

Dickinson, David L. and Todd McElroy (2010) "Rationality Around the Clock: Sleep and Time-of-Day Effects on Guessing Game Responses." **Economics Letters**, 108(2): 245-248. DOI: 10.1016/j.econlet.2010.04.029 Published by Elsevier (ISSN: 0165-1765)

## **Rationality Around the Clock: Sleep and Time-of-Day Effects on Guessing Game Responses**

**David L. Dickinson and Todd McElroy**

### **ABSTRACT**

We administer an online Guessing Game collecting responses across all 24 h of the day. While time-of-day itself does not affect guesses, when including trait-level sleepiness and previous night sleep, adverse sleep states lead to responses significantly farther from equilibrium.

### **ARTICLE**

#### **1. Introduction**

In industrialized countries, a significant public health concern has been the erosion of good sleep habits. At least 25% of adults are considered chronically sleep deprived<sup>1</sup>, and our 24/7 society places increasing demands on individuals at irregular times of the day. Travel schedules, shift work, and poor sleep priorities can all deplete cognitive resources and together may negatively impact rational decision-making. Sleep deprivation has been estimated to cost the U.S. economy \$40 billion annually in lost worker productivity (Stoller, 1997), and public safety is at risk in some professions where sleep loss and irregular scheduling are common (e.g., medical residency and air traffic control). Nevertheless, the behavioral effects of adverse sleep states have been largely ignored by economists. Current sleep research is now examining which components of decision-making are most affected by adverse sleep states, but there is a lack of research on decision tasks of interest to economists.

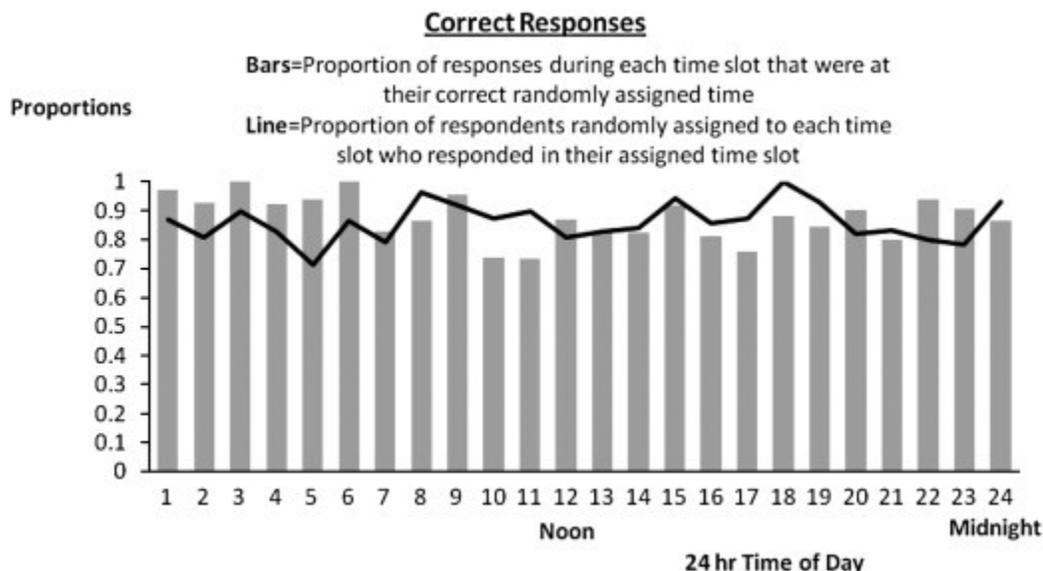
This study explores sleep effects in the p-beauty contest, or "guessing game" (Nagel, 1995). This well-known game is a building block of more complicated environments where iterative reasoning skills are engaged. Decision-making that requires iterative reasoning and anticipation of others' behavior is a high-level skill at risk in sleepy decision-makers. We report results from a unique experimental design in which subjects

are recruited to complete an online survey within a randomly assigned one-hour time period, and a one-shot Guessing Game is implemented with real financial incentives within the survey. Responses are time-stamped for compliance, and other survey questions yield measures of the respondent's sleepiness and previous night's sleep. Sample selection in our experimental design implies that we likely have a conservative estimate of true behavioral effects. Our results show that adverse sleep attributes lead to Guessing Game responses consistent with lower levels of iterative reasoning, and these behavioral effects grow in magnitude and significance for individuals with multiple adverse sleep attributes.

## 2. Experiments

Names from student email lists were randomly assigned to a one-hour response window on either Tuesday, Wednesday, or Thursday of the experiment week. Recruitment emails highlighted the incentives, the survey link, and the randomly assigned response time slot. Fig. 1 shows that most subjects responded during their assigned time slot. A random prize drawing of \$100 (\$300 for midnight to 8 a.m. time slots) was used as an incentive for completing the survey within the assigned response window. The initial recruitment email highlighted that one survey question would also offer the chance to win an additional \$50. The survey question itself also highlighted the additional incentive:

“Please submit an integer number (no decimals) between 0 and 1000 (including 0 and 1000 as possibilities). All survey respondents this semester will be making this same choice. The winner is the person whose chosen number is closest to  $\frac{2}{3}$  of the average number submitted. The winning prize is \$50.(in the event of a tie, the prize money will be equally shared).”



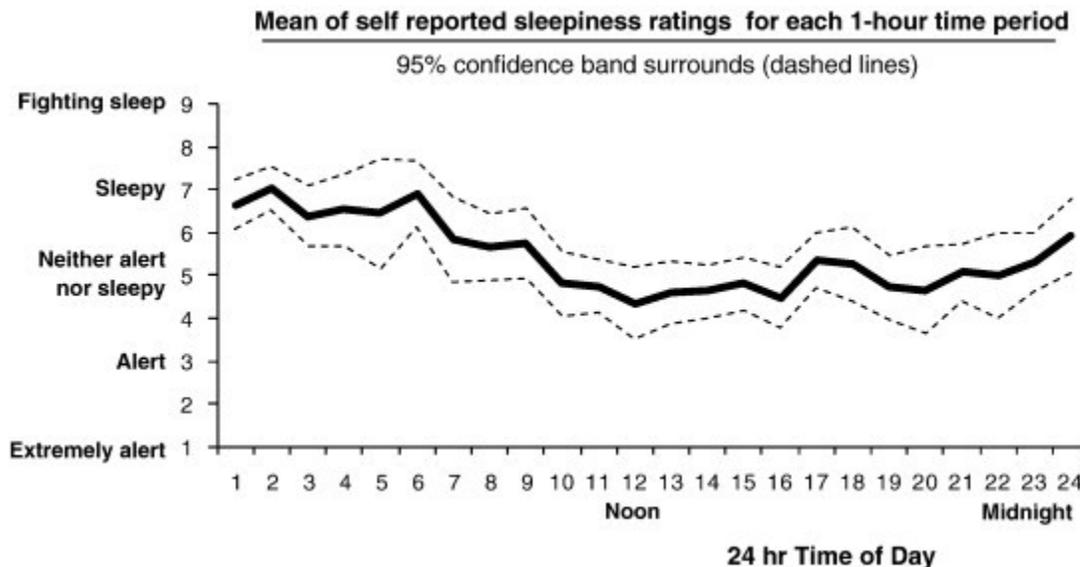
**Fig. 1. Summary of survey responses according to randomly assigned time slot.**

The equilibrium guess, 0, is attained by eliminating dominated responses, assuming others do so as well, and iterating this process infinitely. For example, all responses  $> 667 = 2/3 \cdot 1000$  are dominated and thus eliminated, but a second iteration would assume no one should guess above  $445 = 2/3 \cdot 667$ , and so on. While the survey instructions asked subjects to complete the survey themselves with no aids, we have no way of verifying compliance. However, one of the survey questions asked subjects to indicate the freezing point of vodka ( $-16.51^\circ\text{F}$  for 80-proof vodka). The answer is easily found with a Google search, yet only about 3.5% of all subjects answered correctly. Thus, we feel somewhat confident that subjects respected our survey rules.

### 3. Results

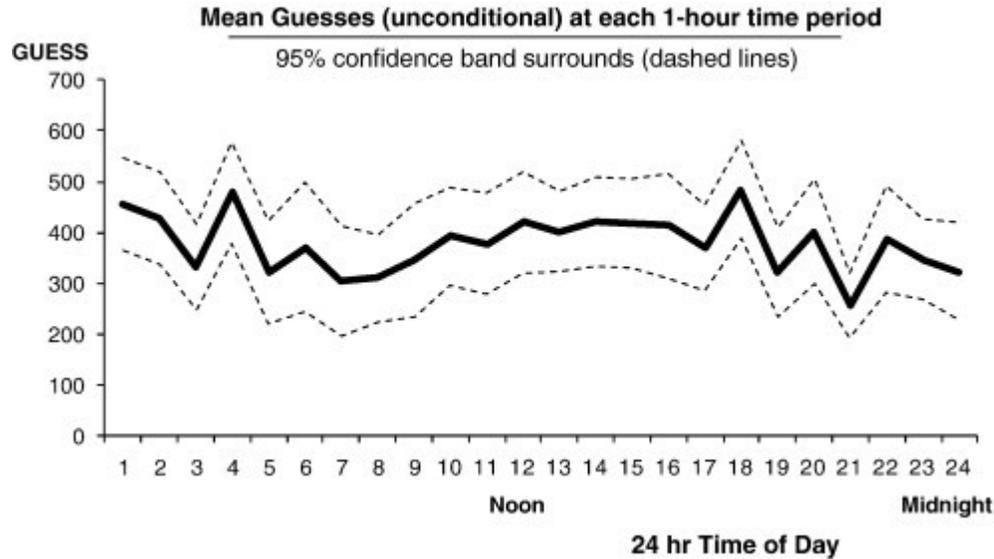
A total of 684 subjects completed the survey, with an average 28.5 responses per each 1-hour time slot of the day (23.5 responses per time slot from 1 a.m.–5 a.m.). Self-reported state-level sleepiness (i.e., responses to “How sleepy do you feel right now?”) in Fig. 2 is generally higher during the night, as expected, even though subjects self-select into the experiment. Fig. 3 shows mean unconditional Guesses during each 1-hour response window.

-----



**Fig. 2. State-level (Karolinska) Sleepiness Scores by time-of-day.**

-----



**Fig. 3. Guesses by time-of-day.**

The key independent variables used to examine Guess levels are the following: previous night sleep, trait-level sleepiness, and good/bad time-of-day. The mean previous night sleep level was 6.5 h (responses were in intervals (4–5 h, 5–6 h, etc.), and coded at the midpoint). We define sleep deprivation as  $SD = 1$  if previous night sleep was less than 6 h and  $SD = 0$  if previous night sleep was greater than 7 h (discarding responses in the interval “6–7 h”, given the greater subjectivity in classifying such average responses as either  $SD = 0$  or  $SD = 1$ ). One's propensity to fall asleep during the day (i.e., “trait-level” sleepiness) is generated from a series of 7 questions with with 0–3 response scales. We define the dummy variable  $TSleepy = 1$  for scores greater than 10, and  $TSleepy = 0$  for scores less than 10 (again, discarding scores of the arbitrary cutoff value of 10).<sup>2</sup> Finally, time-of-day is split into times associated with higher versus lower alertness ratings as established in the literature (see Smith et al., 2002). We therefore score  $BadTime = 1$  for response times from 1:00 to 5:00 a.m., and  $BadTime = 0$  for noon-7 p.m. response times. A correlation of +.18 exist between  $TSleepy$  and  $SD$  (Spearman,  $p = .01$ ) using these definitions. The Appendix shows the relevant online survey questions and Guessing Game instructions that we used to generate our dependent and independent variables.<sup>3</sup>

We proceed using these categorical regressors to examine time-of-day, trait-level sleepiness, and previous night sleep. Preliminary regressions retaining the semi-continuous nature of trait-level sleepiness and previous night sleep variables show they lack predictive power, which may not be surprising given that these self-reported variables likely contain noise. Our results (below) using categorical variables indicate that there are likely important tipping points before these explanatory variables affect behavior.

Table 1 reports mean levels of Guess by the different sleep categories. Two-sample t-tests of mean Guess levels in single-condition comparisons show that  $\text{Guess}_{SD = 1} > \text{Guess}_{SD = 0}$  ( $p = .07$ : one-tailed test), while we cannot reject the hypothesis that mean guess levels are identical for subgroups of  $\text{Badtime} = 0$  v.  $\text{Badtime} = 1$  or  $\text{TSleepy} = 0$  v.  $\text{TSleepy} = 1$ . This result is consistent with recent results in Dickinson and McElroy, 2009 D.L. Dickinson and T. McElroy, Naturally-occurring sleep choice and time of day effects on p-beauty contest outcomes, Working paper, Appalachian State University #09-03 (2009). Dickinson and McElroy (2009), which reported that objectively measured voluntary sleep loss lowers the estimated level of iterative reasoning in the Guessing Game. We find the same basic result here with our SD variable, though in a distinct online (one-shot) experiment with self-reported sleep levels.

**Table 1.**  
**Mean guesses by sleep condition.**

<b>Bad</b>	<b>Mean value</b>	<b>Good</b>	<b>Mean value</b>	<b>t-value (Good–Bad guess difference)</b>
<i>Single conditions comparisons</i>				
Badtime ( $n = 94$ )	397.41	Goodtime ( $n = 242$ )	402.08	-0.16
TSleepy ( $n = 151$ )	398.25	Not TSleepy ( $n = 532$ )	373.83	1.05
Last night Sleep-Deprived (SD) ( $n = 234$ )	396.29	Last night Well-rested (WR) ( $n = 259$ )	363.29	1.47*
<i>Double-condition comparisons</i>				
Badtime and TSleepy ( $n = 22$ )	477.64	Goodtime and Not TSleepy ( $n = 154$ )	394.81	1.34*
Badtime and Last night SD ( $n = 35$ )	443.57	Goodtime and Last night WR ( $n = 95$ )	368.08	1.43*
TSleepy and Last night SD ( $n = 64$ )	414.00	Not TSleepy and Last night WR ( $n = 216$ )	360.82	1.45*
<i>Triple-condition comparisons</i>				
Badtime and TSleepy and Last Night SD ( $n = 12$ )	530.08	Goodtime and Not TSleepy and Last night WR ( $n = 77$ )	364.39	1.75**

\*,\*\* Indicate one-tailed statistical significance at the  $p = .10$  and  $p = .05$  levels, respectively.

For “Double-Condition” comparisons, we compare best-case to worst-case scenarios. For example, the first double-condition comparison examines the subset of data where both Badtime = 1 and TSleepy = 1 to the subset of data where Badtime = 0 and TSleepy = 0. These comparisons show that any pair of “bad” sleep attributes generates Guess levels significantly farther from equilibrium compared to an individual with neither of the “bad” conditions.<sup>4</sup> Though neither BadTime nor TSleepy individually affect Guess levels, an individual with both TSleepy = 1 and BadTime = 1 has mean Guess level farther from equilibrium. Similarly, for the triple-conditions case, Guesses for those rare individuals with all three bad sleep attributes (SD = 1, Badtime = 1, and TSleepy = 1, n = 12) are significantly farther from equilibrium compared to subjects having all three good sleep attributes (SD = 0, Badtime = 0, and TSleepy = 0). As bad sleep conditions accumulate from the double- to triple-bad condition, the magnitude of the behavioral effect grows as well.<sup>5</sup> Thus, our key result is that subjects with more adverse sleep attributes submit guesses consistent with lower levels of iterative reasoning, even in a largely uncontrolled experimental environment.

#### **4. Conclusions**

How sleep choices and time of day affect higher level thinking is of particular importance in modern societies that seem to continually constrain individuals to perform under adverse sleep and time-of-day conditions. Our results show that responses in the Guessing Game are farther from equilibrium if one slept less than 6 h the previous night (compared to more than 7 h) or if several adverse sleep attributes are present. While non-equilibrium responses in this Guessing Game may result from bounded rationality or expectations that others are not rational, more recent research on 2-person Guessing Games concludes that even when the influence of expectations is removed, rationality is still bounded (Grosskopf and Nagel, 2008). Thus, a reasonable interpretation of our results is that sleep loss harms this type of rationality (i.e., iterative reasoning and anticipation). Moreover, we find that a cocktail of several adverse sleep states magnifies the extent of this effect. So while a middle-of-the night decision may not, by itself, significantly affect behavior, when added to one already prone to sleepiness or with little sleep the previous night (or both), the effects on higher level thinking are shown to accumulate.

#### **Acknowledgements**

The authors wish to thank Mike McKee for helpful comments on this manuscript. We also thank the Appalachian State University for providing funding for this research.

## References

Dickinson and McElroy, 2009 D.L. Dickinson and T. McElroy, Naturally-occurring sleep choice and time of day effects on p-beauty contest outcomes, Working paper, Appalachian State University #09-03 (2009).

Grosskopf and Nagel, 2008 B. Grosskopf and R. Nagel, The two-person beauty contest, *Games and Economic Behavior* 62 (2008), pp. 93–99.

Johns, 1991 M.W. Johns, A new method for measuring daytime sleepiness: the Epworth sleepiness scale., *Sleep* 14 (6) (1991), pp. 540–545.

Nagel, 1995 R. Nagel, Unraveling in guessing games: an experimental study, *American Economic Review* 85 (5) (1995), pp. 1313–1326.

Smith et al., 2002 C.S. Smith, S. Folkard, R.A. Schmieder, L.F. Parra, E. Spelten and H. Almirall et al., Investigation of morning–evening orientation in six countries using the preferences scale, *Personality and Individual Differences* 32 (2002), pp. 949–968.

Stoller, 1997 M.K. Stoller, The socio-economics of insomnia: the materials and methods, *European Psychiatry* 12 (Suppl 1) (1997), pp. 41s–482s.

## Notes

1 Based on National Sleep Foundation survey data. These survey data report average nightly sleep levels that are consistent with the authors' objectively measured average nightly sleep levels (see Dickinson and McElroy, 2009).

2 An alternative is to define TSleepy based on splitting the sample at mean response = 8, which produces qualitatively similar results. Trait-level sleepiness questions in the survey are from the well-known Epworth Sleepiness Scale (Johns, 1991), which is comprised of 8 questions/scenarios. We inadvertently omitted one of the scenarios in our question set, leaving us with 7 scenarios.

3 Online Appendix available at [http://www.appstate.edu/~dickinsondl/pdfs/EL28502\\_APPENDIX.pdf](http://www.appstate.edu/~dickinsondl/pdfs/EL28502_APPENDIX.pdf) , which also includes the email text used for recruitment.

4 This analysis, which is akin to split-sample analysis, is different than analyzing Guess for the interaction variables = 0 v. = 1 (where the interaction may equal 0 if either dummy variable equals zero). These key results remain unchanged if one uses dummy variable interaction terms instead.

5 Combined conditions are also a useful way to combat the subject response error likely present in TSleepy and SD.