

August 2019

## **A Distributional and Theoretical Analysis of Reaction Time in the Reversal Task Across Adulthood**

Kaitlynn N. Leclaire  
*University of Wisconsin-Milwaukee*

Follow this and additional works at: <https://dc.uwm.edu/etd>



Part of the [Clinical Psychology Commons](#)

---

### **Recommended Citation**

Leclaire, Kaitlynn N., "A Distributional and Theoretical Analysis of Reaction Time in the Reversal Task Across Adulthood" (2019). *Theses and Dissertations*. 2212.  
<https://dc.uwm.edu/etd/2212>

This Thesis is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact [open-access@uwm.edu](mailto:open-access@uwm.edu).

A DISTRIBUTIONAL AND THEORETICAL ANALYSIS OF REACTION TIME IN THE  
REVERSAL TASK ACROSS ADULTHOOD

by

Kaitlynn N. Leclaire

A Thesis Submitted in  
Partial Fulfillment of the  
Requirements for the Degree of

Master of Science  
in Psychology

at

The University of Wisconsin-Milwaukee

August 2019

## ABSTRACT

### A DISTRIBUTIONAL AND THEORETICAL ANALYSIS OF REACTION TIME IN THE REVERSAL TASK ACROSS ADULTHOOD

by

Kaitlynn N. Leclaire

The University of Wisconsin-Milwaukee, 2019  
Under the Supervision of Professor Driscoll

The frontal lobes are known to atrophy with age (Lockhart & DeCarli, 2014) and integrity of this region has been implicated in maintaining executive functioning (Chayer & Freedman, 2001). Reversal learning tasks are frequently used in experimental paradigms to assess components of executive function. Extant reversal learning literature has largely assessed measures of accuracy, but reaction time (RT) has not yet been well characterized. The current study examines the empirical RT distribution of the reversal task by utilizing distributional and theoretical analyses to better characterize performance and how it changes with age. Participants included 43 young (ages 18-30;  $M = 21.76$ ,  $SD = 2.85$ ) and 139 community dwelling middle-aged adults (ages 40-61;  $M = 49.96$ ,  $SD = 6.14$ ). Results showed a Normal-3 Mixture distribution best fit the sample as a whole, with the ex-Gaussian distribution passing visual inspection. This suggests both models and their parameters should be considered to evaluate group differences for this task. Correlation results showed RT and accuracy are distinct components of reversal learning. Age related significantly to RT and more so to efficient RTs ( $\mu$ ) than overall RT. A generalized regression further revealed that RT adds unique variance to explaining age-related differences in performance. Specifically, middle-aged adults

showed slower, efficient RT and increased intra-individual variability which has been previously linked to poorer frontal lobe processes and age-related cognitive decline (MacDonald, Nyberg, & Backman, 2006). Lastly, four RT-based factors were identified (Mental Efficiency, Intra-Individual Variability, Mental Speed, Inattention) that successfully distinguished groups with fractionated profiles of performance and should be further investigated to explore potential clinical implications in the context of cognitive aging. Overall, these findings highlight the importance of examining the RT distribution and measuring RT as a fractionated construct to further explain age-related differences in reversal learning.

*Keywords:* normal aging, reversal learning, reaction time, Diffusion model

© Copyright by Kaitlyne N. Leclaire, 2019  
All Rights Reserved

## TABLE OF CONTENTS

	<b>Page</b>
<b>List of Figures</b>	vi
<b>List of Tables</b>	vii
<b>Acknowledgements</b>	viii
<b>Background</b>	1
I. Aims & Hypotheses	8
<b>Method</b>	9
I. Participants	9
II. Stimulus Presentation	9
III. Procedure	10
IV. Data Analysis	11
A. Data Preparation	11
B. Distributional Analyses & Testing Group Differences	13
i. Testing Aim 1	13
ii. Testing Aim 2	14
C. Traditional Analyses	14
D. Exploratory Analyses	15
<b>Results</b>	16
I. Distributional Analyses	16
II. Group Differences	16
III. Traditional Analyses for Comparison (e.g., ANOVAs)	17
IV. Exploratory Analyses	20
A. Correlations	20
B. Factor Analysis	21
C. Cluster Analysis	21
D. Generalized Regression	25
<b>Discussion</b>	25
I. Limitations	29
<b>Conclusion</b>	30
<b>References</b>	39

## LIST OF FIGURES

Figure 1. An Illustration of the Ex-Gaussian distribution	31
Figure 2. An Illustration of the Normal-3 Mixture distribution	32
Figure 3. An Illustration of the Ratcliff Diffusion model	33
Figure 4. An illustration of the Reversal Task	34
Figure 5. The Normal-3 Mixture distribution best fit group RT (for all participants)	35
Figure 6. Ex-Gaussian distribution passes visual inspection	36

## LIST OF TABLES

Table 1. Factor loadings for Diffusion model, ex-Gaussian, and Normal-3 Mixture distribution parameters	37
Table 2. Mean values for the factor scores entered into the Hierarchical Cluster Analysis	38



## **ACKNOWLEDGEMENTS**

First and foremost, I would like to thank my advisor, Dr. Ira Driscoll, for her continued support, guidance, and mentorship throughout this process. I would also like to thank my mentor, Dr. David Osmon, for his unwavering assistance and the immense statistical knowledge that he shared with me to help make this project possible. Thank you to my committee member, Dr. Raymond Fleming, for his valuable insight and stimulating questions that challenged me to deeply think and grow as a researcher. Finally, I must express my profound gratitude to my family, friends, and partner, Michael, for providing me with unfailing support and continuous encouragement.

## **Background**

Normal aging is associated with functional decline in many cognitive domains, though not all are affected. Age-related declines in episodic memory, working memory, spatial learning, and certain components of attention are common (for a review see Drag & Bieliauskas, 2009). Alternatively, facets of cognition involving verbal skills, semantic memory, implicit learning, and priming are largely unaffected (Drag & Bieliauskas, 2009; Kausler, 1994). Throughout the inevitable process of aging, the frontal lobes are known to undergo atrophy and are perhaps first to be affected by age-related pathological changes (Drag & Bieliauskas, 2009; Lockhart & DeCarli, 2014). This brain region has historically been implicated in relation to executive function which oversees many daily cognitive processes (Chayer & Freedman, 2001). Previous work has shown that cognitive performance can remain stable through the tenth decade of life, suggesting that normal age-related decline is modest and that more pronounced declines in cognition may represent incipient disease (Storandt, Grant, Miller, & Morris, 2002). However, age remains one of the greatest risk factors for development of many late-life neurodegenerative disorders, including dementia (Lindsay et al., 2002). With increasing longevity and the rapidly increasing number of individuals in the population that may be affected, as well as the broader societal burden (e.g., financial, public health) that co-occurs, there is a growing interest in maintaining cognitive and everyday functioning, as well as preserving quality of life, throughout the lifespan. Therefore, it has become increasingly urgent to develop and utilize methods to sooner and better detect subtle cognitive changes in middle age for early intervention and preventative efforts. These methods should be specific and sensitive enough to detect differences in

changes related to normal or pathological aging, both of which are mild in middle age. Experimental behavioral tasks are promising avenues to assess early changes in executive function. These functions can be delineated in non-human paradigms and then tested in the human population in hopes of establishing consistency for translational testing.

Executive function refers to a set of high-level cognitive processes such as inhibition, sustained attention, decision-making, working memory, and cognitive flexibility (Diamond, 2013). Reversal learning tasks, known to be frontal lobe-dependent, are frequently used in experimental paradigms to assess components of executive function. While reversal tasks have been predominantly utilized in non-human experimental paradigms (Brushfield, Luu, Callahan, & Gilbert, 2008; Lai, Moss, Killiany, Rosene, & Herndon, 1995; Schoenbaum, Nugent, Saddoris, & Gallagher, 2002; Tapp et al. 2003; Voytko, 1999), they have been adapted for human research (typically using a probabilistic paradigm) (Mell, Heekeren, Marschner, Wartenburger, Villringer, & Reischies, 2005; Weiler, Bellebaum, & Daum, 2008) and allow for cross-species comparisons. Reversal learning tasks first require subjects to learn a number of unambiguous (elemental) discrimination contingencies. The elemental discriminations involve learning to respond to 'correct' stimuli that are rewarded as opposed to 'incorrect' stimuli which are not rewarded. Once participants reach a learning criterion (e.g., 11 correct responses on 12 consecutive trials), the task requires a flexible adjustment in behavior when reward contingencies previously learned are reversed. Ultimately, the Reversal Task assesses the ability to modify previously learned associations and is thought to measure cognitive flexibility and perseveration. In a

primate model of human aging, it has been suggested that executive dysfunction, as measured by reversal learning perseverative responding, may exemplify an important attribute of age-related cognitive decline (Lai et al., 1995). While learning simple two-choice discriminations (i.e., elemental discriminations) is often unaffected by aging (Driscoll et al., 2003; Reed & Squire, 1999; Rickard & Grafman, 1998), older adults have more difficulty with and show impairments on tasks that require switching cognitive sets (Wecker, Kramer, Hallam, & Delis, 2005). Therefore, reversal learning may be a sensitive approach for assessing age-related changes in executive functioning. Extant reversal learning literature has historically focused on measures of accuracy (correct vs incorrect responses) or learning (number of trials required to reach criterion) as their primary outcome, but reaction time (RT) has not yet been well characterized. The current study aims to fill this gap by examining distributional properties and a theoretical model of RT to fully investigate the underlying cognitive processes of the reversal learning task.

Reaction time, the time taken to carry out a task, has been a prominent measure of cognitive processes in psychology since its debut by Donders (1869). It's common practice for RT to be analyzed using parametric tests, such as an analysis of variance (ANOVA) using sample means. However, this prevalent approach may not be the appropriate statistical method when considering the characteristics of RT data and the parametric test assumptions of ANOVAs. RT is known to have a positively skewed distribution to the right with a 'fat tail' (Luce, 1986). Therefore, normal distributions (e.g., Gaussian distribution) and central measures of tendency, such as the mean, may not adequately capture important aspects of performance due to being obscured by the

non-normal distribution. Hence, parametric statistics can lead to misinterpretation of RT performance. However, analysis of the RT distribution has a potential to reveal important and interesting findings that are typically masked by parametric test assumptions.

Past psychometric studies have highlighted age-related declines in cognitive function that reflect a shift in (for a review see Salthouse, 2010) and support analysis of the entire distribution. Numerous models can be employed to fit an RT distribution, as its shape can vary considerably depending on the specific task demands and individual variability in performance. Therefore, a single distribution model and its parameters may not be a best fit or most accurate in describing performance on all RT-based tasks. A comprehensive understanding of the appropriate analyses to examine the RT distribution of performance on a specific timed task is essential considering the use of RT distributions in making inferences about underlying psychological processes.

A prominent model in cognitive literature is the ex-Gaussian distribution, which seems to be a good fit for many RT tasks (Heathcote, Popiel, & Mewhort, 1991; Hockley, 1984; Leth-Steenson, Elbaz, & Douglas, 2000; Luce, 1986; Mewhort, Braun, & Heathcote, 1992; Ratcliff, 1978, 1979). The ex-Gaussian model is a convolution of a normal and exponential distribution (see Figure 1) and consists of three parameters: Mu ( $\mu$ ), Sigma ( $\sigma$ ), and Tau ( $\tau$ ). The normal component of the distribution is composed of Mu and Sigma, the mean and variance of the normal component of the ex-Gaussian distribution, respectively. The exponential component corresponds to Tau (i.e., the mean and standard deviation of the right tail). These parameters have been useful in distinguishing fast and slow components in responding. Furthermore, given the

convoluted nature of the ex-Gaussian model, it can fit distributions ranging from near normal to very skewed. Despite the popularity of this model for analyzing RT, ex-Gaussian is not always the best fitting model. For example, a Weibull distribution was found to best fit the empirical RT distribution of a simple RT visual detection task (Maloney & Wandell, 1984). Furthermore, Sternberg & Backus (2015) showed that even when weak assumptions for sameness of ex-Gaussian shape were used, they were violated in all but approximately 21% of cases. Therefore, other distributions must be considered and investigated. For example, a Normal-3 (NL-3) Mixture distribution can be useful in revealing fast, slower, and slowest RTs (Osmon, Kazakov, Santos, & Kassel, 2018). NL-3 fits the empirical RT distribution with the convolution of three normal Gaussian components (see Figure 2). The first component consists of the fastest responses, the second component the slower responses, and the third includes the slowest responses. Each NL-3 component consists of several parameters: location or mean (Location1, Location2, Location3), dispersion or standard deviation (Dispersion1, Dispersion2, Dispersion3) and probability or proportion of responses in each of the three normal distributions (Probability1, Probability2, Probability3).

While distribution parameters can be helpful in characterizing RT performance, a theoretical model is needed to understand the cognitive implications of timed performance. The Diffusion Model (Ratcliff, 1978) is a successful and validated theoretical model that has been increasingly used in analyzing fast two-choice decisions. The model consists of several parameters that allow detailed explanations of behavioral data (i.e., variation in RT and accuracy). Explanations are accomplished by separating the quality of evidence that enters a decision from decision criteria and non-

decision processes (Ratcliff & McKoon, 2008). Ultimately, accuracy, mean RTs, and RT distributions are translated into components of cognitive processes. The following parameters are estimated in the Diffusion model: decision threshold ( $a$ ), drift rate ( $v$ ), and response time constant ( $t_0$ ) (see Figure 3). The decision threshold ( $a$ ) is the width of the interval between decision boundaries (i.e., correct or incorrect response). The drift rate ( $v$ ) is the mean rate at which information is accumulated to make a decision. Lastly, response time constant ( $t_0$ ) reflects the duration of extra-decisional components (i.e., non-decisional components such as stimulus encoding and response execution). Due to the ability of separating components of processing, the Diffusion model has been frequently used to study the effects of age on memory and decision criteria across a variety of two-choice decision tasks (Ratcliff & McKoon, 2015; Ratcliff, Thapar, & McKoon, 2001, 2003, 2004, 2006, 2007, 2010, 2011; Thapar, Ratcliff, & McKoon, 2003) and may be an appropriate fit for the current two-choice discrimination task (i.e., reversal learning task).

A consensus in the literature is that aging has a small effect on item recognition memory, but this is largely based on accuracy measures (Balota, Dolan, & Duchek, 2000; Craik, 1994; Craik & McDowd, 1987; Ratcliff & McKoon, 2015). Another central finding in the literature is that RTs increase and become more variable as people age (Deary & Der, 2005; Morse, 1993). More recent research in aging using the Diffusion model for fast two-choice decisions suggests that as people age their RTs increase, although this increase is associated with little or no change in accuracy (Ratcliff et al., 2010). Despite this relative consistency in accuracy, RT remains an important facet of performance to characterize given the information that can be obtained from numerous

parameters from distributional and theoretical models. For example, using the Diffusion model, Ratcliff and colleagues (2010) suggest the slowdown in RT is due to older adults' setting conservative decision boundaries. Reluctant to commit errors, they set their decision criteria significantly further from the starting point of the decision process than young adults do. However, they remain just as accurate as young adults when making the ultimate decision. Furthermore, older adults are slower in non-decision time, which includes processes outside of decision-making itself (e.g., stimulus encoding and response execution) and contributes to slower RTs (Ratcliff & McKoon, 2015; Ratcliff et al., 2006, 2007, 2010, 2011). However, it remains unclear if physical components, such as changes in the motor system, contribute to this slowing and subsequently impact the quality of RT as a performance measure in old age.

Extant literature presents a dilemma that is two-fold. First, if accuracy is used as the dependent variable for two-choice tasks then aging is not expected to significantly affect cognitive processes. On the other hand, if mean RTs are used as the dependent measure then aging is expected to have a large effect (Ratcliff, 2007). Interestingly, in much of the past experimental literature, either accuracy or mean RTs have been used alone to describe behavioral data. However, it is important to consider the two variables simultaneously to get a clearer picture regarding age-related cognitive changes. For instance, examining accuracy in isolation may suggest little to no cognitive decline, while mean RTs alone may suggest cognitive slowing. The Diffusion model provides a framework to remedy this problem and has several advantages over alternate approaches (e.g., ANOVAs of RT means) (Voss & Voss, 2007). First, the model considers both accuracy and RT simultaneously, which addresses the aforementioned



concern about a single dependent variable. Second, the model consists of several parameters that allow detailed explanations of the behavioral data. Third, data is efficiently utilized such that both RT means are used and the RT distribution is fully analyzed, which allows for a more thorough understanding of reversal learning and how it changes with age.

### **Aims & Hypotheses**

The present study aims to fill a gap in the literature by thorough examination of the empirical reaction time distribution in middle-aged (ages 40-61) and young (ages 18-30) adults, in order to better characterize reversal learning performance and how it changes with age. First, the best statistical fit to the reversal task RT distribution was examined by employing multiple distributional analyses to allow for comparison (Aim 1). Based on the existing literature, I predicted that the ex-Gaussian distribution would be the best fit among the 10 distributions tested. While ex-Gaussian parameters would be retained to compare groups for differences, Normal-3 Mixture (NL-3) distribution parameters would also be retained to examine fast, medium, and long RTs between middle-aged and young adults (1a). Based on prior reports, I also predicted that middle-aged participants would show a higher proportion of responses, with longer and more variable RTs, in the second and third distributions of the NL-3 compared to younger participants (1b). Furthermore, I predicted that middle-aged participants would have 'fatter' tails in the RT distribution with greater values for the Tau parameter of the ex-Gaussian distribution (i.e., slower RTs) (1c). Second, the current study aimed to describe if and how reversal learning changes with age, by examining accuracy and RT measures by employing a Diffusion model analysis to separate components of cognitive

processing (Aim 2). Based on prior literature (Ratcliff et al., 2010), I predicted that middle-aged participants would have a higher threshold ( $a$ ), indicating they would be more conservative about a decision and accumulate more information before making a correct response (2a). In contrast, no group differences were expected in drift rate ( $v$ ), indicating that the quality of information collected by young and middle-aged participants does not differ (2b). Lastly, I predicted that middle-aged participants would be slower in non-decision time ( $t_0$ ) than young participants, indicating slower encoding and/or response execution (2c).

## **Method**

### **Participants**

43 young adult (ages 18-30;  $M = 21.76$ ,  $SD = 2.85$ ; 29 females, 14 males) and 139 middle-aged adult (ages 40-61,  $M = 49.96$ ,  $SD = 6.14$ ; 80 females, 59 males) participants were included in the current study. Young adults were recruited through SONA (an online participation database of the Psychology department at the University of Wisconsin-Milwaukee), while middle-aged adults were recruited from the community. Participants who self-reported psychiatric and neurological disorders, learning disability, and other medical conditions (e.g., head injury, stroke, seizures) were excluded. All participants had normal or corrected vision, which was necessary for completion of the tasks (e.g., to see the consent form and computer monitor). The local Institutional Review Board (IRB) approved all procedures.

### **Stimulus Presentation**

Visual discriminations were presented on a 15.6-inch Dell laptop screen with 1600 x 900-pixel resolution using Presentation® software, a stimulus delivery and

experiment control program (Version 16.0, Neurobehavioral Systems, Inc., Berkeley, CA, [www.neurobs.com](http://www.neurobs.com)). Visual stimuli were presented with a height of 342 pixels by a width of 512 pixels.

## **Procedure**

All data collection occurred on the same day. First, written informed consent was obtained from all subjects in accordance with the University of Wisconsin-Milwaukee IRB. Then, participants completed two visual discrimination tasks (i.e., elemental discriminations and reversal learning) on a laptop computer. On each trial, pairs of non-nameable visual stimuli, randomly generated from a subset of 27, were presented on a computer screen. Each trial began with a fixation cross that was presented in the center of the display for 1s and followed by a stimulus pair presentation. Participants were instructed to choose one of two visual stimuli to find out whether their choice was correct or incorrect by pressing one of two designated keyboard keys that correspond to either the left or the right stimulus element of a pair. The key press cleared the stimulus pair from the display. Response evaluations (e.g., accuracy and RT) were recorded to a file following the key press. Correct responses were followed by a presentation of the word 'Correct' in the center of the display accompanied by a high-pitched tone for 1s. Incorrect responses were followed by the word 'Incorrect' presented in the center of the display accompanied by a low-pitched tone for 1s. After a 2s delay the display was cleared and a 2s intertrial interval followed. Participants were instructed to respond as quickly and accurately as possible based on feedback they received after each trial ended.

Both tasks were presented in six phases using a stepwise approach. Phases 1-3 consisted of elemental discriminations only (see Figure 4). Stimulus pair A+B- was presented during phase 1. During phase 2 stimulus pair C+D- was presented in addition to the A+B- pair. In phase 3 the final elemental discrimination pair, E+F-, was presented in addition to the previous two pairs. At this point, participants would have learned which visual stimuli was 'Correct' in each pair. Then, in the following phases (4-6), the reward contingencies previously learned were reversed (i.e. reversal learning; see Figure 4). Reversal learning was also introduced in a step-wise approach, such that stimulus pair A-B+ was presented during phase 4. During phase 5, stimulus pair C-D+ was presented in addition to the A-B+ pair. In phase 6, the final pair, E-F+, was presented in addition to the previous two pairs. In the elemental discrimination phases, a correct response indicated successful learning and memory of the rewarded stimuli in each pair, while RT measured the participant's speed of processing. A longer (i.e., slower) RT indicates a greater amount of time required to recall or learn the correct stimuli. During the reversal phases, a correct response indicated successful learning of the new (i.e., reversed) contingency, while an incorrect response suggests difficulty switching cognitive sets. Within each phase every fifth trial displayed a previously learned stimulus pair. Training continued for each phase until the participant made 11 correct responses on 12 consecutive trials. A maximum of 400 trials were allotted to complete the task, after which the testing was discontinued.

## **Data Analysis**

**Data preparation.** Data preparation techniques were used to explore outliers and missing values with imputation using JMP techniques (SAS, 2015a). Specifically,

even with distribution-level analyses physiologically 'impossible' RTs are usually excluded. In the current study, RTs less than 150ms were excluded as 'impossible' RTs. Long RTs are often not excluded when the entire RT distribution is analyzed with non-normal techniques; however, long RTs were examined using outlier techniques because such an approach may facilitate distinguishing a 'true' outlier from a long RT due to inattentiveness. In particular, with skewed RT distributions traditional mean-based and distribution-dependent methods such as ANOVA are inappropriate because they are based upon the assumption of a normal distribution. Therefore, multiple outlier analyses were used that are more appropriate for non-normal distributions, including Quantile Range, Robust Fit, Multivariate Robust, and Multivariate k-Nearest Neighbor analyses. Quantile Fit is appropriate because it is a distribution-free approach that sorts data from smallest to largest values and identifies extreme values based upon empirical quantile ranges. Robust Fit uses a robust estimate of the distribution center and spread that for this study was based upon the Cauchy distribution of extreme values. Multivariate Robust calculates the Mahalanobis distance of a multivariate distribution that will include the Diffusion model parameters  $a$ ,  $v$ , and  $t_0$ . Finally, the Multivariate k-Nearest Neighbor uses multiple values of k-nearest neighbors to each data point using a Fibonacci sequence to identify extreme values. Extreme data points were excluded based upon convergence of all above approaches.

Additionally, data were prepared by examining missing values. Given there was a small number of missing values, imputation was used to maximize sample size. In cases where a participant was missing an entire portion of the reversal task (e.g., all reversal phases) that participant was excluded (N=1). Otherwise, missing data was

imputed using the multivariate normal imputation procedure with a shrinkage estimator. This method imputes data based upon the least squares derived covariance matrix of all non-missing cases. Using shrinkage estimators improves the estimation of the covariance matrix.

### **Distributional Analyses & Testing Group Differences**

**Testing aim 1.** RT distributions are known to be right-skewed, obviating the use of parametric statistics that are dependent upon normal, Gaussian distributions. As a result, group level distributional analyses were completed by examining best fit of the following distributions using JMP (SAS, 2015a, 2015b): ex-Gaussian, Normal-3 Mixture, Normal-2 Mixture, LogNormal, GLog, Gamma, Weibull, Extreme Value, Exponential, and Normal. Best fit was determined based on which distribution had the lowest Akaike's Information Criterion-corrected (AIC-c: SAS, 2015b), which is a maximum likelihood estimation method. If other distributions had AIC-c values within a 0 to 2-point range of the model with the lowest AIC-c they were considered a comparable fit (Burnham & Anderson, 2004). Determining best fit characterizes RT performance and allows best choice of the appropriate parameters to evaluate group differences. To examine group differences, a nonparametric analysis was performed because assumptions of ANOVA were not met due to the non-normal RT distribution. Therefore, a Kruskal-Wallis 2-Sample Exact test was utilized, as it satisfies the assumption that the groups have similar distribution shapes (ex-Gaussian). The Kruskal-Wallis test ranks all data from the groups to test whether the samples originate from the same distribution. Specifically, this procedure compares the ranks to examine whether the medians of the groups are different. Best distributional fit was also analyzed at the individual level.

**Testing aim 2.** Diffusion model analyses assume that information accumulates continuously until one of two thresholds is reached (i.e., a correct or incorrect response). Fast-dm is free software for estimating diffusion model parameters (Voss & Voss, 2007). The program uses the partial differential equation (PDE) method to calculate predicted RT distributions. For parameter estimation, the Kolmogorov-Smirnov (KS) statistic was applied. This statistic is the maximal vertical distance of the predicted and empirical cumulative RT distribution. The Diffusion model translates accuracy, mean RTs, and RT distributions into components of cognitive processing to characterize the decision-making process. Accuracy aids in describing response threshold and bias, while RT provides insight into speed and quality of information processing (Ratcliff & McKoon, 2008). The following parameters were estimated in the Diffusion model: decision threshold ( $a$ ), drift rate ( $v$ ), and response time constant ( $t_0$ ). The decision threshold ( $a$ ) is the width of the interval between decision boundaries (i.e., correct or incorrect response). The drift rate ( $v$ ) is the mean rate at which information is accumulated to make a decision (when information accumulation reaches the correct or incorrect boundary,  $a$ ). Lastly, response time constant ( $t_0$ ) reflects the duration of extra-decisional components (including stimulus encoding and response execution among any other non-decisional factors contributing to RT).

**Traditional analyses.** While ANOVAs assume a normal distribution and are not appropriate analytic methods for RT tasks, these traditional analyses remain widely used in the extant literature and are also examined here for comparison purposes to nonparametric analyses and the Diffusion model.

## Exploratory Analyses

To further investigate reversal learning performance, correlations and partial correlations between accuracy and RT were used to determine whether the two measures assess distinct aspects of performance. Generalized regressions addressed what distinct measures were contributing to distinguishing age-related differences in reversal learning. This technique, rather than ANOVAs or multiple regressions, was used for two purposes. First, high dimensional data were utilized with 17 predictors and generalized regression with the adaptive elastic net procedure has the oracle property, which allows zeroing out non-contributory variables. Second, the criterion variable (i.e., age) was distributed as a binomial variable with Poisson distribution characteristics, making traditional ANOVA and multiple regression techniques inappropriate. Generalized regression allows modeling the Poisson distribution to better model the data and find reliable predictors of age. Finally, cluster analysis was used to determine whether combinations of variables would identify unique clusters of participants in 4-dimensional space. Since cluster analysis is susceptible to mere level-of-performance differences, principal components analysis was executed utilizing reaction time and accuracy data from the Reversal Learning task, as well as parameters from the Diffusion model ( $a$ ,  $v$ ,  $t_0$ ), ex-Gaussian ( $\mu$ ,  $\sigma$ ,  $\tau$ ), and Normal-3 (Location1-3, Dispersion1-3, Probability1-3) distribution analyses. Accuracy data from the Reversal Task included trials-to-criterion for both the elemental and reversal phases of the task. Utilizing the factor scores taken in the principal components analysis ensures orthogonal variables to enter into the cluster analysis to generate better groupings of participants' reversal learning performance. As a result, a hierarchical cluster analysis



was utilized, and clusters were interpreted according to the profile of results from the aforementioned parameters to assist in identifying unique clinical groups based on reversal learning performance.

## **Results**

### **Distributional Analyses**

A Normal-3 Mixture distribution was the best fit for group RT (see Figure 5), including all participants, with no other model within 2 AIC-c points, although an ex-Gaussian distribution passed visual inspection (see Figure 6). While the Normal-3 was the best fit for the sample as a whole, very few individual participants had the Normal-3 Mixture as the best fitting distribution. The LogNormal distribution was the best fit for 120/182 participants (89 middle-aged; 31 young). The Normal-2 Mixture distribution was the best fit for 39/182 participants (30 middle-aged; 9 young), while the Normal-3 Mixture was the best fit for 17/182 participants (14 middle-aged; 3 young). Furthermore, 6 middle-aged participants had a GLog (n=1), a Gamma (n=3) and a Weibull (n=2) as best fit. Many participants had multiple best fit distributions according to the range of 2 points for AIC-c values, most of which were Normal-2 or Normal-3 Mixture distributions.

### **Group Differences**

Group differences in the Diffusion model, ex-Gaussian, and Normal-3 Mixture distribution parameters were examined using both the Wilcoxon/Kruskal-Wallis tests due to the non-normal parameter distributions. Additionally, the Kolmogorov-Smirnov 2-Sample test was utilized because it is sensitive to differences in both shape and location of the two group distributions. For sake of comparison to existing literature, traditional

ANOVA methods were also employed to show the importance of using distribution-free statistics in RT tasks where normal distributions are not extant.

For Diffusion model parameters, groups differed only on the  $a$  parameter ( $\chi^2[1] = 11.93, p = .0006$ ;  $KS = .13, p = .0038$ ), indicating middle-aged adults displayed a more conservative decision threshold. Within ex-Gaussian parameters only  $\mu$  was different between groups on both parametric and nonparametric statistics ( $\chi^2[1] = 47.76, p < .0001$ ;  $KS = .27, p < .0001$ ), indicating middle-aged adults showed slower, efficient RTs. In contrast, the ex-Gaussian parameter  $\sigma$  was only different on the Kolmogorov-Smirnov Two-Sample test, indicating that central tendency of the two groups was not different but distribution shape was different ( $KS = .10, p = .0352$ ). Group differences in  $\sigma$  revealed middle-aged adults displayed more variable, efficient RTs. Non-parametric tests were important for the ex-Gaussian  $\tau$  parameter where no difference was evident on ANOVA ( $F[1,180] = 1.20, p = .275$ ) but both the Kruskal-Wallis and Kolmogorov-Smirnov 2-Sample test showed a group difference ( $\chi^2[1] = 4.99, p = .0255$ ;  $KS = .14, p = .0012$ ), with younger adults displaying more slower and inefficient RTs in the exponential component.

### **Traditional Analyses for Comparison (e.g., ANOVAs)**

One-way ANOVAs examining the middle-aged and young groups on the ex-Gaussian parameters ( $\mu$ ,  $\sigma$ ,  $\tau$ ), Normal-3 parameters (Location1-3, Dispersion1-3, Probability1-3), and Diffusion model parameters ( $a$ ,  $v$ ,  $t_0$ ) found the groups to differ on the  $\mu$  parameter [ $F(1,179) = 36.441, p < .001$ ], Location1 [ $F(1,155) = 8.849, p = .003$ ], and Location2 [ $F(1,151) = 5.624, p = .019$ ]. These results suggest that younger adults have faster and more efficient responding ( $\mu = 462.98$ ) than

middle-aged adults throughout the Reversal Task ( $\mu = 633.41\text{ms}$ ), as well as faster RTs in the first ( $L1 = 791.76\text{ms}$ ) and second ( $L2 = 1544.91\text{ms}$ ) components of the NL-3 distribution than middle-aged adults ( $L1 = 912.58\text{ms}$ ;  $L2 = 1853.82\text{ms}$ ).

To further investigate group differences within different phases of the Reversal Task (i.e., elemental vs. reversal), 2x2 ANOVAs were conducted to examine the effect of age and phase on RT, trials-to-criterion, and Diffusion model parameters ( $a$ ,  $v$ ,  $t_0$ ). Results showed a significant main effect of age regarding correct RT responses [ $F(1,358) = 20.262$ ,  $p < .001$ ], as well as significant main effects of age [ $F(1,338) = 15.248$ ,  $p < .001$ ] and phase [ $F(1,338) = 33.361$ ,  $p < .001$ ] for incorrect RT responses. Younger adults responded significantly faster than middle-aged adults when making both correct and incorrect responses. Furthermore, younger adults responded significantly faster than middle-aged adults when making incorrect responses in both the elemental and reversal phases of the task. Of note, 18 middle-aged and 4 younger adults were not included in this incorrect RT analysis as they did not have any incorrect responses during the elemental phase of the task. In regard to accuracy and learning, results of a 2x2 ANOVA showed a significant main effect of phase, such that, regardless of age, all subjects required a significantly lesser number of trials to reach criterion during the elemental phase ( $M = 36.25$ ,  $SD = .69$ ) in comparison to the reversal phase ( $M = 39.14$ ,  $SD = .69$ ),  $F(1,358) = 8.679$ ,  $p = .003$ . Groups did not significantly differ in trials-to-criterion across the task as a whole. To further investigate responding during the reversal phase, the first presentation of reversal learning (Phase 4) was also examined. Results of one-way ANOVAs showed both middle-aged and young adults required similar trials-to-criterion within Phase 4, but younger adults responded

significantly faster ( $M = 877.07$ ,  $SD = 267.76$ ) than the middle-aged adults ( $M = 1044.96$ ,  $SD = 270.48$ ) throughout this first phase of reversal learning. Distributions were normal for ED-RT-correct and ED-trials-to-criterion as well as RL-RT-correct but not RL-trials-to-criterion.

Lastly, Diffusion model parameters were examined with 2x2 ANOVAs to further characterize responding based on age and phase. Results showed a significant interaction between age and phase on the decision boundary parameter [ $F(1,358) = 4.076$ ,  $p = .044$ ], such that middle-aged adults displayed lower conservatism in making a decision (i.e., a lower decision threshold;  $a = 2.76$ ) than young adults ( $a = 2.89$ ) on the easy, elemental trials, but displayed greater conservatism on the harder, reversal trials (middle-aged adults,  $a = 2.22$ ; young adults,  $a = 1.95$ ). Furthermore, results indicated a significant interaction between age and phase on the drift rate parameter [ $F(1,358) = 6.296$ ,  $p = .013$ ], but no main effects of either age or phase. Middle-aged adults displayed a higher drift rate ( $v = 1.55$ ) for the elemental phase, but a lower drift rate ( $v = 1.33$ ) for the reversal phase, in comparison to young adults (ED,  $v = 1.31$ ; RL,  $v = 1.44$ ). This suggests that there is a difference in information quality depending upon difficulty of the task. Finally, results showed a significant main effect of age on the non-decision parameter [ $F(1,358) = 4.227$ ,  $p = .041$ ], indicating middle-aged adults ( $t_0 = .51$ ) were slower in the cumulative total of stimulus encoding and response execution than young adults ( $t_0 = .46$ ) throughout the entire reversal task.

However, distributions were not normal suggesting that ANOVAs are of questionable value since normality is an underlying assumption of this test statistic.

## Exploratory Analyses

**Correlations.** Correlating the Diffusion model ( $a$ ,  $v$ ,  $t_0$ ), ex-Gaussian ( $\mu$ ,  $\sigma$ ,  $\tau$ ), and NL-3 parameters (Location1-3, Dispersion1-3, Probability1-3) with the trials-to-criterion variables from the elemental discrimination (ED) and reversal learning (RL) portions of the Reversal Task revealed little evidence for collinearity. Out of 272 correlations there was only one value  $>.9$  (Probability1 with Probability2), 2 values  $>.8$  (Location2 with both Location3 & Dispersion1), and 3 values  $>.7$  (Location1 with  $\mu$  & Location2 & Dispersion1).

ED trials-to-criterion was not correlated with ED RT-all trials ( $r = -.01$ ,  $p = .90$ ). Likewise, RL trials-to-criterion related nonsignificantly with RL RT-all trials ( $r = .11$ ,  $p = .12$ ).  $\mu$  showed a nonsignificant relationship to ED trials-to-criterion ( $r = -.14$ ,  $p = .07$ ) and likewise with RL trials-to-criterion ( $r = -.03$ ,  $p = .68$ ). These results indicate that speed and accuracy in reversal learning are unrelated, even for efficient RT ( $\mu$ ), and that no speed-accuracy trade-off was present in these data.

The relationship of accuracy and RT to age was also examined. Age related significantly to  $\mu$  ( $r = .36$ ,  $p < .0001$ ), ED RT-all trials ( $r = .21$ ,  $p = .004$ ), and RL RT-all trials ( $r = .30$ ,  $p < .0001$ ). Given the non-normal distributions, Spearman's rho was executed, showing similar, though slightly higher, values for all analyses and are not further discussed. Partial correlations were examined between the two speed measures ( $\mu$  and RT-all trials [separately for ED and RL]) and age.  $\mu$  reduced the relationship of age with the other two RT variables to negligible values ( $r < -.02$  for ED RT-all trials;  $r = .11$  for RL RT-all trials). In contrast, ED/RL RT-all trials did not greatly affect the relationship between  $\mu$  and age (for ED:  $r = .36$  reduced to  $.29$ ; RL:  $r = .36$  reduced to

.22). These results suggest that age relates to RT in the reversal task, and more so to efficient RTs ( $\mu$ ) than overall RT.

Overall, the correlation results indicate that RT and accuracy are distinct components of reversal learning performance and should both be evaluated further with the other Diffusion model, ex-Gaussian, and Normal-3 Distribution parameters to understand the relationship between Reversal Learning and age. As a result, a cluster analysis was utilized to examine unique performance profiles across the factors generated from the principal component analysis. Furthermore, a generalized regression analysis was used to better understand which variables were most predictive of age.

**Factor analysis.** Diffusion model ( $a$ ,  $v$ ,  $t_0$ ), ex-Gaussian ( $\mu$ ,  $\sigma$ ,  $\tau$ ), and NL-3 (Location1-3, Dispersion1-3, Probability1-3) variables used in the principal component analysis yielded four factors with eigenvalues  $>1.0$  ( $4.79 = 27.19\%$ ,  $3.04 = 18.85\%$ ,  $2.25 = 17.38\%$ ,  $1.51 = 13.92\%$ ), accounting for  $77.32\%$  of the variance. The four factors are labeled 'Mental Efficiency', 'Intraindividual Variability', 'Mental Speed', & 'Inattention' (see Table 1 for loadings on each of the four factors). Factor scores were generated for use in the cluster analysis and generalized regression (see Table 2).

**Cluster analysis.** The cubic clustering criterion was maximized at 5 clusters and the 4 factors entered into the cluster analysis were well absorbed with:  $R^2 = .39$  for Mental Efficiency (ME);  $R^2 = .52$  for Intra-Individual Variability (IIV);  $R^2 = .39$  for Mental Speed (MS); and  $R^2 = .61$  for Inattention (IA).

Cluster 1 was the youngest cluster ( $X_{age} = 37$ ) and represented a good performing group with the best or nearly best performance across all reversal learning

factors (ME = -.71; IIV = -.17; MS = -.56; IA = -.17). Furthermore, this cluster, along with cluster 4, had the least conservative Diffusion model criterion ( $a$ ), second fastest drift rate ( $v$ ), and was in the fastest group for non-decision speed ( $t_0$ ). Among the ex-Gaussian parameters, cluster 1 had the fastest  $\mu$ , the lowest  $\sigma$  variability, and the second lowest  $\tau$  behind cluster 4. This cluster's quick mental speed was verified in the NL-3 distribution location parameters, having the fastest times in both the first and second normal component and tied for fastest in the third component with cluster 3. Likewise, variability in each of the three normal components was lowest (first component) or in the lowest group (second and third component). The other four clusters were a mix of good and bad performance, indicating that the cluster analysis was successful in distinguishing groups with fractionated profiles of performance.

Cluster 2 was the second oldest cluster ( $X_{age} = 48$ ) and distinguished by good performance in Intra-Individual Variability (-.54), Mental Speed (-.22), and Inattention (.61), while overall Mental Efficiency (.71) was distinguished as the worst compared to all other clusters. Further characterizing this cluster, it, along with cluster 5, had the most conservative decision boundary ( $a$ ) in the Diffusion model. Additionally, it was behind only cluster 4 for the slowest drift rate ( $v$ ) and was behind only cluster 3 for the slowest non-decision time ( $t_0$ ). In contrast, efficient RT and variability was relatively intact ( $\mu = 586\text{ms}$ ,  $\sigma = 70\text{ms}$ ) and off-task RT was middling ( $\tau = 620\text{ms}$ ). Consistent with relatively intact efficient RT, cluster 2 had the highest percentage of RTs in the first NL-3 component and relatively few RTs in the second and third components of the NL-3 distribution. One may hypothesize that such a pattern of results indicates an older group with slow motor-sensory responding and tentative decision-making owing to

greater consideration of information, which may be indicative of mild cognitive impairment.

Cluster 3 was the second youngest cluster ( $X_{age} = 43$ ) and had relatively good Mental Efficiency (-.60) while Intra-Individual Variability (.62), Mental Speed (1.47), and Inattention (3.83) were poor. Interestingly, cluster 3 had the fastest drift rate ( $v$ ) but the slowest non-decision time ( $t_0$ ) and efficient RT ( $\mu = 759\text{ms}$ , along with cluster 4). Consistent with its high Intra-Individual Variability, efficient responses were associated with the greatest variability ( $\sigma = 115\text{ms}$ , along with cluster 4). This cluster also had the greatest off-task RTs ( $\tau = 1027\text{ms}$ ) and had the longest RTs in the third component of the NL-3 distribution. They also displayed the greatest variability in that third component and the lowest probability of response in the first normal component of the NL-3 distribution, but the greatest probability in the third normal component ( $\text{Probability}_3 = 37\%$ ). One may hypothesize that such a pattern of results suggests an impulsive, relatively young cluster with variable responding owing to being inattentive and off-task.

Cluster 4 displayed relatively low Intra-Individual Variability (-.25) and little Inattention (-.82), but their Mental Efficiency (.28) was marred by slowed Mental Speed (.94). Further analysis of specific parameters showed a relatively conservative boundary for decisions ( $a = 2.28$ ) and the slowest drift rate ( $v = 1.70$ ) and non-decision time ( $t_0 = .58$ , along with cluster 3). It was the oldest cluster ( $X_{age} = 49$ ), had the slowest efficient RT ( $\mu = 771$ , along with cluster 3), and had the most variability among efficient responses ( $\sigma = 115\text{ms}$ , along with cluster 3). In contrast to cluster 3, off-task responding was the least likely in this cluster ( $\tau = 387\text{ms}$ ). Mental Speed was



consistently slow across all three normal components of the NL-3 distribution and variability was great in the first normal component but not the second or third normal components, consistent with low Tau values. This cluster had the highest percentage of responses in the first component (prob1 = 87%, along with cluster 2) and was among the lowest percentages in the second and third component. While this cluster was similar to cluster 2 in age and slowed responding, it showed less impairment because of few inattentive/off-task responses. As a result, it may be hypothesized that this cluster is most representative of normal age-related cognitive changes.

Cluster 5's Mental Efficiency (.62) was marred by the greatest Intra-Individual Variability (1.78) among all of the clusters by a wide margin. Otherwise, their Mental Speed (-.08) was relatively good with few inattentive/off-task responses (-.25). Such variable responding led to a conservative decision boundary ( $a = 2.92$ ), although drift rate was second best ( $v = 1.18$ ) and not different from the best drift rate ( $v = .845$ ). Additionally, non-decision time ( $t_0 = .393$ ) and efficient responding ( $\mu = 574\text{ms}$ ) were relatively fast, as well as responding in the first and second normal component of the NL-3 distribution (Location1 = 847ms; Location2 = 1736ms) but not the third normal component (Location3 = 4098ms). Because of the variability in RT, this cluster had a high proportion of responses in the second normal component (Probability2 = 40%) but not the third normal component (Probability3 = 3%). This cluster was largely middle-aged with the 90th percentile at 55.8 years; and, therefore, may represent a non-impaired group with trait-level variable responding, considering the good efficient responding with good ability to accumulate information and stay on-task.

**Generalized regression.** In order to identify the best predictors of age among the reversal learning parameters, an adaptive elastic net generalized regression with Poisson distribution modeling and AIC penalty was utilized. Variables were included from the Diffusion model ( $a$ ,  $v$ ,  $t_0$ ), ex-Gaussian distribution ( $\mu$ ,  $\sigma$ ,  $\tau$ ), NL-3 distribution (Location1-3, Dispersion1-3, Probability1-3), and the ED and RL trials-to-criterion variables. A Generalized  $R^2=.83$  demonstrated good prediction of age with 7 of the original 17 variables being significant predictors, including, in order of significance,  $\mu$ ,  $\sigma$ , Dispersion1, Location1, Dispersion2,  $t_0$ , and  $\tau$ . Additionally,  $v$ ,  $t_0$ , and  $\tau$  were marginally significant predictors. Only two of those predictors had Independent Resampled Variable Importance values greater than .1 ( $\mu = .645$ ; Dispersion1 = .221), with slower efficient RTs ( $\mu$ ) and more variable efficient RTs (Dispersion1) being associated with older age.

## **Discussion**

To the best of my knowledge, the present study is the first within the reversal learning literature to examine the distributional properties of the Reversal Task RT distribution and to apply a theoretical model to put forth a nuanced account of reversal learning performance and how it changes with age. Historically, the literature has largely focused on measures of accuracy as the primary outcome, which is likely due to the daunting statistical considerations of non-normal distributions. The present study contributes to the accuracy literature by demonstrating that thorough examination of RT can reveal important age differences in speed of responding for reversal learning performance.

As predicted, the RT distribution of the Reversal Task was a non-normal distribution, obviating the use of the traditionally applied ANOVA models because they are dependent upon a normal distribution. My first aim was to examine the best statistical fit to the tasks RT distribution to characterize performance and determine the appropriate parameters to evaluate group differences. Contrary to predictions and much of the RT literature on two-choice tasks, the ex-Gaussian distribution was not the best fit for group RT. However, it did pass informal visual inspection suggesting its parameters may still offer important characterization of performance. Rather, a NL-3 distribution was the best fit for group RT (including all participants) and further supports the notion that the ex-Gaussian distribution cannot always be assumed to be the best fitting model. Furthermore, individual analysis demonstrated that clinical examination of reversal learning results at the single subject level of analysis cannot be reliably accomplished using group level statistical findings. That is, while the group level distribution fit a Normal-3 model, very few had this distribution as a best fit on an individual level, which may have contributed to the lack of significant group differences in NL-3 parameters. This suggests that a more individualized approach may be most appropriate for this task. If the Reversal Task is to be used in clinical settings, future development of normative data for the task and its parameters would be necessary to better interpret individual level performance. Normative data would aid in identifying individual level strengths and weaknesses in cognitive processes and assist in identifying age-related cognitive decline by comparison to group level performance.

Notably, exploratory results revealed meager correlations between RT and accuracy, and partial correlations showed that the relationship between RT and age is

largely unaffected when parsing out contributing variance associated with accuracy of performance in this reversal learning task. The results indicate that RT and accuracy are distinct components of reversal learning performance. Using RT predictors of age in a generalized regression further revealed that RT adds unique variance to explaining age-related differences in reversal learning performance. Specifically,  $\mu$  (mean efficient RT) and Dispersion1 (variable efficient RT) were the only two variables that explained the variance in performance between middle-aged and young adults. Interestingly, middle-aged adults showed slower, efficient RT and increased intra-individual variability, which has previously been linked to poorer frontal lobe processes and age-related cognitive decline (MacDonald, Nyberg, & Backman, 2006). In line with these findings, middle-aged adults displayed a more conservative decision boundary (a) further supporting their slowed but efficient responding. Additionally, while results showed that younger adults responded faster and less variably when making efficient responses, they displayed more responses in the 'tail' of the distribution indicating slower and inefficient RTs, which may be due to taking the task less seriously than the middle-aged adults. Overall, this suggests it is important to measure RT as a fractionated construct and highlights the importance of utilizing appropriate analytical methods to characterize and interpret RT data. Undifferentiated traditional measures of RT (i.e., mean RT of all trials) frequently give a false picture of the relationship between RT and accuracy. Such methods tend to overlook intra-individual variability, do not account for the non-normal distribution and, therefore, do not optimize the information that can be gained from thoroughly examining the RT distribution and its appropriate parameters. Therefore, future studies should consider measuring RT as a fractionated

construct and investigating differences in intra-individual variability to further explain age-related differences in reversal learning tasks.

Given this wealth of information, I identified and put forth 4 clinically relevant factors, comprised solely of RT variables, that further characterize reversal learning performance. The first factor was labeled Mental Efficiency and consisted of numerous loadings across the Diffusion model, ex-Gaussian, and NL-3 distribution parameters. As such, it is viewed as representing a general factor of speed and efficiency of information processing, consistent with the tendency of the principal component analysis to maximize loadings of variables onto the first factor. The second factor was deemed Intra-Individual Variability as it consisted of more specific loadings from parameters of the NL-3 distribution that represent variability in responding. Specifically, the highest loadings indicated a large number of variable responses from the second normal component of the NL-3 distribution but not the third component. These loadings occur in combination with strong negative loadings suggesting few responses that were not variable in the first component. Additionally, a small negative loading on the Diffusion model parameter  $v$  indicates slow accumulation of information. Variable responding is suggested as the mechanism that slows RT, but not so much as to indicate extreme slowing due to frank lapses of attention or loss of mental set that occurs when responding slips into the third normal component. Such variability in responding is then associated with slow accumulation of information and few efficient responses in the first normal component. Mental Speed consisted of strong loadings from speed measures for basic non-decision RT, efficient responding from the normal component of the ex-Gaussian distribution, and fast RTs in the first component of the NL-3 distribution.

Hence, the factor was labeled Mental Speed. Inattention consisted of strong loadings from parameters that reflect extremely long RTs, making this factor different from the Intra-Individual Variability factor where RTs were only moderately long. Additionally, a negative loading from drift rate ( $v$ ) indicated that when long RTs occur then information accumulation is slow, as well as unlikely to be in the first component of the NL-3 distribution. This profile of loadings seems to indicate poor responding that likely reflects loss of mental set or some other pathological process that makes quick or even mildly slow responding unlikely. Ultimately, the aforementioned factors proved successful in distinguishing groups of individuals with fractionated profiles of performance which may be of clinical importance within the context of cognitive aging. Future studies should further investigate these RT factors and their ability to detect clinically relevant groups.

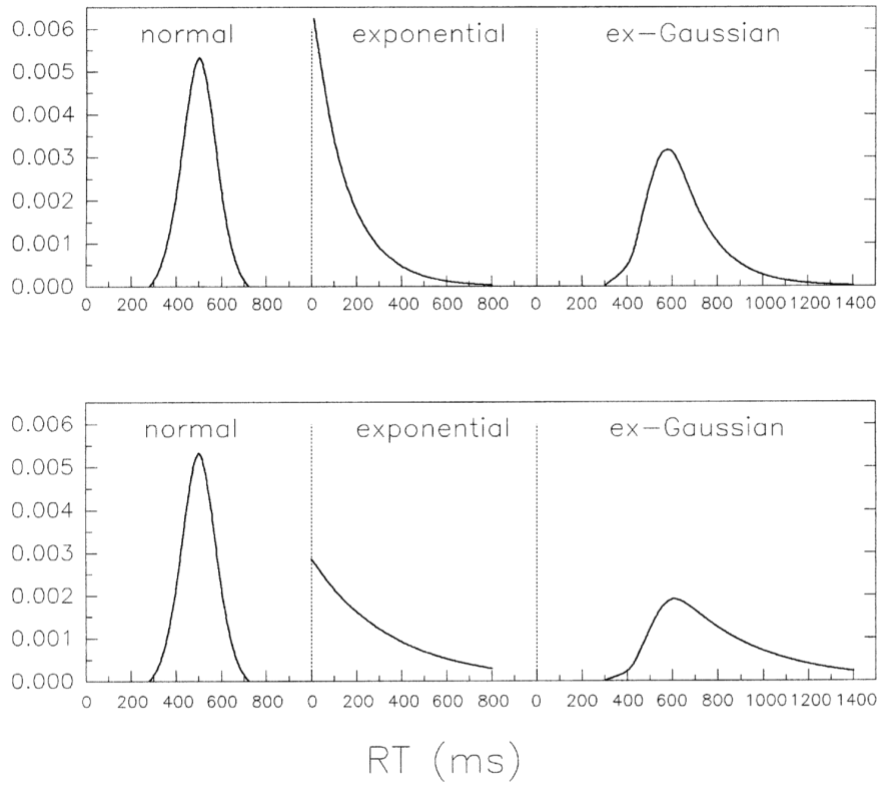
### **Limitations**

There are several limitations to the current study. First, given the simplistic nature of the Reversal Task employed here and its application to a healthy aging sample (as it was originally designed for pathological aging), it is possible participants performed at ceiling and, therefore, did not show much variability. However, the procedure design puts forth minimal burden on working memory and should therefore better reflect other components of executive function that may impact accuracy and speeded performance. Nonetheless, future studies may consider including groups of participants that are representative of both normal and pathological aging to assist in refining detection methods. Second, our “older” sample is considerably younger than most older samples in the literature and consists mostly of middle-aged adults. However, this may also be viewed as a strength as most aging studies omit middle age. It is imperative to study

middle age to better understand early signs of cognitive dysfunction, well before the onset of any clinical symptoms, which may aid in early intervention and prevention efforts. Conversely, it is possible the rate of slowing in processing speed and cognitive dysfunction was not as considerable in middle age as it may be in older populations. Therefore, future studies may consider examining healthy young, middle-aged, and older adults to determine if performance in middle age is distinct from or similar to one of the other groups. Third, while participants were screened for psychiatric and neurological disorders, they may not have specified existence of a motor disorder. As such, we cannot with certainty rule out lower RTs due to motor system dysfunction in some participants.

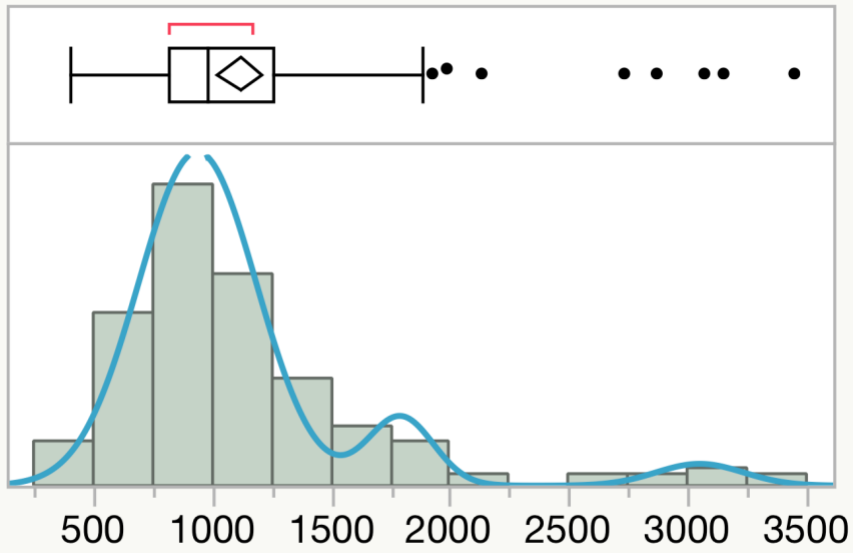
### **Conclusion**

In summary, a thorough analysis of the Reversal Task response time distribution goes above and beyond purely descriptive statistics (e.g., accuracy and mean RT) to characterize speeded performance. While distributional analyses are an initial step in identifying parameters to characterize performance, theoretical models are needed to better understand the cognitive processes underlying RT performance, and non-parametric statistics are needed to fully appreciate the wealth of information RT can provide.



*Figure 1.* The ex-Gaussian distribution is a convolution of a Gaussian (normal) and exponential distribution (figure from Leth-Steensen et al., 2000). The top and bottom portions of the figure display two visual examples of the ex-Gaussian model.





*Figure 2.* A Normal-3 Mixture distribution fits the empirical RT distribution of an individual with the convolution of three normal Gaussian components (figure from Osmon et al., 2018).

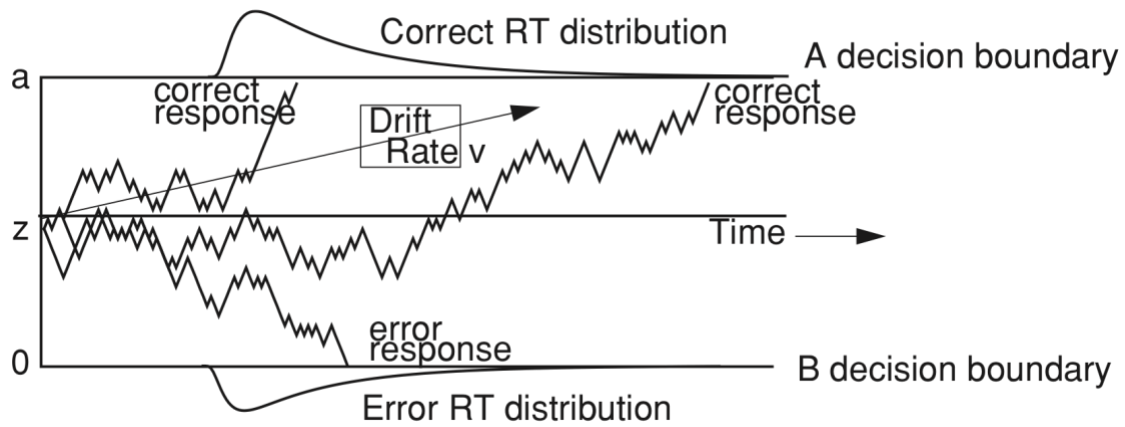
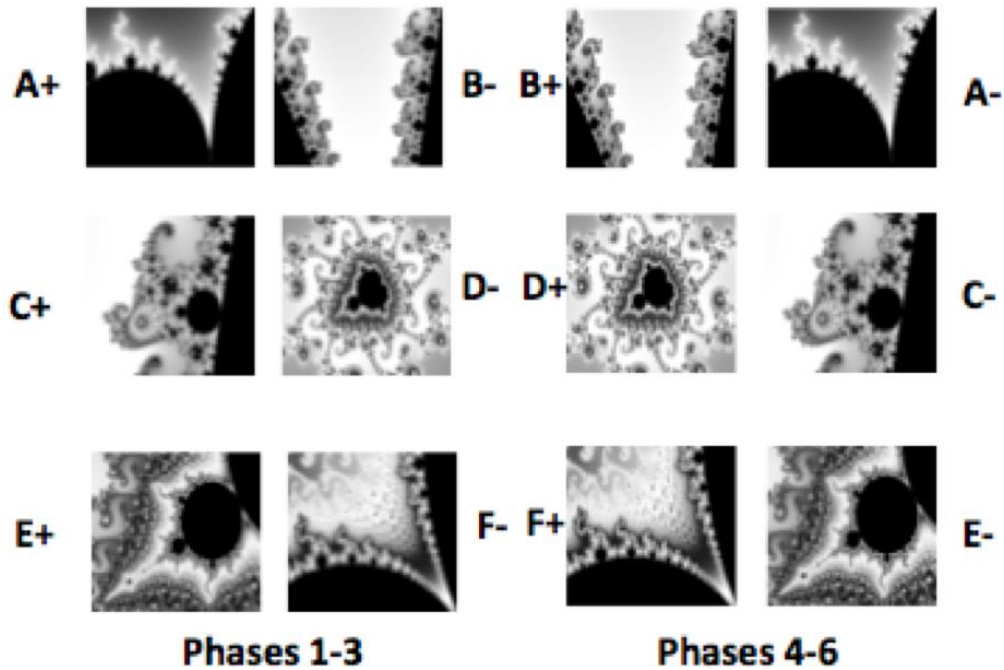
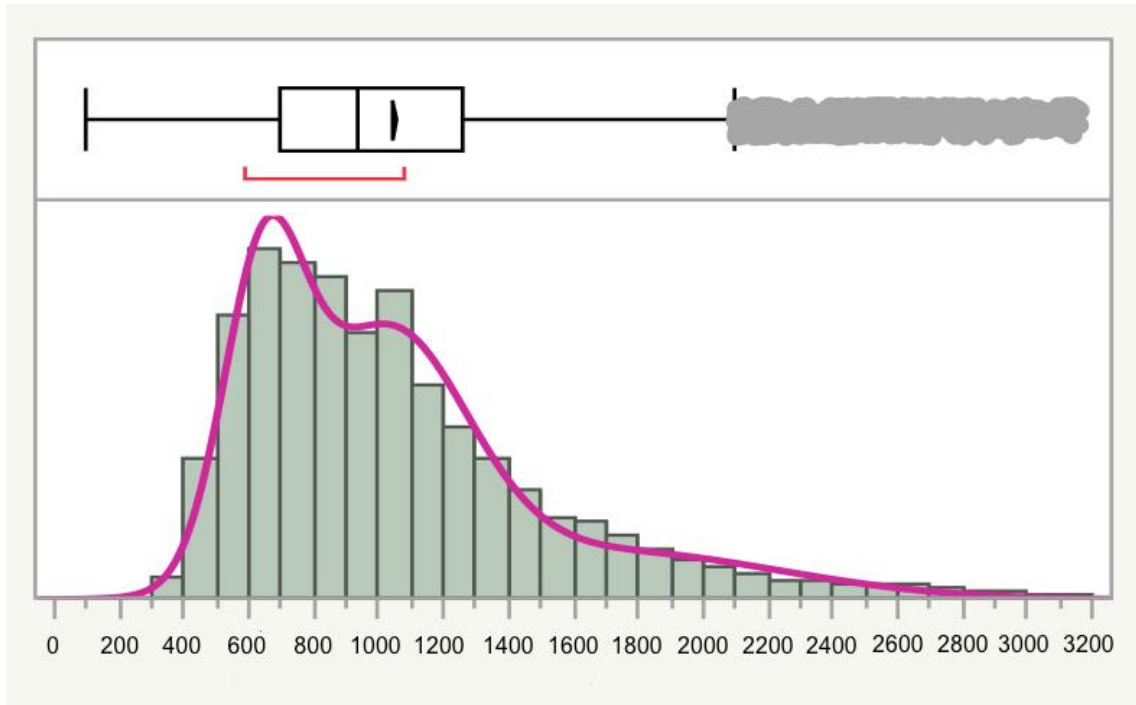


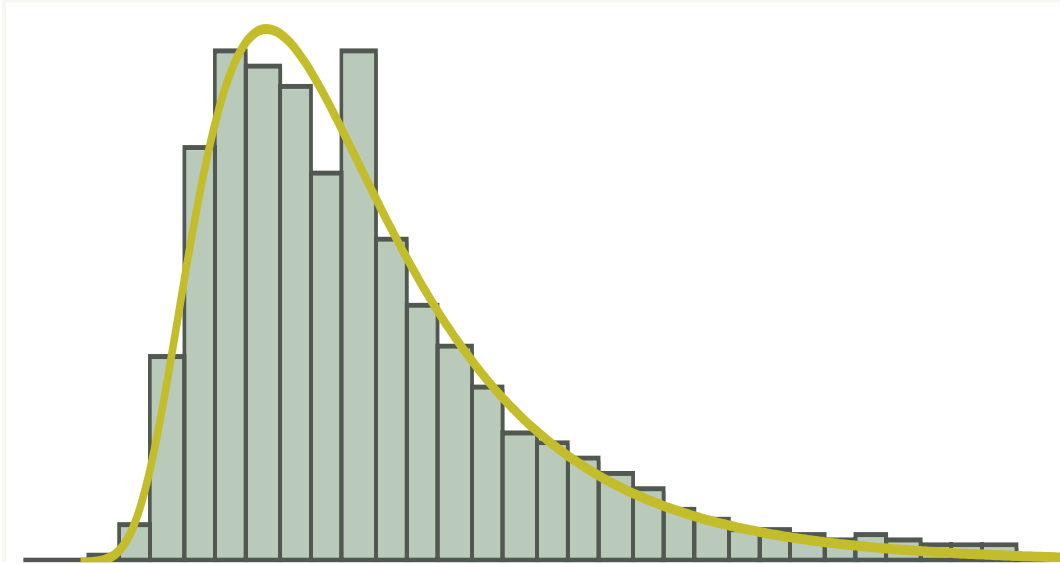
Figure 3. The Ratcliff Diffusion model (figure from Ratcliff, 1978) consists of several parameters:  $z$  (the starting point at which the diffusion process begins),  $v$  (drift rate),  $t_0$  (non-decision time),  $a$  (decision threshold),  $A$  &  $B$  (correct and error response boundaries).



*Figure 4.* An illustration of the Reversal Task used in the current study. Phases 1-3 consist of elemental discriminations, while phases 4-6 are the reversal phases. The rewarded (i.e., correct) response is indicated by the “+” sign, and the incorrect response is indicated by the “-” sign.



*Figure 5.* The Normal-3 Mixture distribution best fit group RT (for all participants).



*Figure 6.* The ex-Gaussian distribution passes visual inspection (for all participants), despite not meeting best-fit criterion, and appears to follow the shape of the reversal task distribution. This suggests its parameters may still offer important characterization of performance.

Table 1. Factor loadings for Diffusion model, ex-Gaussian, and Normal-3 Mixture distribution parameters.

	<b>Mental Efficiency</b>	<b>IIV</b>	<b>Mental Speed</b>	<b>Inattention</b>
a	0.624049	0.164950	-0.174955	0.187493
v	0.076084	-0.342509	0.190213	-0.533621
t <sub>0</sub>	-0.170846	-0.050997	0.800733	-0.059006
Mu	0.241455	-0.066891	0.917949	-0.058406
Sigma	-0.020711	-0.006895	0.720714	-0.019307
Tau	0.388844	0.129123	-0.218435	0.718340
Location1	0.648723	-0.315106	0.599753	0.131810
Location2	0.873291	-0.335329	0.172570	0.099380
Location3	0.945443	0.017413	0.059562	-0.067368
Dispersion1	0.763657	-0.529630	0.020369	0.113133
Dispersion2	0.368218	0.866534	-0.093152	0.043958
Dispersion3	0.397326	0.054469	0.041874	0.729555
Probability1	0.448178	-0.758636	-0.033541	-0.390853
Probability2	-0.279339	0.906403	-0.092083	0.093353
Probability3	-0.459769	-0.124039	0.269023	0.710631

*Table 2.* Mean values for the factor scores entered into the Hierarchical Cluster Analysis.

Cluster	Count	Mental Efficiency	IIV	Mental Speed	Inattention
1	67	-0.71	-0.17	-0.56	-0.17
2	43	0.71	-0.54	-0.22	0.61
3	5	-0.60	0.62	1.47	3.83
4	42	0.28	-0.25	0.94	-0.82
5	25	0.62	1.78	-0.08	-0.25

## References

- Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. *The Oxford Handbook of Memory*, 395-409.
- Brushfield, A. M., Luu, T., Callahan, B., & Gilbert, P. E. (2008). A comparison of discrimination and reversal learning for olfactory and visual stimuli in aged rats. *Behavioral Neuroscience*, 122(1), 54-62.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33, 261-304.
- Chayer, C., & Freedman, M. (2001). Frontal lobe functions. *Current Neurology and Neuroscience Reports*, 1(6), 547-552.
- Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science*, 3(5), 155-158.
- Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(3), 474-479.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135-168.
- Deary, I., & Der, G. (2005). Reaction time, age, and cognitive ability: Longitudinal findings from age 16 to 63 years in representative population samples. *Aging, Neurobiology, and Cognition*, 12(2), 187-215.
- Donders, F. C. (1869). On the speed of mental processes. In W.G. Koster (Ed.), *Attention and Performance II. Acta Psychologica*, 30, 412-431.
- Drag, L. L., & Bieliauskas, L. A. (2010). Contemporary review 2009: Cognitive



- aging. *Journal of Geriatric Psychiatry and Neurology*, 23(2), 75-93.
- Driscoll, I., Hamilton, D. A., Petropoulos, H., Yeo, R. A., Brooks, W. M., Baumgartner, R. N., Sutherland, R. J., 2003. The aging hippocampus: Cognitive, biochemical and structural findings. *Cerebral Cortex*, 13(12), 1344-1351.
- Heathcote, A., Popiel, S. J., Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, 109(2), 340-347.
- Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(4), 598-615.
- Lai, Z. C., Moss, M. B., Killiany, R. J., Rosene, D. L., & Herndon, J. G. (1995). Executive system dysfunction in the aged monkey: Spatial and object reversal learning. *Neurobiology of Aging*, 16(6), 947-954.
- Leth-Steensen, C., Elbaz, Z. K., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica*, 104(2), 167–190.
- Lindsay, J., Laurin, D., Verreault, R., Hebert, R., Helliwell, B., Hill G. B., & McDowell, I. (2002). Risk factors for Alzheimer's disease: A prospective analysis from the Canadian study of health and aging. *American Journal of Epidemiology*, 156(5), 445-453.
- Lockhart, S. N., & DeCarli, C. (2014). Structural imaging measures of brain aging. *Neuropsychology Review*, 24(3), 271-289.
- Luce, R. D. (1986). *Response times*. New York: Oxford University Press.

- Maloney, L. T., & Wandell, B. A. (1984). A model of a single visual channel's response to weak test lights. *Vision Research*, 24(7), 633–640.
- Mell, T., Heekeren, H. R., Marschner, A., Wartenburger, I., Villringer, A., Reischies, F.M. (2005). Effect of aging on stimulus-reward association learning. *Neuropsychologia*, 43(4), 554-563.
- Mewhort, D. J., Braun, J. G., & Heathcote, A. Response time distributions and the Stroop task: A test of the Cohen, Dunbar, and McClelland (1990) model. *Journal of Experimental Psychology. Human Perception & Performance*, 18(3), 872-882.
- Morse, C. K. (1993). Does variability increase with age? An archival study of cognitive measures. *Psychology and Aging*, 8(2), 156-164.
- Osmon, D. C., Kazakov, D., Santos, O., & Kassel, M. T. (2018). Non-Gaussian distributional analyses of reaction times (RT): Improvements that increase efficacy of RT tasks for describing cognitive processes. *Neuropsychology Review*, 28(3), 359-376.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59-108.
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 86(3), 446-461.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20, 873–922.
- Ratcliff, R., & McKoon, G. (2015). Aging effects in item and associative recognition memory for pictures and words. *Psychology and Aging*, 30(3), 669-674.
- Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, 16(2), 323-341.

- Ratcliff, R., Thapar, A., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & Psychophysics*, 65(4), 523-535.
- Ratcliff, R., Thapar, A., McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*, 50, 408-424.
- Ratcliff, R., Thapar, A., McKoon, G. (2006). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin & Review*, 13(4), 626-635.
- Ratcliff, R., Thapar, A., & McKoon, G. (2007). Application of the diffusion model to two-choice tasks for adults 75-90 years old. *Psychology and Aging*, 22(1), 56-66.
- Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology*, 60(3), 127-157.
- Reed, J. M., & Squire, L. R. (1999). Impaired transverse patterning in human amnesia is a special case of impaired memory for two-choice discrimination tasks. *Behavioral Neuroscience*, 113(1), 3-9.
- Rickard, T. C., & Grafman, J. (1998). Losing their configural minds: Amnesic patients fail on transverse patterning. *Journal of Cognitive Neuroscience*, 10(4), 509-524.
- Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society*, 16(5), 754-760.
- SAS Institute Inc. (2015a). *JMP® 12 basic analysis*. Cary: SAS Institute Inc.
- SAS Institute Inc. (2015b). *JMP® 12 specialized models*. Cary: SAS Institute Inc.
- Schoenbaum G., Nugent, S., Saddoris, M. P., & Gallagher, M. (2002). Teaching old rats new tricks: Age-related impairments in olfactory reversal learning. *Neurobiology of Aging*, 23(4), 555-564.

- Sternberg, S., & Backus, B. T. (2015). Sequential processes and the shapes of reaction time distributions. *Psychological Review*, *122*, 830–837.
- Storandt, M., Grant, E. A., Miller, J. P., & Morris, J. C. (2002). Rates of progression in mild cognitive impairment and early Alzheimers disease. *Neurology*, *59*(7), 1034-1041.
- Tapp, P. D., Siwak, C. T., Estrade, J., Head, E., Muggenburg, B. A., Cotman, C. W., & Norton, M. W. (2003). Size and reversal learning in the beagle dog as a measure of executive function and inhibitor control in aging. *Learning & Memory*, *10*(1), 64-73.
- Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging*, *18*(3), 415-429.
- Version 16.0, Neurobehavioral Systems, Inc., Berkeley, CA, [www.neurobs.com](http://www.neurobs.com)
- Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767-775.
- Voytko, M. L. (1999). Impairments in acquisition and reversals of two-choice discriminations by aged rhesus monkeys. *Neurobiology of Aging*, *20*(6), 617-627.
- Wecker, N. S., Kramer, J. H., Hallam, B. J., & Delis, D. C. (2005). Mental flexibility: Age effects on switching. *Neuropsychology*, *(19)*3, 345-352.
- Weilner, J. A., Bellebaum, C., & Daum, I. (2008). Aging effects acquisition and reversal of reward-based associative learning. *Learning & Memory*, *15*, 190-197.