Retentive Vacillation: Accounting for Variability in Human Memory

Eash V. Aggarwal

University of Pennsylvania

Abstract

While there have been several studies conducted to account for the variability in memory that occurs in a single person over a substantial extent of time, there are no significant data that aim to account for fluctuations in memory ability that occur in relatively shorter time spans. This study was designed to determine if short-term variability in human memory (within minutes or hours) can be computationally accounted for by external factors, or if there simply exist certain endogenous states of the brain that are better suited for memory than others, and the brain internally fluctuates between these favorable and unfavorable states in a given amount of time. Participants were repeatedly situated in a room devoid of external stimuli, distractions, and agitations while completing a ninety-minute free recall experiment to measure their changes in memory ability throughout the task. This variance was then computationally analyzed to determine if a model of explanatory variables could be generated to account for a significant portion of its existence. It was predicted that much of this inconsistency in memory would be inexplicable though the remaining external variables not controlled for by the experiment, thus resulting in the conclusion that the portions of the brain responsible for memory oscillate among internal states that are either more or less effective than others, regardless of changes in the environment. This research will allow for the advancement of memory revitalization therapies in patients with neurodegenerative diseases, and will provide new insights into the study of human memory systems.

Keywords: variability; memory; free recall; predictive model; episodic memory

Personal Statement

Without delving too deeply into all of the philosophical aspects of my interests, the human brain is a subject that fascinates me greatly due to its extreme complexity and seemingly impossible functions. The brain is who each of us is, and controls everything about us. From a broad viewpoint, understanding the brain is the key to understanding aspects of humanity that seem inexplicable alone.

How does a high school student acquire such an intense passion for a subject like neuroscience? Before I had discovered my interest in the brain, I was doing what almost every high school student does to some degree: showing up to classes and studying for my tests to get good grades. It was not until I took AP Psychology in my junior year of high school that I discovered a genuine interest in the subject material *beyond* what was written in the textbook. Soon, I found myself looking up neuroscience articles about the brain and staying after class to ask my teacher questions about material that was far outside the syllabus.

Meanwhile, I was participating in a program through the Fox Chase Cancer Center called the TRIP Initiative (a name that I am proud to take credit for), doing research about olfactory responses in cancerous fruit flies. This program helped me expand my practical science experiences, as I felt the desire to find a more application-based environment in which to learn science than a high school classroom. Here, I found in myself a passion for research as I stood proudly in front of a symposium of scientists and presented the findings of my research.

Fueled by the combination of my newfound interest in the brain and my passion for research, I decided to seek out a research opportunity in neuroscience. In reading about neuroscience research, I came across the Computational Memory Lab at UPenn. The computational means of brain research intrigued me, as I already possessed a programming background, so I personally contacted the PI to inquire about in internship over the summer. While the lab does not normally accept high school students, he was particularly impressed with my qualifications and interview, and was happy to invite me into his lab.

When I began my research project in the lab, I had taken AP Statistics and AP Calculus AB, scoring a 5 on both exams. Little did I know that I would have to learn almost all of the statistics and math I used in my project for the first time on my own. From multiple regression models and one-sample t-tests to confidence intervals and standard deviations, the statistics I used to complete this project were either completely new to me or never practiced enough for me to use. Therefore, I spent a lot of time on the internet learning or asking grad students for their help running statistical analyses of my data.

While I considered math to be fairly fun (as nerdy as that sounds), I had never really taken an interest in the subject because I was simply never convinced of the real world applications for the concepts we learned. However, in completing this research project, I found that I had to use my calculus skills in creating geometric and arithmetic series equations to computationally loop through giant datasets. Statistics came alive for me when the coefficient of determination of a multiple regression model I created helped me make conclusions about the effects that sleep and time of day have on memory.

I think the most common pitfall for high school students is that they care more about grades rather than learning. I know, because I was the same way. My advice to all high school students is to take periodic breaks from obsessing over grades and really understand what it is they are learning and how it applies to the world. Doing so allows for the discovery of a genuine passion, something that one would want to immerse his or herself into even if it were not required. My advice to someone who has discovered their passion and wants to undertake a project combining math and science would be to be open to learning things they may never have even heard of before. I would advise to not get discouraged if you find that the research you want to do involves topics beyond the scope of your high school curriculum, and instead to embrace this as a challenge to learn much more than you could ever learn in class. Do not just stick to topics you are comfortable with in doing your research; go above and beyond to learn more about the science and math you will be integrating in your project, and you will find it to be invaluable.

Research Section

Human memory is undeniably variable. It is evident that no two people can have the exact same memory capacity; however, it is also palpable that a single person's individual memory is extremely inconsistent. It has been shown that there are several factors that could account for such variation; for example, differences in a person's memory across his or her lifespan could be accounted for by age and neuroplasticity, while differences across years could be accounted for by factors such as maturation and experience (Maylor 1998; Stebbins et al., 2002; Tisserand, McIntosh, van der Veen, Backes, & Jolles, 2005). On the other hand, memory discrepancies become more difficult to account for when variability occurs from one hour to the next, or furthermore, from minute to minute. With near constant external conditions in such short amounts of time, it is possible that there exist certain concrete external variables that can account for such differences. At the same time, it could also be the case that such differences are inexplicable by external factors, and that the brain is simple better suited for cognitive tasks at certain times than others. The question becomes, are there concrete underlying factors that can neatly account for such variability, or do there simply exist certain endogenous states of the brain which are in turn better or worse suited for memory formation? The answer to this question has several pragmatic uses in science, and it also exposes the extent of human knowledge about the brain. If there truly do exist certain factors which, when working in unison, affect a person's memory abilities, such information would be useful to both the average person and to psychologists, who are motivated to pursue the most ideal conditions for encoding and retrieval. They would be able to directly control certain environmental factors, which, in turn, would predictably affect memory. Furthermore, this information could be applied in the treatment of patients with neurodegenerative diseases in their memory revitalization therapies, to make sure they are constantly in situations that promote good memory. If it is the case that, regardless of all external conditions holding constant, the brain still possesses certain effective and ineffective states, then this discovery warrants much further research into what exactly causes these shifts. Analyses were also conducted to determine whether memory variability over a slightly longer interval, from day to day, could be accounted for by different external variables. Studies have shown that amount of sleep is related to quality of memory in terms of encoding and retrieving abilities (Fowler, Sullivan, & Ekstrand, 1973; Ekstrand, Barrett, West, & Maier, 1977). Barrett and Ekstrand (1972) also studied the effects of time of day of sleep cycles on memory. Due to the presence of different factors affecting a person's encoding and retrieving abilities from across these larger time periods, it was more likely variability could be more significantly accounted for by some of these variables. Thus, we wanted to see controlling for certain day-to-day factors compares to controlling the effects of a much larger portion of the possible minute-to-minute factors affecting memory. Participants completed 24 free recall sessions across several days. In each session, they were presented with twenty-four different lists of words, and asked to freely recall as many of the words as they could after each list. The data for each person were then analyzed to determine variability in list performance within one session, and also variability in performance across all sessions. See 'Experimental Paradigm' section for a more comprehensive explanation of the experiment. A study conducted at the Rotman Research Institute of Baycrest proved that, especially in older adults, auditory distraction stimuli have disruptive effects on cognitive tasks, such as word recall (Craik 2014). Similarly, much research has shown that memory is impaired while multitasking and also suffers due to external interference (Clapp, Rubens, Sabharwal, & Gazzaley, 2011; Clapp, & Gazzaley, 2012; Clapp, Rubens, & Gazzaley, 2009; Wais, Martin, & Gazzaley, 2012). As a consequence, our research aims to eliminate almost every possible external distraction that could potentially explain memory variability. However, for across session variability, which exists due to factors that change in a participant's environment or lifestyle from day to day, there are a multitude of such possible variables. Many that can be quantifiably accounted for include practice effects, hours of sleep the night before, alertness levels, time of day, and which day of the week the session occurs on; these are but a few of the factors that change from day to day and may affect memory. Accounting for variables that could cause list-to-list variability within one session, however, becomes more difficult as most of the environmental distractions and variations were held constant. Possible causes of variability could arise from memory fatigue, so performance would be expected to decrease as the experiment proceeds. Older participants were also given a break after every eight lists, so it was hypothesized that performance would decrease every eight lists, and then spike. Additionally, individual qualities about the words themselves which are being presented could account for this variability, if certain words were better suited for memorization and recall than others. Moreover, a participant's performance on the prior list could be used as a predictor for performance on the next list; for example, if a participant was aware that his or her recall on the previous list was low, the participant might resolve to try a lot harder on the next list. All of these factors are potential explanatory variables of the disparity that exists in short term human memory, and this research aims to determine if these factors, when combined into the predictive model, can account for variability.

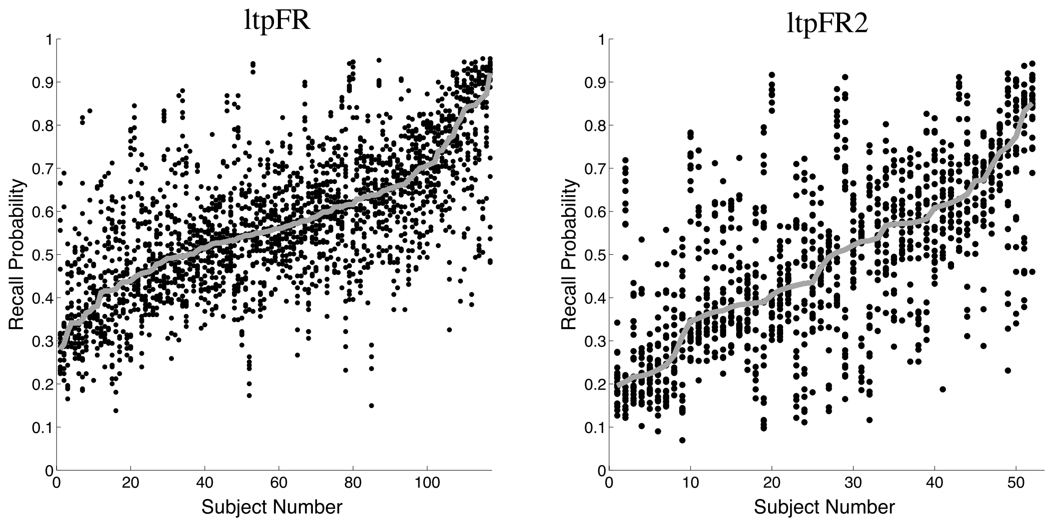
Results

We observed a range of performance between different sessions, with some being better or worse than others. This was expected, as there are many factors that contribute to a subject performing differently from session to session. In Figure 1A, the graph for ltpFR2 shows the exact proportions of words recalled from each of the 24 sessions completed by one person, with the lowest dot representing the session with lowest proportion of words recalled in a session, and so on. The subjects were ordered from lowest to highest median recall. The graph for ltpFR1 represents the same information. As is evident, some participants were very consistent in their memory across sessions, thus having shorter ranges, while other participants were variable in their performance from session to session. These graphs represent the initial analyses conducted on the participants' data to simply quantify and visualize the extent of variability that exists from session to session.

Additionally, the graph of average performance with respect to lists within a session also shows variability among lists, with participants consistently demonstrating an observable difference between their average performance on their worse lists and their average performance on their better lists in a session. In Figure 1B, the graph for ltpFR2 shows the exact averages of the 24 lists in each session per person, with the lowest dot representing the average of the lists with the lowest proportion of words recalled across all 24 sessions, and so on. The subjects were ordered from lowest to highest median recall. The graph for ltpFR1 represents the same information, and the subjects were again ordered by median recall, from low to high. The variability is visible by the vertical range of dots for each subject. Again, it is clear that some participants were more consistent in their memory memory abilities from list to list than others, but there clearly exists a significant amount of variability from lowest recall to highest recall within a session for every participant.

Figure 3 shows that, of the entire wordpool that the 1638 words that were presented to participants were chosen from, each word has certain qualities that distinguish it from

(A) Variability Across Session



(B) Variability Within Session

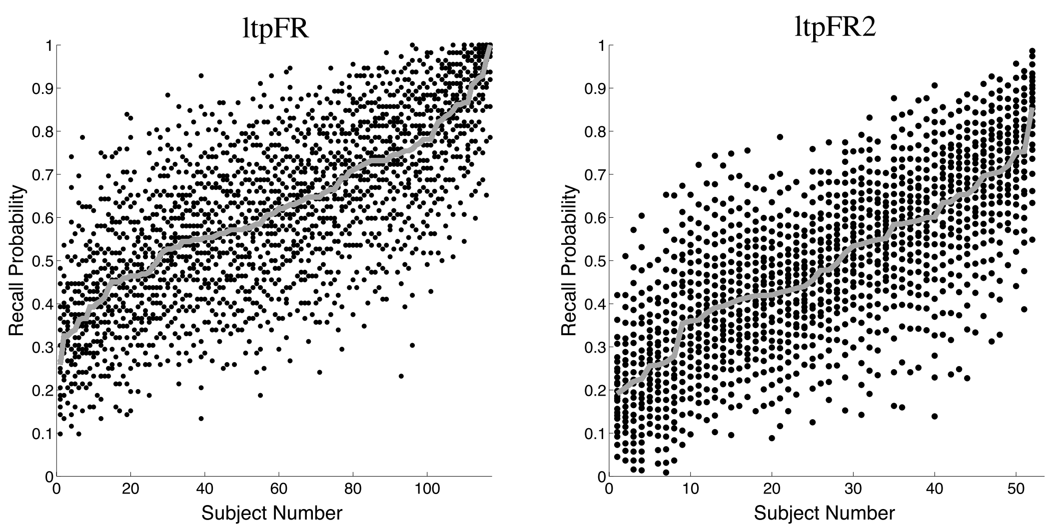


Figure 1. **Graphic representation of within session recall variability.** Subjects exhibited variability in probability of recall in both experiments, across sessions (A) and within sessions (B) respectively. The grey line denotes the median per subject. (A) Dots represent the total proportion of words recalled in one session for each subject's 24 sessions. (B) Dots represent the across session average of recall probability per list, after rank ordering lists by overall recall performance. One point is an average of 24 lists, one per session, where each list is in the same position on a rank from worst list to best list in that session.

Analysis of Experimental Word List

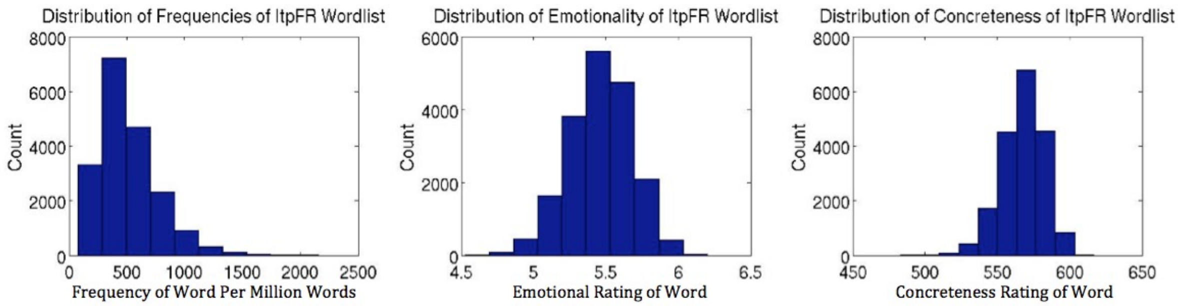


Figure 2. **Histograms representing the distributions of the experimental word list across three different categories.**

the other words. The number of encounters of each word in every million words found in the English language is one such factor, and may influence the relative ease with which the word is remembered. For example, a word that is less frequent in the English language may be more easily remembered and recalled. Another such characteristic is emotional valence, or how positive or negative of a feeling the word conveys. In a previous large-scale study, participants were shown portions of a very large pool of words one at a time and asked to rate the emotional valence according to the following scale: negative (ratings 1-4), neutral (ratings 4-6), or positive (ratings 6-9); all of the words in this experiment's word pool were drawn from this experiment and thus could be analyzed for emotional valence (Long 2015). Similarly, a characteristic that may influence recall is the word's concreteness, and quantitative measures of the concreteness of the words in the pool were also available from a previous study, in which participants rated words from the range of most abstract (100) to most concrete (700) (Steyvers 2004). Each word in the experimental word pool was matched with its corresponding frequency, emotional valence, and concreteness values as denoted in the reference databases, and a histogram was created for each quality to determine the spread of the words across these ratings. Each of these quantitative values for each word were averaged per list, to determine if the average frequency, emotional valence, and concreteness of the words in each list could be used to predict probability of recall.

After multiple regression analyses, we observed that, in the case of across session variability, the β values reported in Table 1 associated with none of the variables accounted for were significantly different from zero ($p < .05$). For variables like sleep and alertness, significant correlations were not found due to participants self-reporting similar numbers each session despite variable performances. Variability is also harder to account for with factors like time of day, which have subjective influences, as shown in previous findings that memory in older subjects is better in the morning than in the afternoon, while younger subjects show good memory in the afternoon and evening (May, Hasher, & Stoltzfus, 1993). Additionally, with respect to within session variability, the β values reported in Table 1 associated with all of the variables were significantly different from zero ($p < .05$), with the largest effects produced by session number, list number, and prior list's recall in ltpFR2. The distribution of the R Squared values represented by Figure 3A-D of each person's prediction model shows that the specific models generated for the participants based on the multiple regression of their data vary in their predictive ability; however, the singular model generated does not significantly account for the variability in memory. This is to be expected for the across session prediction model, as only a small fraction of the possible factors that may account for this variability was considered in the regression model. However, even after almost all of the possible external variables that could account for within session variability were regressed out by the model and observed in Figures 4B and 5B, the within session variability left unaccounted for by the model is not significantly less than the actual variability.

Table 1

ltpFR2: Summary of Simultaneous Regression Analysis for Variables Predicting Probability of Recall (N = 52)

Variable	*M* **β**	*SE* **β**
Across Session		
Session Number (1-24)	0.017	0.014
Amount of Sleep (Hours)	0.001	0.003
Alertness (Scale of 1-5)	-0.005	0.004
Start Time (Military Time)	-0.000	0.003
Day (Monday (1) to Friday (5))	0.007	0.004
Within Session		
Session Number (1-24)	-0.020*	0.001
List Number (1-24)	-0.017*	0.001
Frequency (Per Million Words)	0.005*	0.000
Emotionality (Negative (1) to Positive (9))	0.003*	0.000
Concreteness (Abstract (100) to Concrete (700))	0.002*	0.000
Prior List Recall	-0.008*	0.001

Note. Multiple Regression Analyses were conducted both across session and within session. Each participant's data for all variables were regressed against probability of recall to create a significant prediction model.
*$p < .05$

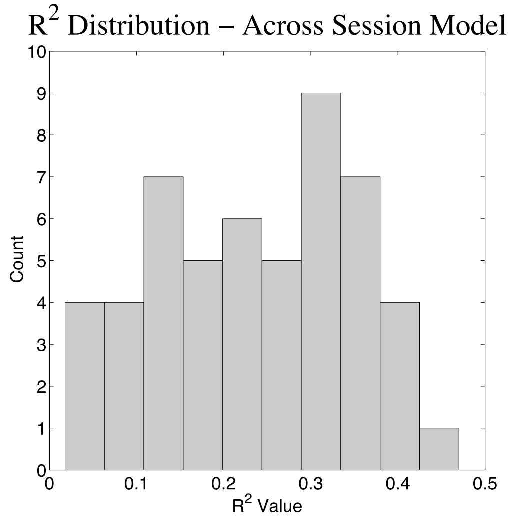
Table 2

ltpFR: Summary of Simultaneous Regression Analysis for Variables Predicting Probability of Recall (N = 117)

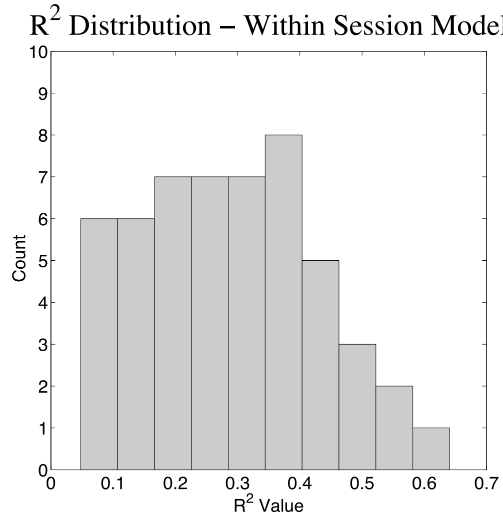
Variable	*M* **β**	*SE* **β**
Across Session		
Session Number (1-16)	-0.012	0.009
Amount of Sleep (Hours)	0.000	0.007
Alertness (Scale of 1-5)	-0.001	0.007
Start Time (Military Time)	0.003	0.008
Day (Monday (1) to Friday (5))	0.009	0.007
Within Session		
Session Number (1-16)	0.002	0.003
List Number (1-16)	-0.006	0.003
Frequency (Per Million Words)	0.003	0.002
Emotionality (Negative (1) to Positive (9))	0.003	0.002
Concreteness (Abstract (100) to Concrete (700))	0.004*	0.002
Prior List Recall	-0.018*	0.002

Note. Multiple Regression Analyses were conducted both across session and within session. Each participant's data for all variables were regressed against probability of recall to create a significant prediction model.
*$p < .05$

ltpFR2 Prediction Models

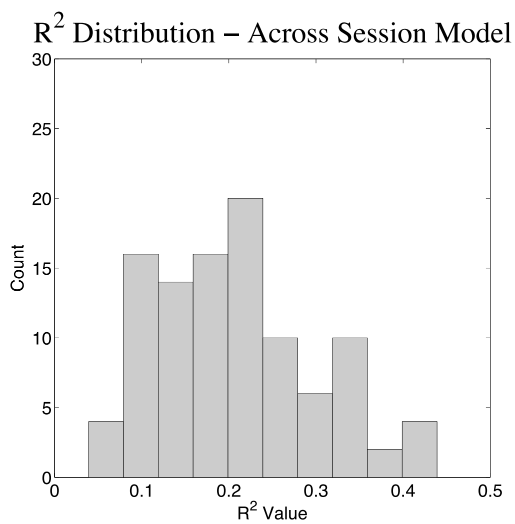


(a) $\mu = 0.2337$, $\sigma = 0.1139$, n = 52

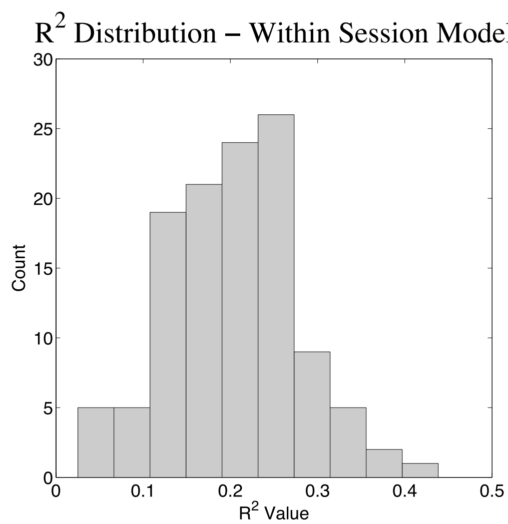


(b) $\mu = 0.2879$, $\sigma = 0.1435$, n = 52

ltpFR Prediction Models



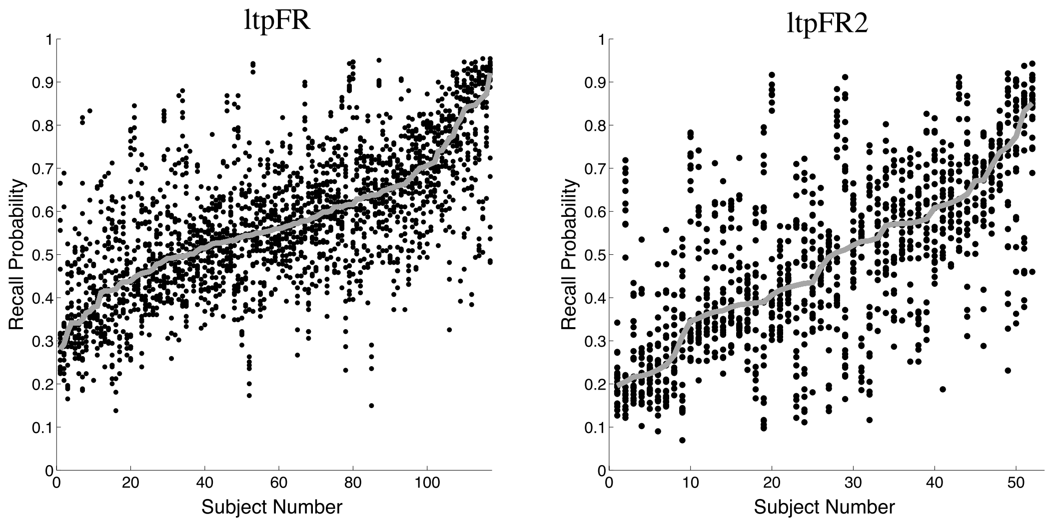
(c) $\mu = 0.2066$, $\sigma = 0.0907$, n = 117



(d) $\mu = 0.2015$, $\sigma = 0.0764$, n = 117

Figure 3. **Histograms representing the distributions of R Squared values of each of the prediction models.** We observed variability in the predictive value of the models generated for each of the subjects from his or her individual data, both across sessions (A) and within sessions (B). A predictive model was generated for each participant through a multiple regression of their data. A model like this was created for each participant and his or her individual model's R Squared value is represented in the histograms above.

(A) Variability Across Session - Before Prediction Model



(B) Variability Across Session - After Prediction Model

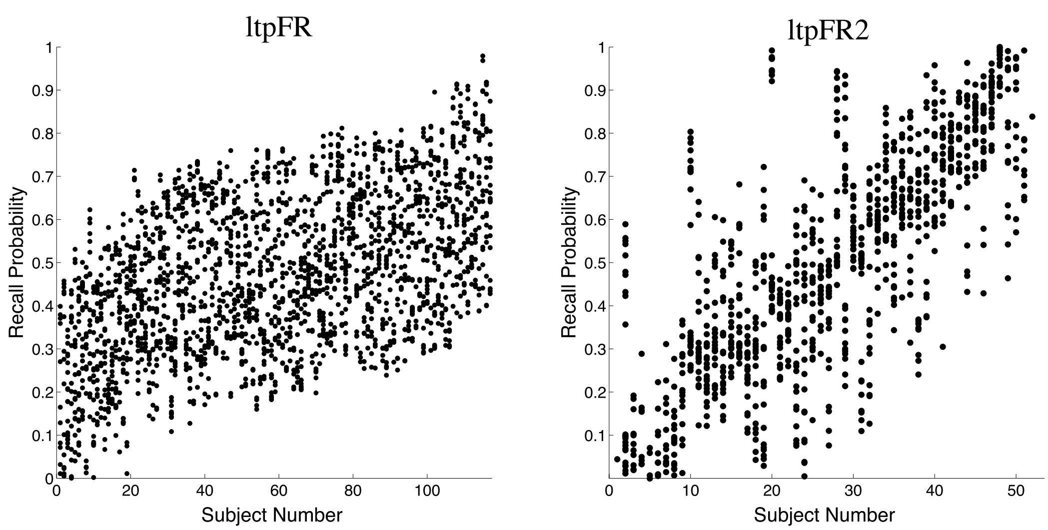
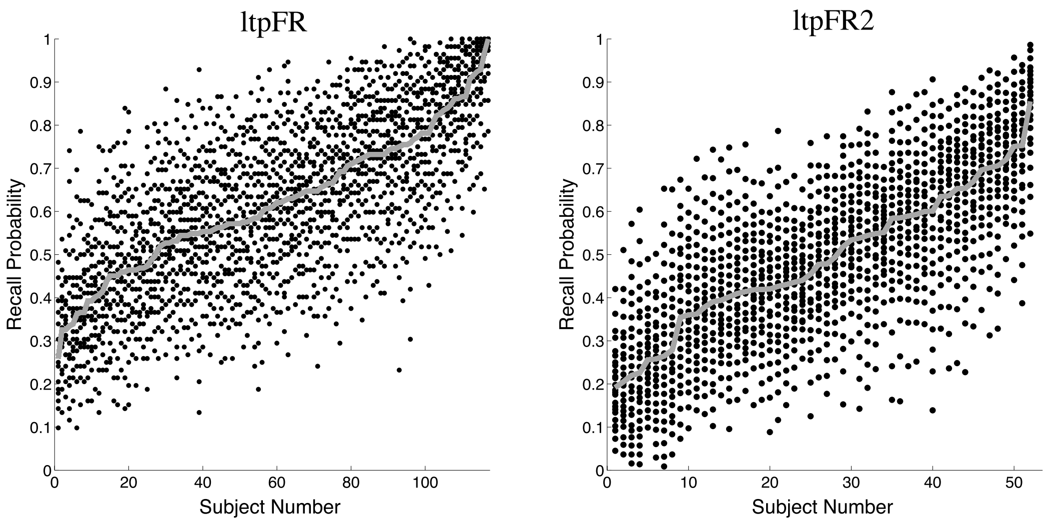


Figure 4. **Graphic juxtaposition of the effects of the predictive model on initial across session recall variability (A).** The final predictive model that was generated was used to predict recall based on the raw data for each variable associated with the sessions. For each session, the corresponding session number, amount of sleep, alertness, start time, and day of the experiment were regressed to determine a predicted probability of recall for that entire session. We then determined the error of this predicted value from the actual probability of recall for the sessions. Dots (B) represent the predictive error added to the subjects' average session performance across the entire experiment.

(A) Variability Within Session - Before Prediction Model



(B) Variability Within Session - After Prediction Model

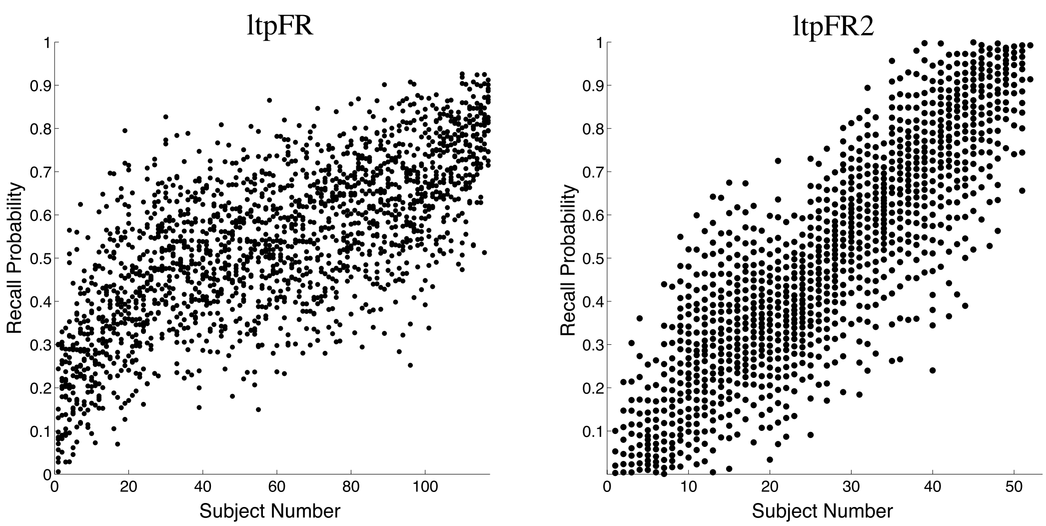


Figure 5. **Graphic juxtaposition of the effects of the predictive model on initial within session recall variability (A).** We generated a final predictive model which we used to predict list recall based on the raw data for each variable associated with the list. For each list, the corresponding session number, list number, average frequency, emotionality, concreteness, and prior list's recall were regressed to determine a predicted probability of recall for that list. The error of this predicted value was then determined from the actual probability of recall for the list. Dots (B) represent the predictive error added to the subjects' average list performance across the entire experiment.

Discussion

A constantly changing environment can indeed explain variability in memory; in 2014, Craik explored the effects of distractions on cognition and concluded that memory abilities are weakened by unwanted external stimuli. However, if almost all external distractions and agitations are accounted for, what explanation is there for the variability in human memory ability that still persists?

To emphasize the controlled environment of this experiment, participants were placed in a soundproof room with constant temperature and lighting, with plain walls and no distractions. They were not allowed to bring their cell phones inside with them, and were offered snacks beforehand to control for the possibility of any uncomfortable states. In the experiment, participants were asked to simply form a picture of the word in their mind as clearly as possible while the word was being presented, and then do the same for the next word without trying to remember the previous item. While it is possible that certain words may have individual meaning to a participant and thus increase the probability that said word is recalled, the items were chosen so that individual meaning alone cannot account for why a person might remember, hypothetically, 20% of a list one time and then perhaps 70% of the next list.

This study aims to control for many external stimuli, and to use all of the remaining factors that could result in memory differences in a short amount of time in a predictive model to try to account for the remaining variability. *See Figure 1A-B.* The factors that were uncontrolled for by the experiment and that could result in memory variability include session and list number, which account for mental fatigue over time. Additionally, another inconsistent portion of the experiment was the words being presented, but as they were selected to have a small spread of frequency, emotional valence, and concreteness values with no significant outliers, it was unlikely for these qualities to be significant predictors. *See Figure 3.*

The results of the multiple regression analyses, as delineated in Table 1, proved that only some of the variables in the model were significant, so only these variables were used in the prediction model. Theoretically, if the model were accurate, it would produce predicted values close to the actual recall values, and the points on Figures 5B and 6B would be very tightly centered on an average for each person. However, after regressing out the effects of these variables with the model, it was shown that variability did not significantly decrease, and that the model is an inaccurate predictor of recall. *See Figures 5 and 6.*

Thus, even when all possible explanations of this variability that exists in human memory in a short period of time, shown in Figure 1A-B, are accounted for, variability still persists. This suggests that external conditions alone cannot significantly account for variability in a person's memory abilities, that there exist certain endogenous states of the brain that are different in their memory capabilities, and that the brain internally fluctuates between these states without respect to the external environment. It would be edifying to investigate what exactly causes these shifts in the brain; for example, it could be accounted for by differences in electrical activity in the brain or metabolic processes that occur in the body, or perhaps by a completely unknown cause. These results show the need for further investigation into the physiology of the brain to explore the true complexity with which it operates.

References

Barrett, T. R., & Ekstrand, B. R. (1972). Effect of sleep on memory: III. Controlling for time-of-day effects. Journal of experimental psychology, 96(2), 321.

Clapp, W. C., & Gazzaley, A. (2012). Distinct mechanisms for the impact of distraction and interruption on working memory in aging. Neurobiology of aging, 33(1), 134-148.

Clapp, W. C., Rubens, M. T., & Gazzaley, A. (2009). Mechanisms of working memory disruption by external interference. Cerebral Cortex, bhp150.

Clapp, W. C., Rubens, M. T., Sabharwal, J., & Gazzaley, A. (2011). Deficit in switching between functional brain networks underlies the impact of multitasking on working memory in older adults. Proceedings of the National Academy of Sciences, 108(17), 7212-7217.

Craik, F. (2014). Effects of distraction on memory and cognition. Frontiers in Psychology, 5.

Ekstrand, B. R., Barrett, T. R., West, J. N., & Maier, W. G. (1977). The effect of sleep on human long-term memory. Neurobiology of sleep and memory, 419-438.

Fowler, M. J., Sullivan, M. J., & Ekstrand, B. R. (1973). Sleep and memory. Science, 179(4070), 302-304.

Long, N., Danoff, M., & Kahana, M. (2015). Recall dynamics reveal the retrieval of emotional context. Psychonomic Bulletin & Review Psychon Bull Rev, 1328-1333.

May, C. P., Hasher, L., & Stoltzfus, E. R. (1993). Optimal time of day and the magnitude of age differences in memory. Psychological Science, 4(5), 326-330.

Maylor, E. (1998). Retrieving names in old age: Short- and (very) long-term effects of repetition. Mem Cogn Memory & Cognition, 309-319.

Stebbins, G. T., Carrillo, M. C., Dorfman, J., Dirksen, C., Desmond, J. E., Turner, D. A., ... & Gabrieli, J. D. (2002). Aging effects on memory encoding in the frontal lobes. Psychology and aging, 17(1), 44.

Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In A. F. Healy (Ed.), Cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer. Washington, DC: American Psychological Association.

Tisserand, D. J., McIntosh, A. R., van der Veen, F. M., Backes, W. H., & Jolles, J. (2005). Age-related reorganization of encoding networks directly influences subsequent recognition memory. Cognitive brain research, 25(1), 8-18.

Wais, P. E., Martin, G. M., & Gazzaley, A. (2012). The impact of visual distraction on episodic retrieval in older adults. Brain research, 1430, 78-85.

Wilson, M.D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. Behavior Research Methods, Instruments & Computers, 20(1), 6-11.