the main text. It is also supportive of a hypothesis that there is compensatory effort engaged following sleep deprivation that helps maintain choice accuracy.