TIP-OF-THE-TONGUE PHENOMENA ACROSS THE ADULT LIFESPAN:
THE ROLE OF CONCURRENT STRESS

by

CHRISTOPHER J. SCHMANK

B.A., Saint Louis University, 2010

A thesis submitted to the Graduate Faculty of the
University of Colorado Colorado Springs
in partial fulfillment of the
requirements for the degree of
Master of Arts
Department of Psychology
2016
This thesis for the Master of Arts degree by

Christopher J. Schmank

has been approved for the

Department of Psychology

by

Lori James, Chair

Michael Kisley

Elizabeth Daniels

Date 11/15/2016
ABSTRACT

The present study extended previous research (James, Schmank, Castro, Hendricks, & Buchanan, 2016) that indicated high-stress testing conditions (i.e., evaluative observation while performing challenging tasks) caused increases in tip-of-the-tongue states (TOTs) compared to low-stress testing conditions. In the current study, young (ages 18-29), middle-aged (ages 30-60), and older adult (ages 61-80 participants were randomly assigned to a high- or low-stress testing condition. Stress was manipulated using the Trier Social Stress Task (TSST; Kirschbaum, Pirke, & Hellhammer, 1993) in the high-stress condition and a placebo-TSST (Het, Rohleder, Schoofs, Kirschbaum, & Wolf, 2009) in the low-stress condition. Following the TSST or placebo-TSST components, participants completed a naming task by providing the full name of a presented famous face (name stimuli) or an English word (non-name stimuli), indicating any TOTs they experienced. For all age groups, participants in the high-stress condition had more TOTs than those in the low-stress condition, using various calculations of TOT rate. There were no main effects of age group, nor did age group ever interact with stress condition. Results indicate that concurrent, psychosocial stress impairs word retrieval irrespective of participant age. Our prediction concerning stress generating transmission deficits leading to increased TOT experiences was supported. Our findings did not support the prediction that TOT experiences would increase with
age. Critically, our results indicate that stress did impact TOT rate, but contrary to our prediction, this impact was same across each age group in the current study.
# TABLE OF CONTENTS

## CHAPTER

I. **INTRODUCTION** ..................................................................................................................1
   - Cognitive Aging ..................................................................................................................1
   - TOT States .........................................................................................................................7
   - Theoretical Basis of Tip-of-the-Tongue States .............................................................18
   - Effects of Stress on Cognitive Process ...........................................................................25
   - TSST and Language Production Research .....................................................................27
   - The Present Experiment .................................................................................................29
   - Predictions ......................................................................................................................30

II. **METHOD** .........................................................................................................................31
   - Participants ......................................................................................................................31
   - Materials .........................................................................................................................33
   - Procedure ........................................................................................................................37
   - Scoring of TOTs ..............................................................................................................40

III. **RESULTS** ........................................................................................................................42
   - Subsidiary Analyses ........................................................................................................46

IV. **DISCUSSION** ....................................................................................................................53

REFERENCES ..........................................................................................................................62

APPENDIX ...............................................................................................................................68
LIST OF TABLES

TABLE

1. Demographic Information of Participants Across Age Groups ..................... 32
2. Mean Trait Anxiety Scores Reported from the STAI (SDs in parentheses) .................................................................................................... 43
3. Mean SRLE Scores (SDs in parentheses) ................................................................. 43
4. Mean Number of Familiar Stimuli (SDs in parentheses) ........................................ 44
5. Mean TOT Responses (SDs in parenthesis) ................................................................. 44
6. Mean TOT Responses as a Proportion of Known Information (SDs in parentheses) ......................................................................................... 45
7. Mean TOT Responses as a Proportion of Unsuccessful Retrievals (SDs in parentheses) ......................................................................................... 45
8. Mean TOT Responses for Name Stimuli (SDs in parentheses) ................................ 46
9. Mean TOT Responses for Names as a Proportion of Known Name Stimuli (SDs in parentheses) ......................................................................................... 47
10. Mean TOT Responses for Names as a Proportion of Unsuccessful Retrievals for Name Stimuli (SDs in parentheses) ......................................................................................... 47
11. Mean TOT Responses for Non-Name Stimuli (SDs in parentheses) ............... 48
12. Mean TOT Responses for Non-Names as a Proportion of Known Non-Name Stimuli (SDs in parentheses) ......................................................................................... 49
13. Mean TOT Responses for Non-Names as a Proportion of Unsuccessful Retrievals for Non-Name Stimuli (SDs in parentheses) ......................................................................................... 49
14. Mean Responses from 1-Item Stress Measure Reported Following TSST (SDs in parentheses) ......................................................................................... 50
15. Mean State Anxiety Scores Reported from the STAI (SDs in parentheses). .51
LIST OF FIGURES

FIGURE

CHAPTER I

INTRODUCTION

Language production is one of the main adaptive processes that set human beings apart from all other animal life. We are able to use language both externally (e.g., in order to communicate our thoughts and feelings to others) and internally (e.g., generating our verbal stream of consciousness). This innate ability typically begins to emerge at a very young age and continues to develop across adolescence and throughout adulthood. Yet we occasionally find ourselves in situations where language production—more specifically word retrieval—becomes difficult, halting, and even faulty. The purpose of the current research was to expand the existing knowledge concerning the tip-of-the-tongue (TOT) state, a memory retrieval failure that occurs during language production (as described previously by Brown & McNeill, 1966; Burke, MacKay, Worthley, & Wade, 1991). The nature of the current project was two-fold. Primarily, we were interested in determining the impact of acute, concurrent social stress on TOTs. Secondly, we were interested in determining if different rates of TOT experiences occurred across samples of individuals representing young, middle-age, and older adulthood. Critically, we were interested in determining whether or not stress demonstrated differential effects on TOTs for people of different ages.

Cognitive Aging

The impact of aging, both physically and mentally, has been well documented (see Craik, 2000, Craik & Byrd, 1982; Craik & Jennings, 1992; Park, 2000; Park &
McDonough, 2013; Salthouse, 1988, 1996, 2000, 2003; 2004; 2010a; 2010b). It is a common belief that the detrimental impact of aging on cognitive abilities does not occur until later during one’s lifetime (Salthouse, 2004). In the early 1920s to the mid-1930s, researchers began to notice striking differences between the cognitive abilities of younger and older adults (see Foster & Taylor, 1920; Jones & Conrad, 1933). These researchers noticed that younger adults were more adaptable and able to perform better than older adults on particular cognitive tasks involving creativity or novelty (e.g., sentence construction, memory for abstract materials); whereas older adults were much more adept compared to younger adults at performing tasks that relied on previously accumulated knowledge (e.g., low-frequency definition generation, logical error detection).

Salthouse (2010a) differentiates between the process and product of individuals’ cognitive abilities. The former term refers to how efficiently one uses cognitive abilities at a specific moment in time (e.g., memory, reasoning, new association formation, spatial visualization, problem solving, and reaction times), whereas the latter term refers to the aggregate knowledge of the many processes completed throughout one’s life (e.g., vocabulary and general knowledge information). Salthouse’s study established that measures of the process of cognitive abilities tend to demonstrate age-related decline beginning in individuals’ mid-20s. This decline revealed that by the time adults reach the age of 70 they score between 0.5 and 2 standard deviations below adults in their 20s and 30s. Additionally, when Salthouse (2010a) included childhood-aged participants this trend became relatively quadratic, demonstrating rapid increases from early childhood to the individuals’ early 20s followed by a nearly linear, progressive reduction in performance throughout middle-age to late-adulthood.
On the other hand, scores on measures assessing the product of cognitive abilities tended to remain relatively stable across the lifespan (Salthouse, 2010a). In measures of general vocabulary (e.g., Wechsler Vocabulary tests) and antonym vocabulary, gradual increases occurred in the early 20s up through the 60s prior to gradual decline through later life. Similar vocabulary measures were collected for children, and performance on these product variables rapidly increased beginning in childhood and continued to increase through the individuals’ early 20s. This gradual increase from early childhood to young adulthood has typically been followed by stability or minimal decline across individuals’ lifespan.

In summary, age-related declines in measures of cognitive functioning can be large: participants in their 20s typically scored within the 75\textsuperscript{th} percentile of the sampled population, whereas individuals in their 70s scored near the 20\textsuperscript{th} percentile (Salthouse, 2004; 2010a). Salthouse demonstrated that age-related declines tend to begin in one’s early 20s and increase across the lifespan. However, the overall impact of aging tends to not be noticed until individuals were in their 40s or 50s (Salthouse, 2010b). Age-related declines were demonstrated in several different cognitive abilities (e.g., process variables such as memory) but not in others (e.g., product variables such as vocabulary).

Park (1997, 1999) argued that the underlying issue in the decline in cognitive abilities for older adults was related to the amount of available resources (e.g., processing speed, working memory function, inhibitory mechanisms, and sensory function). A study conducted by Park and her colleagues (1996) demonstrated that measures of general knowledge (i.e., crystallized intelligence) tended to remain stable across the lifespan; while processing speed, working memory ability and capacity, and cued and free recall
performance tended to exhibit linear declines with age. Park’s findings concerning
cognitive decline in speed of processing and working memory were corroborated more
recently by Hertzog, Dixon, Hultsch, and MacDonald (2003). Additionally, recent
studies have also included measures of problem solving and reasoning, as well as
measures of long-term memory (see Park & Reuter-Lorenz, 2009; Salthouse, 1996). The
general findings from these studies indicate decline with age across all these measures.

Hasher and Zacks originally developed a theory to explain how working memory
performance becomes impaired with age (Inhibitory Deficit Hypothesis [IDH]; Hasher &
Zacks, 1988; Hasher, Zacks, & May, 1999; Zacks & Hasher, 1997). Specifically, these
researchers were interested in interference during working memory tasks. The IDH has
been used more recently to explain broad characteristics of cognitive aging. Under the
IDH, inhibitory processes play a large part in active suppression of any and all
unnecessary information during retrieval. Problems tend to occur under this system when
alternative information actively competes with goal directed information, reducing the
efficiency of retrieval of the targeted information (e.g., Baddeley, 1982; Brown, 1979;
Roediger, 1974; Roediger & Neely, 1982). The IDH suggests that aging causes a
reduction in the ability to inhibit irrelevant information, so that older adults experience
even more negative effects of competing information than young adults. One critique of
the inhibitory view is our understanding of whether or not the differences proposed by
this hypothesis should be ascribed to inhibitory processes or to a failure of proper
activation (Lustig, Hasher, & Zacks, 2007). In spite of this, the IDH remains a popular
hypothesis that has been used to explain a wide range of empirical effects in various areas
of cognitive aging.
Several lines of research have indicated impairments and age-related declines for tasks that involve recollection (Jacoby & Dallas, 1981; Rabinowitz, 1984). Specifically, age related differences have been demonstrated to be small for recognition, yet larger for recall (Perlmutter, 1979). Craik and McDowd (1987) concluded from their experimentation that older adults’ recall performance was worse than younger adults’ when these researchers held recognition scores constant across age groups. In a review paper published by Light (1991), it was concluded that older adults demonstrated impaired responding when compared to younger adults on direct measures of memory (e.g., recall, recognition). On the other hand, Light, Singh, and Capps (1986) established that activation processes remained stable across adulthood. This has been corroborated by research utilizing semantic and repetition priming (e.g., Balota, Duchek, & Paullin, 1989; Graf & Mandler, 1984) which has indicated that activation remains stable across the adult lifespan. Further, when older adults were asked to complete an indirect measure of memory, like repetition priming, no significant differences between the age groups emerged; further bolstering the idea that activation processes are stable across the lifespan (Light, LaVoie, Valencia-Laver, Owens, & Mead, 1992).

Bowles and Poon (1981) were interested in whether aging differentially affects one’s ability to make lexical decisions for high-frequency and low-frequency words. The researchers utilized a lexical decision paradigm in which participants were presented with a string of letters. Following this presentation, participants were required to determine whether the string of letters made up an English word or not. These researchers demonstrated that there were no differences between older and young adults’ speed and accuracy of making a lexical decision. For both age groups, participants responded to
low-frequency words more slowly than high-frequency words. These findings indicated that young and older adults did not differ in their performance during the lexical decision task, and that the only apparent differences were between high- and low-frequency words.

An extension from the previous literature concerned the speed with which adults made lexical decisions. Bowles and Poon (1985) tested how older and younger adults differed on a word retrieval task. This study established that differences between word retrieval performance of older and younger adults was dependent on the prime condition presented to participants (i.e., orthographically related, neutral, unrelated, and semantically related stimuli). Young adults produced words faster than older adults after presentation of each prime condition although the differences were not always statistically significant. Young adults also responded more accurately to stimuli that followed presentation of neutral, unrelated, and semantically related primes. Older adults were fully capable of processing the information within their semantic and lexical networks; however, they moved from these semantic networks to correct lexical access necessary for successful word retrieval and subsequent production slower than young adults. In a review of laboratory studies looking at word-retrieval or naming tasks, Bowles, Obler, and Poon (1989) determined that groups of young and older adults experienced facilitation effects during naming tasks after being presented with orthographically-related cue words. This experiment also demonstrated that interference effects occurred for both semantically related and unrelated cue words in both age groups; however, the older adults’ naming performance was impacted to a greater degree when material was semantically related, consistent with the IDH (described earlier).
TOT States

The current research focused on aspects of word retrieval rather than broad patterns of cognitive aging. Specifically, we were interested in whether TOT experiences occurred differently for three age groups representing the adult lifespan and how psychosocial stress impacted this phenomenon.

Historical Background. Brown and McNeill (1966) were the first to study the TOT phenomenon empirically. However, William James portrayed this particular phenomenon in a highly eloquent manner:

Suppose we try to recall a forgotten name. The state of our consciousness is peculiar. There is a gap therein; but no mere gap. It is a gap that is intensely active. A sort of wraith of the name is in it, beckoning us in a given direction, making us at moments tingle with the sense of our closeness and then letting us sink back without the longed-for term. If wrong names are proposed to us, this singularly definite gap acts immediately so as to negate them. They do not fit into its mould. And the gap of one word does not feel like the gap of another, all empty of content as both might seem necessarily to be when described as gaps. […] The rhythm of a lost word may be there without a sound to clothe it; or the evanescent sense of something which is the initial vowel or consonant may mock us fitfully, without growing more distinct. (James, 1893, pp. 163-164)

This description of the TOT state, albeit flowery, parallels the definition put forth by Brown and McNeill (1966) that TOT states represent conditions under which individuals have knowledge of a word or name they wish to retrieve that ends in a state of agonizing retrieval, prior to momentary production failure.
Brown and McNeill (1966) induced TOT states in the laboratory by utilizing a task that mimicked a reverse dictionary search. Participants listened as definitions of low frequency, English words interspersed with descriptions of famous people and places were read aloud by the experimenter. Upon hearing each definition or description, participants responded with the target word if they knew it. If participants did not know the word they were instructed to do nothing until the next stimulus was verbally presented. Alternatively, if participants felt they were in a TOT state, they were to answer specific questions regarding the target word (e.g., How many syllables does the word have? What is the first letter of word?). Participants were also asked to write down any words that shared the same meaning or contained similar sounds as the word the participants were currently attempting to retrieve.

The results from this groundbreaking study were two-fold: first, it was determined that participants in a TOT state were capable of supplying experimenters with partial information about sought after target words; and second, participants were capable of generating words similar either in sound or meaning to the desired target word. The participants’ ability to retrieve this type of information demonstrated that participants may have semantic information from the target word available during a TOT state, enabling them to generate words with similar semantic properties; more impressively, however, participants were able to generate words that shared similar phonological properties of the word they were attempting to retrieve. One interpretation of reporting this type information would be that participants were in such a state of “knowing” that although successful retrieval was not accomplished, information about the target word was available. The more partial information within the intended target word that
participants produced, the greater the success rate participants reported in resolving the TOT state.

Later a group of researchers collaborated to further the methodology for the empirical study of TOT states. Burke et al. (1991) provided a comprehensive examination of TOT states, and were also the first to explore how aging impacts TOTs. Burke and colleagues utilized three different types of methodology in order to study TOTs. For acquiring information regarding naturally occurring TOTs they utilized retrospective questionnaires and generated a diary study. In order to study TOTs in laboratory settings, word retrieval tasks were generated. The purpose of the research conducted by Burke et al. was to examine TOTs in young and older adults (middle-aged adults were included only in the diary and retrospective questionnaires), while they simultaneously worked to establish a theory that could be used to explain the presence of TOTs (reviewed in a later section of this proposal).

The initial phase of the research conducted by Burke et al. (1991) consisted of a retrospective (i.e., after the fact) questionnaire about the rate of occurrence of TOTs in the lives of young, middle-age, and older adults. Participants were instructed to estimate how frequently they experienced TOT states in their day-to-day lives. These questionnaires were utilized during the second phase of the study, in which the same group of 130 adults kept journals of the frequency of experienced TOT states during a 4-week period. Participants logged their experiences with TOTs while also reporting information concerning the types of words (e.g., proper names, verbs, adjectives, etc.) that elicited the TOT state and other features of these target words (e.g., initial/final syllable, syllabic stress, number of syllables), if available. Utilizing these techniques
Burke et al. were able to determine: (a) the types of words for which participants reported the most TOTs, (b) the different features of TOT words that participants could recall, and (c) helpful strategies for resolution of the TOT state (i.e., successful word retrieval after experiencing a TOT state).

The results from Burke et al. (1991) demonstrated that the frequency of TOT states increased across the age groups. Young adult participants reported significantly fewer TOT states than both middle-aged and older adult participants. Additionally, a hierarchical multiple regression analysis was conducted in order to see if certain variables related to education were better predictors than age. Specifically, they utilized participants’ scores on a standardized vocabulary and digit span test and their overall number of years of completed education, in addition to their ages. This analysis indicated that age was the only significant predictor of naturally occurring TOTs, accounting for 14% of the variance in scores on TOTs. This study also provided information on the different types of words that yielded the most TOT states for their sample. Proper names elicited the highest frequency of reported TOTs, followed by abstract words, then common object names. Within proper names it was determined that names of acquaintances generated the greatest number of TOT states for all age groups. One interesting trend of this diary data indicated that the amount of TOT occurrences for proper names and common object names tended to increase across age, whereas the amount of TOT occurrences for abstract words tended to decrease across age.

Burke et al. (1991) also established that their participants were both familiar with and confident in their knowledge of the word they were attempting to produce during TOT states: middle-aged adults demonstrated the highest degree of confidence followed
by older and young adults, respectively. Participants were generally able to retrieve certain bits of phonological information regarding the word causing the TOT state (e.g., initial phonemes, number of syllables). Younger adults reported greater amounts of partial word information than the middle-aged and older adult participants. Across all three age groups the most helpful method for resolving the TOT state was pop-ups (i.e., the target word came to participants seemingly from nowhere, after a period of time), followed by consultations with others or some type of research, and systematic memory search strategies (i.e., going through each letter of the alphabet). However, resolution method also exhibited an interaction with age. Specifically, older adults resolved more TOTs by pop-ups than middle-aged or young adult participants, whereas middle-aged adults reported a marginal tendency to resolve more TOT states by searching their memories than either young or older adults.

With the results from their diary study in mind, Burke et al. (1991) generated a paradigm to study TOTs in laboratory settings, similar to Brown and McNeill (1966). A subset of participants ($N = 42$) from the previous diary study were chosen to participate in a study that consisted of 100 TOT target definitions, selected from five specific categories (e.g., abstract nouns, object names, adjectives/verbs, names of famous people, and names of famous places) that commonly elicited TOTs during the diary study. One purpose was to determine which of these categories produced the highest frequency of TOT states among the young and older adults.

The procedure was described to participants as similar to a trivia game presented on a computer. Participants were presented with a specific question (e.g., “What is the old name of Taiwan?”) and reported their familiarity with the sought after target word
and the certainty with which they would correctly recall this word. At this point participants would respond either that they knew the word, that they did not know the word, or that they were in a TOT state for that particular word. Each response type led to another set of questions; if participants claimed to know the answer they were to provide the correct target word; however, if their response was incorrect or they claimed not to know the answer, the participants completed a 4-item, forced-choice recognition task. If participants reported being in a TOT state, then a series of questions were presented (e.g., What is the first letter or group of letters in the word? How many syllables are in the word?). This series of questions was utilized to determine if participants had any partial word knowledge available during his or her TOT experience. After filling in partial word information, participants in a TOT state attempted to guess the answer to the target definition; if these participants responded incorrectly the forced-recognition task was presented to participants with a fifth answer choice labeled \textit{None of the above}. If participants responded incorrectly to this question, the computer program would inform them of their incorrect choice, and the correct target word. If participants correctly selected the target word from this recognition task, they were informed of their correct word choice.

Across the board, older adults rated their familiarity with the targets and certainty of recall as higher than younger participants. Likewise, older adults produced larger percentages of “know” and “TOT” responses to target words. However, the amount of “don’t know” responses was highest for young adults. The age differences established by Burke et al. (1991) in their diary study were partially supported in laboratory settings. Specifically, older adults, when compared to younger adults, demonstrated greater
frequencies of TOT states with fewer persistent alternates. Additionally, older adults had
the most TOT states for targets that were descriptions of famous people, and tended to
report higher frequencies of TOT states for object names and certain abstract words (e.g.,
adjectives and verbs) than other abstract words (e.g., non-object nouns) and place names.
However, the data for the young adult sample indicated that the target information that
elicited the most TOT experiences were words that represented object names followed by
abstract words (e.g., non-object nouns, adjectives, and verbs, respectively), and famous
people; place names elicited the fewest TOT states among young adults.

Persistent alternates, examples of which were alluded to previously, are non-target
words that share sounds or meaning with a sought after target word: during a word
retrieval task participants attempting to retrieve the word *pasteurized* may be unable to do
so due to a persistent alternate that shares either semantic and phonological information
with the target word (e.g., *homogenized*). Burke et al. (1991) described these alternates
as instances where a particular word repeatedly reenters one’s mind during retrieval of a
related, target word. Supplementary findings included a decrease in the number of
persistent alternates with increasing age (Burke et al., 1991).

Two particular studies were conducted to establish the role cue words play
regarding TOT states (Jones, 1989; Jones & Langford, 1987). Specifically, these
researchers wanted to establish whether alternate words during TOT experiences fit a
blocking hypothesis originally established by Woodworth (1929). Anecdotally, blocking
is what most individuals believe happens during a TOT experience. Under Woodworth’s
blocking hypothesis, TOT frequencies increase when related information (i.e.,
phonological or semantic) enters the individuals’ attention causing them to feel like
correct information has been blocked from successful production. The studies conducted by Jones (1989) and Jones and Langford (1987) found that TOT frequency was dependent on the type of alternate (e.g., phonological, semantic, phonological and semantic, no relation) presented to participants. Specifically, phonological alternates tended to increase TOT states. This appeared to demonstrate support for the blocking hypothesis.

More recently, researchers replicated the experiments by Jones and Langford in order to determine the correct results regarding alternates during TOT states. Meyer and Bock (1992) were interested in explaining whether TOT states more correctly represented: a) momentary inabilities to access known words due to inhibitory mechanisms or b) weak activation of the targeted, yet inaccessible, word. The first study conducted by Meyer and Bock demonstrated that phonological alternates presented after a definition led to higher correct response frequencies, with no statistical impact on TOT states. However, the trend of the data indicated that the most TOTs occurred after semantically related alternates, followed by unrelated, and then phonological alternates. These data supported a weak activation model of TOT occurrences more than a blocking hypothesis. Specifically, facilitation occurred when phonologically related cues were presented prior to word retrieval, compared to when participants were primed with semantic or unrelated cues; blocking did not occur when participants were presented words of similar sound prior to retrieval. These researchers further established that the results reported by Jones (1989) regarding the blocking hypothesis were biased due to lack of proper stimulus counterbalancing. Meyer and Bock (1992) concluded that the word definitions utilized in the phonological alternate condition of Jones (1989) were
especially susceptible to word retrieval issues: the definitions presented with phonological alternates were the hardest to recall across the entire sample of stimuli, regardless of the presence of phonologically related words.

Similarly, Perfect and Hanley (1992) replicated the protocol of Jones and Langford (1987) with a control group that received no alternates during the TOT task (Experiment 1) and a completely counterbalanced set of word definitions (Experiment 2). The results for the Experiment 1 indicated that the presence of phonological alternates during a TOT task increased TOT states more so than the presence of semantic and unrelated alternates; however, Perfect and Hanley (1992) interpreted these results differently than Jones (1989) and Jones and Langford (1987) because of the results acquired from the group that did not receive alternates. Participants who received no alternates during presentation of this target word list demonstrated approximately equal TOT rates as participants who received the alternates during the same phonological alternate condition. Likewise, the results for Experiment 2 indicated that when word lists were counterbalanced across conditions the presence of phonologically related alternates had no impact on the number of TOTs reported by participants. Thus, the trend established by Jones and Langford (1987) was an artifact of how the definition list was presented to participants, and not due to the presence of phonologically related alternates. The lack of counterbalancing led to having the most difficult targets in the phonologically similar condition, yielding misattributed support for blocking by Jones (1989) and Jones and Langford (1987).

Little research has tested TOTs in middle-aged adult participants. Burke et al. (1991) is one of the only studies to date to collect naturally occurring TOT data from a
group of middle-aged participants. Their data indicated that both older and middle-aged participants reported more TOT experiences over a 4-week period than the younger participants. There were not statistically significant differences between the TOT experiences of middle-age and older adults. These middle-aged individuals also tended to report that the words that produced the highest amounts of natural TOTs were proper names, followed by abstract words, and object nouns. An earlier study was conducted by Cohen and Faulkner (1986) which examined how participants experienced TOTs regarding proper names during their day to day lives. The results from this study indicated that older adults tended to have a greater frequency of misremembered names than participants in either middle-aged or young adult groups. Cohen and Faulkner also demonstrated that older adults tended to have these issues for names that were easy for them to remember or were well known to them. There were no differences between these variables when middle-aged and young adults were compared. These findings were corroborated by one final study conducted by Heine, Ober, and Shenaut (1999). No differences were determined between young adults and a group of young-old adults (i.e., 60-74 years of age), however, both of these age groups responded with statistically significantly fewer TOTs than a group of older adults (i.e., 80-92 years of age). Together, these studies indicate that outside of laboratory settings, TOTs tend to increase with age, however, the results are somewhat varied.

Three studies have included middle-aged adult samples in order to determine what the TOT experiences of middle aged adults are like in laboratory settings (Dahlgren, 1998; Evrard, 2002; Juncos-Rabadán, Facal, Rodríguez, & Pereiro, 2010). Evrard (2002) collected TOT information from a group of young, middle-age, and older adults and
corroborated findings that proper names elicited greater TOT frequencies than common words for all participants (cf., Burke et al., 1991). The only difference in TOT rates was between young and older adult participants—middle-aged adults reported TOT rates in-between these age groups, but were not statistically different from either group.

Dahlgren (1998) conducted a similar study on a group of young, middle-age, and older adults. Dahlgren expected to establish aging effects on the TOT phenomenon like those demonstrated in Burke et al. (1991), while also predicting that older adults’ increased amount of knowledge would facilitate correct retrieval, thereby reducing TOTs. Dahlgren’s general findings did not corroborate the aging results established by Burke and colleagues (1991; i.e., that TOT experiences increase with age). Instead, she found that older adults demonstrated more TOT experiences than members of either the middle-aged or young adult groups, both of which had roughly the same rate of TOTs. However, when TOT experiences were controlled for knowledge (using participants’ WAIS-R vocabulary scores), the main effect of age was nullified. After controlling for knowledge the trajectory of TOT experiences indicated that older adults had the most occurrences, followed by young adults, and finally middle-aged adults, however this trend was not statistically significant.

Lastly, prompted by the idea put forth by Dahlgren (1998), Juncos-Rabadán et al. (2010) studied participants’ word retrieval using a free-recall task on the computer. They used a sample of participants representing the adult lifespan from young (19-26 years of age) to older adulthood (three age groups: 50-59; 60-69; 70-82). Juncos-Rabadán et al. were primarily interested in the relationship between TOT responses and vocabulary knowledge across age groups. Unlike the results of Evrard (2002) and Dahlgren (1998),
Juncos-Rabadán and colleagues (2010) demonstrated that participants in their 50s had higher TOT rates than young adults, and demonstrated similar responding to all older adult age groups. Specifically, older adults produced more correct responses than young adults, and also had larger TOT rates calculated using various conditional equations. The differences between the findings of the Burke et al. (1991), Dahlgren (1998), and Juncos-Rabadán et al. (2010) papers suggest a need for more work testing TOTs across the adult lifespan, with a specific focus on middle-aged participants.

**Theoretical Basis of Tip-of-the-Tongue States**

Following Brown and McNeill (1966), four directions of research were pursued by: (a) memory researchers, (b) theorists concerned with language production, (c) metamemory researchers (i.e., those who focus on the cognitive processes used to monitor memory retrieval), and (d) philosophical researchers (i.e., those who use the TOT state to infer the processes of consciousness; for a summary see Brown, 2012). Researchers concerned with memory have been primarily interested in explaining the limits of the semantic memory system and describing the means of retrieval from the semantic memory store (see Baddeley, 1982; Collins & Loftus, 1975). The current research only focused on the approach from a language production perspective because of the importance of language units (e.g., phonology) to the TOT experience.

The node structure theory (NST) resembles a spreading activation model (cf., Collins & Loftus, 1975) of language production, representing both aspects of language production and perception of language, in addition to memory and other cognitive processes. Like other spreading activation theories of language production (e.g., Dell,
1986) this theory includes a set of mental processing units, called nodes, which are situated hierarchically (See Figure 1).

MacKay (1987) purported that when speaking was concerned within the NST, the hierarchy is followed in a top-down manner: starting with semantic knowledge of the concepts and words we want to produce, followed by the lexical representations and then the complete phonology of each word, finishing with the correct muscle movements necessary to articulate the information. On the other hand, perception of spoken language occurs in a bottom-up manner, with initial input coming from another set of nodes called sensory analysis nodes—used to detect meaningful sounds in one’s environment. These sounds map onto specific phonological units, which when combined form meaningful lexical units (i.e., words). These lexical units become linked to pieces of semantic information under this model, which provides information about the meaning of the words or names being comprehended (MacKay, 1987).

MacKay (1987) distinguished between five specific properties of nodes and their connections: activation, priming, self-inhibition, satiation, and linkage strength. Activation must occur in order for individuals to consciously retrieve the necessary information that a particular node represents for perceptual recognition and/or word production. MacKay (1987) describes activation as a process that occurs in an all-or-none fashion without spreading: the intensity of activation never changes because it either occurs or it does not. Unlike activation, priming does vary in intensity, and prepares all connected nodes for activation. Priming is the transmission of excitation across interconnected nodes and once a node becomes activated it simultaneously primes all of the nodes that are attached both top-down and bottom-up. Without the necessary amount
Figure 1. Nodes representing Frisbee in the semantic, phonological, and muscle movement systems. Reprinted from “On the Tip of the Tongue: What Causes Word Finding Failures in Young and Older Adults?” by D. M. Burke et al., 1991.
of priming, certain phonological nodes for language production do not activate, causing word production failures like the TOT phenomenon. Nodes can never become activated without a sufficient amount of priming, thus priming is a necessary precursor to all of perception and production. Self-inhibition and satiation occur after priming has led to repeated activation, however as they are not relevant to TOT experiences, they will not be discussed further. Lastly, linkage strength refers to the connectivity between nodes. In order to increase linkage strength of a particular node repeated priming and subsequent activation must occur. When nodes are used more frequently and with greater recency, their linkage strength increases. Nodes with strong connections tend to represent highly mastered knowledge and behaviors that can be accessed and activated quickly and efficiently.

The semantic and phonological language production systems are the most relevant of NST components for the current study. The semantic system is made up of propositional and lexical nodes: respectively, these nodes consist of the concepts that are used to give meaning to the words that make up phrases and the precise set of concepts that the target word represents. The lexical system connects top-down with phonological nodes of the semantic system; nodes in this system represent consonant and vowel sounds, which form syllables within the language system. These phonological nodes are interconnected within this theory so that information can flow in either a top-down or bottom-up manner. More specifically, the phonological system connects with the semantic system bottom-up and connects top-down with the muscle movement system (Burke et al., 1991).
Figure 1 presents an example of how the NST might represent the word *Frisbee*. If one directs their attention to the middle set of interconnected phonological nodes (i.e., circles representing the sounds /fr/, /is/, /b/, and /ee/), these nodes converge in a bottom-up fashion at the lexical node for *Frisbee*. When individuals attempt to produce the muscle movements necessary for production of *Frisbee* this information occurs sequentially, in a top-down fashion so that the phonology of *Frisbee* is produced in the correct order. Thus, information flows in both directions under the NST.

In order to successfully produce a name (e.g., Burke et al., 1991; James, 2006), participants must first process the target stimulus (e.g., a famous face or low-frequency word definition). The information presented to participants and any other information that participants know about the target word or name transmits priming between nodes in the semantic network. Priming from the semantic network flows in a top-down fashion and activates the correct lexical representation of the stimulus. Finally, successful transmission of priming from the lexical network to the phonological system activates the complete phonology of the target stimulus, which must occur prior to successful production of the target word or name.

TOTs occur when the correct lexical representation of the stimulus is activated, however the phonology of the target stimulus is incompletely activated. Figure 1 can also be used to demonstrate how TOT experiences might occur both in real life and in laboratory settings. If this was a target word utilized in a typical TOT paradigm, a definition might be used like: What is the name of a circular disc made out of plastic, which is typically thrown between multiple people using a flick of the wrist technique to spin the object? Typically, a TOT occurs if participants are successful in processing the
semantic information they were given and that they know about the target word. Success at this stage activates the lexical node for *Frisbee*, generating a strong feeling that participants know the sought after word. However, if the lexical representation of *Frisbee* does not transmit sufficient top-down priming to each of the connecting phonological nodes, activation of the complete phonology will not be completed, leading to the faltering production.

There is a specific hypothesis instantiated within the NST regarding how aging leads to increased TOT rates: the Transmission Deficit Hypothesis (TDH; see Burke et al., 1991; MacKay, 1987). Burke and colleagues (1991) originally proposed the TDH to explain the general cause of word retrieval failures like the TOT phenomenon. Under the TDH, insufficient transmission of priming from the semantic system to the phonological system causes word retrieval failures; a possible result of this failure is a TOT experience. The TDH proposes that as individuals age, the linkage strength between nodes across the entire network decrease. The TDH also postulates that transmission deficits occur due to decreases in the frequency of specific node usage and the recency of node usage. This reduction in priming due to the influences of recency, frequency, and aging can make node activation impossible; reduced priming across nodes due to infrequent or non-recent usage or general aging is hypothesized to generate the retrieval impairments typical during TOT states. Although transmission deficits are common, they have the most dramatic impact on one-to-one, non-redundant connections. An example of this type of connection can be seen in Figure 1. If there is insufficient priming summating from the lexical node for the target word (*Frisbee*) to any of the phonological nodes within the target word (e.g., the first vowel group /is/), production of *Frisbee*
cannot occur due to a failure in activation. More specifically, production cannot occur if one-to-one connections from the lexical representation of the word to any of the segments of phonological information are weak and do not allow the phonological node to be activated. For the example presented above, no information converges top-down from the lexical representation of *Frisbee* to the initial vowel group, increasing the chance of a language production failure. On the other hand, when we use words frequently, linkage strength tends to increase, making words that are used recently and frequently easier to retrieve and subsequently produce.

Under the TDH, TOT states occur when semantic information (i.e., the definition of a target word or picture of a famous person) activates a specific lexical node in the language system; however, the complete phonology of the targeted word is not available to these individuals. The failure in activation occurs at the phonological level, however, some pieces of phonological information may be activated while others are not. Two possible consequences that are known to occur during TOTs are retrieval of partial phonological information and persistent alternates. Access to these types of partial word information is one of the most prominent aspects of the TOT state under the TDH (Burke et al., 1991). As an example, if participants are given the definition for the word “omnivore” (e.g., “What do you call a creature that eats both animal and plant foods?”) he or she may be unable to produce the complete phonology of the word. Many individuals experiencing a TOT state are able to report the first letter of the target word and/or the number of syllables within the word, information which is represented within the phonological system. However, because of insufficient priming of other phonological nodes, the complete phonology of the word may be absent from consciousness. This
would potentially lead participants to retrieve the information that the target word starts with an “o-”, or that it contains three syllables, but not the entire phonology. Instead, these individuals might only be able to report information that they do know, like that the similar words for meat-eater and plant-eater are “carnivore” and “herbivore”, respectively.

Semantic nodes (e.g., definitions of words or descriptions of people) typically summate top-down to one specific lexical node (e.g., defined word or described person); however, certain aspects of semantic and phonological information can lead to summation of priming to a word related to the target word (i.e., persistent alternate). Persistent alternates are typical during TOTs because when participants experience a TOT state certain phonological nodes and any known semantic information is retrieved; retrieval of this information can cause other words related either in sound or meaning to spontaneously become activated. Due to activation of related but incorrect target words, participants might experience a halt in the production of the targeted word due to the persistent alternate repeatedly emerging in the participants’ consciousness.

**Effects of Stress on Cognitive Processes**

Anecdotally, people report feeling that under high degrees of stress their language production tends to falter. Research has demonstrated negative effects of stress on memory where participants had problems inhibiting irrelevant information during a memory task (McEwen & Sapolsky, 1995). Buchanan, Tranel, and Adolphs (2006) also demonstrated the inhibitory effect of stress on participants’ word recall ability (i.e., ability to produce words previously learned from a list), but not their recognition ability (i.e., ability to recognize words previously learned from a list). On the other hand,
Schwabe and Wolf (2010) demonstrated that stressful components of experimentation reduced the participants’ ability to perform both recall and recognition tasks. Other experiments have demonstrated facilitative effects of stress on memory consolidation (Buchanan & Lovallo, 2001; Roozendaal, 2000; Wolf, 2008). One of the known variables that impacts how stress affects memory is the timing when participants experience the stressor. If stress is implemented immediately after a learning task, participants’ memory is enhanced on subsequent tests (Roozendaal, 2000; Wolf, 2008). When stressful components were implemented prior to a task involving memory, participants experienced detriments to their retrieval abilities (Buchanan et al., 2006; deQuervain, Roozendaal & McGaugh, 1998; Schwabe & Wolf, 2009).

The Trier Social Stress Test (TSST) is a laboratory technique that has been used as a reliable method of increasing salivary cortisol levels (hormone levels that indicate stress) and other physiological signs of stress (Kirschbaum, Pirke, & Hellhammer, 1993). The TSST was developed to induce psychosocial stress responses in laboratory settings, and consists of three separate tasks: (a) a speech preparation task, (b) the speech task, and (c) a mental math task. During the preparation period participants draft a speech regarding a stressful situation (e.g., an important job interview) for a 5-10 min period of time. After participants complete the preparation period, they give their prepared speech for 5 min in front of two confederate observers. Next, participants complete a mental math task, in which they sequentially subtract the number 13 starting with the number 1,022 as quickly and correctly as possible. Using confederates adds aspects of social observation and evaluation, which make both the speech task and mental math task even more difficult and stressful.
One recent study assessed how participants would react to a placebo TSST paradigm (Het, Rohleder, Schoofs, Kirschbaum, & Wolf, 2009). Participants who experienced the placebo version of the TSST performed a simple speech task (e.g., talk about a recently viewed movie; describe a recent vacation spot), followed by mental addition starting at zero. These tasks were completed without confederate observers or the experimenter present. The researchers believed this would remove the effects of social observation and evaluation, as well as reduce the difficulty of the tasks believed to cause elevated stress during the traditional TSST. Het and colleagues demonstrated that their placebo TSST condition was well-suited as a control condition for paradigms utilizing the TSST. This was established utilizing both a between- and within-subjects design. Participants experiencing the placebo TSST reported lower physiological stress (e.g., salivary cortisol levels), greater feelings about their abilities and the control they had during the procedure, and decreased feelings of subjective threat and challenge in both research designs than participants in traditional TSST conditions.

**TSST and Language Production Research**

Previous research utilizing TSST and placebo TSST procedures has tested how stress affects participants’ performance during language production tasks (Buchanan, Laures-Gore, & Duff, 2014; James, Schmank, & Castro, 2013; James, Schmank, Castro, Hendricks, & Buchanan, 2016). Buchanan et al. (2014) were interested in assessing the impact of psychosocial stress on general speech production (e.g., fluency, pauses, word productivity). Buchanan and colleagues employed a modified TSST condition (i.e., stress condition) and a placebo TSST condition. They found reduced speech fluency for participants in the stress condition compared to the placebo group. These speech fluency
impairments also became exacerbated across time during the 5 min task. Buchanan et al. interpret their findings as corroborating anecdotal evidence concerning stress and language production: as stress increases, language production becomes less efficient.

James et al. (2016) conducted a study concerning the impact of stress on word retrieval failures, specifically the TOT phenomenon. James and colleagues used a similar TSST as Buchanan et al. (2014): a high-stress condition received TSST instructions and a low-stress group received placebo TSST instructions. James et al. (2016) predicted that stress would impact word retrieval similarly to how aging impacts word retrieval under the TDH, specifically, by reducing priming to nodes across the participants’ entire network (i.e., decreasing the transmission of priming reduces the ability to successfully recall certain known information). After completion of the TSST procedure, participants completed a naming task consisting of 60 low-frequency English word definitions. They found an increase in TOT frequencies for participants who experienced the high-stress condition compared to the low-stress condition. In an extension of this work, James et al. (2013) compared the effects of these same stress conditions between young and older adult participants. Although correct responding was not impacted by stress condition or age group, participants who were in the high-stress condition reported TOTs more frequently than participants in the low-stress condition. The overall size of the stress effect on TOTs appeared to be the same across age groups in the study conducted by James and colleagues (2013). The present experiment was designed to extend this research.
**The Present Experiment**

To date, the James et al. (2013) study is the only one to test how stress impacts TOT responses for different age groups. There is very little research on middle-aged individuals’ experience with TOTs in general, and no research testing the effects of stress on word retrieval in middle-aged adults. The present research fills these gaps in the literature by including a group of participants between ages 30-60. Using a between-participants design (2 Stress Condition [High, Low] x 3 Age Group [Young, Middle-Age, Older]), the overall purpose of the current research study was to experimentally manipulate psychosocial stress (using the TSST to create a high-stress condition vs. placebo TSST to create a low-stress condition) to determine how stress impacted TOT responses across three age groups representing the adult lifespan. As in James et al. (2016), the present study used the TSST to manipulate stress levels prior to and throughout a naming task. High- and low-stress conditions were compared to assess how salient, concurrent stress impacts TOT incidence across three age groups representing the adult lifespan.

Previous research has demonstrated that pictures of famous people and definitions of low-frequency words effectively elicit TOT states for participants completing a naming task (Burke et al., 1991; James, 2006; Rendell, Castel, & Craik, 2005) and that TOT rates were generally greater for proper names compared to low-frequency words (Burke et al., 1991; Evrard, 2002, Rastle & Burke, 1996). In order to ensure that TOT rates were reasonably high, the current experiment included both pictures of famous people (i.e., name stimuli) and definitions of words (i.e., non-name stimuli).
Predictions

Based on previous research findings and on the TDH account of TOTs (Burke et al., 1991), it was predicted that as individuals increase in age their TOT responses would increase, yielding a main effect of age group on TOTs. Specifically, significant differences were expected between young and older adults, with middle-aged adults’ proportion of TOTs falling in between the young and older adult groups. Based on previous findings and on the extension of the TDH account of TOTs to include a mechanism for the negative impact of stress on TOTs (James et al., 2013), it was predicted that participants in the high-stress condition would report a larger proportion of TOT responses than participants in the low-stress condition, yielding a main effect of stress condition. Additionally, we predicted an interaction between age group and stress condition: within the low-stress condition older adult participants would report the largest proportion of TOT responses followed by middle-aged, and finally young adult participants, and an identical, albeit exacerbated, trend was expected for proportion of TOT responses across age groups for participants in the high-stress condition. In other words, we predicted that increased age would exacerbate the negative effects of the high-stress condition.
CHAPTER II

METHOD

Participants

A sample of 76 adults ranging in age from 18 to 80 ($M = 42.62$, $SD = 21.08$) were conveniently sampled from the Colorado Springs, Colorado area. Participants in the young adult group ranged from 18 to 29, in the middle-aged adult group ranged from 30 to 60, and in the older adult group ranged from 61 to 80. A subset of participants was sampled from the University of Colorado Colorado Springs (UCCS) Psychology department and another subset of participants were sampled from the UCCS Gerontology Center Participant Registry. Participants across three age groups were randomly assigned to either a high- or low-stress condition. The first participant from each age group was assigned a condition determined by a coin flip, with alternating stress conditions for subsequent participants. For example, if the first participant was assigned to the high-stress condition, the second participant would be assigned to the low-stress condition. There were equal numbers of participants in low- and high-stress conditions ($n = 38$) overall and participants were tested individually. Table 1 presents basic demographics for participants in each age group.

Fifty-five participants reported their ethnicity as Caucasian; 12 reported being Hispanic, 3 reported African American, 3 Asian or Pacific Islander, and 3 Other (e.g., 1 of the participants used White/Hispanic/Asian; another responded they were Native
Table 1.

Demographic Information of Participants Across Age Groups

<table>
<thead>
<tr>
<th></th>
<th>Young Adults (n = 28)</th>
<th>Middle-Aged Adults (n = 24)</th>
<th>Older Adults (n = 24)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>M</strong></td>
<td><strong>SD</strong></td>
<td><strong>M</strong></td>
<td><strong>SD</strong></td>
</tr>
<tr>
<td>Age*</td>
<td>21.71</td>
<td>3.17</td>
<td>40.50</td>
</tr>
<tr>
<td>Education</td>
<td>14.79</td>
<td>1.57</td>
<td>15.50</td>
</tr>
<tr>
<td>Vocabulary*</td>
<td>28.46</td>
<td>3.88</td>
<td>32.83</td>
</tr>
</tbody>
</table>

*Note. * \(p < .001\). Each age group mean was significantly different from both of the others for both age in years and vocabulary score.

American/Caucasian, and the last responded Mixed-Hispanic/Caucasian). Seven participants reported having a native tongue different from English: five were originally Spanish speaking, one Kurdish, and one Japanese. All 7 of these participants reported that presently English was the primary language spoken in their household, and indicated having spoken English for at least 10 years prior to experimentation.

Participants in the older adult group completed a measure of cognitive ability in order to determine if they were eligible as cognitively healthy older adult participants. The range of scores from this inventory was between 27 and 30 \((M = 28.55, SD = 0.86)\), and elimination criteria were scores less than 26 (Folstein, Folstein, & McHugh, 1975). No participants were excluded for having low scores on this measure (more regarding this inventory will be in the Materials section).

Participants sampled from UCCS were compensated with 2.5 points of extra credit to be applied to a course of their choice. Participants that were sampled from the UCCS Gerontology Center Participant Registry were compensated $15 for their participation. All participants completed informed consent prior to experimentation.
Materials

Laboratory space. Two laboratory spaces were used during experimentation. Room A was equipped with a desk situated against one wall and a small table against the other where participants were seated while filling out paperwork as well as preparing for upcoming tasks to be completed in the other laboratory space (Room B). Room B contained a table situated in the middle of the room with two chairs. One chair was placed in the middle of the table facing a set of blinds and the other was placed at one end of the table facing the opposite end. Behind the blinds was a one-way mirror used only during the high-stress condition to increase participants’ feelings of social observation and the possibility of evaluation or analysis.

Stress manipulation and recording devices. The stress manipulation utilized in this study closely followed that of James and colleagues (2013; 2016). A Philips LFH 0655 digital audio recorder was used during both high- and low-stress conditions to record participants’ speeches and TOT naming task responses electronically. In the high-stress condition a Sony Handycam DCR-DVD308 video recording device and an auxiliary Harman/Kardon Model HK695-01 computer speaker served as props to increase the degree of stress experienced by participants in this condition; the Handycam was never used to record any part of the experiment. Additionally, experimenters during the high-stress condition wore a white lab coat and responded to participants using a stoic and expressionless manner, to further increase participants’ stress experience.

Psychological measures and surveys. A consent form informed participants of the gist of the current research paradigm. However, the true purpose of the research was not revealed to participants until the end of the experiment. The consent form also
ensured participants that their participation would remain anonymous and confidential. Participants self-reported demographic information concerning: their age on the date of participation; gender; handedness; number of years of school completed and their highest degree awarded; whether participants are native English speakers (i.e., ensuring that English was the participants’ first language); ethnicity; and health and medical conditions (e.g., stroke, heart attack, major brain injury, and hearing or vision problems).

**State-Trait Anxiety Inventory.** The State-Trait Anxiety Inventory (STAI) for adults created by Spielberger (1983) was used to establish participants’ state- and trait-anxiety levels. State-anxiety refers to how anxious or stressed participants feel at that moment in time. This 20-item inventory contains questions that are worded both positively and negatively (e.g., “I feel secure,” “I am worried”) which participants respond to using a 4-point Likert type scale ranging from 1 (Not at all) to 4 (Very much so). Thus, the lowest score possible is 20 and the greatest score possible is 80, reflecting extremely low to extremely high state-stress. Trait-anxiety refers to how participants feel generally using 20 items (e.g., “I am ‘calm, cool, and collected,’” “I feel like a failure”). Participants respond to each item using a 4-point Likert type scale ranging from 1 (Almost never) to 4 (Almost always). Again scores range from 20 (extremely low trait-anxiety) to 80 (extremely high trait-anxiety). Metzger (1976) established high test-retest reliability for the trait form ($r = .97$) and medium test-retest reliability for the state form ($r = .45$) of the original version of the STAI (Spielberger, Gorsuch, & Lushene, 1970). Similarly, Ramanaiah, Franzen, and Schill (1983) demonstrated that both forms of the revised STAI (Spielberger, 1983) had high internal consistency measured using
Cronbach’s alpha (\(\alpha = .90\)) with high intercorrelations (\(r = .70\)) between the state-anxiety and trait-anxiety inventories.

**Shipley Vocabulary Test.** The Shipley Institute of Living Scale (SILS; Shipley, 1946) contains a 40-item vocabulary test. Participants completed this inventory by reading a specific word and selecting its closest synonym from a list of one related and three unrelated words (e.g., for *PLAGIARIZE* the correct response is *appropriate*; it is neither *intend*, *revoke*, nor *maintain*). Shipley (1946) tested the reliability of the SILS by obtaining responses from 322 army recruits. The internal consistency established from this study for the vocabulary test was .87, demonstrating strong internal consistency. Goodman, Streiner, and Woodward (1974) established that the test-retest reliability of the vocabulary portion of the SILS did not change with multiple administrations across individuals: the test-retest reliability coefficient was calculated as .87.

**TOT naming task.** The naming task was presented to participants using Microsoft PowerPoint on a MacBook Pro. This set of stimuli contained 40 definitions of low-frequency English words phrased as questions (i.e., non-name stimuli; based on studies conducted by Burke et al., 1991 and James et al., 2013) and 40 pictures of famous people (i.e., name stimuli; based on a study conducted by James, 2006). These items were presented in an intermixed, alternating format (i.e., participants either viewed a non-name stimulus item first followed by a name stimulus item or vice versa). Four stimulus sets were constructed, two began with a non-name stimulus item presented before a name stimulus item; and two beginning with a name stimulus item before a non-name stimulus item. Within each set, order of items was randomly determined once and was the same for all participants.
**Familiarity ratings.** After completion of the naming task participants judged their familiarity with each of the low-frequency words and famous faces previously presented to them in the naming task. Using a procedure similar to Burke et al. (1991) participants responded to each target non-name stimulus item or name stimulus item of the naming task on a 6-point Likert-type scale ranging from 1 (*Not at all*) to 6 (*Extremely familiar*). For example, if participants had never heard of the non-name stimulus item *Kleptomaniac* they would mark their familiarity as a 1.

**Mini Mental State Exam.** The Mini Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975) is the inventory, mentioned previously, that measures the cognitive abilities of older adults. Specifically, participants respond to information concerning: orientation with present time and current location (e.g., “What is the current year,” “In what state are we in”), attention and calculation techniques (e.g., ask participants to serially subtract the number 7 from 100), word recall, phrase repetition (e.g., have participants listen and repeat phrase “No ifs, ands or buts”), and comprehension of and subsequent execution of complex commands (e.g., naming of familiar household items, sentence generation, and figure copying). Scores range from 0 to 30. Mitrushina and Satz (1991) described evidence of convergent validity from correlations of the MMSE with 19 other tasks that measured intelligence quotients, verbal ability, logical and critical thinking, and memory tasks. The MMSE correlated with tests that measured these abilities ranging from small to medium positive associations (*rs* = .28 – .46).

**One-item stress ratings.** A 1-item stress rating was implemented six times across both the high- and low-stress conditions of the current study to access participants’
perceived stress throughout the experiment. The item asked participants, “How stressed do you feel currently?” Participants responded using a 7-point Likert type scale ranging from 1 (Extremely low stress) to 7 (Extremely high stress).

**Survey of Recent Life Experiences (SRLE).** The SRLE was developed by Kohn and MacDonald (1992) to assess hassles that participants have experienced over the past month. For example, participants answer items regarding five specific factors: social and cultural difficulties (e.g., Being let down or disappointed by friends), work (e.g., Dissatisfaction with work), time pressure (e.g., Too many things to do at once), finances (e.g., Financial burdens), and social acceptability (e.g., Being ignored). Participants responded to each item based on how often that experience had occurred over the last month using a 4-point Likert type scale ranging from 1 (Not at all part of my life) to 4 (Very much part of my life). de Jong, Timmerman, and Emmelkamp (1996) calculated the internal consistency of the complete version of the SRLE to be highly satisfactory (α = .89) and each of the five factors exhibited Cronbach’s α values greater than .70.

**Procedure**

All participants entered Room A. Participants received a consent form to establish whether or not they would participate in a study entitled “Non-Verbal Behavior During Language Production.” This title was utilized to conceal the true purpose of the study from participants. After providing informed consent, participants completed the demographic form, SILS Vocabulary form, and the trait-form of the STAI. Participants were then given a blank piece of paper to prepare either for the high- or low-stress condition. The main differences between the stress conditions were: the specific directions for the TSST; the use of prop recording devices; presence of an experimenter
dressed in a white lab coat; and the supposed presence of an evaluative observer in the high-stress condition.

**High-stress condition.** Participants began the TSST by preparing for an important job interview by drafting a speech for a 5 min period. Participants were told that their speeches would have to last for 5 min. No specific job was given to participants. If participants asked what job they were applying for, they were prompted to choose a job that would be vital to their personal growth. A stopwatch feature on an iPhone was used to track the 5 min period allotted for this portion of the task. After participants finished the preparation period their speech draft was collected and participants completed the first 1-item stress rating before they followed the researcher to Room B.

Once in Room B, participants remained standing while the experimenter arranged the room: exposing the one-way mirror, setting up the audio recorder, computer, and video recorder. Participants delivered their speech standing, facing the one-way mirror while the experimenter sat in the back corner of the room. Prior to the speech task a recording of a man’s voice was played via iTunes from the computer. It informed participants that an expert on nonverbal behavior was on the other side of the one-way mirror observing the participants’ movements throughout the remainder of the experiment. After the recording was played, it was reiterated that the participants’ speech should last for 5 min. If participants failed to speak for the entire 5 min period, they were informed that there was still time remaining to complete their speech. If participants stopped speaking for a second time before the 5 min period ended, the experimenter waited approximately 20 s before pre-arranged questions were asked (e.g., What are your
personal strengths? What do you think about teamwork? What are your major shortcomings?). After the speech task, participants were informed of the mental arithmetic portion of the TSST. Participants serially subtracted the number 13 from 1022 mentally and responded out loud as quickly and correctly as possible. If participants made a mistake during this task they were informed of their error and referred to their most recent correct response. This task lasted 5 min. After the TSST, participants completed their second 1-item stress rating prior to the naming task.

**Low-stress condition.** Similar to the procedure followed by the participants in the high-stress condition, participants in the low-stress condition completed a 5 min preparation period followed by a 5 min speech. Unlike the high-stress condition, participants drafted a speech describing their favorite vacation spot and the types of activities they typically do there. The same stopwatch feature was used to measure all timed aspects of this condition. Prior to moving to Room B participants completed their first 1-item perceived stress rating.

Once in Room B, participants sat while the experimenter set up the audio recorder. After participants received instructions for the speech task the researcher left the room for the duration of the speech. Participants delivered their speech for 5 min, uninterrupted and with no intervention if they stopped early. Afterwards, the same mathematics task used in the high-stress condition was completed using paper and pen. Again, the experimenter removed himself during completion of this task. Participants were free to use whatever strategies made this task easiest for participants in the low-stress condition. After the arithmetic task was completed participants responded to their second 1-item stress rating before beginning the naming task.
**Naming task.** The naming task was presented electronically on a personal laptop computer. Participants read instructions regarding the naming task from a PowerPoint presentation. The task required participants to respond with the correct word corresponding to either a presented non-name stimulus item (e.g., responding with the word *abacus* when presented with “What do you call an instrument for performing calculations by sliding beads along rods or grooves?”) or the full name of the person pictured for name stimulus items (e.g., responding *John Travolta* when presented with a picture of his face). Participants were informed that during this naming task they might experience a TOT state, which was described as a state where they know the particular response but they cannot currently produce the correct answer. During the naming task participants in both the high- and low-stress conditions completed the last four 1-item stress ratings after responding to items numbered 20, 40, 60, and 80. Participants in the high-stress condition completed the naming task under supposed observation.

Finally, participants completed the state-anxiety form of the STAI, the SRLE, and reported familiarity ratings for each name and non-name stimulus item presented during the experiment. All adults over the age of 60 additionally completed the MMSE. These forms were removed and participants read through the debriefing form, before being thanked, and compensated for their participation.

**Scoring of TOTs**

Participants’ responses were scored and then calculated as proportions. Correct responding occurred when participants reported target name or non-name information without mentioning TOT states. Each non-name stimulus was correctly responded to using one predetermined English word, while each name stimulus was correctly
responded to using both the complete and correct first and last name of the individual. TOTs were scored when participants reported being in a TOT state, and later indicated that their TOT was indeed for the intended name or non-name target (i.e., we included only what are sometimes called “positive TOTs”). Using each individual participant’s familiarity ratings for each stimulus presented during the naming task, we removed target items with which a participant was not familiar. Specifically, we retained only items for which a participant reported being at least Somewhat familiar (i.e., any stimulus for which a participant reported being only Slightly familiar or Not at all familiar was omitted from the analysis). The mean number of items removed on this basis was 20.49 (SD = 15.32), or 25.61% of trials. This enabled us to conduct analyses of non-name and name retrieval using only stimuli that were reported as familiar by the participant, on a participant-by-participant basis. Thus, proportion of TOTs was the number of TOTs for each participant divided by the number of trials with familiar stimulus items for that participant. Two other dependent measures were computed for TOTs: 1) TOTs as a proportion of known information, calculated by dividing TOTs by the sum of correct and TOT responses and 2) TOTs as a proportion of unsuccessful retrievals, calculated by dividing TOTs by the total number of familiar stimulus items for each participant minus their number of correct responses (see Juncos-Rabadán et al., 2010; James & Burke, 2001 for examples of research using these conditional TOT measures).
CHAPTER III
RESULTS
In order to determine whether random assignment yielded similar participant
groups in terms of baseline stress, separate (2: Stress Condition x 3: Age Group) factorial
analyses of variance (ANOVAs) were conducted on STAI trait-anxiety scores and SRLE
scores, with descriptive statistics for these measures presented in Tables 2 and 3. No
differences were established between stress conditions for trait-anxiety scores, \( F(1, 69) =
1.40, p = .24, \eta^2_p = .02 \) (1 participant failed to complete the trait-anxiety portion of the
STAI, reflected in the reduced degrees of freedom for this analysis), or SRLE scores, \( F(1,
56) = 1.19, p = .28, \eta^2_p = .02 \) (14 participants failed to complete the SRLE, reflected in the
reduced degrees of freedom for this analysis). Main effects were established for age
group on both trait-anxiety, \( F(2, 69) = 5.52, p = .006, \eta^2_p = .14 \), and SRLE scores, \( F(2,
56) = 4.98, p = .01, \eta^2_p = .15 \). Post hoc comparisons for the main effect of age group were
corrected using the Gabriel procedure due to the small group sizes. These indicated that
older adult participants reported lower trait-anxiety and lower SRLE scores than middle-aged
\( p = .03 \) and \( p = .02 \), respectively) and young adult participants \( p = .007 \) and \( p =
.02 \), respectively), with no differences between young and middle-aged participants.
There was no interaction between stress condition and age group for trait-anxiety scores,
\( F(2, 69) = 0.28, p = .75, \eta^2_p = .008 \); nor was there an interaction between stress condition
and age group for the SRLE, \( F(2, 56) = 0.23, p = .79, \eta^2_p = .008 \).
Table 2.

*Mean Trait Anxiety Scores Reported from the STAI (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>39.67 (9.18)</td>
<td>35.82 (13.26)</td>
<td>29.75 (8.95)</td>
<td>35.42 (11.00)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>41.46 (11.41)</td>
<td>41.69 (15.41)</td>
<td>31.32 (6.59)</td>
<td>38.52 (12.54)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>40.50 (10.12)</td>
<td>39.00 (14.46)</td>
<td>30.50 (7.77)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.

*Mean SRLE Scores (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>72.21 (14.50)</td>
<td>71.29 (29.87)</td>
<td>58.30 (9.57)</td>
<td>67.52 (18.41)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>79.58 (16.67)</td>
<td>79.91 (29.40)</td>
<td>58.81 (7.13)</td>
<td>74.34 (22.10)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>75.52 (15.67)</td>
<td>76.56 (29.01)</td>
<td>58.53 (8.34)</td>
<td></td>
</tr>
</tbody>
</table>

A subsequent factorial ANOVA was conducted on total number of stimulus items that were rated as familiar by participants, which indicated a main effect of age group, $F(2, 70) = 7.50, p = .001, \eta^2_p = .18$. Young adults reported being familiar with fewer of the stimulus items used in the present study when compared to both middle-aged ($p = .002$) and older adult ($p = .01$) participants, these older two groups were not statistically different from one another. There was neither a main effect of stress condition ($F(1, 70) = 0.20, p = .66, \eta^2_p = .003$) nor an interaction between stress condition and age group ($F(2, 70) = 1.59, p = .21, \eta^2_p = .04$). Descriptive statistics for this analysis are presented in Table 4.
Table 4.

*Mean Number of Familiar Stimuli (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>53.13 (10.51)</td>
<td>62.00 (17.57)</td>
<td>66.50 (14.44)</td>
<td>59.92 (14.88)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>49.38 (14.50)</td>
<td>68.69 (14.09)</td>
<td>59.25 (13.73)</td>
<td>59.11 (15.94)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>51.39 (12.43)</td>
<td>65.62 (15.79)</td>
<td>62.88 (14.27)</td>
<td></td>
</tr>
</tbody>
</table>

**Tests of Predicted Effects**

Factorial ANOVAs were conducted on TOTs, on TOTs as a proportion of known information, and on TOTs as a proportion of unsuccessful retrievals. The 2 x 3 ANOVA on TOTs indicated a main effect of stress condition, $F(1, 70) = 8.06, p = .006, \eta^2_p = .10$. There was neither a main effect of age group, $F(2, 70) = 0.11, p = .89, \eta^2_p < .01$, nor an interaction between stress condition and age group for TOTs, $F(2, 70) = 0.01, p = .99, \eta^2_p < .001$. Descriptive statistics for this analysis are presented in Table 5.

Table 5.

*Mean TOT Responses (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>.06 (.04)</td>
<td>.05 (.03)</td>
<td>.06 (.06)</td>
<td>.06 (.04)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>.03 (.03)</td>
<td>.03 (.04)</td>
<td>.03 (.02)</td>
<td>.03 (.03)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>.05 (.04)</td>
<td>.04 (.04)</td>
<td>.04 (.04)</td>
<td></td>
</tr>
</tbody>
</table>

The 2 x 3 ANOVA on TOTs as a proportion of known information indicated a main effect of stress condition, $F(1, 70) = 6.52, p = .01, \eta^2_p = .09$. There was neither a main effect of age group, $F(2, 70) = 1.60, p = .21, \eta^2_p = .04$, nor an interaction between stress condition and age group for TOTs as a proportion of known information, $F(2, 70) = 0.99, p = .38, \eta^2_p = .03$. Descriptive statistics for this analysis are presented in Table 6.
Table 6.

Mean TOT Responses as a Proportion of Known Information (SDs in parentheses).

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>.13 (.13)</td>
<td>.07 (.04)</td>
<td>.08 (.07)</td>
<td>.09 (.09)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>.05 (.06)</td>
<td>.04 (.06)</td>
<td>.05 (.04)</td>
<td>.05 (.05)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>.09 (.10)</td>
<td>.05 (.05)</td>
<td>.06 (.06)</td>
<td></td>
</tr>
</tbody>
</table>

Finally, the 2 x 3 ANOVA on TOTs as a proportion of unsuccessful retrievals indicated a main effect of stress condition, $F(1, 70) = 9.36, p = .003, \eta_p^2 = .12$. There was also a marginal main effect of age group, $F(2, 70) = 2.87, p = .06, \eta_p^2 = .08$. Pairwise comparisons conducted using the Bonferroni correction indicated marginal differences between young and middle-aged ($p = .07$) adults; no differences existed between middle-aged and older adults, nor between young and older adults. There was also no interaction between stress condition and age group for TOTs as a proportion of unsuccessful retrievals, $F(2, 70) = 2.32, p = .11, \eta_p^2 = .06$. Descriptive statistics for this analysis are presented in Table 7. To summarize, all three TOT measures indicated increased TOTs in the high stress compared to low stress condition, but no statistically significant effects of age group nor interactions between stress condition and age group.

Table 7.

Mean TOT Responses as a Proportion of Unsuccessful Retrievals (SDs in parentheses).

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>.14 (.10)</td>
<td>.33 (.26)</td>
<td>.17 (.17)</td>
<td>.20 (.19)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>.10 (.11)</td>
<td>.11 (.09)</td>
<td>.11 (.15)</td>
<td>.11 (.11)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>.12 (.10)</td>
<td>.21 (.22)</td>
<td>.14 (.16)</td>
<td></td>
</tr>
</tbody>
</table>
Subsidiary Analyses

**Name stimuli only.** To determine whether our main results were obtained for both name and non-name stimuli, we also conducted factorial ANOVAs on the three TOT proportions presented above, separately for name and non-name stimuli. The 2 x 3 ANOVA for name TOTs indicated a main effect of stress condition, $F(1, 70) = 11.36, p = .001, \eta^2_p = .14$. There was neither a main effect of age group, $F(2, 70) = 0.27, p = .76, \eta^2_p = .01$, nor an interaction between stress condition and age group for TOTs, $F(2, 70) = 0.05, p = .95, \eta^2_p = .001$. Descriptive statistics for this analysis are presented in Table 8.

Table 8.

**Mean TOT Responses for Name Stimuli (SDs in parentheses).**

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>.08 (.07)</td>
<td>.08 (.06)</td>
<td>.07 (.09)</td>
<td>.08 (.07)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>.03 (.06)</td>
<td>.04 (.06)</td>
<td>.02 (.03)</td>
<td>.03 (.05)</td>
</tr>
<tr>
<td>Average across</td>
<td>.05 (.07)</td>
<td>.06 (.07)</td>
<td>.05 (.07)</td>
<td></td>
</tr>
<tr>
<td>Stress Conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The 2 x 3 ANOVA on name TOTs as a proportion of known information indicated a main effect of stress condition, $F(1, 70) = 14.36, p < .001, \eta^2_p = .17$. There was neither a main effect of age group, $F(2, 70) = 1.29, p = .28, \eta^2_p = .04$, nor an interaction between stress condition and age group for TOTs out of known information, $F(2, 70) = 1.60, p = .21, \eta^2_p = .04$. Descriptive statistics for this analysis are presented in Table 9.
Table 9.

*Mean TOT Responses for Names as a Proportion of Known Name Stimuli (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>.20 (.16)</td>
<td>.11 (.09)</td>
<td>.12 (.13)</td>
<td>.14 (.13)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>.04 (.08)</td>
<td>.06 (.10)</td>
<td>.03 (.05)</td>
<td>.05 (.08)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>.12 (.15)</td>
<td>.08 (.10)</td>
<td>.08 (.11)</td>
<td></td>
</tr>
</tbody>
</table>

Finally, the 2 x 3 ANOVA on name TOTs as a proportion of unsuccessful retrievals indicated a main effect of stress condition, $F(1, 70) = 12.10, p = .001, \eta^2_p = .15$.

There was no main effect of age group, $F(2, 70) = 2.23, p = .12, \eta^2_p = .06$. and no interaction between stress condition and age group for TOTs as a proportion of unsuccessful retrievals, $F(2, 70) = 2.03, p = .14, \eta^2_p = .06$. Descriptive statistics for this analysis are presented in Table 10. In summary, analyses using only name stimuli replicated the patterns we obtained for all stimuli: There were only main effects of stress condition, and no effects of age group nor interactions between stress condition and age group.

Table 10.

*Mean TOT Responses for Names as a Proportion of Unsuccessful Retrievals for Name Stimuli (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>.13 (.13)</td>
<td>.35 (.24)</td>
<td>.24 (.31)</td>
<td>.23 (.24)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>.09 (.19)</td>
<td>.10 (.15)</td>
<td>.05 (.11)</td>
<td>.08 (.15)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>.11 (.16)</td>
<td>.22 (.23)</td>
<td>.15 (.25)</td>
<td></td>
</tr>
</tbody>
</table>
**Non-name stimuli only.** Factorial ANOVAs were also conducted for non-name stimuli using the TOT measures above. The 2 x 3 ANOVA for non-name TOTs indicated no main effects of stress condition, $F(1, 70) = 0.25, p = .62, \eta^2_p = .004$, or age group, $F(2, 70) = 0.71, p = .50, \eta^2_p = .02$. There was also no interaction between stress condition and age group, $F(2, 70) = 0.46, p = .63, \eta^2_p = .01$. Descriptive statistics for this analysis are presented in Table 11.

Table 11.

*Mean TOT Responses for Non-Name Stimuli (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>.05 (.04)</td>
<td>.03 (.04)</td>
<td>.03 (.03)</td>
<td>.04 (.04)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>.03 (.06)</td>
<td>.03 (.02)</td>
<td>.03 (.04)</td>
<td>.03 (.04)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>.04 (.05)</td>
<td>.03 (.03)</td>
<td>.03 (.04)</td>
<td>.03 (.04)</td>
</tr>
</tbody>
</table>

The 2 x 3 ANOVA on non-name TOTs as a proportion of known information indicated no main effects of stress condition, $F(1, 70) = 0.70, p = .41, \eta^2_p = .01$, or age group, $F(2, 70) = 1.80, p = .17, \eta^2_p = .05$. Again, no interaction was found between stress condition and age group for this TOT measure, $F(2, 70) = 1.51, p = .23, \eta^2_p = .04$. Descriptive statistics for this analysis are presented in Table 12.

Finally, the 2 x 3 ANOVA on non-name TOTs as a proportion of unsuccessful retrievals indicated a marginally significant main effect of stress condition, $F(1, 70) = 2.94, p = .09, \eta^2_p = .04$. There was no main effect of age group, $F(2, 70) = 0.56, p = .57, \eta^2_p = .02$. There was also no interaction between stress condition and age group for TOTs as a proportion of unsuccessful retrievals, $F(2, 70) = 1.32, p = .27, \eta^2_p = .04$. Descriptive statistics for this analysis are presented in Table 13. In summary, analyses using only
Table 12.

*Mean TOT Responses for Non-Names as a Proportion of Known Non-Name Stimuli (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>.10 (.13)</td>
<td>.03 (.05)</td>
<td>.04 (.05)</td>
<td>.06 (.09)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>.05 (.08)</td>
<td>.04 (.03)</td>
<td>.05 (.05)</td>
<td>.04 (.06)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>.08 (.11)</td>
<td>.04 (.04)</td>
<td>.04 (.05)</td>
<td></td>
</tr>
</tbody>
</table>

Table 13.

*Mean TOT Responses for Non-Names as a Proportion of Unsuccessful Retrievals for Non-Name Stimuli (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>.19 (.25)</td>
<td>.27 (.34)</td>
<td>.11 (.12)</td>
<td>.19 (.25)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>.09 (.14)</td>
<td>.10 (.07)</td>
<td>.14 (.23)</td>
<td>.11 (.15)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>.14 (.21)</td>
<td>.17 (.23)</td>
<td>.12 (.18)</td>
<td></td>
</tr>
</tbody>
</table>

non-name stimuli failed to replicate the patterns we obtained for all stimuli: There were no main effects of stress condition for this subset of stimuli. As with name stimuli, there were no main effects of age group, and no interactions between stress condition and age group.

**Manipulation Checks.** There were main effects of stress condition in the overall analyses of all three TOT measures, and also in the analyses including only name stimuli. To determine whether participants reported experiencing increased stress in response to our manipulation, we analyzed their responses to the 1-item stress measure given throughout the experiment and to the state portion of the STAI given at the end of the
session. Single-item stress ratings were analyzed using six 2 x 3 factorial ANOVAs, one for each measurement taken throughout the procedure. Across all six of these analyses, there was only one significant main effect of stress condition, \( F(1, 70) = 4.14, p < .05, \eta^2_p = .06 \), which occurred just after participants completed the speech and arithmetic tasks, and prior to beginning the naming task. Descriptive statistics for this measurement are presented in Table 14. No other ANOVA on these ratings yielded any main effects or interactions.

Table 14.

*Mean Responses from 1-Item Stress Measure Reported Following TSST (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Stress</td>
<td>3.60 (1.60)</td>
<td>3.73 (1.90)</td>
<td>4.25 (1.42)</td>
<td>3.84 (1.62)</td>
</tr>
<tr>
<td>Low-Stress</td>
<td>3.46 (1.20)</td>
<td>3.15 (0.99)</td>
<td>2.92 (1.51)</td>
<td>3.18 (1.23)</td>
</tr>
<tr>
<td>Average across Stress Conditions</td>
<td>3.54 (1.40)</td>
<td>3.42 (1.47)</td>
<td>3.58 (1.59)</td>
<td></td>
</tr>
</tbody>
</table>

The 2 x 3 factorial ANOVA on state-anxiety scores found no main effect of stress condition, \( F(1, 70) = 0.18, p = .68, \eta^2_p = .003 \), nor a main effect of age group, \( F(2, 70) = 2.15, p = .12, \eta^2_p = .06 \). There was also no interaction between stress condition and age group, \( F(2, 70) = 0.05, p = .95, \eta^2_p = .001 \). Descriptive statistics for this measure are presented in Table 15. To summarize, our manipulation checks indicate that participants did not report differences in their stress levels in response to our manipulation, with the exception of immediately following the TSST speech and math tasks.
Table 15.

*Mean State Anxiety Scores Reported from the STAI (SDs in parentheses).*

<table>
<thead>
<tr>
<th></th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
<th>Average across Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High-Stress</strong></td>
<td>40.53 (11.61)</td>
<td>36.09 (14.00)</td>
<td>35.83 (10.98)</td>
<td>37.76 (12.04)</td>
</tr>
<tr>
<td><strong>Low-Stress</strong></td>
<td>40.54 (7.63 )</td>
<td>34.85 (9.14)</td>
<td>33.92 (11.43)</td>
<td>36.50 (9.68)</td>
</tr>
<tr>
<td><strong>Average across Stress Conditions</strong></td>
<td>40.54 (9.79)</td>
<td>35.42 (11.36)</td>
<td>34.88 (11.00)</td>
<td></td>
</tr>
</tbody>
</table>

**Correlation Analyses.** Bivariate correlations were conducted using Pearson product correlation coefficients ($r$) for all participants across the experiment. Regarding demographic variables, age was positively correlated with both years of education ($r = .35, p = .002$) and vocabulary scores ($r = .69, p < .001$), and years of education and vocabulary scores correlated positively ($r = .44, p < .001$). Negative correlations were found between age and trait-anxiety ($r = -.46, p < .001$), state-anxiety ($r = -.29, p = .01$), and SRLE scores ($r = -.45, p < .001$). Years of education were negatively correlated with state-anxiety ($r = -.25, p = .03$). Additionally, vocabulary scores were negatively correlated with trait-anxiety ($r = -.37, p = .001$), state-anxiety ($r = -.31, p = .007$), and SRLE scores ($r = -.34, p = .007$). Trait-anxiety demonstrated the largest correlation with SRLE ($r = .85, p < .001$), followed by the correlation between trait-anxiety and state-anxiety ($r = .51, p < .001$), and then the correlation between state-anxiety and SRLE ($r = .39, p = .002$).

Correlations between age and the three TOT measures indicated no statistically significant relationships: Age was unrelated to TOTs ($r = -.05, p = .67$), TOTs as a proportion of known information ($r = -.15, p = .20$), and TOTs as a proportion of unsuccessful retrievals ($r = .02, p = .85$). Ratings on the 1-item stress measure taken
following the speech/mathematics task correlated with TOTs ($r = .25, p = .03$) and with TOTs as a proportion of unsuccessful retrievals ($r = .37, p = .001$).

When stimuli were separated by stimulus type (names vs. non-names), additional correlations were established. For name stimuli, TOTs as a proportion of known information were negatively correlated with vocabulary ($r = -.27, p = .002$) and marginally negatively correlated with age ($r = -.19, p = .10$). One of the 1-item stress measures correlated with TOT measures: The rating following the speech/mathematics task correlated with TOTs as a proportion of familiar name information ($r = .29, p = .01$) and TOTs as a proportion of unsuccessful retrieval of name information ($r = .39, p = .001$). For non-name stimuli, TOTs as a proportion of known information were positively correlated with state-anxiety scores ($r = .26, p = .03$) and negatively correlated with vocabulary scores ($r = -.23, p = .05$).

Lastly, we separated the data set by stress condition and re-ran all correlations. For participants in the low-stress condition, no correlations were found between TOT measures and any demographic variables. For participants in the high-stress condition, TOTs as a proportion of known information were positively correlated with both trait-anxiety ($r = .36, p = .04$) and state-anxiety ($r = .40, p = .01$) scores. Also, TOTs as a proportion of known information were negatively correlated with vocabulary ($r = -.35, p = .03$) and TOTs as a proportion of unsuccessful retrievals were positively correlated with vocabulary ($r = .39, p = .02$).
CHAPTER IV
DISCUSSION

The purpose of the present research was to experimentally manipulate psychosocial stress using the TSST to create a high-stress condition versus the placebo TSST to create a low-stress condition to determine how stress impacts TOT responses across three age groups representing the adult lifespan. We predicted main effects for stress condition and age based on the TDH (Burke et al., 1991; MacKay, 1987). We believed that individuals in the high-stress condition would demonstrate larger proportions of TOTs compared to participants in the low-stress conditions, because stress would transiently impact participants’ word retrieval ability by decreasing the overall transmission of priming. We also predicted that older adults would demonstrate the largest proportion of TOT responses, followed by middle-aged, and finally young adult participants, due to the general reduction in the transmission of priming associated with increased age suggested by the TDH model. Finally, an interaction between age group and stress condition was predicted, whereby within in the low-stress condition older adult participants would report the largest proportion of TOT responses followed by middle-aged, and then young adult participants and increased age would compound the negative effects of the high-stress condition, yielding a similar but exacerbated trend for TOT responses across age groups.

Our prediction regarding the impact of psychosocial stress on TOTs was generally supported across our analyses. We corroborated evidence presented in James et al.
(2016) regarding the effect of stress condition on TOT responses. Both of these experiments involve the TSST combined with the supposed presence of an unseen evaluator during the entirety of an experimental naming task for high-stress participants compared to the placebo TSST combined with no evaluator for low-stress participants. When data were analyzed across the entire naming task controlling for participants’ stimulus familiarity, marginal group means indicated that participants in the high-stress condition reported approximately twice as many TOTs as participants in the low-stress condition (see marginal means in Tables 5, 6, and 7). Additionally, analyses including only name stimuli showed the same trends in the data to an exacerbated degree (see marginal means in Tables 8, 9, and 10). The main effect of stress was not found for any of the TOT measures for analyses including only non-name stimuli. The trend of the results was, however, in the predicted direction for each analysis: Participants in the high-stress condition reported more instances of TOTs than participants in the low-stress condition but the mean difference was smaller than overall or name only stimuli (see Tables 10, 11, and 12). Although the base rates for TOT experiences were relatively low in general, the fact that the negative effect of stress on TOTs, previously found by James et al. (2016), was replicated by the current research demonstrates support for evaluative, psychosocial stress increasing TOT states. This supports the hypothesis put forth by James et al. (2016) that within the TDH model of TOTs, psychosocial stress transiently increases transmission deficits that increase the likelihood of participants experiencing TOTs.

The prediction we made regarding how aging would impact TOT occurrences was not supported across any of the analyses. We were able to extend the research conducted
by James et al. (2013), who tested the impact of psychosocial stress on TOTs across young and older adults, by collecting data from a sample of both young and older adults while also including a group of middle-aged adults. We predicted, based on the TDH and previous research on aging and TOTs (see Burke et al., 1991; Dahlgren, 1998; Evrard, 2002; Juncos-Rabadán et al., 2010), that older adults would demonstrate the largest proportion of TOT responses followed by young adults, and that middle-aged participants would fall into the middle of this range. There were no statistically significant age differences for any of the three TOT measures across the entire naming task, nor when stimuli were separated by item type. However, the trend of our results was quite variable: young adults demonstrated the largest mean proportion of TOT experiences for five of the nine total analyses; middle-aged adults had the largest mean proportion of TOT experiences for the remaining four conditional TOT analyses. Although these were non-significant trends, they do indicate that more research should be conducted including middle-aged adults, because the pattern of responding in the present research contradicts previous literature and the effect of aging on TOT experiences as predicted by the TDH. Additionally, we did not find the predicted interaction between stress and aging on TOTs. One reason why no interaction was found could have been due to older adults’ overall reaction to our stressor. These older adults reported lower trait anxiety and reported suffering from fewer life stressors over the previous month than our other age groups. Also, although not statistically significant, older adults rated their state anxiety as the lowest of all age groups in both the high- and low-stress conditions. Additionally, only the middle-aged adults and the older adults in the low-stress condition reported lower state anxiety than our sample of older adults who participated in the high-stress condition.
Taken together this suggests that older adults’ very low baseline stress ratings led them to be less impacted by our stress manipulation, and perhaps their very low stress is why we did not find an age-related increase in TOTs in this study.

The current TOT measures were all based on stimulus sets that were individually conditionalized to only include familiar stimuli. This refined procedure is not common and may alter the typical age differences. As participants increased in age, TOT occurrences did not increase as expected, regardless of how TOT proportions were calculated. It is possible that the trends existing in our marginal age group means were due to calculating TOT proportions by using the number of familiar stimulus items in the denominator, which differed across age groups. Burke et al. (1991) also collected data on familiarity, however, they did not conditionalize their TOT measures using these scores. The overall trend of the familiarity ratings presented by Burke et al. were corroborated by our data: Young adult participants were familiar with fewer stimulus items, compared to the other age groups; middle-aged adults reported being familiar with the most stimulus items followed by older adults, and the difference between these two groups was negligible. We had initially compiled the naming task stimuli with the intent that they would be similarly familiar to participants across the adult lifespan. However, this was not successful based on the number of items participants in each age group endorsed as familiar. Under the TDH, TOT experiences can only occur for a word or name that is familiar to a participant. This familiarity conditionalization gives us an accurate baseline from which to examine TOT rates. However, that could indicate that part of the age group differences that have been identified in previous TOT literature may have been
driven by the inclusion of some stimulus items that were unfamiliar to participants, which would bias the proportions demonstrated across age groups.

Some studies, like Oberle and James (2013) have employed methods to conditionalize TOT dependent measures to the stimuli known to participants on a participant-by-participant basis. Oberle and James found marginal age differences for TOTs between young and older adults after conditionalization of their data. Prior to this conditionalization, the main effect of age on TOTs was not significant. Age-related increases in TOTs were not supported by the current research, regardless of conditionalization. We believe that this trend may have been due to middle-aged and older adults’ overall familiarity with the stimuli. These age groups were familiar with more items than the young adult groups. Lack of familiarity with the stimulus items could have affected the TOT proportions of the current research. Overestimation might have occurred for participants with lower familiarity to the stimuli set, due to a general reduction in the denominator used to compute these proportions. Overall overestimation due to unfamiliarity might explain why our data did not replicate previous research regarding TOTs and aging.

One issue with the present study was the lack of statistically significant differences between the state stress ratings of participants in the high- and low-stress conditions. It was expected that state-anxiety STAI scores would be higher for participants in the high-stress condition compared to participants in the low-stress condition. The lack of difference between state-anxiety scores in our high- and low-stress conditions indicates that either: 1) the manipulation of stress was not strong enough between the high- and low-stress conditions to yield differences on self-report state-
anxiety ratings; 2) the state-anxiety portion of the STAI was not successful at gauging the specific type of stress participants experienced during the experiment; or 3) our participants did not successfully self-report their stress levels either because they were unaware of the stress they experienced or they purposefully misreported their anxiety levels. On the other hand, one indication that the high-stress condition was in fact more stressful than the low-stress condition was demonstrated by the statistically significant main effect of stress condition demonstrated on the 1-item stress measure immediately following the speech and mathematics tasks. This follows a particularly stressful portion of the TSST procedure in the high-stress condition, and suggests that participants in the high-stress condition were able to perceive and report their increased stress. However, we expected increased stress ratings from participants in the high-stress condition throughout the experiment from this time point on. Overall, our manipulation of stress was successful in eliciting more TOT responses for high- compared to low-stress conditions. However, participants’ self-reported stress levels did not reflect the manipulation of stress.

The stress measures used in the present study may not have been entirely appropriate to measure the specific type of stress elicited by the TSST. The state form of the STAI includes inventory items that address stress under circumstances unlike the what would be expected during the TSST (e.g., “I am presently worrying over possible misfortunes”; “I feel indecisive”). Therefore, it is possible that this inventory measures stress unrelated to the type of stress caused by the TSST. Perhaps a different anxiety inventory measuring stress consistent with the psychosocial, evaluative aspects present in our modified TSST and observation components of the naming task would be more
appropriate to determine differences between high- and low-stress conditions. One measure that might be appropriate assesses communication apprehension (CA; see McCroskey, 1982). This scale (Personal Reports of Communication Apprehension) consists of 24 questions that gauge how individuals perceive their anxiety regarding communicating in different contexts (i.e., group discussions, meetings, interpersonal communications, and public speaking). This measure is typically utilized to demonstrate CA as a psychological trait, however, the inventory could be modified in order to acquire state-based responses from participants. We would then expect differences reported between high- and low-stress conditions on this measure, where high-stress participants demonstrate greater CA than low-stress participants. Up to this point in time no research has been conducted looking at how TOTs fluctuate across different levels of CA. It is possible that individuals who demonstrate high CA may also suffer from greater transmission deficits leading to TOTs and that a stress measure looking at CA would be more appropriate for this research than general stress measures.

The overall size of the stress effect on TOTs was similar across age groups, which corroborates findings from a previous study on TOTs and stress in aging (James et al., 2013). The overall stress effect has now been replicated across two experiments, indicating that increased stress levels manipulated by using the TSST techniques increase TOT occurrences. The TDH claims that TOTs occur due to transmission deficits on the one-to-one, top-down connections from the lexical node (representing the target name or non-name item) to each distinct part of the phonology necessary for generating the word in speech. These findings imply that the transmission deficits that lead to TOT experiences were increased under increased stress. These stress effects have been
hypothesized to be transient, meaning they cause a temporary increase in word finding problems, that persist for a period of time before dissipating. Future extensions of this research would benefit from the use of a within-subjects design. A within-subjects design would remove any unknown participant differences between the stress conditions, which would reduce the overall error in the analysis, producing a better idea of TOT experiences that specific participants experience under high- or low-stress. Additionally, the transient nature of stress could be evaluated by analyzing TOT proportions at different times during the naming task. If stress has a transient impact on TOTs then we would expect to see the largest TOT proportions paralleling the highest stress ratings, with reduced TOT proportions as stress alleviates. There is also a need to develop a more sensitive self-report stress measure to give during the naming task procedure, or research should employ physiological measures (e.g., heart rate, Galvanic skin response) to provide a more precise measure of experienced stress that can also be measured across time. Additionally, raising the stakes regarding the topic of the 5 min TSST speech would help future researchers increase the stress experienced by older adult participants. The current study had all individuals prepare and deliver a speech regarding an important job interview, and middle-aged and older adults may have more experience in this domain, which could reduce the stress they experienced due to the task. An extension that might raise the stakes of the speech task would involve an argument task where participants would prepare their position for or against a particular issue presented to them during a preparation period. Prior to delivering their speech it would be revealed that another “participant” would be viewing their argument in order to deliver a counterargument to what was presented. This speech topic would involve both social and
evaluative aspects of the traditional TSST procedure, and should produce the predicted stress effects. We believe that by making the argument task focus on participants delivering a persuasive type speech will also raise the stress across age ranges, not demonstrated in the current research.

The purpose of the present research was to demonstrate how TOT occurrences differ across three different age groups under conditions experimentally manipulated to be high-stress or low-stress. We demonstrated that the evaluative, psychosocial stress caused by the TSST led to increased TOTs during the naming task compared to the placebo TSST. The current findings supported our prediction concerning how stress generates transmission deficits that lead to increased TOT experiences. However, age differences that have been previously demonstrated were not found, nor was the predicted interaction between stress condition and age group. The findings of the current study did not support the prediction made by the TDH that TOT experiences increase with age. Critically, we aimed to establish whether or not stress demonstrated differential effects on TOTs for people of different ages. We found that stress does tend to impact transmission deficits, however, it causes these deficits to the same degree across the three age groups in the current study.
REFERENCES


APPENDIX

IRB APPROVAL

University of Colorado
Colorado Springs

Institutional Review Board (IRB) for the Protection of Human Subjects

Date: 10/19/2015

IRB Review

IRB PROTOCOL NO.: 16-064
Protocol Title: Non-Verbal Behavior During Language Production
Principal Investigator: Christopher Schmank
Faculty Advisor if Applicable: Lori James
Application: New Application
Type of Review: Expedited
Risk Level: No more than Minimal Risk
Renewal Review Level (If changed from original approval) if Applicable: N/A No Change
This Protocol involves a Vulnerable Population: N/A (No Vulnerable Population)
Expires: 18 October 2016

*Note, if exempt: If there are no major changes in the research, protocol does not require review on a continuing basis by the IRB. In addition, the protocol may match more than one review category not listed.

Externally funded: ☐ No ☑ Yes
OSP #: ☐
Sponsor: ☑

Thank you for submitting your Request for IRB Review. The protocol identified above has been reviewed according to the policies of this institution and the provisions of applicable federal regulations. The review category is noted above, along with the expiration date, if applicable.

Once human participant research has been approved, it is the Principal Investigator’s (PI) responsibility to report any changes in research activity related to the project:
- The PI must provide the IRB with all protocol and consent form amendments and revisions.
- The IRB must approve these changes prior to implementation.
- All advertisements recruiting study subjects must also receive prior approval by the IRB.
- The PI must promptly inform the IRB of all unanticipated serious adverse events (within 24 hours). All unanticipated adverse events must be reported to the IRB within 1 week (see 45CFR46.103(d); failure to comply with these federal mandates results in suspension or termination of the project.
- Renew study with the IRB prior to expiration.
- Notify the IRB when the study is complete.

If you have any questions, please contact Research Compliance Specialist in the Office of Sponsored Programs at 719-255-5903 or irb@uccs.edu

Thank you for your concern about human subject protection issues, and good luck with your research.

Sincerely yours,

Michelle Okuma, PhD
IRB Reviewer

irb@uccs.edu