12-1-2015

Evaluation of Equivalence Relations: Models of Assessment and Best Practice

Dawn Amber Seefeldt
Southern Illinois University Carbondale, dawn.seefeldt@forcescience.org

EVALUATION OF EQUIVALENCE RELATIONS: MODELS OF ASSESSMENT AND BEST PRACTICE

by

Dawn A. Seefeldt

B.A., University of South Dakota, 2009
M.A., Minnesota State University, Mankato, 2011

A Dissertation
Submitted in Partial Fulfillment of the Requirements for the
Doctor of Philosophy
with a Specialization in Behavior Analysis and Therapy

Rehabilitation Institute
in the College of Education and Human Services
Southern Illinois University
December 2015

DISSERTATION APPROVAL

EVALUATION OF EQUIVALENCE RELATIONS: MODELS OF ASSESSMENT AND BEST PRACTICE

by

Dawn A. Seefeldt

A Dissertation Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in the field of Rehabilitation with a specialization in Behavior Analysis and Therapy

Approved by:

Dr. Mark Dixon, Chair

Dr. Ryan Redner

Dr. Tom Upton

Dr. Stephanie Dollinger

Dr. Jonathan C. Baker

Graduate School
Southern Illinois University Carbondale
October 26th, 2015

AN ABSTRACT OF THE DISSERTATION OF

Dawn A. Seefeldt, for the Doctorate of Philosophy degree in Rehabilitation with a specialization in Behavior Analysis and Therapy, presented on October 26, 2015, at Southern Illinois University Carbondale.

TITLE: EVALUATION OF EQUIVALENCE RELATIONS: MODELS OF ASSESSMENT AND BEST PRACTICE

MAJOR PROFESSOR: Dr. Mark Dixon

Due to changing age demographics in the United States, by 2050, an estimated 62.1 Americans will be over the age of 65 and the number of Americans with cognitive impairment, such as Alzheimer's disease, will increase drastically as well (Alzheimer's Association, 2014; Ortman, Velkoff, & Hogan, 2014). Once a diagnosis or behavioral indicators of cognitive impairment are present, it would be beneficial to apply a treatment package that promotes the maintenance or re-establishment of stimulus control in the environment. From a behavioral perspective, stimulus control aids in learning and memory through both respondent and operant conditioning. In the current study, stimulus equivalence training was completed and compared to cognitive and functional assessments scores with older adult participants with and without cognitive impairment as a systematic replication of Gallagher and Keenan (2009). Formation of equivalence relations after exposure to linear series (LS) training with 2 3-member stimulus classes across arbitrary, familiar, and stimuli from Gallagher and Keenan (2009) was compared via trials to criterion, accuracy per relation, and session length to scores on the Mini-Mental Status Exam (MMSE), Saint Louis University State Exam (SLUMS), and Barthel ADL Index. Several methodological changes were applied to a second study to examine the impact of increased programmed stimuli, training changes, and the use of one-to-many (OTM) and many-to-one (MTO) training structures with 3 3-member stimulus classes on equivalence formation.

Composite performance scores were created for accuracy during LS, OTM, and MTO training. Non-parametric analyses were conducted between assessment and composite scores. The SLUMS and Barthel ADL Index were not correlated with any composite scores. However, MMSE scores and LS composite scores were correlated. The SLUMS was more sensitive to the detection of cognitive impairment as judged by classification and diagnoses. OTM and MTO composite scores also had a strong, positive correlation. Overall, more participants demonstrated higher levels of accurate responding during LS training than during OTM and MTO training. In contrast to previous research, only 25% of the sample demonstrated equivalence formation. Limitations and future research directions are discussed.

ACKNOWLEDGEMENTS

I would like to extend a special thank you to my advisor, Dr. Jonathan Baker, and all of the mentors and teachers who came before him and helped shape me into the behavioral scientist I am today. Dr. Baker's behavioral gerontology research laboratory was instrumental in the development of research designed to better understand and benefit older adults with various forms of cognitive impairment. I greatly appreciated the collaboration and suggestions from students in the lab. I also appreciated the support of my colleagues during the completion of this study.

Thank you to my dissertation committee for undertaking the role of guiding and refining this project. The interdisciplinary and eclectic backgrounds of the committee have greatly enhanced the methodology of the current study.

PREFACE

Age demographics are changing in the United States. By 2050, an estimated 62.1 Americans will be over the age of 65 (Ortman, Velkoff, & Hogan, 2014). For the first time in history, there will be more elderly adults than individuals under the age of 18, a phenomenon referred to as the "graying" of America (Moody & Sasser, 2015; Ortman, Velkoff, & Hogan, 2014). The prevalence of cognitive impairment, such as Alzheimer's disease, will increase along with this demographic change (Alzheimer's Association, 2014).

AD is now categorized under two umbrella terms, Minor Neurocognitive Disorder and Major Neurocognitive Disorder, in the Diagnostic and Statistical Manual (5th ed.) (DSM-5; American Psychiatric Association, 2013). Minor Neurocognitive Disorder may negatively impact one major life area or instrumental activities of daily living (IADLs), such as using a telephone or managing a checkbook (American Psychiatric Association, 2013). Major Neurocognitive Disorder significantly impacts an individual's day-to-day functioning and some forms result in continued impairment leading to the inability to complete grooming routines, recognize objects, or speak (Potkin, 2002).

Diagnoses of Mild or Major Neurocognitive Disorder usually result from gene testing, scores on cognitive assessment instruments, as well as the presence of biomarkers observed during brain scans (American Psychiatric Association, 2013). However, not all individuals with a gene related to AD will develop the disease (Greenwood & Parasuraman, 2012). Similarly, diagnostic issues arise with the use of cognitive assessments. Some assessments have been criticized for lack of sensitivity to early forms of cognitive impairment (Woodruff-Pak, 2001). Sensitivity is crucial given that the majority of individuals with mild forms of cognitive impairment will later develop AD or dementia (Gauthier et al., 2006).

In addition to diagnostic issues, current procedures do not inform treatment recommendations. A measure that could serve as an assessment and subsequent teaching tool would be beneficial. From a behavioral perspective, promoting the maintenance or re-establishment of stimulus control in the environment could aid learning and memory. Stimulus control can be achieved through both respondent and operant conditioning.

Respondent conditioning research has demonstrated age differences in the eye blink response, which may be related to age-related changes in learning (Woodruff-Pak & Thompson, 1988). The eye blink response has also been associated with cognitive functioning in older adult participants (Woodruff-Pak, 2001). Results suggest that eye blink conditioning may serve as an early detector of future development of forms of Major Neurocognitive Disorder making the procedure more sensitive than popular cognitive assessments. Despite increased sensitivity, the procedure has not evolved into clinical application.

Aside from respondent conditioning, stimulus control is also achieved through operant conditioning. The relationship between stimuli and responses is complex, as multiple stimuli may evoke one response or a single stimulus may evoke numerous responses (Michael, Palmer, & Sundberg, 2011). These convergent and divergent relations can result from direct training or emerge untrained. Stimulus equivalence procedures involve the emergence of untrained relations, or derived relational responding, through conditional discrimination training, or matching-to-sample (Urcuioli, 2013).

Derived relational responding is a marker of learning and memory, which deteriorates with advancing cognitive impairment (Arntzen & Steingrimsdottir, 2014; Bódi, Csibri, Myers, Gluck, & Kéri, 2009; Gallagher & Keenan, 2009; Saunders et al., 2005; Sidman, 2013; Steingrimsdottir & Arntzen, 2011a;b). Stimulus equivalence formation is associated with near perfect cognitive assessment scores (Gallagher & Keenan, 2009). Individuals who score in the typically functioning range on some cognitive assessments are not able to engage in derived relational responding, which suggests the increased sensitivity of conditional discrimination procedures as a diagnostic tool. Similarly to other behavioral procedures, stimulus equivalence has not been widely adopted. More research on stimulus equivalence with older adult participants may help drive clinical application.

The purpose of the current studies is to first systematically replicate Gallagher and Keenan (2009) by analyzing the association between equivalence class formation and cognitive assessment for older adult participants by utilizing stimuli from previous research as well as familiar and arbitrary stimuli. The first study will also extend previous research by examining the potential correlation between functional assessment scores and stimulus equivalence formation. The second study will advance previous research by utilizing previous recommendations to attempt to create an optimal treatment package for participants who do not demonstrate equivalence class formation during study one through the use of larger stimulus classes and different training structures.

TABLE OF CONTENTS

CHAPTER **PAGE**

ABSTRACT i

ACKNOWLEDGMENTS iii

PREFACE iv

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTERS

CHAPTER 1 – Introduction 1

CHAPTER 2 – Conditional Discrimination Training & Stimulus Equivalence 9

CHAPTER 3 – General Methods 24

CHAPTER 4 – Methods: Study I 29

CHAPTER 5 – Results: Study I 36

CHAPTER 6 – Discussion: Study I 41

CHAPTER 7 – Methods: Study II 45

CHAPTER 8 – Results: Study II 53

CHAPTER 9 – Discussion: Study II 60

CHAPTER 10 – General Discussion 63

REFERENCES 86

APPENDICES

Appendix A – Saint Louis University Mental State Exam 95

Appendix B – Barthel ADL Index 96

Appendix C – Procedural Integrity Evaluation 97

VITA 98

LIST OF TABLES

TABLE **PAGE**

Table 1 – All Participants by Order of Recruitment and Demographics 76

Table 2 – Study I Results: Assessment Scores & Equivalence Responding 77

Table 3 – Study I Results: Trials to Criterion, Accuracy, and Session Length 78

Table 4 – Study II: Participant Order of Exposure by Training Modality & Stimulus Type 79

Table 5 – Study II Results: Assessment Scores & Equivalence Responding 80

Table 6 – Study II Results: Trials to Criterion, Accuracy, and Session Length 81

LIST OF FIGURES

FIGURE **PAGE**

Figure 1 – Study I Stimuli.....82

Figure 2 – Classification Analysis.....83

Figure 3 – Study II Stimuli.....84

Figure 4 - Mastered Relations by Stimulus Type and Training Structure.....85

CHAPTER 1

INTRODUCTION

The population distribution in the United States used to be more triangular, with the smallest portion comprised of individuals over 65, but each year the distribution is becoming more rectangular in nature. In 2012, the United States Census Bureau recorded 43.1 million Americans over the age of 65, a number expected to nearly double by 2050 (Ortman, Velkoff, & Hogan, 2014). The baby boomers began turning 65, the legal age of older adulthood, in 2011 and are a large contributor of the “graying” of America (Colby & Ortman, 2014). As the population of older adults continues to grow, so will the prevalence of cognitive impairment.

Currently, the most diagnosed form of cognitive impairment is Alzheimer’s disease (AD). Approximately 5.2 million older adults have AD, or 1 out of every 9 Americans age 65 and higher (Alzheimer’s Association, 2014). AD is now categorized under two umbrella terms, Major Neurocognitive Disorder and Minor Neurocognitive Disorder, in the Diagnostic and Statistical Manual (5th ed.) (DSM-5; American Psychiatric Association, 2013). Individuals suffering from Major Neurocognitive Disorder demonstrate a significant reduction in performance from a previous level of functioning on cognitive-related tasks. Decreases in cognitive performance impact instrumental activities of daily living (IADLs) such as a) independently attending appointments, b) appropriately administering medications, and c) balancing finances (American Psychiatric Association, 2013). IADLs are required for independent living, but are not essential for survival.

As the insidious deterioration continues, individuals will experience greater difficulty learning new information as well as remembering previously learned information. If individuals survive long enough with AD, the great majority of the ability to communicate and complete activities of daily living (ADLs) will be lost (Potkin, 2002). ADLs comprise a list of daily tasks that must be accomplished in order to survive. Examples of ADLs include (a) bathing, (b) dressing, (c) transferring, (d) ambulating, (e) toileting, and (f) eating (Potkin, 2002).

The Alzheimer's Association and U.S. Census Bureau aid in the determination of the increasing prevalence of cognitive impairment and aging Americans, but neither statistic provides information on the amount of Americans who are likely to develop Major Neurocognitive Disorder in their lifetime. Individuals with Minor Neurocognitive Disorder, or what used to be called Mild Cognitive Impairment, may provide insight into older adults at risk of developing AD, although some individuals may already have a diagnosis of AD. AD is now a cause of both major and minor forms of cognitive impairment, thereby allowing different diagnoses for the gradual decline in functioning (American Psychiatric Association, 2013). However, individuals with cognitive impairment whose symptoms do not meet the definition of Major Neurocognitive Disorder and do not yet have a diagnosis of AD are associated with developing AD and dementia at higher rates with onset of diagnosis beginning after approximately five years (Gauthier et al., 2006). Mild Neurocognitive Disorder consists of the same symptoms as Major Neurocognitive Disorder with the exception of lack of significant impact on ability to carry out daily activities (American Psychiatric Association, 2013).

Diagnoses of Mild or Major Neurocognitive Impairment can result from gene testing, scores on cognitive assessment instruments, as well as the presence of biomarkers observed during brain scans (American Psychiatric Association, 2013). The Apolipoprotein E (APOE) gene is related to cognition and aging (Greenwood & Parasuraman, 2012). Research has demonstrated that a variation of the gene, e4, is related to the development of Alzheimer's disease and has been widely studied (Saunders et al., 1993). However, only half of e4 homozygotes with a genotype prone to develop into Alzheimer's actually do and individuals may develop the disease without the presence of an e4 allele (Greenwood & Parasuraman, 2012).

Cognitive assessments are commonly used to identify the presence of cognitive impairment, but may not be sensitive enough to diagnose mild forms (Woodruff-Pak, 2001). The Mini-Mental Status Exam is commonly used to determine global cognitive functioning (MMSE; Folstein, Folstein, & McHugh, 1975). The MMSE may be a popular instrument as it is relatively inexpensive and takes 5 – 10 minutes to administer. However, the MMSE does not assess all cognitive domains and may yield results indicative of impairment once a person has atrophy in the medial temporal lobe, which impacts memory (Stewart, O'Riley, Edelstein, & Gould, 2012; Woodruff-Pak, Ewers, & Vogel, 2001).

The Saint Louis Mental Status Examination has been used as an alternative to the MMSE (SLUMS; Tariq, Tumosa, Chibnal, Perry, & Morley, 2006). The SLUMS is publically available and demonstrates higher levels of sensitivity than the MMSE (Stewart, O'Riley, Edelstein, & Gould, 2012). The SLUMS incorporates a delayed recall task, similar to the MMSE, but also utilizes a recall task based on information in a short story. However, more research has been recommended in the area of a correlation between the SLUMS and ADL performance (Stewart, O'Riley, Edelstein, & Gould, 2012). Results from gene testing, the MMSE, and the SLUMS are limited, as they are diagnostic in nature and rarely inform treatment options.

Once a diagnosis or behavioral indicators of cognitive impairment are present, it would be beneficial to apply a treatment package that promotes the maintenance or re-establishment of stimulus control in the environment. From a behavioral perspective, stimulus control aids in learning and memory through both respondent and operant conditioning. Respondent, or classical and higher order, conditioning involves the pairing of a neutral stimulus with an unconditioned or conditioned stimulus that elicits an unconditioned or conditioned response (Skinner, 1938). After repeatedly pairing the stimuli, the neutral stimulus becomes a conditioned stimulus that elicits the conditioned response without the presence of the unconditioned or conditioned stimulus used during pairing. The organism responds to the presence of the stimulus and the organism's behavior is, therefore, under stimulus control (Skinner, 1938).

Respondent conditioning research has demonstrated age differences in the eye blink response, which may be related to age-related changes in learning (Woodruff-Pak & Thompson, 1988). Young, middle age, and older adult participants were presented with a 80 dB tone, the conditioned stimulus, for 500 ms, followed by a 10-psi airpuff of oxygen, the unconditioned stimulus, 400 ms after tone onset. Eye blinks, the conditioned response, of 0.5 mm or greater amplitude occurring between 25 – 399 ms after tone onset were recorded (Woodruff-Pak & Thompson, 1988). Eye blinks were selected as the target conditioned response, as previous research found that said response occurs less often, with lower magnitude, and takes longer to condition in middle age and older adult non-human animals (Thompson & Woodruff-Pak, 1987) as well as in older adult humans (Jerome, 1959). Woodruff-Pak and Thompson (1988) extended previous research with the finding that conditioned responding begins to slow and decreases in magnitude in the fourth decade of human life. This finding was increasingly more pronounced across older age groups (i.e., 50 – 59, 60 and higher).

The eye blink response has also been tied to cognitive functioning in older adult participants (Woodruff-Pak, 2001). Typically developing older adults who do not continuously engage in conditioned responding are at an increased risk of developing probable Alzheimer's disease within 2 – 3 years, whereas participants that perform above the criterion level do not develop probable Alzheimer's disease within in the same time frame (Downey-Lamb &

Woodruff-Pak, 1999; Ferrante & Woodruff-Pak, 1995). Within a longitudinal study, Woodruff-Pak, Ewers, and Vogel (2001) examined the performance of one participant, whose conditioned responding decreased to poor levels of performance over a span of two years, but the participant continued to receive a normal functioning score on the MMSE. A score indicating moderate cognitive impairment did not occur for another eight years. For all participants, this longitudinal study found similar results along with more errors while drawing a clock face and increased mortality for poor conditioners (Woodruff-Pak, Ewers, & Vogel, 2001). These results suggest that eye blink conditioning may serve as an early detector of future development of forms of Major Neurocognitive Disorder. Despite the sensitivity of eye blink conditioning procedures, the procedure has not been widely adopted as a means to predict future cognitive impairment.

Aside from respondent conditioning, stimulus control is also achieved through operant conditioning. Certain conditions, such as motivating operations, setting events, antecedents, and consequences, increase or decrease the probability of the occurrence of a response (Skinner, 1953). Individuals are said to learn when responding occurs more often in the presence of a discriminative stimulus, which signals the availability of reinforcement, and less often in the presence of a stimulus delta, which signals the unavailability of reinforcement (Skinner, 1938; 1953; Urcuioli, 2013). A reinforcer is established when responding occurs under similar situations in the future (Skinner, 1953).

Related to the current research, operant conditioning encompasses the act of remembering, in order to appropriately respond in the future (Donahoe & Palmer, 2004; Palmer, 1991). For instance, an individual learns to engage in discriminative responding when presented with a visual stimulus of an apple and a orange and presented with the auditory stimulus, "Point to apple," followed by the general conditioned reinforcer, "Good job!" when the individual makes the appropriate selection. If, “Good job!” functions as a reinforcer, the appropriate pointing response should increase in the future. Stimulus control is also involved when tacting an image of an apple in a different context or stating, “Apple,” if someone asked, “What type of fruit was in the image I previously showed you?” Stimulus control impacts performance when varying degrees of associated stimuli are present.

Responses may occur less often over time, which may be referred to as the behavior of forgetting. However, radical behaviorists argue that forgetting is not simply due to the passage of time, and rather, responses may weaken in presence of new stimuli, competing responses, and setting changes (Palmer, 1991). A response may not generalize to new stimuli, such as when different therapists are used to facilitate an intervention. A response may come under the control of new stimuli due to serendipitous presentation of unrelated, but physically similar stimuli or infrequent presentation of the previous discriminative stimulus (Palmer, 1991). Additionally, contextual changes may inhibit generalization of responding in new settings, such as a response that is likely to occur in an occupational setting, but less likely to occur while at home.

The relationship between stimuli and responses is complex. A single response may be evoked by multiple stimuli and numerous responses can also be evoked by a single stimulus. These phenomena are referred to as convergent and divergent multiple control (Michael, Palmer, & Sundberg, 2011). Said relations can form during direct training, but may also emerge without direct training (Urcuioli, 2013). Stimulus equivalence procedures involve the emergence of untrained relations through conditional discrimination training, or matching-to-sample (Urcuioli, 2013). A sample stimulus is presented along with comparison stimuli that are presented simultaneously or after a brief delay (Saunders, Chaney, & Marquis, 2005). Participants are typically provided with feedback for correct responses. Participants are exposed to direct training for two or a few stimuli, but begin to demonstrate untrained relations as well (Urcuioli, 2013).

These emergent relations are an indication of learning and memory and, as such, decrease with advancing cognitive impairment (Arntzen & Steingrimsdottir, 2014; Bódi, Csibri, Myers, Gluck, & Kéri, 2009; Gallagher & Keenan, 2009; Saunders et al., 2005; Sidman, 2013; Steingrimsdottir & Arntzen, 2011a;b). Stimulus equivalence formation is associated with near perfect MMSE scores (Gallagher & Keenan, 2009). Individuals who score in the low-normal range on the MMSE (e.g., 24 – 27) are not always able to form equivalence, which suggests equivalence class formation may be more sensitive than cognitive assessments. In this respect, stimulus equivalence has been regarded as a useful diagnostic tool and potentially helpful in the early detection of pathological conditions, although, similar to the conditioned eye blink response, it has not been widely adopted (Gallagher & Keenan, 2009). Additionally, previous research has not examined stimulus equivalence as a teaching tool for older adults. Data collection is sometimes performed in a single day, which yields little evidence for the use of stimulus equivalence training as a teaching technology.

The purpose of the current trilogy of studies is to first replicate previous research in Study I by analyzing the association between stimulus equivalence class formation and MMSE scores as conducted by Gallagher and Keenan (2009). Individuals with and without cognitive impairment will be recruited for the study to better examine the lines of fracture during equivalence performance. Secondly, Study I extends Gallagher and Keenan (2009) by also analyzing the association between SE class formation and SLUMS scores as well as ADL assessment scores. Lastly, Study II will extend previous research by examining the impact of methodological changes by incorporating recommendations from the literature and evaluating stimulus control for individuals who did not form equivalence relations during Study I.

CHAPTER 2

CONDITIONAL DISCRIMINATION TRAINING & STIMULUS EQUIVALENCE

Different forms of discrimination training involve various amounts and types of stimuli. Discriminations may involve a single stimulus during successive training, concurrent presentation of sample and comparison stimuli during simultaneous training, as well as delayed presentation of comparison stimuli after presentation of sample stimuli during delayed, discrimination training (Urcuioli, 2013). A simple, successive discrimination involves the presentation of a single sample stimulus followed by a participant response, but may also involve individually pairing a sample stimulus to numerous comparison stimuli (Urcuioli, 2013).

Training becomes conditional when simultaneously presenting another sample stimulus and manding for the individual to tact the correct response in verbal behavior procedures (Urcuioli, 2013). Conditional discrimination procedures can also involve training non-verbal behavior, such is the case with stimulus equivalence formations (Hall & Chase, 1991). An individual may be presented with one arbitrary, sample stimulus and two arbitrary, comparison stimuli. The individual selects one arbitrary comparison stimulus and receives an indication of correct responding, usually in the form of praise, “Correct,” (Gallagher & Keenan, 2009). Conditional discrimination training may also incorporate a delay between presentation of sample and comparison stimuli (Saunders, Chaney, & Marquis, 2005).

Within stimulus equivalence, there are three types of relations: a) reflexivity, b) symmetry, and c) transitivity (Sidman, 1971; Sidman & Tailby, 1982). Reflexivity involves the identification of an identical comparison stimulus when presented with a sample stimulus. For example, selecting comparison stimulus A when presented with sample stimulus A (A-A), instead of choosing stimulus B, C, D, etc would be desirable responding. Symmetry develops by first training an individual to match two arbitrary stimuli, such a stimulus A to stimulus B. Thus, when an individual is presented with stimulus A, stimulus B will be selected (A-B). Once matching is demonstrated, symmetry can be tested by presenting stimulus B. If the individual selects comparison stimulus A in the presence of sample stimulus B, symmetry has formed (B-A). Lastly, transitivity develops by training a second matching relation, such as stimulus B to stimulus C (B-C), and securing symmetry (C-B). Then, if an individual selects comparison stimulus A in the presence of sample stimulus C, transitivity has been achieved (A-C). The strongest measure of equivalence is when symmetry and transitivity are demonstrated (Sidman, 1994). This occurs when the participant selects comparison stimulus A in the presence of sample stimulus C (C-A).

The above example uses a linear series (LS), or serial, training model. In serial training stimulus A is matched to stimulus B and stimulus B is matched to stimulus C. Matching continues along this linear path until each stimulus has been matched to the preceding stimulus (A→B→C→D→E→F, etc.). LS training allows for the examination of both transitivity (A-C relations) and equivalence (C-A relations). The number of nodes increases along with the number of members in a stimulus class, which can result in more trials to criterion (Fields, Adams, Verhave, & Newman, 1990). Several articles have suggested that LS training is less effective than many-to-one (MTO) or one-to-many (OTM) training formats (Arntzen & Hansen, 2011; Eilifsen & Arntzen, 2015; Saunders et al., 2005). However, older adult participants have demonstrated derived relations with the use of LS procedures (Gallagher & Keenan, 2009; Pérez-González & Moreno-Sierra, 1999).

Unlike LS training, in OTM procedures, a single stimulus is matched to all other stimuli in the class (Saunders et al., 2005). For instance, stimulus A is matched to stimulus B, then stimulus A is matched to stimulus C, followed by stimulus D, and so on. OTM procedures have been used with a wide variety of populations ranging from children (Sidman & Tailby, 1982), adolescents to middle aged adults (Eilifsen & Arntzen, 2015), typically functioning adults (Arntzen & Hansen, 2011), to older adults (Wilson & Milan, 1995). As an example, in one of the hallmark studies on equivalence, 6 of 8 children formed three 4-member classes after exposure to simultaneous OTM conditional discrimination training (Sidman & Tailby, 1982). Interestingly, one participant was able to demonstrate derived relational responding without naming the stimuli. This finding suggests that verbal behavior may be unnecessary, albeit helpful, during the formation of equivalence classes.

In contrast to OTM procedures, MTO training involves training multiple stimuli to a single stimulus (Saunders et al., 2005). For training, stimulus B, stimulus C, and stimulus D are all matched to stimulus A. MTO procedures are apparent in the literature across the lifespan. MTO procedures have been used with college students (Rehfeldt & Hayes, 2000), adults with intellectual disabilities (Rehfeldt & Root, 2004), and cognitively intact older adults (Aggio & Domeniconi, 2012). Each of the mentioned studies will be reviewed below.

All forms of training require, at minimum, two 3-member stimulus sets in order to assess for stimulus equivalence formation. However, many studies have examined the impact of multiple sets of stimuli consisting of 3-member or higher classes with typically functioning children, children with visual impairments (Toussaint & Tiger, 2010), adolescents and adults (Eilifsen & Arntzen, 2015), solely adults (Arntzen & Hansen, 2011; Rehfeldt & Hayes, 2000), and older adults (Aggio & Domeniconi, 2012; Saunders et al., 2005) as well as adults with intellectual disabilities (Rehfeldt & Root, 2004). Toussaint and Tiger (2010) used stimulus equivalence to train school-aged children to form relations between Braille letters, printed letters, and their spoken names. Participants were exposed to 1 – 5 classes containing 4 – 6 members. Emergence of untrained Braille-printed letter relations occurred for all participants.

Stimulus equivalence, and other forms of conditional discrimination training have utilized tabletop and computer-based training modalities. Dymond, Rehfeldt, and Schenk (2005) provide guidelines when using a tabletop format in order to minimize human error. Safeguarding against threats to validity may include (a) minimizing outside distractions in the experimental setting, (b) presenting stimuli using a placement board or laminated sheet to ensure consistent presentation and distance between stimuli, (c) providing consistent instructions, (d) concealing visual access to the data sheet, (e) training and evaluation of procedural integrity for research assistants, (f) utilizing and adhering to clear response definition, (g) providing scheduled reinforcers, (h) maintaining intertrial intervals, and (i) collecting and evaluating interobserver reliability (Dymond, Rehfeldt, & Schnek, 2005).

However, with the increase in touch screen technologies, many researchers are moving to a computer-based modality, including research with older adults (Bódi, Csibri, Myers, Gluck, & Kéri, 2009; Engel, 2005; Saunders et al., 2005; Steingrimsdottir & Arntzen, 2014). Some researchers use a reflexivity or orientation task with stimuli unrelated to testing to assess if the participant is able to complete tasks via a touch screen computer monitor (Gallagher & Keenan, 2009; Plaza et al., 2012; Rehfeldt & Root, 2004). However, some researchers incorporate card sorting, to identify pre-existing, unprogrammed relations between the stimuli, and an orientation task into the methods (Steingrimsdottir & Arnzten, 2014).

Once equivalence has been demonstrated, researchers have questioned if performance maintains over the passage of time. Researchers have also questioned if trained relations are present and necessary to continue to form equivalence during retention tests. Thus, previous research suggests mixed results for maintenance of derived and/or directly trained relations after initial testing (Arntzen & Hansen, 2011; Eilifsen & Arntzen, 2015; Fienup & Dixon, 2006; Rehfeldt & Hayes, 2000; Rehfeldt & Root, 2004).

Maintenance of both equivalence relations with arbitrary stimuli and generalized equivalence relations with arbitrary stimuli differing along a continuum of hue were tested after college age participants completed MTO stimulus equivalence training with three sets of stimuli consisting of four member classes (Rehfeldt & Hayes, 2000). Participants were required to talk aloud during MTO training. Stimuli consisted of compound samples and a single comparison stimulus during training. Equivalence and generalization tests involved examination of relations between unitary stimuli, instead of stimuli compounds with unitary stimuli. Rehfeldt and Hayes (2000) found that of eight participants, two individuals demonstrated maintenance of equivalence relations and three demonstrated continued accurate performance during the generalization test 2 -3 months after initial testing.

Maintenance of stimulus equivalence formation has also been examined in adult participants with intellectual disabilities (Rehfeldt & Root, 2004). A MTO matching-to-sample format was used to attempt to train equivalence for three 3-member classes. A follow-up equivalence and generalization test took place 1 - 4 months after initial training. Generalization was based on hue similar to Rehfeldt and Hayes (2000). All four participants demonstrated most if not all equivalence relations after initial testing and maintained the majority of emergent responding during the retention test. Participants diagnosed with intellectual disabilities may perform similarly to older adults with cognitive impairment, suggesting that such participants may be able to form and maintain equivalence relations with sufficient training.

Arntzen and Hansen (2011) sought to continue the equivalence maintenance research conducted by Rehfeldt and Hayes (2000) with typically developing participants by conducting a follow-up test after exposure to LS, MTO, or OTM equivalence training procedures with 27 of the original 30 adult participants. Average age of participants ranged from 26 - 36.8 years old for each condition during initial testing. The training conditions included three sets of stimuli composed of either three or six member classes. A retention test was conducted between 2 – 4 weeks after initial training and testing. Approximately half of the participants that demonstrated equivalence during initial testing maintained equivalence class formation at follow-up (i.e., 9/16 participants). Arntzen and Hansen (2011) note that none of the participants in the LS condition demonstrated equivalence during the retention test, whereas three participants from the MTO condition and six participants from the OTM condition maintained equivalence.

Eilifsen and Arntzen (2015) examined maintenance effects after exposing 18 participants (age 16 – 46 years) to LS, MTO, and OTM training. Participants were trained using three classes of stimuli containing five members. No participants in the LS condition demonstrated equivalence, whereas five participants in the MTO and five participants in the OTM condition demonstrated emergent relations. If participants demonstrated 90% or higher accuracy for directly trained or emergent relations, they moved to re-training plus re-testing or re-testing alone at three retention points (i.e., 12 participants). Re-testing occurred one day, two weeks, and four weeks after initial testing. Eilifsen and Artzen (2015) found that directly trained and emergent relations remained intact within a month of initial training and testing regardless of exposure to re-training for individuals that initially demonstrated equivalence formation. However, participants that initially only formed directly trained relations demonstrated weaker stimulus control over the course of a month. Given the age disparity among participants, it would have been interesting to examine age effects on initial and subsequent equivalence performance.

Stimulus equivalence research has been conducted with typically developing older adult participants, yielding varying results and minimal maintenance data (Aggio & Domeniconi, 2012). Conditional discrimination training has been conducted both abroad (Aggio & Domeniconi, 2012; Pérez-González & Moreno-Sierra, 1999; Steingrimsdottir & Arntzen, 2014) and within the United States. (Gallagher & Keenan, 2009; Saunders et al., 2005; Wilson & Milan, 1995). Some research has begun to examine a potential relationship between stimulus equivalence formation and performance on cognitive assessments (Gallagher & Keenan, 2009; Wilson & Milan, 1995). Best practice for training emergent relations in older adults has also been investigated in terms of training structure, simultaneous vs. delayed presentation of stimuli, and use of arbitrary vs. identical stimuli (Engel, 2005; Saunders et al., 2005; Steingrimsdottir & Arntzen, 2014).

In Spain, Pérez-González and Moreno-Sierra (1999) conducted conditional discrimination training with a child, young adult, two middle aged adults, and four older adults to examine potential age differences in stimulus equivalence formation across the lifespan. Tabletop LS training with two 3-member classes consisting of different shapes were used to test equivalence performance. Interestingly, all participants were related, which allowed for a more homogeneous sample for cross-sectional analysis. Older participants required more training trials to reach criterion, but were able to demonstrate equivalence formation.

In Brazil, Aggio and Domeniconi (2012) used a computer-based mode for training stimulus equivalence in typically functioning older adult participants (60 – 75 years young), as judged by scores on the MMSE and WAIS-III (Wechsler, 1981). Eight participants were split between two conditions consisting of either three 3-member stimulus classes or three 6-member classes of stimuli. MTO simultaneous matching-to-sample procedure with an increasing number of comparison stimuli was used to train relations. Participants were first presented with a single comparison and then advanced to presentation of 2 - 3 comparison stimuli during training for matching. All participants demonstrated stimulus equivalence, however, number of trials to criterion decreased as the number of stimuli in a class increased and five participants were exposed remedial training. A maintenance test occurred six weeks after initial testing. Participants exposed to six member classes during training performed better by making fewer errors on the maintenance test.

Wilson and Milan (1995) examined age-related differences in equivalence formation between younger and older adults using three 3-member stimulus classes and OTM format. Both groups of participants were able to form directly trained and emergent relations within a similar amount of trial blocks. However, fewer older adults formed equivalence and, those who did, had a longer response latency than younger adults. Wilson and Milan (1999) also examined the impact of MMSE score on the development of emergent relations for older adults participants, but did not find a correlation.

Gallagher and Keenan (2009) measured direct and derived relations with LS conditional discrimination training using a tabletop format with non-sense syllabus and symbols. Participants (age 57 – 94 years young) were exposed to two 3-member classes and provided praise for appropriate responding. Similar to Wilson and Milan (1995), MMSE scores were gathered and compared to equivalence class formation. Gallagher and Keenan (2009) found that the majority of older adults achieved emergent relations who scored high on the MMSE (range 27 – 30), and some participants scored in the low-normal range, but were unable to demonstrate equivalence.

The study suggested the potential utility of using stimulus equivalence as a diagnostic tool for cognitive impairment.

Saunders et al. (2005) examined a variety of methodological changes across two studies and argue that a linear series training structure, such as that used in Gallagher and Keenan (2009) may be less effective than OTM or MTO training methods for older adult participants. Older adult participants were exposed to simultaneous and delayed LS, OTM, and MTO training structures and 2 – 4 stimulus classes composed of 3 – 4 members on a computer. All stimuli were arbitrary symbols. Participants required fewer trials to criterion and demonstrated more emergent relations during delayed, rather than simultaneous, MTS procedures. Saunders et al. (2005) also argue that increasing the number of members in a stimulus class does not impact faster or stronger equivalence performance. However, Saunders et al. (2005) did not examine maintenance effects, whereas previous research has recommended larger stimulus classes result in better maintenance six weeks after initial training and testing (Aggio & Domeniconi; 2012).

In Norway, Steingrimsdottir and Arntzen (2014) also examined simultaneous vs. delayed presentation of stimuli. Participants in the delay condition were able to form relations in fewer trials and were more likely to continue participation in the study. Steingrimsdottir and Artnzen (2014) further extended Saunders et al. (2005) by exposing older adult participants to both identity and arbitrary matching to examine potential exposure effects. Exposure to identical matching before arbitrary matching may hinder the development of relations, as participants may continue to search for similarities among the stimuli (Sidman, 1994). Participants in Wilson and Milan (1995) who did not demonstrate equivalence reported searching for similarities among stimuli.

In contrast to such suggestions, Steingrimsdottir and Arntzen (2014) found that identity matching before arbitrary matching did not hinder performance during the latter training and participants performed similar during simultaneous identical-arbitrary (ID-AR) matching as in 0s delay ID-AR matching. Additionally, seven participants acquired stimulus equivalence during 0s delay identical-arbitrary training, whereas fewer participants (i.e., 5/8) acquired equivalence formations during the simultaneous identical-arbitrary training. Five participants also demonstrated equivalence after exposure to either simultaneous and 0s delay AR-ID training. Presentation of identical stimuli prior to arbitrary stimuli as well as use of a 0s delay may improve older adult equivalence formation, but more research needs to be conducted to parse a part these variables.

In an unpublished master's thesis, Engel (2005) demonstrated stimulus equivalence formation using stimuli varying along a continuum of related properties. Stimuli varied between 1 – 4 class-consistent properties. Older adults (67 – 87 years young) required more training trials than younger adults (age 18 – 20), but both groups of participants were able to demonstrate equivalence formation. A transfer of stimulus function test was also conducted based on shared features of the stimuli. Both younger and older adults demonstrated a transfer of function to class-consistent, yet untrained stimuli with more accurate responding for stimuli that shared a higher number of class-consistent features.

More recently, discrimination training and stimulus equivalence research has been conducted with older adults participants suffering from cognitive impairment around the globe, with participants from Hungary (Bódi et al., 2009), Norway (Arntzen & Steingrimsdottir, 2014; Steingrimsdottir & Arntzen, 2011a;b), and Spain (Plaza, López-Crespo, Antúnez, Fuentes, & Estévez, 2012). Sidman (2013) also offers suggestions for how to identify areas of stimulus control deficits and strengths for individuals with Alzheimer's disease based on research largely conducted in the United States.

Bódi et al. (2009) used a simultaneous conditional training procedure to form face-fish associations. Matching was trained for A1-X1, A2-X1, and A1-X2 and acquired equivalence was measured for the A2-X2 relation, a many-to-many training format. These procedures were repeated for another set of four stimuli (B1, B2, Y1, and Y2). Participants consisted of 20 typically functioning older adults (MMSE, 29.4_{M} and 0.7_{SD}) and 25 older adults with mild AD (MMSE score 24.0_{M}, 1.3_{SD}) residing in Hungary. After computer-based testing, participants received cards of stimuli and were asked to pair the appropriate face to the corresponding fish stimulus in order to examine the formation of associations without a forced-choice.

Bódi et al. (2009) found that 22 of the 25 participants will mild AD were able to form directly trained relations, although those participants made more errors than cognitively intact participants. Number of errors further increased during the equivalence test as well as during the card-pairing test for participants with mild AD. Participants with mild AD formed equivalence relations at chance levels. The results suggest that development of emergent relations and stimulus generalization is more difficult for individuals with cognitive impairment.

Arbitrarily related, yet familiar, stimuli were employed by Bódi and colleagues (2009) to examine stimulus equivalence formation using simultaneous conditional discrimination training. Steingrimsdottir and Arntzen (2011a) sought to examine the impact of using identical and arbitrary stimuli during simultaneous and 0, 3, 6, and 9s delay MTS procedures for an older adult with cognitive impairment. Adding a delay between presentation of the comparison and sample stimuli is one way to measure short-term memory (Arntzen, 2006). A serialized and MTO conditional discrimination procedure was used for arbitrary matching. Programmed consequences faded from 100 – 0% across phase two. Stimuli consisted of South Park cartoon characters.

The participant in Steingrimsdottir and Arntzen (2011a) had a diagnosis of Alzheimer's disease, an MMSE score of 20, indicating moderate cognitive impairment, and a prescription of 10mg Aricept taken once nightly. The participant formed matching relations in the 2-choice identity MTS, but performed at chance levels when presented with three comparison stimuli. The participant continued to respond in accordance to similarities between stimuli during arbitrary MTS after being exposed to identical MTS, which negatively impacted performance. After returning to identity matching using increasing delays, the participant demonstrated matching without programmed consequences (89 – 100% accuracy). However, the participant began taking an additional 5 mg of Aricept each morning, which correlated with a significant reduction in number of training trials needed to form matching (1,696 vs. 117).

Implications from Steingrimsdottir and Arntzen (2011a) are that individuals with cognitive impairment may benefit from a limited amount of comparison stimuli during training and may be impacted by exposure to identity matching prior to arbitrary matching. Steingrimsdottir and Arntzen (2014) found that typically functioning older adults were not impacted by exposure to identity matching prior to arbitrary matching, so the difference in performance may be related to cognitive functioning. Due to using identical MTS, Steingrimsdottir and Arntzen's (2011a) findings are limited to directly trained relations. Inclusion of a second participant with similar MMSE score who was first exposed to arbitrary MTS may have lead to interesting findings of directly trained and emergent relations.

Steingrimsdottir and Arntzen (2011b) also examined the impact of simultaneous and delayed identity matching on ability to form relations, but with a participant with severe cognitive impairment, as indicated by a MMSE score of 10. The participant was exposed to identity matching with two and three comparison stimuli. The participant was able to form matching relations when presented with two comparison stimuli, but performed at chances levels when presented with three choices during simultaneous MTS. The participant was also able to demonstrate matching using 0s delay MTS with praise delivered on a FR1 schedule, but performance decreased as the schedule of programmed consequences thinned.

To further examine the impact of delayed MTS procedures during identity matching, a titrated delay based on participant performance was used for an older adult with dementia (Arntzen, Steingrimsdottir, & Antonsen, 2013). The participant was the same woman who participated in Steingrimsdottir and Arntzen (2011a). Initially, the delay altered between 10,000 ms during the first and third phases and 12,000 ms during phase two. The participant was unable to demonstrate identity matching at the 12,000 ms delay, but was successful at the 10,000 ms delay. During the subsequent study, the delay began at 0s and increased or decreased by 250 ms for accurate and inaccurate responding on six consecutive trials, respectively. The participant was able to appropriately respond when the delay ranged between 7,500 – 12,250 ms. Therefore, Arntzen, Steingrimsdottir, and Antonsen (2013) argue that a titrated model based on participant performance may be appropriate for the promotion of learning in older adults with impairment.

Arntzen and Steingrimsdottir (2014) extended previous findings by examining a potential delay threshold and the impact of a 500 ms or 100 ms increase in delay during identity matching procedures for a 62-year-old man with dementia. The participant received an MMSE score of 23, indicating mild cognitive impairment. Differential feedback was provided for accurate and inaccurate responding. The participant responded with more stability and was also able to respond accurately at longer delays in the 100 ms step condition compared to the 500 ms step condition (6, 800 vs. 4,500 ms). Arntzen and Steingrimsdottir (2014) provided further evidence for the use of a titrated delay MTS procedure with a smaller step size for an individual with cognitive impairment.

Performance on delayed MTS using identical matching procedures was compared between 16 older adult participants that were typically functioning ($N = 8$, MMSE 29_M and 1_{SD}) and those diagnosed with AD (MMSE 17_M and 3_{SD}) (Plaza et al., 2012). Stimuli consisted of two sets of four photographs of men wearing suites that were associated with different conditions. Four comparison stimuli were presented 5s or 25s after presentation of the sample stimulus. The conditions differed by the consequence provided. Appropriate answers in the differential outcome condition were followed by the same consequence, “You may win an umbrella!” whereas appropriate answers in the non-differential condition were followed by the statement “You may win a(n) _______ (umbrella, mug, scarf, perfume)!” Therefore, appropriate responses in the non-differential outcome condition were not associated with a specific reward item.

The differing conditions assessed the stimulus control of a single vs. varied outcome (Plaza et al., 2012). Participants with cognitive impairment made more errors, but still engaged in 80% accurate responding overall. Responding was similar across differential and non-differential outcomes (89% vs. 85%) and delays (84% for the 25s delay and 89% for the 5s delay) across participants. However, responses for participants with AD were more negatively impacted by the increased delay and positively impacted by the differential outcome condition. These results suggest that short-term memory is impaired in individuals with AD and the act of remembering or responding under stimulus control may be improved by using the same consequence, instead of varied consequences for accurate responding.

Sidman (2013) discussed the potential utility of MTS procedures for helping to identify deficits and determine behaviors to shape in order to strengthen stimulus control for individuals with cognitive impairment. Sidman (2013) poses an insightful vignette related to stimulus equivalence. A father is asked to say his daughter's name, and cannot, yet is given his daughter's name and asked to identify her out of an array of photographs, and is able to make the appropriate selection. If the task is formulated one way, the association appears to be non-existent. If the task is posed another way, the association is demonstrated. Sidman (2013) argues that with the assistance of non-verbal conditional discrimination procedures, individuals with stimulus control deficits may be taught to engage in naming, such as with the potential generalization of autobiographical stimuli.

The purpose of the current studies is to first replicate previous research in Study I by analyzing the association between stimulus equivalence class formation and MMSE scores as conducted by Gallagher and Keenan (2009). Individuals with and without cognitive impairment will be recruited for the study to better examine the lines of fracture during equivalence performance. Secondly, Study I extends Gallagher and Keenan (2009) by also analyzing the association between SE class formation and SLUMS scores as well as ADL assessment scores. Study I also extends previous research by including stimuli from Gallagher and Keenan (2009) as well as arbitrary and familiar stimuli to examine the impact of stimulus type of equivalence formation.

Lastly, Study II will extend previous research by examining the impact of methodological changes by incorporating recommendations from the literature, such as increasing stimulus class size, changing the training modality, and evaluating increased programmed stimuli for individuals who did not form equivalence relations during Study I.

CHAPTER 3

GENERAL METHODS

Overview

The following studies served to systematically replicate and extend previous research related to the use of MTS procedures with older adults with and without cognitive impairment. A potential association between cognitive and functional assessment scores and stimulus equivalence performance using MTS procedures was analyzed. The procedures align with stimulus equivalence training as developed and refined by Sidman (1971; 1994).

Participants

Older adult participants within the same age range as Gallagher & Keenan (2009; Range: 57 - 94) were recruited for participation in the present study. Older adults were recruited via a site administrator at an area assisted living facility. The administrator first contacted the participant, if there was no known cognitive impairment, or the medical power of attorney, if cognitive impairment was suspected, to secure consent to be contacted by researchers. Participants engaged in the study on a voluntary basis after going through the informed consent process if they met criteria.

Inclusion criteria include a Mini-Mental Status Exam score of 16 – 30 (MMSE; Folstein, Folstein, & McHugh, 1975), diagnoses of Alzheimer's disease, dementia, stroke, traumatic brain injury, or Parkinson's disease or no cognitive-related diagnoses, and being 57 – 94 years of age. The MMSE is described in detail below, see Procedures. Exclusion criteria include being age 56 or younger at the inception of the study and visual or manual dexterity limitations that hinder ability to use a touch screen interface. Visual acuity was assessed via the Lighthouse Near Visual Acuity Test (2^{nd} ed.) (LNVAT; New York). Additionally, visual acuity and manual dexterity were met if the participant demonstrated the ability to touch the images during reflexivity training and testing, see Procedures. All participants demonstrated acceptable visual acuity and manual dexterity for study inclusion.

Originally, 12 participants were recruited for the study (age range: 68 – 94 years young, visual acuity range: 20/40 – 20/60 with correction, 3 males). After going through informed consent, one participant discontinued participation before beginning the assessment phase due to family considerations, one participant was excluded due to a low MMSE score (11), another discontinued after completing the assessment phase due to medical complications, and one participant did not demonstrate reflexivity after a 45 minute training session and was excluded from participation in the remainder of the study. See Table 1 for participants by order of recruitment and specific demographic data.

Apparatus, Setting, and Stimuli

Gallagher and Keenan (2009) used a tabletop format and 20 cm x 14 cm laminated stimuli to execute conditional discrimination training and testing. However, multiple researchers have demonstrated that older adults can use advanced technology during conditional discrimination procedures (Bódi, Csibri, Myers, Gluck, & Kéri, 2009; Engel, 2005; Saunders, Chaney, & Marquis, 2005, Steingrimsdottir & Arntzen, 2014). Digital stimuli were presented on a HP 13.3 inch Split x2 Touch-Screen 2-in-1 laptop. Training and testing procedures were programmed using Visual Express 2013 for Windows Desktop. Session data were automatically transferred to a note pad on the laptop and subsequently transferred to a HIPPA-secured laboratory computer. All sessions were videotaped using a Macintosh 13 inch MacBook Pro5,5 in a quiet room within each participant’s residence or day center.

Assessments

Cognitive Assessment

Participants meeting the age, visual, and manual dexterity requirements advanced to the assessment phase prior to Study I. Two cognitive assessments were administered, the MMSE (Folstein, Folstein, & McHugh, 1975) and the Saint Louis University Mental State Examination (SLUMS; Tariq, Tumosa, Chibnall, Perry III, & Morley, 2006; Appendix A). The MMSE is not included in the appendices as it is under copyright. First, participants completed the 5 – 10 min MMSE and had to achieve a score ranging from 16 – 30 in order to match participants from the Gallagher and Keenan (2009) study. Scores on the MMSE can range from 0 – 30, with scores lower than 24 denoting mild cognitive impairment and scores lower than 10 suggesting severe impairment (Folstein, Folstein, & McHugh, 1975). Actual score can be contrasted with expected score when adjusting for age and education (Tombaugh, McDowell, Kristjansson, & Hubley, 1996). However, Gallagher and Keenan (2009) found that only participants with scores ranging from 27 – 30 were able to demonstrate the formation of equivalence classes. This suggests the MMSE may not be sensitive enough to detect some forms of learning impairment given that participants were not able to demonstrate stimulus equivalence in the low-normal range on the MMSE (25 – 26).

The administration of the MMSE in the current study replicates the methodology of Gallagher and Keenan (2009) while also extending the cognitive assessment phase to include the SLUMS in order to provide further testing of a possible convergence between the two assessments with an increased sensitivity of the SLUMS to cognitive decline (Stewart, O'Riley, Edelstein & Gould, 2012). The sensitivity of the SLUMS may also more strongly correlate with stimulus equivalence performance than the MMSE. Similar to the MMSE, SLUMS scores range from 0 – 30, but have a slightly different scoring categorization. For individuals with a high school education or higher, scores on the SLUMS may be more in line with stimulus equivalence performance. Scores below 27 indicate mild neurocognitive impairment and scores below 21 indicate dementia (Tariq et al., 2006). Tariq and colleagues (2006) suggest that for individuals with less than a high school education, scores below 25 indicate mild neurocognitive disorder and scores below 20 indicate dementia. The MMSE and SLUMS have been administered on the same day, but were presented to the participants at least 24 hours apart due to item overlap related to orientation to time and recall of objects (Stewart, O'Riley, Edelstein & Gould, 2012).

ADL Assessment

After the cognitive assessment phase, participants or caregivers will complete a brief functional assessment of activities of daily living (ADL). The Barthel Index of Activities of Daily Living uses a descriptive scale to categorize ability to complete the following 10 tasks (a) voiding, (b) grooming, (c) toileting, (d) feedback, (e) transferring, (f) ambulating, (g) dressing, (h) using the stairs, and (i) bathing (see Appendix B) (Barthel ADL Index; Mahoney & Barthel, 1965). Scales range from 2 – 4 options depending upon the participant's level of functioning. For example, grooming ranges from 0 (needs assistance with personal care) to 1 (independent) and mobility ranges from 0 (immobile) to 3 (independent). The assessment is based on behavior the participant exhibits as opposed to behavior the participant may be capable of performing. Lower scores indicate more impairment.

The Barthel ADL Index was chosen as it includes more content items than competing ADL assessments, is widely used to assess functioning for older adults, and has high psychometric properties (Hartigan, 2007). Collins and colleagues (1998) demonstrated acceptable levels of agreement between different modes of administration: self-report, nurse's clinical observation, nurse's direct observation, and physiotherapist's direct observation. Incorporation of an ADL assessment is another extension beyond the methodology of Gallagher and Keenan (2009). The Barthel ADL Index was administered to examine a potential correlation between cognitive, functional, and stimulus equivalence performance, which has yet to be examined in the literature.

Procedural Integrity

A secondary observer reviewed 30% (n = 16) of training and testing sessions across Study I and II to evaluate the primary data collector's adherence to study procedures, which resulted in 96% treatment integrity. The Procedural Integrity Evaluation (PIE; see Appendix C for the complete PIE) form was used to measure treatment integrity. The PIE consists of 7 yes or no questions involving setting, directions delivered to participant, appropriate phase and stimuli used during session, appropriate session length, sole use of programmed differential consequences during training conditions, standardized experimenter statements for participant questions during training and testing, and adherence to termination criteria. Procedural items that could not be determined due to lighting or a technical error were scored as a zero to maintain a stringent procedural integrity analysis.

The secondary observer also reviewed 30% training and testing sessions to evaluate accurate coding, formulas, and data output from Visual Express into notepad, resulting in 100% programming integrity.

CHAPTER 4

METHODS: STUDY I

Purpose and Hypotheses

The purpose of the first study was to systematically replicate Gallagher and Keenan (2009) using a simultaneous LS conditional discrimination procedure with the same stimuli, as well as arbitrary and familiar stimuli, with older adult participants with and without cognitive impairment. Participants were administered the MMSE, as in Gallagher and Keenan (2009), and were also administered the SLUMS and Barthel ADL Index. The null hypotheses were that assessment score performance would be unrelated to performance during matching-to-sample (MTS) procedures, the SLUMS would not be more sensitive to the detection of cognitive impairment than the MMSE, and that use of familiar or Gallagher and Keenan (2009) stimuli would not result in more accurate responding than arbitrary stimuli. The main alternative hypothesis for the current study was that higher assessment scores would be related to better performance during tests of emergent relations, including accurate responding and fewer trials to criterion. Secondly, SLUMS scores would be more sensitive to the detection of cognitive impairment than MMSE scores, which would be correlated with performance during MTS training. Lastly, participants would engage in more accurate responding when presented with familiar stimuli and stimuli from Gallagher and Keenan (2009) than with arbitrary line drawings due to the participants' learning histories.

Participants

In total, eight participants began study one (age range: 68 – 88 years young, visual acuity range: 20/40 – 20/60 with correction, one male). The sample included typically functioning older adults as well as those with diagnoses of Alzheimer's disease and dementia. Some participants also suffered from diabetes, neuropathy, COPD, cataracts, incontinence, hypertension, high blood pressure, neuropathy, anxiety, and depression. Childhood and mid-life onset diabetes may interfere with cognitive functioning in middle age and late life (Nunley et al., 2015; Roberts et al., 2014). One participant discontinued participation after completing training with the stimuli from Gallagher & Keenan (2009) due to scheduling conflicts. Another participant discontinued participation after completing training with the Gallagher & Keenan (2009) and arbitrary stimuli due to frequent dialysis and fatigue.

Stimuli

The same symbols and non-sense syllables were used for training and testing in the current study (Gallagher & Keenan, 2009). Stimuli: A1 "Paf", B1 "&", C1 "○", A2 "Zid", B2 "@", and C2 "☐". Additionally, familiar and arbitrary stimulus sets were also incorporated to assess the possible impact of type of stimuli on the development of emergent relations given that the stimuli in Gallagher and Keenan (2009) involved familiar shapes and symbols. See Figure 1 for all stimulus sets. Thus, participants were exposed to six classes of stimuli with three members in each stimulus class. Stimuli appeared smaller (11 cm x 7 cm) than those used in Gallagher and Keenan (2009) in order to fit on the touch screen monitor.

Assessment

Reflexivity Training and Test

Following Gallagher and Keenan (2009), participants had to demonstrate reflexivity before moving to conditional discrimination training. During reflexivity training, participants were presented with a sample stimulus in the top middle of the screen while the same stimulus and another comparison stimulus were located below and to the left and right of the sample stimulus. For example, the participant had to successfully select comparison stimulus A1 when presented with sample stimulus A1. Participants were presented with a visual and auditory consequent stimulus for correct responses and a black screen for incorrect responses. The process was repeated for the remaining stimuli twice, so comparison stimuli could be randomly presented on the right and left of the sample stimulus.

After the mastery criterion of 90% or higher responding occurred during training, participants were exposed to reflexivity testing. The appearance of stimuli was the same as during reflexivity training. Feedback was not provided for incorrect or correct responses during the test. Participants that matched stimuli with 90% or higher accuracy moved to the matching phase of the study. In order to achieve this criterion, participants could only miss one trial during the reflexivity test (i.e., 11/12 = 92%). The reflexivity test also functioned as a guard against visual and manual dexterity limitations.

Procedures

Research Design

A linear series, or serial, conditional discrimination method was used to train relations (Gallagher & Keenan, 2009) across three types of stimuli. First, A-B relations were trained and subsequently tested for symmetry (i.e., B-A). Second, B-C relations were trained and also tested for symmetry (i.e., C-B). Participants were then tested for transitivity (i.e., A-C) and equivalence (i.e., C-A). Formation of equivalence relations as well as matching, symmetry, and transitivity was determined based on a 90% accuracy criterion. Data were analyzed based on progression during stimulus equivalence testing for (a) reflexivity, (b) symmetry, (c) transitivity, and (d) equivalence in comparison to MMSE, SLUMS, and ADL test scores similar to Gallagher and Keenan (2009). Trials to criterion were also analyzed (Saunders et al., 2005) and compared to both progression during testing and assessment scores.

Participant Instructions

Gallagher and Keenan (2009) did not specify instructions provided to participants. The present study incorporated the directions provided by Saunders et al. (2005) during training sessions. The participants were seated with the touch screen monitor turned away from the center mass of the participant. The experimenter then turned the monitor towards the center mass of the participant and the participant read, “Your task is to use the computer feedback to learn the relationships between the images” and pressed a button reading, “Begin,” to start the program. The screen displayed a sample stimulus. The experimenter instructed the participant to touch the image, at which time, two comparison stimuli appeared. The experimenter then said, “Touch another image” followed by, “Please continue.” For the current study, if participants asked for further directions during the session, the experimenter responded, “We can talk at the end.” Instructions were only provided before the first trial of each training session.

Matching and Symmetry

A two choice simultaneous match-to-sample (MTS) procedure was used to train A-B and B-C matching with programmed consequences for accurate responding. Stimulus placement was identical to the reflexivity test. When the desired stimulus was selected, an auditory and textual stimulus of, “Correct” was displayed for 1500 ms. No feedback was provided for incorrect answers. Matching training began with A-B relations in which A1 was presented as the sample stimulus, B1 and B2 were presented as comparison stimuli, and when B1 was selected in A1’s presence “Correct” was displayed for 1500 ms. The same procedure was used to train the A2-B2 relation. If participants failed to demonstrate improvement towards the matching mastery criterion after 45 minutes of training, their participation in experiment one was terminated. Previous research suggests that older adult participants form relations via linear MTS procedures within 9 – 173 trials with simultaneous presentation of stimuli and 9 – 266 trials with delayed presentation of comparison stimuli (Saunders, Chaney, & Marquis, 2005).

Training was followed by a symmetry test if 90% or higher accuracy was achieved. The test included 20 counterbalanced trials of B1-A1 and B2-A2 presentations with no programmed feedback. Participants had to correctly identify the comparison stimulus 90% or more of the time (i.e., 18/20 trials) to move to B-C matching. The procedures and accuracy criterion described above also applied to B-C matching training and symmetry testing. After the C-B symmetry test, the participant advanced to the transitivity test.

Test transitivity

Participants were presented with sample stimuli A1 or A2 and comparison stimuli C1 and C2. Presentation form was the same as in other training and testing conditions. Comparison stimuli were counterbalanced between left and right positions. A1-C1 and A2-C2 test trials were also counterbalanced over 20 trials. No feedback was provided during the test. Participants with 90% or higher accuracy were deemed to have demonstrated transitivity (i.e., 18/20 correct trials).

Test equivalence

The test involved the same procedures and format as the transitivity test, except the sample stimuli consisted of C1 and C2 with A1 and A2 serving as comparison stimuli. Participants demonstrated equivalence (i.e., C1-A1 and C2-A2 relations) if 90% or more trials were correct (i.e., 18/20 correct trials).

Data Analysis

A Shapiro-Wilks test of normality was conducted for each assessment. A classification analysis based on cut-off scores, for the MMSE and SLUMS, was conducted to test the null hypothesis that scores on the SLUMS would not be more sensitive to the detection of cognitive impairment than the MMSE.

Spearman's Rho correlations were conducted to examine potential associations between each assessment (MMSE, SLUMS, and Barthel ADL Index) and each participant's composite score from LS training after a Shapiro-Wilks test of normality was conducted for the LS composite score. Conducting correlations between assessment scores and composite LS performance tests the null hypothesis that assessment score performance is unrelated to performance during MTS procedures. Composite LS training scores were created based on highest level of participant performance for each stimulus type. Participants received 0 points for failing to demonstrate matching, 1 for demonstrating matching, 2 for demonstrating symmetry, 3 for demonstrating transitivity, and 4 for demonstrating equivalence for each stimulus type. Then, each stimulus type score was summed into one composite LS training score for a possible range of 0 – 12 points for each participant. Composite scores are ordinal, as scores are rank ordered, but are not reflective of equally spaced intervals.

The non-parametric tests were conducted with the alternative hypothesis that higher assessment scores were related to the emergence of more untrained relations and the null hypothesis that assessment scores were unrelated to performance during matching-to-sample (MTS) procedures. Non-parametric tests are appropriate for data that is not normally distributed and ordinal. Performance outcome data were also visually analyzed for each participant via a table similar to the analysis in Gallagher and Keenan (2009).

Extended Analysis: Study Two

All individuals that completed Study I moved to Study II, see chapter 7. Study II combined several methodological changes to attempt to identify best practice for training relations with older adults.

CHAPTER 5

RESULTS: STUDY I

Statistical Analyses

Test of Normality and Descriptives

A Shapiro-Wilk test of normality was conducted for MMSE, SLUMS, and Barthel ADL Index, and LS composite score variables. MMSE scores were significant ($S\text{-}W = .817$, $df = 9$, $p < .05$), thus indicating a non-normal distribution. SLUMS ($S\text{-}W = .885$, $df = 9$, *ns*), Barthel ADL Index ($S\text{-}W = .911$, $df = 9$, *ns*), and LS composite variables ($S\text{-}W = .960$, $df = 6$, *ns*) were not significant. Visual analysis of histograms indicated the MMSE ($M = 25.18$, $SD = 5.72$), SLUMS ($M = 21.8$, $SD = 5.69$), and Barthel ADL Index ($M = 16.36$, $SD = 2.46$) had a negative skew, meaning the majority of scores were higher than the mean.

Spearman's Rho

Non-parametric correlations were then conducted given the non-normal distribution of MMSE scores, skewness of assessment scores variables, and small sample size. Association between composite performance during linear series (LS) training ($M = 3$, $SD = 2.6$) and assessment score variables were tested. MMSE and composite LS performance were positively correlated with a large effect size ($r_s = .790$, $p < .05$). Composite LS performance was not significantly associated with SLUMS ($r_s = .277$, *ns*) or Barthel ADL Index scores ($r_s = .085$, *ns*). The findings reject the null hypothesis in regard to the correlation between MMSE scores and composite LS scores and fail to reject the null hypothesis in relation to a lack of significant association between SLUMS or Barthel ADL Index score and LS composite performance. There is some evidence that MMSE scores are related to performance during MTS procedures.

Level of Impairment Classification

A categorization analysis based on cut-off scores between the MMSE and SLUMS was conducted. Scores on the MMSE can range from 0 – 30, with scores lower than 24 denoting mild cognitive impairment and scores lower than 10 suggesting severe impairment (Folstein, Folstein, & McHugh, 1975). Scores on the SLUMS can also range from 0 - 30, with scores below 27 indicating mild neurocognitive impairment and scores below 21 indicating dementia for individuals with a high school diploma or higher level of education (Tariq et al., 2006). Two participants (25%) fell within the normal range on both assessments, four participants (50%) fell within the Mild Neurocognitive Disorder range on the SLUMS and normal range on the MMSE, and two participants (25%) scored within the dementia range on the SLUMS and normal or mild cognitive impairment range on the MMSE. Categorization between the SLUMS and MMSE differed for 75% of the present sample (see Figure 2). Difference scores ranged from 1 – 15 points. The SLUMS correctly classified 67% (n = 2) whereas the MMSE correctly classified 0% of the sample that had a diagnosis of dementia or Alzheimer's disease (n = 3). The classification analysis provides evidence for rejecting the null hypothesis that SLUMS scores are not more sensitive to the detection of cognitive impairment than MMSE scores given overall lower scores on the SLUMS than the MMSE that classify participants as having some form of cognitive impairment.

Equivalence Formation Performance by Stimulus Type

Equivalence formation performance data are presented in a table for visual inspection. Table 2 includes highest level of mastery: reflexivity, symmetry, transitivity, or equivalence and assessment scores. See Table 3 for trials to criterion for each directly trained relation, percentage of accuracy of responses during tests for emergent relations, and overall session time. These tables provide an overall illustration of each participant's performance. Gallagher and Keenan (2009) suggested equivalence is formed for 90% of individuals with an MMSE score of 27 or higher. The tables help to visually illustrate if the current findings converge or diverge with Gallagher and Keenan's (2009) results with the addition of scores on the SLUMS and Barthel ADL index.

Gallagher and Keenan (2009) Stimuli

Of the 12 recruited participants, eight were tested using a linear series (LS) training format for at least one type of stimulus, see Table 2. Of these eight participants, ID 3 and ID 12 demonstrated equivalence formation (25%). Both of these participants only demonstrated equivalence with the stimuli from Gallagher and Keenan (2009), had the highest MMSE scores (29, 30), two of the highest SLUMS scores (27, 25), and different Barthel ADL Index scores (14, 17). The participant with a perfect score on the MMSE (30) was able to form equivalence in approximately 452 s (7 min, 32 s) and the participant with a 29 demonstrated equivalence after 447 s (7 min, 27 s). Both participants formed equivalence with only 29 trials to criterion for directly trained relations.

Session lengths ranged from 850 (14 min, 10 s) – 2,982 s (49 min, 42 s) and trials to criterion ranged from 119 - 539 for the remaining six participants, see Table 4. One participant, with a MMSE score of 22, demonstrated CB symmetry with the stimuli from Gallagher and Keenan (2009). Three participants with MMSE scores of 27 – 28 and SLUMS scores of 23, 24, and 28 demonstrated matching with the stimuli from Gallagher and Keenan (2009). One of those participants, ID 8, (MMSE = 28, SLUMS = 23) almost met mastery criteria for BA symmetry (85%) and transitivity (75%) of emergent relations. Two other participants with MMSE scores in the normal range (27, 28) did not form matching relations (range 767 – 847 for matching trials at session termination) after 45 minutes of training for the Gallagher and Keenan (2009) stimuli. These two participants had disparate SLUMS (12, 23) and Barthel ADL Index scores (15, 11), respectively.

Familiar Stimuli

None of the participants from Study I formed equivalence using familiar stimuli, see Table 2. Trials to criterion ranged from 39 – 90 and session length ranged from 594 (9 min, 54 s) – 2,163 s (36 min, 3 s) for participants that demonstrated matching, see Table 3. Two participants were able to demonstrate CB symmetry and one participant was able to form matching using familiar stimuli. The participant who demonstrated matching relations almost met criteria for symmetry (85%) and had a perfect MMSE score. However, three participants failed to form matching relations for familiar stimuli (MMSE range 22 – 27, matching trials range 64 – 676 at session termination).

Arbitrary Stimuli

Similar to the results while using familiar stimuli, none of the participants demonstrated equivalence formation using arbitrary stimuli during Study I, see Table 2. Trials to criterion ranged from 29 – 61 and session length ranged from (9 min, 29 s) – 2,220 s (37 min) for participants who formed matching, see Table 3. One participant demonstrated transitivity (MMSE = 28) and another demonstrated BA and CB symmetry (MMSE = 30) for arbitrary stimuli. The participant who demonstrated transitivity formed AB and BC matching, but was not able to demonstrate BA and CB symmetry. The participant with a perfect MMSE score had near 0% accuracy for transitivity and equivalence suggesting the formation of the exact opposite relations of those trained. The five remaining participants failed to demonstrate matching relations with arbitrary stimuli (MMSE range 22 – 29, matching trials range 172 – 642 at session termination).

Performance Comparison by Stimulus Type

During Study I, 67% (*n* = 4, ID numbers 2, 8, 10, 12) of participants demonstrated a higher level of performance when presented with stimuli from Gallagher and Keenan (2009) than when presented with familiar or arbitrary stimuli. Two participants only formed matching (ID numbers 2 and 10), one participant only demonstrated higher than chance levels of responding during transitivity (ID number 8), and one participant only demonstrated equivalence (ID number 12) with Gallagher and Keenan (2009) stimuli. Two other participants demonstrated matching or equivalence relations for Gallagher and Keenan (2009) stimuli (25%, ID numbers 1 and 3), but did not complete the remainder of Study I, so their performance cannot be compared to equivalence formation when using familiar or arbitrary stimuli. Arbitrary stimuli were the optimal stimulus type for a single participant (12.5%, *n* = 1, ID number 9).

Participant 9 demonstrated fewer trials to criterion (*n* = 61) for matching when presented with arbitrary stimuli than for Gallagher and Keenan (2009) stimuli (*n* = 300) or familiar stimuli (*n* = 79). Participant 9 also only formed transitivity (90%) with arbitrary stimuli. Another participant's performance did not indicate an optimal stimulus type (12.5%, *n* = 1, ID number 5). Participant 5 was unable to form matching across all three stimulus types. Thus, the results provide support for rejecting the null hypothesis that use of Gallagher and Keenan (2009) stimuli would not result in more accurate responding than arbitrary stimuli, but fail to reject the null hypothesis that familiar stimuli would not result in more accurate responding than arbitrary stimuli.

CHAPTER 6

DISCUSSION: STUDY I

The current study was unable to replicate the findings from Gallagher and Keenan (2009). Only 25% (*n* = 2) participants were able to demonstrate equivalence using the previous study's stimuli, and no participants were able to demonstrate equivalence with arbitrary or familiar stimuli. The two participants who were able to form equivalence had perfect and near-perfect MMSE scores (29 and 30). In Gallagher and Keenan (2009), 90% of participants demonstrated equivalence class formation with MMSE scores of 27 and higher. However, Steingrimsdottir and Artzen (2014) found that only 66% of participants with MMSE scores of 28 and above demonstrated equivalence. These two participants also had two of the highest SLUMS scores (25 and 30) and were able to rapidly form equivalence with only 9 – 20 trials to criterion per relation in approximately 7 min 30 s. Previous research has found that participants have achieved equivalence in as little as 9 trials, but others have taken 173 trials in a simultaneous training format (Saunders, Chaney, & Marquis, 2005).

Despite a lack of equivalence demonstration, performance was indicative of learning throughout training and testing using the simultaneous LS training format and Gallagher and Keenan (2009) stimuli. Six participants demonstrated AB matching and five participants demonstrated BC matching. Two participants formed BA symmetry, although another participant was at 85% accuracy. Three participants formed CB symmetry, two participants formed transitivity, and two more performed at higher than chance levels (55 and 65%) during the transitivity test. Only 25% (*n* = 2) of the participants were unable to demonstrate AB and BC matching.

When presented with arbitrary stimuli, two participants formed AB and BC matching, one participant (ID 12) also formed BA and CB symmetry, and the other (ID 9) formed transitivity. Participant 12 demonstrated near zero percent accuracy during transitivity and equivalence tests, suggesting the formation of the opposite relations as those trained. Participant 12 spoke aloud during some training and testing sessions and also reported a strategy for remembering relations, see Other Considerations, located below, for more information. Transitivity and not equivalence likely formed for participant 9 due to the formation of AB and BC relations without the formation of BA and CB symmetry. The majority of participants were unable to form matching within the pre-determined session length or before self-terminating the session.

Half of the participants were able to demonstrate AB and BC matching when familiar stimuli were used during MTS training. Two participants formed CB symmetry and a third participant (ID 12) almost achieved both BA and CB symmetry. Again, participant 12 achieved 0% accuracy during transitivity and equivalence testing, suggesting the formation of the opposite relations as those trained or interference from covert and overt verbal behavior. Formation of opposite relations may have occurred due to sample/S- control, in which the participant selects against an inaccurate stimulus, instead of selecting for the appropriate stimulus (Carrigan & Sidman, 1992). No participants performed at higher than chance levels during the transitivity and equivalence tests. Fifty percent of participants were unable to form AB matching after 45 min of training.

Overall, participants performed better when using stimuli from the Gallagher and Keenan (2009) study. Therefore, accuracy during training and testing of Gallagher and Keenan (2009) stimuli greatly influenced participant composite LS performance scores. For instance, participant

2 earned two points for demonstrating CB symmetry with Gallagher and Keenan (2009) stimuli and earned zero points for failing to demonstrate matching with arbitrary or familiar stimuli. Participant 12 was awarded four points for demonstrating equivalence with Gallagher and Keenan (2009) stimuli and only two points for demonstrating symmetry with arbitrary stimuli and just one point for demonstrating matching with familiar stimuli.

Participants may have performed better when trained with Gallagher and Keenan (2009) stimuli than familiar or arbitrary stimuli given that the shapes and symbols were non-arbitrary, but the relations between the stimuli were arbitrary. Participants may have been able to easily come under the stimulus control of these stimuli given their learning history without interference due to how the stimuli were non-arbitrarily related within the MTS program. However, learning histories may have interfered with the development of equivalence class formation for familiar stimuli. Participants may have been attempting to group stimuli using rule-governed behavior, instead of programmed contingencies, as has been noted in previous research (Wilson & Milan, 1995). For example, one participant asked if the stimuli were related by indoor or outdoor location after a training session. Distractor stimuli were not used, yet the participants who did not form matching continued to engage in responses that resulted in a black screen, as opposed to a screen displaying “Correct!”

Differences in training modality may have impacted participant performance in the current study. Gallagher and Keenan (2009) used a tabletop format to deliver equivalence training and testing. The current study used a computer-based training and testing interface. Tabletop administration can be an effective means to shape learning, but may increase the likelihood of threats to internal validity (Dymond, Rehfeldt, & Schenk, 2005). Experimenters may inadvertently provide feedback to participants by the way accurate and inaccurate responses are recorded on a data sheet or through facial gestures or body language while using a tabletop format.

Additionally, trial as well as intertrial time may be longer when using a tabletop format, which may result in greater exposure to stimuli during the trial and a longer delay before the next sample stimulus is presented. A slower training pace may impact performance. In the present study, some participants touched the screen rapidly. Participants may have not attended to the sample or comparison stimuli in order to make an accurate selection, and instead, may have simply touched the screen to move to the next trial or to have the effect of making stimuli disappear or appear. Some participants also alternated between selecting the comparison stimuli to the right and left of the sample stimulus, regardless of the stimuli on the screen. Other participants continued to touch the screen while the consequent stimulus was displayed.

In contrast, stimulus control may have been strengthened via the tabletop training format in Gallagher and Keenan (2009). Participant responses may have come under the stimulus control of experimenter feedback better than computer feedback. Experimenters said, “Correct!” for accurate responding and did not provide verbal feedback for inaccurate responding in Gallagher and Keenan (2009). These procedures were replicated as best as possible in the current study using a computer screen. A white screen with black text read, “Correct!” following accurate responses or went to all black for inaccurate responses. A computer screen could not exactly replicate the behavior of an experimenter. Participants may have been confused by the intertrial black screen not realizing it followed inaccurate responding, and may have thought the computer malfunctioned or that feedback indicating they were correct occurred on a variable ratio schedule. Measures to enhance stimulus control and evaluate different training structures were incorporated into methodological changes for Study II.

CHAPTER 7

METHODS: STUDY II

Purpose and Hypotheses

The purpose of Study II was to examine the impact of methodological changes to the MTS procedures used in Study I to assess for a best practice approach for training relations among older adults with and without impairment. The null hypotheses were that participants would not benefit from methodological changes during Study II, as such higher assessment scores would not be related to equivalence formation after methodological changes, and OTM and MTO training structures would not relate to improved performance over equivalence formation using an LS training structure in Study I. The alternative hypothesis was that participants with higher MMSE, SLUMS, and ADL assessment scores who did not benefit from linear series simultaneous MTS procedures would demonstrate an improvement in stimulus equivalence formulation after delayed MTO and OTM training in Study II, including more accurate responding and fewer trials to criterion.

Participants

All individuals that completed Study I, regardless of demonstration of equivalence relations, directly advanced to Study II. All demographic information remained the same for the participants (see Table 1).

Apparatus, Setting, and Stimuli

The apparati and settings were the same from Study I to Study II, although only arbitrary and optimal stimuli were used. Arbitrary stimuli were similar to those used in Study I and are in line with recommendations from previous research that stimuli should not be easily recognizable or familiar (Saunders et al., 2005; Steingrimsdottir & Arntzen, 2014). However, if another type of stimulus was optimal during Study I that type was also tested in Study II. For instance, during Study I, participant performance was highest for the stimuli from Gallagher and Keenan (2009) for 67% of the participants who took part in Study II. Therefore, those four participants underwent stimulus equivalence training and testing for both arbitrary stimuli and stimuli similar to Gallagher and Keenan (2009), as the previous study only used 2 3-member classes. Non-sense syllables and common symbols and shapes were utilized in Study II to imitate the stimuli from Gallagher and Keenan (2009). The two remaining participants were only presented with arbitrary stimuli given one participant's inability to form A-B matching for all three types of stimuli during Study I and another participant's optimal performance occurred with arbitrary stimuli. Familiar stimuli were not used in Study II due to low-level performance when using said stimuli during Study I. See Figure 3 for Study II arbitrary and optimal stimuli.

Participants were exposed to three 3-member stimulus sets for each type of stimulus resulting in exposure to either three or six total stimulus sets. Participants who were only presented with arbitrary stimuli were trained with three total stimulus sets and participants who trained with both arbitrary and optimal stimulus sets were presented with a total of six 3-member stimulus sets. A 3-choice MTS procedure was used for older adult participants to determine if higher levels of class formation result from training more stimulus sets in a single session.

Procedures

Research Design and Overview

Participants were exposed to 0s delay one-to-many (OTM) and many-to-one (MTO) conditional discrimination procedures with differential feedback during training. A 0s delay has been found to be equivalent to (Steingrimsdottir & Arntzen, 2014) and more effective (Saunders et al., 2005) than simultaneous presentation. Presentation of stimuli on the touch screen monitor were the same in Study II as in Study I, with the sample stimulus located in the top middle and comparison stimuli situated below and to the left and right of the sample stimulus. Once participants select the sample stimulus it will disappear and two sample stimuli will appear (i.e., 0s delay).

If the appropriate selection is made, a green background with the audible and textual stimulus, "Correct!" were displayed on the screen for 1500 ms. If participants select the incorrect comparison stimulus, a red background with an audible and textual stimulus, "Wrong" were displayed for 1500 ms. Gallagher and Keenan (2009) provided an audible and textual stimulus for appropriate responses, but did not differentiate the color of the background, or provide a programmed response for inappropriate responses. Previous research has utilized differential feedback with older adults (McHugh & Reed, 2007; Pérez-González & Moreno-Sierra; 1999; Saunders et al., 2005; & Steingrimsdottir & Arntzen, 2014).

The orders of exposure were counterbalanced across participants with different stimuli for each type of training as can be seen in Table 4. OTM and MTO training formats were incorporated as previous research has found both to be effective in helping participants form relations. OTM has been recommended by Arntzen and Holth (1997; 2000), although younger participants and simultaneous presentation of sample and comparison stimuli were used. With older adult participants, OTM and MTO were effective formats for training equivalence relations, but MTO further superseded linear series formatting for number of training sets meeting the criteria for equivalence class establishment (Saunders et al., 2005). A combination of linear and MTO training was also successful (Steingrimsdottir & Arntzen, 2014). Linear series training has also been successful (Gallager & Keenan, 2009; Pérez-González & Moreno-Sierra, 1999). However, given that current participants were not able to form stimulus equivalence using a linear series format, MTO and OTM procedures were used in the present study.

Participant Directions

Directions were similar to Study I, but incorporated a supplemental prompt. Before training on the first set of stimuli, the experimenter turned the monitor towards the center mass of the participant. The participant read, “Your task is to use the computer feedback to learn the relationships between the images” with the additional prompt, “The pictures do not have to match,” and press a button reading, “Begin,” to start the program. As during Study I, the participant was instructed to, “Touch an image,” followed by, “Touch another image.” The experimenter stated, “We can talk at the end,” if the participant asks procedural or other questions during the session.

The addition of another instruction has been used when participants are first exposed to identity MTS and subsequently presented with arbitrary MTS procedures (Steingrimsdottir & Arntzen, 2011b). Participants were first exposed to identical matching during the reflexivity test during study I, which may have hindered ability to form equivalence. The additional prompt may help clarify the current relations.

Assessment

The reflexivity test was omitted in Study II, as participants demonstrated reflexivity during Study I. Additionally, Sidman (1994) as well as Steingrimsdottir and Arntsen (2011b) mentioned that exposure to identical MTS prior to arbitrary MTS may hinder performance, as participants continue to respond in accordance with similarities among stimuli, instead of in relation to current contingencies. The reflexivity test is similar to identical MTS, therefore, eliminating the test, in addition to the additional prompt described above, helped guard against potential interference during conditional discrimination training. However, a recent study with one older adult participant did not find a blocking effect when arbitrary MTS was followed by identical MTS, indicating potential idiosyncratic effects (Steingrimsdottir & Arntzen, 2014).

OTM Matching and Symmetry

Participants were exposed to a 3-choice, 0s delay MTS using OTM training. Participants began with A1-B1, A2-B2, and A3-B3 trials, in which sample stimulus A was presented at the top, middle of the screen and B and C stimuli were presented at the bottom of the screen after a 0s delay. Selecting B1 in the presence of A1, B2 in the presence of A2, and B3 in the presence of A3 resulted in the programmed consequences for an appropriate response, as described above. If an inaccurate stimulus was selected, participants were exposed to the corresponding consequences, also described above. Trial blocks consisted of 15 trials. If participants failed to demonstrate improvement towards the matching mastery criterion after 45 minutes of training (i.e., 90% accuracy, 14/15 trials), they moved to MTO matching training, described below, during the next session.

If 90% accuracy was demonstrated during the matching training for A-B stimulus pairs, participants were exposed to A-C matching across all stimulus sets in the same format as A-B matching training. Once the A-C matching mastery criterion was reached (90%, 14/15 trials) participants advanced to a combined matching training. Combined matching training was used to help guard against order effects and time lapse between the initial and most recently trained AB and AC pairs. Participants had to demonstrated mastery levels of accuracy during combined matching training (90%, 27/30 trials) before advancing to the combined symmetry test. Sessions were terminated if participants did not demonstrate combined matching after 45 minutes of exposure to combined matching trials.

The symmetry test involved 30 counterbalanced trials of B-A and C-A relations across all three stimulus sets without programmed consequences. If participants were able to correctly identify the comparison stimulus during 27 trials (i.e., 90% accuracy), they met symmetry mastery criteria. After the symmetry test, participants advanced to the OTM equivalence test.

OTM Equivalence Test

Participants were exposed to 30 counterbalanced trials of untrained relations (i.e., B1-C1, C1-B1, B2-C2, C2-B2, B3-C3, and C3-B3) without programmed feedback. Stimulus presentation was the same as during concurrent matching training and symmetry testing. Participants demonstrated equivalence if 90% or more trials were correct (i.e., 27/30).

MTO Matching and Symmetry

Similar procedures were used for MTO training as with OTM training, with the exception of the trained relations. B1 and C1 were directly trained to A1 throughout matching training (e.g., B1-A1 and C1-A1) using trial blocks composed of 15 trials. Participants first demonstrated mastery with B-A relations before exposure to C-A matching trials (90%, 14/15 trials). Participants were also exposed to the programmed, differential consequences described above. If matching was demonstrated with 90% or higher accuracy, participants completed combined matching training. Combined matching training was used to help guard against order effects and time lapse between the initial and most recently trained BA and CA pairs. Participants had to demonstrated mastery levels of accuracy during combined matching training (90%, 27/30 trials) before advancing to the combined symmetry test. Sessions were terminated if participants did not demonstrated combined matching after 45 minutes of exposure to combined matching trials.

The combined symmetry test consisted of 30 counterbalanced trials of A-B and A-C relations across all three stimulus sets. Participants demonstrated mastery of forming symmetry if they answered with 90% or higher accuracy (27/30 trials). After the symmetry test, participants advanced to the equivalence test.

MTO Equivalence Test

As with the OTM equivalence test, participants were presented with 30 counterbalanced trials of untrained relations (i.e., B1-C1, C1-B1, B2-C2, C2-B2, B3-C3, and C3-B3) without feedback. Stimulus presentations were the same as during matching training and symmetry testing. Participants demonstrated equivalence if 90% or more trials were correct (27/30).

Data Analysis

A Shapiro-Wilks test of normality was conducted for each composite score. Spearman's Rho correlations were conducted to examine the potential association between assessment scores and composite OTM and MTO training scores to test the null hypothesis that higher assessment scores would not be related to equivalence formation after methodological changes from those used in Study I. A Spearman's Rho correlation was also conducted on OTM, MTO, and LS training scores test the null hypothesis that OTM and MTO training structures would not relate to improved performance over equivalence formation using an LS training structure in Study I. Composite training scores were created based on highest level of participant performance for each stimulus type. Participants received 0 points for failing to demonstrate matching, 1 for demonstrating matching, 2 for demonstrating symmetry, and 3 for demonstrating equivalence for each stimulus type. Then, each stimulus type score was summed into one composite OTM and MTO training score for a possible range of 0 – 3 points for participants who were only presented with arbitrary stimuli and 0 – 6 points for participants who were presented with arbitrary and optimal stimuli. Composite scores are ordinal, as scores are rank ordered, but are not reflective of equally spaced intervals. The outcome data are also displayed for each participant via tables similar to previous research for visual inspection (Gallagher & Keenan, 2009; Saunders et al., 2005).

CHAPTER 8

RESULTS: STUDY II

Statistical Analyses

Test of Normality and Spearman's Rho

OTM ($S\text{-}W = .702$, $df = 6$, $p < .01$) and MTO ($S\text{-}W = .611$, $df = 6$, $p < .01$) scores were significant, thus indicating a non-normal distribution. Association between composite performance during one-to-many (OTM) ($M = .83$, $SD = 1.33$) and many-to-one (MTO) training ($M = 1$, $SD = 2$) and assessment score variables was conducted. Composite OTM performance was not significantly associated with MMSE ($r_s = .679$, *ns*), SLUMS ($r_s = .429$, *ns*) or Barthel ADL Index scores ($r_s = -.600$, *ns*). Composite MTO performance was not significantly associated with SLUMS ($r_s = .600$, *ns*) or Barthel ADL Index scores ($r_s = -.429$, *ns*). However, the association between composite MTO performance and MMSE score was positively trending towards significance ($r_s = .783$, $p = .065$). These findings lend support to fail to reject the null hypothesis that higher assessment scores relate to equivalence formation.

Spearman's Rho was also conducted between LS, OTM, and MTO composite scores. Composite LS scores were not significantly correlated with OTM ($r_s = .541$, *ns*) or MTO ($r_s = .676$, *ns*) composite scores. Composite OTM and MTO performance scores were positively and strongly correlated ($r_s = .920$, $p < .01$) suggesting that participants either performed well or poorly across both training structures. These findings fail to reject the null hypothesis that methodological changes from Study I to Study II, including adoption of MTO and OTM training structures, do not relate to improved equivalence formation performance. However, the findings support the need for further exploration of the methodological changes that may have positively impacted performance, given that some participants demonstrated matching, symmetry, and equivalence, and the need to evaluate performance between Studies I and II, see below.

Determination of Optimal Stimulus Set Type

During Study I, 67% (*n* = 4) of participants demonstrated a higher level of performance when presented with stimuli from Gallagher and Keenan (2009) than when presented with familiar or arbitrary stimuli. Therefore, stimuli similar to those used in Gallagher and Keenan (2009) served as optimal stimuli for participants 2, 8, 10, and 12 during Study II. Arbitrary stimuli were the optimal stimulus type for participant 9. Thus, participant 9 was only exposed to arbitrary stimuli during Study II, as she demonstrated transitivity with arbitrary stimuli, and did not demonstrate matching with Gallagher and Keenan (2009) or familiar stimuli during Study I. Participant 5's performance during Study I did not indicate an optimal stimulus type. Participant 5 was also only exposed to arbitrary stimuli during Study II due to failing to form matching across arbitrary, familiar, or Gallagher and Keenan (2009) stimuli during Study I.

Equivalence Formation Using OTM Training

As mentioned previously, see Table 4 for participant order of exposure for stimulus type and training structure and Figure 2 for images of stimuli used during Study II. Highest level of equivalence class formation with percentage accuracy of responses and assessment scores by stimulus type and training structure can be located in Table 5. See Table 6 for trials to criterion, percentage accuracy for single and combined stimulus sets, and session length by stimulus type and training structure.

Optimal Stimuli

Two participants (ID 12 and 8) demonstrated matching with stimuli similar to those used in Gallagher and Keenan (2009). Matching trials to criterion ranged from 30 – 89 for each trained relation and 58 – 298 for combined matching training. Participant 12, who had a perfect MMSE score, almost met the mastery criterion for symmetry (87%) and equivalence (87%). A stimulus set analysis revealed that participant 12 failed to form symmetry and equivalence for the second set of stimuli, but was able to form equivalence for the first and third stimulus sets. In contrast Participant 12 was able to form equivalence with these stimuli during Study I. Participant 8 also almost met mastery criteria for demonstration of symmetry (83%). However, failed to form symmetry with the first set of stimuli (50%), but demonstrated perfect symmetry for sets 2 (100%) and 3 (100%). Participant 8 also almost met mastery BA symmetry criteria during Study I with similar stimuli (85%). Session length for participants 12 and 8 ranged from 1851 (30 min, 51 s) – 2,197 s (36 min, 37 s). Participants 2 and 10 failed to form matching after 502 – 523 trials and had the lowest SLUMS scores of all participants (12, 14). However, these two participants were able to form matching with Gallagher and Keenan (2009) stimuli using a simultaneous LS training structure during Study I. Overall, 75% of participants performed better using these stimuli during Study I. These data support failing to reject the null hypothesis that an OTM training structure would not relate to improved performance over equivalence formation using an LS training structure in Study I.

Arbitrary Stimuli

The same two participants (8 and 12) that formed matching with optimal stimuli also formed matching with arbitrary stimuli. Matching trials to criterion ranged from 45 – 390 for each trained relation and 28 – 90 for combined matching training. Session length ranged from 1,473 (24 min, 33 s) – 3,278 s (54 min, 38 s) for participants 8 and 12.

Participant 8, who had a high MMSE score (28), met the mastery criterion for symmetry (90%) and demonstrated higher than chance levels of equivalence formation (73%). A stimulus set analysis revealed that participant 8 failed to form symmetry (70%) for the first set of stimuli, but perfectly demonstrated symmetry for sets 2 and 3 (100%). Participant 8 failed to form equivalence with set 1 (40%), nearly formed equivalence with the second of stimuli (80%), and was able to form equivalence with the third stimulus set (100%). In contrast, participant 8 was unable to form matching after 642 trials using arbitrary stimuli during Study I.

Participant 12 (MMSE = 30) almost met mastery criteria for demonstration of symmetry (83%). However, failed to form symmetry with the third set of stimuli (50%), but demonstrated perfect symmetry for sets 1 (100%) and 2 (100%). Participant 12 also almost met mastery criteria for demonstration of equivalence (83%) with higher than chance levels of equivalence for sets 1 (70%) and 3 (80%) and demonstration of equivalence for set 2 (100%). Participant 12 was able to demonstrate BA symmetry (100%) and almost demonstrated CB symmetry (85%) during Study I. However, participant 12 performed at near zero levels for transitivity (15%) and equivalence (0%) during Study I.

Participants 2, 5, 9, and 10 failed to demonstrate performance improvements indicating movement towards matching after 304 – 704 trials and the 2,700 s (45 min) time limit, except participant 2 self-terminated the training session after 2,289 s (38 min, 9 s). MMSE scores ranged from 22 – 28, SLUMS scores ranged from 12 – 24, and Barthel ADL Index scores ranged from 15 – 20 for participants 2, 5, 9, and 10. Three of these participants (ID numbers 2, 5, 10) were also unable to form matching using arbitrary stimuli during Study I, except for participant 9, who demonstrated matching after 61 trials.

These data provided mixed results for the methodological changes in Study II. Two participants demonstrated higher than chance levels of responding for emergent relations while the performance of three other participants remained unchanged, and one participant's performance worsened during Study II. These data are inconclusive for either rejecting or failing to reject the null hypothesis that OTM and MTO training structures would not relate to improved performance over equivalence formation using an LS training structure in Study I.

Equivalence Formation Using MTO Training

Optimal Stimuli

Two participants (ID 12 and 8) demonstrated matching with stimuli similar to those used in Gallagher and Keenan (2009). Matching trials to criterion ranged from 14 - 210 for each trained relation and 28 – 210 for combined matching training. Participant 12, who had a perfect MMSE score, met the mastery criterion for symmetry (100%) and demonstrated higher than chance levels of equivalence (73%). A stimulus set analysis revealed that participant 12 failed to form equivalence for the first (70%) and second (50%) stimulus sets. In contrast, Participant 12 was able to form equivalence with these stimuli during Study I.

Participant 8 almost met mastery criteria for demonstration of symmetry (83%) and equivalence (80%). A stimulus set analysis revealed participant 8 failed to form symmetry with the second set of stimuli (60%), but demonstrated symmetry for sets 1 (90%) and 3 (100%). Participant 8 failed to demonstrate equivalence with stimulus sets 1 (60%) and 2 (80%), but demonstrated equivalence with set 3 (100%). Participant 8 also demonstrated higher than chance levels of accuracy in respect to emergent relations (85% BA symmetry; 75% transitivity), but chance levels of equivalence demonstration (50%) during Study I. Session length for participants 12 and 8 ranged from 801 s (13 min, 21 s) – 2,618 s (43 min, 38 s).

Participants 2 and 10 failed to form matching after 394 – 671 trials and the 2,700 s (45 min) time limit. However, these two participants were able to form matching with Gallagher and Keenan (2009) stimuli using a simultaneous LS training structure and participant 2 demonstrated

CB symmetry (100%) during Study I. Overall, 75% of participants performed better using these stimuli during Study I. These data support failing to reject the null hypothesis that an MTO training structure would not relate to improved performance over equivalence formation using an LS training structure in Study I.

Arbitrary Stimuli

ID 12 was the only participant to meet mastery criteria for demonstration of symmetry (93%) and equivalence (100%). Even though participant 12 was able to demonstrate overall symmetry, the participant did not meet mastery criteria for stimulus set 2 (80%). Participant 12 did not demonstrate equivalence during Study I. Participants 2, 9, and 10 failed to demonstrate performance improvements indicating movement towards matching after 479 – 684 trials and the 2,700 s (45 min) time limit during Study II, which matches performance during Study I. Participants 5 and 8 self-terminated the session before the time limit (range 1,530 – 2,206 s) and completed 199 and 384 trials, respectively. MMSE scores ranged from 22 – 28, SLUMS scores ranged from 12 – 24, and Barthel ADL Index scores ranged from 15 – 20 for participants 2, 5, 9, and 10. Overall, 25% of participants benefitted from a 0 s delay MTO training structure using arbitrary stimuli in Study II. The remaining 75% continued to demonstrate similar performance, providing some support to reject the null hypothesis that an MTO training structure would not relate to improved performance over equivalence formation using an LS training structure in Study I.

Overall Comparison Between Study I and II

The majority of participants demonstrated higher levels of mastery on test of direct relations as well as higher than chance levels of accurate responding during tests of emergent relations during Study I when using Gallagher and Keenan (2009) stimuli, see Figure 4. Within the scope of stimuli similar to Gallagher and Keenan (2009), this data provides evidence to fail to reject the null hypothesis that an OTM or MTO training structure would not relate to improved performance over equivalence formation using an LS training structure. However, some participants demonstrated improved emergent responding when presented with arbitrary stimuli with both OTM and MTO training structures, suggesting a potential benefit of the methodological changes and inconclusive support for rejecting said null hypothesis. Overall, accuracy of responding was lower for both studies in comparison to previous research that found between 66% - 90% of participants with a near or perfect MMSE score demonstrated mastery of responding during tests of equivalence.

CHAPTER 9

DISCUSSION: STUDY II

Several methodological changes were made between Studies I and II. Stimulus control was enhanced in Study II by programming a green background and visual and auditory stimulus of “Correct!” for accurate responding and a red background and visual and auditory “Incorrect” stimulus for inaccurate responding. Previous researchers found that a delayed MTS format may maintain matching performance (Steingrimsdotti & Arntzen, 2011a) and increase equivalence formation (Saunders et al., 2005). Researchers have argued that participants may engage in precurrent behavior (e.g., overtly or covertly naming a sample stimulus) during the delay in order to increase the likelihood of a response being reinforced. Thus, a delayed MTS procedure was used during Study II. Scholars have also argued that using more than two choice stimuli can help guard against sample/S- control (Carrigan & Sidman, 1992), which may help participants to form intended stimulus classes. Thus, a 3-choice presentation was used during Study II. Lastly, previous research has also indicated higher levels of accuracy and fewer trials to criterion when using OTM and MTO training structures in comparison to LS training (Saunders et al., 2005). Thus, MTO and OTM training models were used in Study II.

Participant performance was bimodal during Study II, participants either performed at higher than chance levels during tests of equivalence (ID 8 and 12) or failed to demonstrate matching relations (ID 2, 5, 9, 10). These findings were also reflected by a strong, positive correlation between OTM and MTO composite performance scores. If participants performed well or poorly using OTM training than they were also likely to perform well or poorly using MTO training, respectively. Interestingly, participants 8 and 12 performed better on equivalence tests during Study II. Participant 8 demonstrated 70 – 80% accuracy in Study II and chance level responding during Study I (45 – 50%) equivalence testing. Participant 12 demonstrated 73 – 100% accuracy during Study II and 0 – 100% during Study I equivalence testing.

The use of a 3-choice comparison presentation may have guarded against potential sample/S- control that can occur with 2-choice comparison procedures; however, more comparison stimuli may have limited the successful formation of equivalence classes during Study II (Carrigan & Sidman, 1992). Aside from increasing comparison stimuli, Study II incorporated another 3-member stimulus class. An increase in directly trained relations may have interfered with equivalence formation due to an increased amount of discriminations that had to first be established and subsequently maintained during training.

Saunders et al. (2005) noted that increasing the amount of stimuli and number of classes may either have no effect on performance or may hinder performance, except for variable results using an MTO training structure. Previous research has not found that increasing classes and amount of stimuli promoted the formation of equivalence classes over 2 classes containing 3 – 4 members (Saunders et al., 2005). The current study replicates those findings and helps make a stronger case for reduction of number of classes when working with older adults with less than perfect performance on cognitive assessments. Many older adults that are receiving rehabilitation services are likely to score imperfectly on cognitive and functional assessments.

Individuals with various levels and forms of cognitive impairment may benefit from a decrease in amount of stimuli simultaneous presented and amount of relations trained in a particular session. Researchers using three comparison stimuli (Steingrimsdottir & Arntzen, 2014) found older adult participants formed fewer relations than researchers who used a 2-choice MTS format (Gallager & Keenan, 2009; Pérez-González & Moreno-Sierra; 1999).

Individuals with cognitive impairment may also benefit from being provided multiple matching training sessions or sessions of shorter duration. For example, Saunders et al. (2005) held two to three sessions per week. Conditioned responding is weaker for individuals with cognitive impairment (Downey-Lamb & Woodruff-Pak, 1999; Ferrante & Woodruff-Pak, 1995), which may result in a need for a higher number of trials to research criterion.

The following recommendations are made in order to alter the training method to meet the functioning level of the participants and to facilitate better equivalence performance. Training could begin with a conditional discrimination procedure for 1 3-member stimulus class. Once equivalence has formed, a second 3-member stimulus class could be introduced. After equivalence demonstration, the 2 3-member stimulus classes could appear in tandem as is typical for stimulus equivalence training procedures. This type of hierarchal and then combined introduction of stimuli also mimics training matching and testing symmetry for relations separately and then combining the relations into a single matching training block and subsequent testing block. Then, 3-member stimulus classes could continue to be systematically introduced and combined with ongoing training and testing of the original stimuli. This would somewhat parallel an additive research design, in which amount of stimulus sets in a training paradigm continue to increase, and mini-reversals, in which training is conducted with only one stimulus set periodically.

CHAPTER 10

GENERAL DISCUSSION

The purpose of the current set of studies was to replicate previous research in Study I by analyzing the association between stimulus equivalence class formation and MMSE scores for individuals with and without cognitive impairment using a linear series (LS) training structure (Gallagher & Keenan, 2009). Then, Study I extended Gallagher and Keenan (2009) by analyzing the association between stimulus equivalence performance and two other assessments, the SLUMS and Barthel ADL Index. A potential correlation between assessment scores was also tested. The secondary alternative hypothesis was that SLUMS scores would be more sensitive to the detection of cognitive impairment than MMSE scores, which would be reflected by performance during matching-to-sample (MTS) training.

The purpose of Study II was to extend previous research by examining the impact of making several methodological changes in order to create a stimulus equivalence package that incorporated recommendations from the literature and may enhance stimulus control for individuals who did not form equivalence relations during Study I. Methodological changes included additional visual and auditory differential consequences, the use of one-to-many (OTM) and many-to-one (MTO) training structures, a delayed matching-to-sample (MTS) procedure, additional 3-member stimulus set, and 3-choice comparison presentation during MTS procedures. Assessment scores were also compared to performance across OTM and MTO training structures.

The data from Studies I and II were gathered with several alternative hypotheses, First, with the alternative hypothesis that higher assessment scores would be related to the emergence of more untrained relations, including higher accuracy and fewer trials to criterion. Most participants did not respond in line with the formation of emergent relations regardless of assessment score. This means that participants with similarly high assessment scores did not perform equally across match-to-sample tests. Overall, the data support failing to reject the null hypothesis that higher assessment scores were unrelated to the emergence of more untrained relations, but provide limited evidence to support rejecting said null hypothesis. Only two participants demonstrated mastery of responding during tests of equivalence during the first study and one of the pair also responded in line with equivalence class formation during Study II. These two participants had the highest MMSE scores (29, 30) and two of the highest SLUMS scores (25, 27). However, all but one other participant (score of 22) also had high MMSE scores (27, 28), although all failed to demonstrate responding reflective of the formation of equivalence.

Participants performed better with the previous study's stimuli than arbitrary or familiar stimuli, which contributes to composite performance score and lends caution to the generalization of the findings from this study to a larger sample. The limited evidence that fails to support rejecting the null hypothesis is the significant correlation between LS composite scores and MMSE scores. This means that participants with higher levels of accurate responding on LS testing also received higher scores on the MMSE. The correlation must be interpreted with caution given the limited sample size, non-normal distribution of the data, and that composite scores were largely driven by performance during the presentation of stimuli from Gallagher and Keenan (2009).

Also, the data were collected with the alternative hypothesis that SLUMS scores would be more sensitive to the detection of cognitive impairment than MMSE scores, which would be correlated with performance during MTS training. If the alternative hypothesis were true, results from the SLUMS would indicate that more participants had cognitive impairment, which would converge with diagnoses and performance on MTS training and testing. The data support rejecting the null hypothesis that the SLUMS scores are not more sensitive to cognitive impairment than MMSE scores. In other words, the data provide support for the alternative hypothesis. Only 25% matched on classifications between the MMSE and SLUMS and some assessment scores differed by 15 points, which is the mid-way point on both assessments (range 0 – 30 for total assessment score). The participants that matched on both assessments scored in the normal range and also demonstrated mastery levels of responding on tests of equivalence. An additional five participants scored in the normal range and only one participant in the mild cognitive impairment range on the MMSE. However, four of these participants scored within the mild neurocognitive disorder (NCD) range and two within the dementia range on the SLUMS. Further, sensitivity of the SLUMS was also captured in relation to diagnoses. The SLUMS appropriately classified two participants (67%) as having dementia that had an actual diagnosis of dementia or Alzheimer's disease, while the remaining participant scored in the mild NCD range. In contrast, the MMSE was unable to detect the presence of dementia or Alzheimer's disease in each of these three participants, and instead, the participants fell in the normal (n = 2) or mild cognitive impairment (n = 1) range.

Additionally, participants would engage in more accurate responding when presented with familiar stimuli and stimuli from Gallagher and Keenan (2009) than with arbitrary stimuli due to the participant learning histories with non-arbitrary stimuli. The data support failing to reject the null hypothesis that participants would not perform better when presented with familiar stimuli, but provide support for rejecting the null hypothesis that participants would not perform better when presented with stimuli from Gallagher and Keenan (2009). In other words, participants did not perform well when presented with familiar stimuli, which contradicts the alternative hypothesis, but performed better with stimuli that were used in the previous study, which aligns with the alternative hypothesis. Three participants failed to demonstrate mastery during matching training; one participant responded with mastery levels during tests of matching; and two others demonstrated mastery on tests of symmetry during Study I when presented with familiar stimuli. However, no participants demonstrated responding in line with mastery during tests of equivalence during Study I with familiar stimuli. In fact, during Study II, familiar stimuli were not presented, because they did not serve as the optimal stimulus type for a single participant in Study I.

Participants performed much better when presented with the stimuli from Gallagher and Keenan (2009). The majority of participants demonstrated mastery of responding during tests of direct relations (75%) during Study I, and 50% during Study II across OTM and MTO training structures. More participants (38%) demonstrated mastery of emergent relations when presented with Gallagher and Keenan (2009) stimuli using a LS training structure during Study I than with any other stimulus type across both studies (Range 0 – 33%). Emergent relations included responding at mastery levels during tests of symmetry, transitivity, or equivalence.

Lastly, data were collected with the null hypothesis that participants who did not respond in accordance with the formation of equivalence would not benefit from the methodological changes in Study II. This was true for half of the participants, but two other participants demonstrated higher than chance levels of accurate responding during tests of emergent relations during Study II. Generally, the findings fail to support rejecting the null hypothesis that methodological changes would not improve performance during Study II. Two participants (50%) failed to form matching when presented with optimal stimuli using a OTM and MTO training structure, four participants (67%) failed to demonstrate said responding when using arbitrary stimuli and a OTM training structure, and five participants failed to demonstrate said responding using arbitrary stimuli and a MTO training structure (83%) during Study II. However, two participants demonstrated higher than chance levels of responding during Study II, therefore, a more systematic approach to methodological changes should be examined in future research.

In general, the present project failed to replicate the findings of Gallagher and Keenan (2009), including the association between added assessments and stimulus equivalence performance, and when procedural modifications were included to explore the outcomes of different training types of MTS, OTM and MTO training structures were correlated with each other, but did not improve overall performance.

The following section discusses the results of the study and implications related to general study components. Study components are discussed in the following order: (1) General assessment score discussion, (2) Other considerations, (3) Limitations, and (4) Future research directions.

General Assessment Score Discussion

Assessment scores and equivalence formation performance in the present study both converge and diverge from results of previous research. Stewart and colleagues (2012) found a strong, positive correlation between scores on the MMSE and SLUMS. Classification scores on MMSE and SLUMS did not converge in the present study. This finding is supported by previous research that noted that the SLUMS was more sensitive to cognitive impairment, as measured by higher cut-off scores and more overlap between categorization and a clinician's diagnosis than the MMSE (Stewart et al., 2012; Tariq et al., 2006). Participants in the current study were likely to score in the cognitive impairment range on the SLUMS, but within the normal range on the

MMSE, which aligns with previous research (Stewart et al., 2012). Lastly, previous research called for an examination of the potential association between SLUMS and ADL performance (Stewart et al., 2012). Scores on the SLUMS as well as the MMSE were not correlated with Barthel ADL Index scores in the present study. An elaboration of assessment scores and equivalence performance is located below.

MMSE

The vast majority (88%) of participants who engaged in at least one training session during Study I scored in the normal functioning range on the MMSE (Range: 27 – 30). Only one participant scored in the low cognitive impairment range (MMSE = 22), which may be indicative of Minor Neurocognitive Impairment. Higher MMSE scores resulted in a non-normal score distribution with a negative skew. MMSE scores in the current study were similar to scores in Gallagher and Keenan (2009). However, current findings contradict with previous research.

Gallagher and Keenan (2009) found that participants with a 27 or higher on the MMSE were able to demonstrate equivalence. In the current study, only participants with a 29 or 30, near-perfect and perfect assessment scores, were able to form equivalence when using the same stimuli as Gallagher and Keenan (2009). More so, the current study was unable to replicate the formation of equivalence when using different stimuli (familiar and arbitrary stimulus sets).

However, upon more in-depth analyses using participant composite performance scores across all three stimulus types during LS training, a significant, positive correlation was found between equivalence performance and MMSE scores. This association indicates that equivalence performance using LS training and 2 3-member stimulus classes increases along with MMSE scores. For instance, a participant who met mastery equivalence criteria would have a higher MMSE score than a participant who was only able to demonstrate matching, symmetry, or transitivity. However, MMSE scores were not significantly correlated with equivalence performance during OTM or MTO training during Study II.

SLUMS

All participants, except one, scored lower on the SLUMS than the MMSE resulting in only 25% of participants falling into the same category on both assessments. One participant that was categorized as having normal cognitive functioning demonstrated equivalence with the Gallagher and Keenan (2009) stimuli while the other participant demonstrated matching and then discontinued participation in the study to engage in other activities. Categorization differed for the majority of the participants (75%), which likely contributed to the lack of a significant correlation between scores on the SLUMS and the MMSE.

Lower scores on the SLUMS could be indicative of increased sensitivity for the detection of cognitive impairment or as false positive classifications of impairment. The earlier statement is supported in the present study given that participants who scored a 24 or less on the SLUMS did not formed equivalence. One participant that scored in the Mild Neurocognitive Disorder range on the SLUMS (25) and normal range on the MMSE (30) demonstrated equivalence for Gallagher and Keenan (2009) stimuli during Study I and for arbitrary stimuli within the MTO training structure during Study II. This participant also demonstrated higher than chance levels of equivalence across stimuli within the OTM training structure and the Gallagher and Keenan (2009) stimuli in the MTO training structure. Another participant, with the highest SLUMS score (27), formed equivalence using Gallagher and Keenan's (2009) stimuli. Unfortunately, that participant, along with another with a high SLUMS score (28) dropped out of the study during Study I, making it difficult to assess high SLUMS scores and equivalence formation.

The participant (ID 8) with the next best equivalence performance during Study II scored a 23 on the SLUMS, indicating Mild Neurocognitive Disorder. Participant 8 scored a 28 on the MMSE. Unfortunately, two other participants with similar SLUMS (23, 24) and MMSE (27, 28) scores failed to form matching relations. Therefore, a possible connection between higher equivalence performance and higher SLUMS scores is limited due to attrition and mixed performance results.

Barthel ADL Index

Based on visual analysis, functional impairment was not related to formation of equivalence or the demonstration of more emergent relations. The Barthel ADL Index ranges from 0 – 20 with higher scores indicating increased independence for completion of ADLs. During Study I, the top 50% of performers scored between 11 and 20. Two individuals that demonstrated equivalence during Study I had scores of 14 and 17, whereas some individuals who were unable to form matching relations had very high scores of 18 and 20. The top two performers during Study II scored an 11 and 17 on the Barthel ADL Index. These findings indicate that ability to complete ADLs does not correlate with equivalence class formation for the current sample. This finding was also reflected by the lack of significant non-parametric correlations between Barthel ADL Index scores and composite LS, OTM, and MTO performance scores.

Other Considerations

Verbal Behavior

Two participants (ID 8 and 12) with higher levels of accuracy than other participants demonstrated rule-governed behavior throughout Studies I and II. Participants 8 and 12 self-reported covert verbal behavior and also overtly engaged in verbal behavior during training and testing sessions. These two participants created descriptions and labels for stimuli and sometimes tacted and then echoed these descriptions through training and testing. For arbitrary and familiar stimuli in Study I, participant 12 continuously correctly repeated her self-identified labels aloud during matching and symmetry, but switched the relations during transitivity and equivalence. Upon this switch, the participant scored between 0 – 15% accurate on both tests. Thus, participant 12's responding appeared to be under the control of rules, instead of contingencies.

In Study II, participant 12 identified the stimuli by a characteristic within the 2D line drawings. Arbitrary stimuli became "girl", "envelope", and "W". Participant 8 reported a connection between a Gallagher and Keenan (2009) stimulus and a syllable in the surname of one of her doctors, thus reporting that it was "easy to remember" relations for that particular stimulus. Participant 8 was only able to demonstrate equivalence with the set of stimuli containing the familiar stimulus (set 2, OTM training structure). Gallagher and Keenan (2009) did not report participant verbal behavior during or after training sessions. Equivalence formation in the previous study may have been rule-governed as opposed to contingency-shaped. Future research could examine the use of talk aloud procedures, analysis of overt behavior during a session, as well as self-report of how participants selected a comparison stimulus for individuals that are still able to engage in a sufficient level of verbal behavior. For instance, a participant may name, label, or otherwise describe stimuli when making a selection and be asked to overtly express their covert behavior.

Limitations

Medications

Interestingly, none of the participants in Gallagher and Keenan (2009) were on memory medications, and only one of the current participants was prescribed a memory medication.

Participant 2 was prescribed Rivastigmine, a cholinesterase inhibitor (MMSE 22, SLUMS 14). This participant was unable to form matching across all training structures and stimulus types except for AB matching, BC matching, and CB symmetry for Gallagher and Keenan (2009) stimuli in Study I. Other medications may have interfered with performance in the current study, as participants were prescribed medications to treat Parkinson's disease, anxiety, and depression.

Order Effects

Participant performance may have been impacted by the order of exposure to stimuli and training structure. For example, if participants form equivalence with a with a LS training structure during Study I, they may be more likely to also form equivalence with OTM or MTO training during Study II. In the present sample, participant 12 failed to demonstrate equivalence across the first two LS training sessions utilizing familiar and arbitrary stimuli. Then, participant 12 was able to demonstrate equivalence with stimuli from Gallagher and Keenan (2009) during day 3. Participant 12 continued to perform well across training structure (OTM, MTO) and stimulus type (Gallagher and Keenan, arbitrary) during Study II. Participant 8 failed to form matching on the first training day of Study II with arbitrary stimuli and an MTO training structure. However, participant 8 demonstrated higher than chance levels of accuracy on equivalence testing for the subsequent training structure and stimulus combinations.

Given that most participants were unable to form equivalence during Study II, it is difficult to separate success due to order effects from success due to a specific training structure or stimulus type in the present study. However, this limitation would be beneficial in an applied setting, where once performance came under the control of a certain form of learning, performance could generalize to similar training and testing situations.

Generalization

Generalization of correlational analyses is limited due to the small sample size and skewed data distributions. Future researchers should examine correlations between assessments as well as correlations between assessment scores and equivalence performance with a larger sample size. Likewise, the correlations between MMSE scores and composite LS performance as well as between OTM and MTO composite performance scores should be interpreted with caution given the limited amount of equivalence formation. The OTM and MTO association is difficult to interpret because only two participants were able to form relations higher than AB matching. However, overall, participants tended to score higher on the MMSE and performed better during study one. Similarly, participants either performed very well or very poorly on both OTM and MTO training.

The current study was not designed to address potential medication interference or order effects and future research should address such considerations. Limited success, in relation to the formation of equivalence, for both studies as well as multiple methodological changes between Study I and Study II necessitate future research that systematically introduces or eliminates elements to better examine how to promote stimulus equivalence class formation in older adult participants.

Future Research

Equivalence formation was limited, although some participants scored near the mastery criterion. These participants may have demonstrated equivalence if provided with more training. For example, Study I could have included combined AB and BC matching as was completed after a methodological change during Study II. Combined matching training may have resulted in more accuracy during equivalence testing. Also, participants could have been exposed to matching training if unable to form symmetry or equivalence during Study I and Study II. This would be important when viewing stimulus equivalence as a teaching methodology, as participants are provided with feedback and multiple exposures to the material when mastery criterion is unmet. However, this should be weighed against the use of stimulus equivalence as a diagnostic tool. There may be validity in not re-exposing participants to matching training, or perhaps is re-exposure and a continued lack of demonstration of emergent relations that better indicates the presence of cognitive impairment.

A hierarchical approach to the presentation of stimulus classes was previously discussed. Presentation of classes singly before exposing participants to numerous stimuli in tandem may prove beneficial for older adults with cognitive impairment. Also previously discussed, breaking training sessions into multiple mini-sessions to guard against participant fatigue and maximize cognitive resources over a short period of time may help individuals with impairment. Performance for older adults in previous research was added by two – three training sessions per week (Saunders et al., 2005).

Another challenge that can be of focus in future research is the incorporation of reinforcer assessments to identify functional differential consequences for accurate and inaccurate responding. Additional programmed stimuli from Study I to Study II were not strong enough to evoke equivalence formation. The auditory and visual stimuli, “Correct!” and “Incorrect” should function as generalized conditioned reinforcer/punishers (Skinner, 1957). However, generalized conditioned reinforcers and punishers may not be salient enough to shape operant behavior within stimulus equivalence training paradigms. Future research may consider the incorporation of idiographic reinforcer or preference assessment to provide salient, differential consequences.

An area with a current paucity of literature is the examination of equivalence formation maintenance effects for both arbitrary and non-arbitrary stimuli with older adult participants.

Numerous studies with older adults with cognitive impairment use non-arbitrary stimuli (Bodi et al., 2009; Plaza et al., 2012; Steingrimsdottir & Arntzen, 2011a&b), however, results from the current study suggest familiar stimuli that are arbitrarily related may be contraindicated for equivalence class formation given that some participants reported searching for connections between the stimuli based on learning history. Familiar stimuli may be appropriate when socially significant to the participants, such as the case with names and faces of family members, friends, and caregivers.

As Sidman (2013) called behavior analysts to action, applied behavior analysis has the ability to touch the lives of many individuals with memory loss. More research needs to be conducted utilizing stimulus equivalence, the identification of idiosyncratic and strong learning channels, and socially significant stimuli. Therefore, examining maintenance effects after stimulus equivalence class formation may lead to use of stimulus equivalence procedures as a teaching technology for social significant information for older adults.

Table 1. All Participants by Order of Recruitment and Demographic Data

ID	**Sex**	**Age**	**Visual Acuity**	**Diagnoses and Conditions**	**Assessment Scores**			**Participation**
					MMSE	SLUMS	Barthel ADL	
1	F	88	*	Early Alzheimer's disease, Hypertension, Incontinence; Neuropathy, Depression	27	28	18	Partial Study I
2	F	83	20/50 with glasses	Dementia, Borderline Diabetic, Low blood pressur, Bladder Issues, COPD	22	14	18	Studies I and II
3	F	78	---	Muscle weakness, Neuropathy NOS, Renal Disease, Hypertension, Type II Diabetes Mellitus, Depressive Disorder	29	27	14	Partial Study I
4	F	77	---	Type II diabetes, Hypertension, Neuropathy, COPD, Cataracts	---	26	15	DROPPED
5	F	73	20/60 with glasses	Alzheimer's disease, Type II Diabetes Mellitus, Glasses, HOH, Hypertension, Renal Failure, Incontinence, COPD, Anxiety/Bipolar Disorder	27	12	15	Studies I and II
6	M	72	20/60 with glasses	Diabetes Mellitus, Hypertension, Incontinence, History of cataract in right eye	19	16	18	EXCLUDED
7	F	94	---	---	11	---	---	EXCLUDED
8	F	77	20/60 with glasses	Type II Diabetes Mellitus, Blind in right eye (glasses), HOH, Hypertension, Bladder Issues, Neuropathy, Depression, Anxiety	28	23	11	Studies I and II
9	F	81	20/40 with glasses	Dementia, Glasses, High blood pressure, Bladder Issues, Tremors,Depression, Anxiety	28	23	20	Studies I and II
10	M	78	20/60 with glasses	Glasses; Pacemaker; High blood pressure; Minimial hearing	27	24	17	Studies I and II
11	M		---	---	29	---	17	DROPPED
12	F	68	20/40 wo/ correction	MS, Diabetes Mellitus, High blood pressure, Thyroid disorder	30	25	17	Studies I and II

----Unable to gather data due to study exclusion, attrition, or *schedule conflicts before attrition

Table 2. Study I Results: Assessment Scores & Equivalence Responding

ID	Assessment Scores			Order	Equivalence Responding Using LS Training		
	MMSE	**SLUMS**	**Barthel ADL**		**G&K (2009) (Gk)**	**Familiar (F)**	**Arbitrary (Ar)**
12	30	25	17	FArGk	100%	AB & BC Matching	BA & CB Symmetry
3	29	27	14	GkAr	100%	DROPPED	Failed to Form AB Matching
9	28	23	20	FGkAr	Failed to Form AB Matching	CB Symmetry	Transitivity
8	28	23	11	ArFGk	AB & BC Matching	CB Symmetry	Failed to Form AB Matching
2	22	14	18	GkFAr	CB Symmetry	Failed to Form AB Matching	Failed to Form AB Matching
10	27	24	17	ArGkF	AB & BC Matching	Failed to Form AB Matching	Failed to Form AB Matching
1	27	28	18	Gk	AB Matching	DROPPED	DROPPED
5	27	12	15	FGkAr	Failed to Form AB Matching	Failed to Form AB Matching	Failed to Form AB Matching

Table 3. Study I Result: Trials to Criterion, Accuracy, and Session Length

ID	Trials to Criterion		Percentage Accuracy				Session Length
	G&K (2009) Outcomes for Linear Series Training						
	AB Matching	BC Matching	BA Symmetry	CB Symmetry	AC Transitivity	CA Equivalence	(mm:ss)
12	10	2	95%	100%	100%	100%	7:32
3	20	9	90%	90%	95%	100%	7:27
9	NA (847)	---	---	---	---	---	MAX
8	49	70	85%	25%	75%	50%	14:10
2	339	200	35%	100%	65%	55%	49:42
10	300	30	60%	40%	55%	50%	31:08
1*	640	---	55%	---	---	---	37:02
5	N/A (767)	---	---	---	---	---	MAX
	Familiar Stimuli Outcomes for Linear Series Training						
ID	**Trials to Criterion**		**Percentage Accuracy**				**Session Length**
	AB Matching	BC Matching	BA Symmetry	CB Symmetry	AC Transitivity	CA Equivalence	**(mm:ss)**
12	80	10	85%	85%	0%	0%	11:22
9	29	50	45%	100%	35%	50%	36:03
8	30	9	50%	90%	45%	45%	9:54
2	N/A (646)	---	---	---	---	---	MAX
10	N/A (676)	---	---	---	---	---	MAX
5	N/A (653)	---	---	---	---	---	MAX
	Arbitrary Stimuli Outcomes for Linear Series Training						
ID	**Trials to Criterion**		**Percentage Accuracy**				**Session Length**
	AB Matching	BC Matching	BA Symmetry	CB Symmetry	AC Transitivity	CA Equivalence	**(mm:ss)**
12	19	10	100%	85%	15%	0%	9:29
3*	N/A (172)	---	---	---	---	---	10:59
9	29	32	20%	10%	90%	65%	37:00
8*	N/A (642)	---	---	---	---	---	37:07
2	N/A (610)	---	---	---	---	---	MAX
10	N/A (353)	---	---	---	---	---	MAX
5	N/A (604)	---	---	---	---	---	MAX

*Participant self-initiated session termination before exceeding the time limit; N/A denotes trials to criterion not applicable; (###) signifies number of trials completed before exceeding time limit; MAX indicates time limit elapsed

Table 4. Study II: Participant Order of Exposure by Training Modality & Stimulus Type

ID	Order
12	MTO-Ar; OTM-Gk; MTO-Gk; OTM-Ar
8	MTO-Ar; OTM-Ar; OTM-Gk; MTO-Gk
9	OTM-Ar; MTO-Ar
10	MTO Gk; OTM-Ar; MTO-Ar; OTM-Gk
5	OTM-Ar; MTO-Ar
2	OTM-Gk; MTO-Ar; OTM-Ar; MTO-Gk

Table 5. Study II Results: Assessment Scores & Equivalence Responding

ID	Assessment Scores			Equivalence Responding			
				OTM		MTO	
	MMSE	SLUMS	Barthel ADL	Arbitrary (A)	Optimal (Gk)	Arbitrary (A)	Optimal (Gk)
12	30	25	17	83%	87%	100%	73%
8	28	23	11	73%	70%	Failed to Form BA Matching	80%
9	28	23	20	Failed to Form AB Matching	---	Failed to Form BA Matching	---
10	27	24	17	Failed to Form AB Matching	Failed to Form AB Matching	Failed to Form BA Matching	Failed to Form BA Matching
5	27	12	15	Failed to Form AB Matching	---	Failed to Form BA Matching	---
2	22	14	18	Failed to Form AB Matching	Failed to Form AB Matching	Failed to Form BA Matching	Failed to Form BA Matching

Table 6. Study II Results: Trials to Criterion, Accuracy, and Session Length

ID	Trials to Criterion			Percentage Accuracy								Session Length (mm:ss)
Gallagher and Keenan (2009) Stimuli Outcome for OTM Training												
	Matching			BA/CA Symmetry				BC/CB Equivalence				
	AB	AC	AB/AC	Set 1	Set 2	Set 3	All	Set 1	Set 2	Set 3	All	
12	89	60	58	100%	60%	100%	87%	100%	70%	90%	87%	30:51
8	30	60	298	50%	100%	100%	83%	50%	90%	70%	70%	36:37
2	N/A (502)	---	---	---	---	---	---	---	---	---	---	MAX
10*	N/A (523)	---	---	---	---	---	---	---	---	---	---	41:59
Gallagher and Keenan (2009) Stimuli Outcome for MTO Training												
ID	Trials to Criterion			Percentage Accuracy								Session Length (mm:ss)
	Matching			AB/AC Symmetry				BC/CB Equivalence				
	BA	CA	BA/CA	Set 1	Set 2	Set 3	All	Set 1	Set 2	Set 3	All	
12	14	30	28	100%	100%	100%	100%	70%	50%	100%	73%	13:21
8	45	210	210	90%	60%	100%	83%	60%	80%	100%	80%	43:38
2	N/A (671)	---	---	---	---	---	---	---	---	---	---	MAX
10	N/A (394)	---	---	---	---	---	---	---	---	---	---	MAX
Arbitrary Stimuli Outcomes for MTO Training												
ID	Trials to Criterion			Percentage Accuracy								Session Length (mm:ss)
	Matching			AB/AC Symmetry				BC/CB Equivalence				
	BA	CA	BA/CA	Set 1	Set 2	Set 3	All	Set 1	Set 2	Set 3	All	
12	15	89	87	100%	80%	100%	93%	100%	100%	100%	100%	36:46
8*	N/A (384)	---	---	---	---	---	---	---	---	---	---	25:30
9	N/A (684)	---	---	---	---	---	---	---	---	---	---	MAX
2	N/A (479)	---	---	---	---	---	---	---	---	---	---	MAX
10	N/A (547)	---	---	---	---	---	---	---	---	---	---	MAX
5*	N/A (199)	---	---	---	---	---	---	---	---	---	---	38:34
Arbitrary Stimuli Outcomes for OTM Training												
ID	Trials to Criterion			Percentage Accuracy								Session Length (mm:ss)
	Matching			BA/CA Symmetry				BC/CB Equivalence				
	AB	AC	AB/AC	Set 1	Set 2	Set 3	All	Set 1	Set 2	Set 3	All	
12	45	45	90	100%	100%	50%	83%	70%	100%	80%	83%	24:33
8	390	60	28	70%	100%	100%	90%	40%	80%	100%	73%	54:38
9	N/A (704)	---	---	---	---	---	---	---	---	---	---	MAX
2*	N/A (450)	---	---	---	---	---	---	---	---	---	---	38:09
10	N/A (547)	---	---	---	---	---	---	---	---	---	---	MAX
5	N/A (304)	---	---	---	---	---	---	---	---	---	---	MAX

*Participant self-initiated session termination before exceeding the time limit; N/A denotes trials to criterion not applicable; (###) signifies number of trials completed before exceeding time limit; MAX indicates time limit elapsed

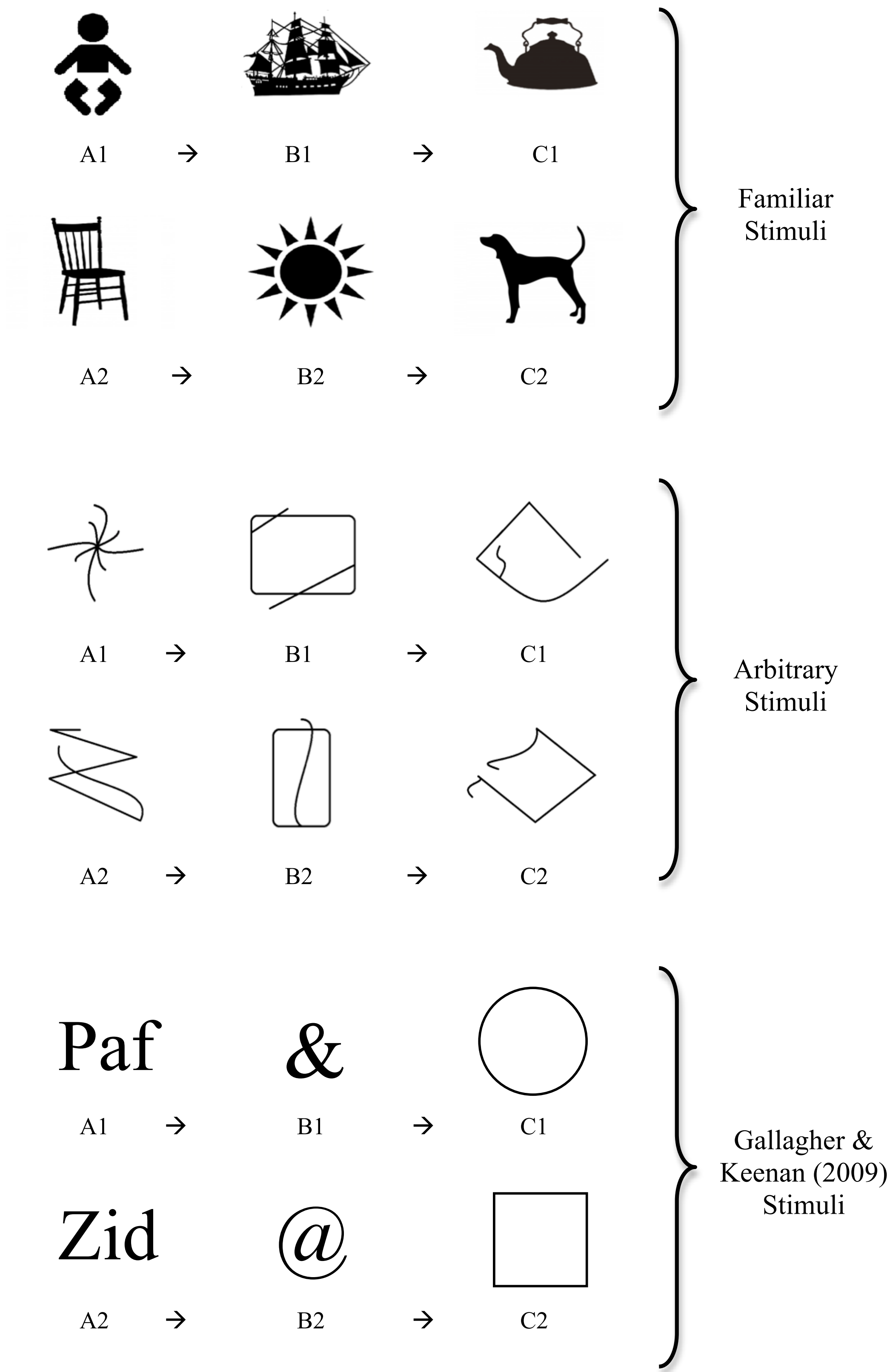


Figure 1. Study I Stimuli

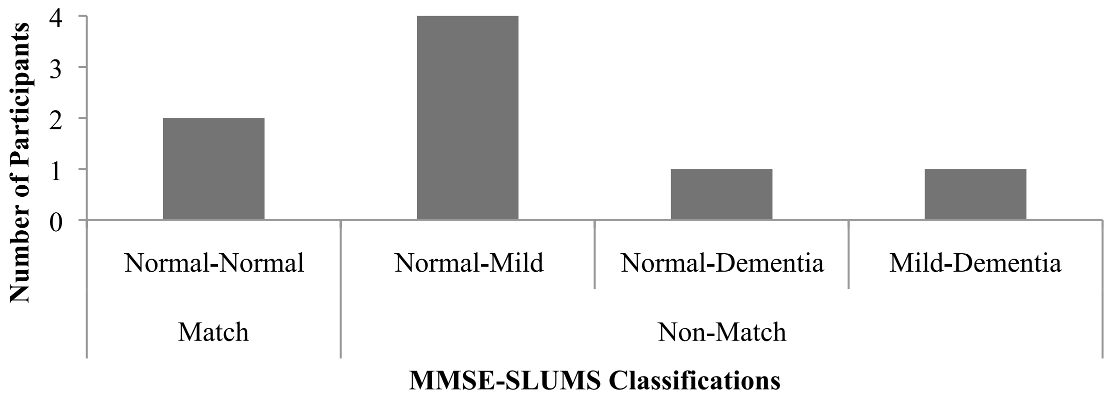


Figure 2. Classification Analysis

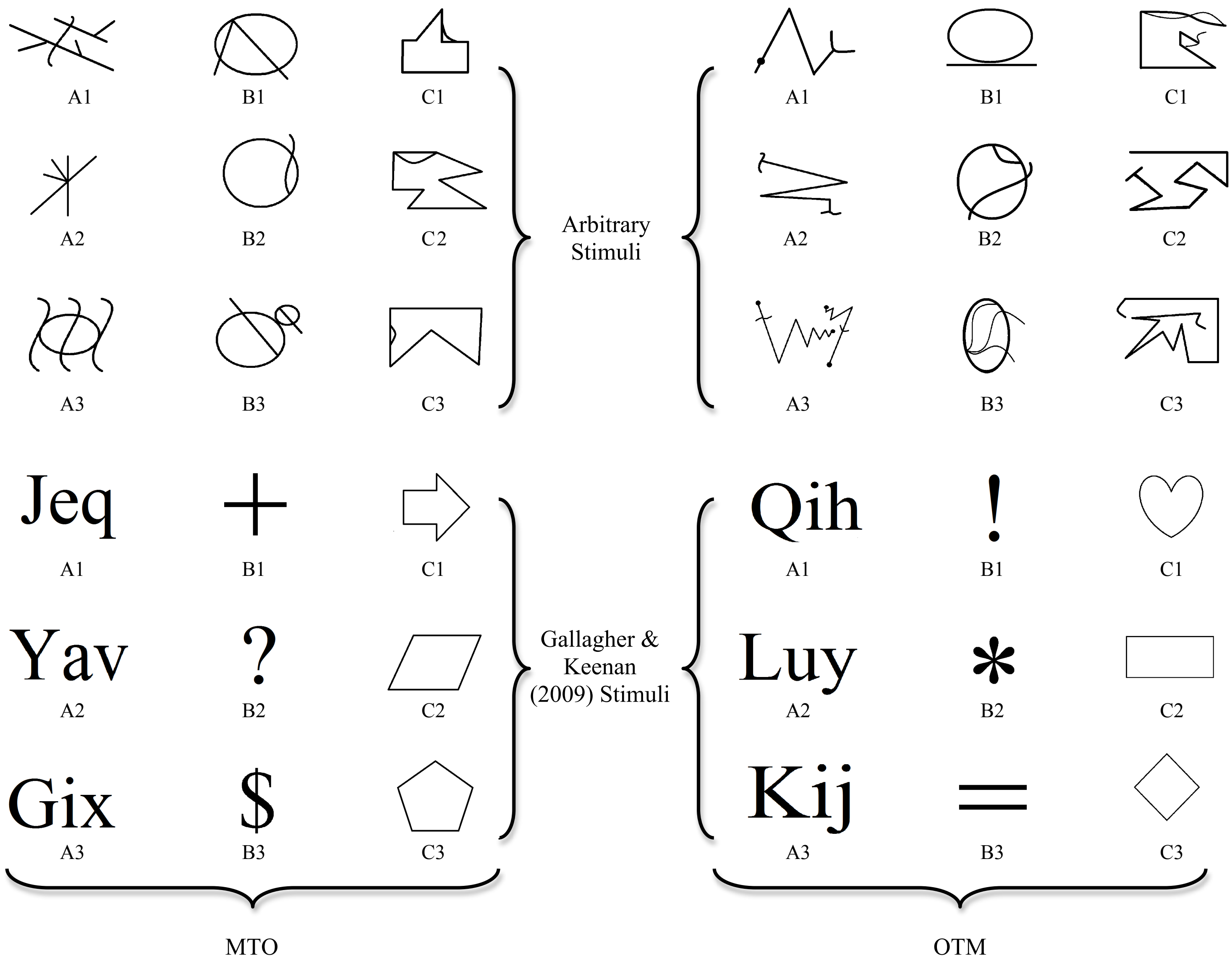


Figure 3. Study II Stimuli

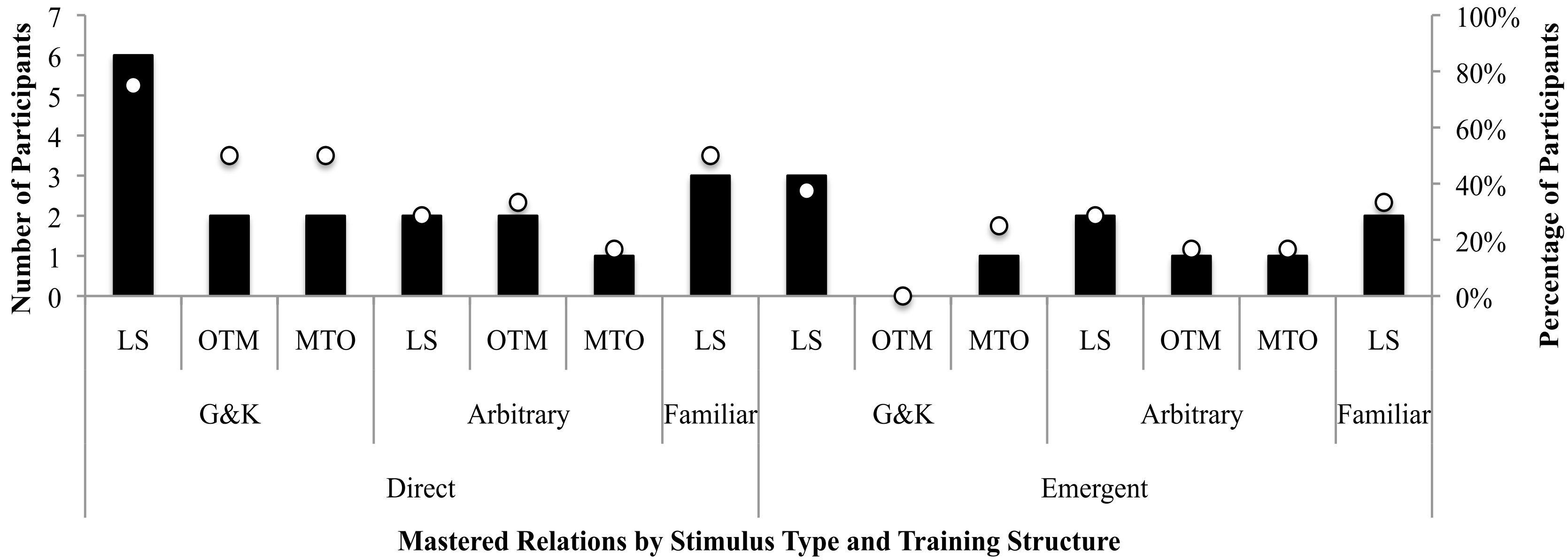


Figure 4. Mastered Relations by Stimulus Type and Training Structure

REFERENCES

Aggio, N. M., and Domeniconi, C. (2012). Formation and maintenance of equivalence classes: A study with elderly participants. *Acta Comportamentalia, 20*, 29 – 43. Retrieved from: www.researchgate.net

Alzheimer's Association (2014). Alzheimer's disease facts and figures. *Alzheimer's and Dementia, 10*, 1 – 80. Retrieved from: http://www.alz.org/downloads/facts_figures_2014.pdf

American Psychological Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.).* Washington, DC: Author.

Arntzen, E. (2006). Delayed matching to sample: Probability of responding in accord with equivalence as a function of different delays. *The Psychological Record, 56*, 135 – 167. Retrieved from: www.ebscohost.com

Arntzen, E., & Hansen, S. (2011). Training structures and the formation of equivalence classes. *European Journal of Behavior Analysis, 12*, 483 – 503. Retrieved from: http://www.ejoba.org/

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47*, 309-320. Retrieved from: www.ebscohost.com

Arntzen, E., & Holth, P. (2000). Equivalence outcome in single subjects as a function of training structure. *The Psychological Record, 50*, 603-628. Retrieved from: www.ebscohost.com

Arntzen, E., & Steingrimsdottir, H. S. (2014). On the use of variations in a delayed matching-to-sample procedure in a patient with neurocognitive disorder. In M. H. Swahn, J. B. Palmier, & S. M. Braunstein (Eds.), *Mental disorder* (pp. 123 – 138). iConcept Press. Retrieved from: https://www.iconceptpress.com/books/mental-disorder/01-0042-001/

Arntzen, E., Steingrimsdottir, H. & Antonsen, A. B. (2013). Behavioral studies of dementia: Effects of different types of matching-to-sample procedures. *Norwegian Journal of Behavior Analysis, 40*, 17 – 29.

Bódi, N., Csibri, E., Myers, C. E., Gluck, M. A. & Kéri, S. (2009). Associative learning, acquired equivalence, and flexible generalization of knowledge in mild Alzheimer disease. *Cognitive and Behavioral Neurology, 22*, 89 – 94. Retrieved from: http://memory.rutgers.edu/pdf/Bodi_AD_FISH_09.pdf

Carrigan, P. F., & Sidman, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 183 – 204. doi: 10.1901/jeab.1992.58-183

Collin, C., Wade, D., Davies, S., & Horne, V. (1988). The Barthel ADL Index: A reliability study. *International Disability Studies, 10*, 61 – 63. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/

Donahoe, J. W., & Palmer, D. C. (2004). *Learning and complex behavior*. Richmond, MA: Ledgetop Publishing.

Downey-Lamb, M. M., & Woodruff-Pak, D. S. (1999). Early detection of cognitive deficits using eyeblink classical conditioning. *Alzheimer's Reports, 2*, 37 – 44. Retrieved from: https://illiad.lib.siu.edu

Eilifsen, C., & Arntzen, E. (2015). Effects of training structure and the passage of time on trained and derived performance. *The Psychological Record, 65*, 1 – 12. doi: 10.1007/s40732-014-0067-2

Engel, C. M. (2005). Aging and categorization: *Using generalized equivalence classes and their characteristics to compare older and younger adults*. (Unpublished master's thesis). The University of North Carolina Wilmington, Wilmington, North Carolina.

Ferrante, L. S., & Woodruff-Pak, D. S. (1995). Longitudinal investigation of eyeblink classical conditioning in the old-old. Journal of Gerontology: *Psychological Science, 50B*, 42 – 50. doi: 10.1093/geronb/50B.1.P42

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53*, 345 – 358. doi: 10.1901/jeab.1990.53-345

Fienup, D. M., & Dixon, M. R. (2006). Acquisition and maintenance of visual-visual and visual-olfactory equivalence classes. *European Journal of Behavior Analysis, 7*, 87 – 98. Retrieved from: http://www.ejoba.org/

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 221 - 231. Retrieved from: www.ebscohost.com

Gallagher, S. M., & Keenan, M. (2009). Stimulus equivalence and the Mini Mental Status Examination in the elderly. *European Journal of Behavior Analysis, 10*, 159 – 165. http://www.ejoba.org/

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K. & …Winblad, K. (2006). Mild cognitive impairment. *Lancet, 367*, 1262-1270. doi: 10.1016/S0140-6736(06)68542-5

Greenwood, P. M., & Parasuraman, R. (2012). Nurture via nature: Genetics, environment, and Cognition. In *Nurturing the older brain and mind* (pp. 187 – 208). Cambridge, MA: The MIT Press.

Hall, G. A. & Chase, P. N. (1991). The relationship beween stimulus equivalence and verbal behavior. *The Analysis of Verbal Behavior, 9*, 107 – 119. Retrieved from: http://www.ncbi.nlm.nih.gov.proxy.lib.siu.edu/pmc/journals/

Hartigan, I. (2007). Comparative review of the Katz ADL and the Barthel Index in assessing the activities of daily living of older people. *International Journal of Older People Nursing, 2*, 204 – 212. Retrieved from: www.ebscohost.com

Jerome, E. A. (1959). Age and learning: Experimental studies. In J. E. Birren (Ed.), *Handbook of aging and the individual* (pp. 655-699). Chicago: University of Chicago Press.

Mahoney, F., & Barthel, D. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal, 12*, 61 – 65. doi: 10.1037/t02366-000

Moody, H. R., & Sasser, J. R. (2015). *Aging: Concepts and controversies (8th ed.)*. Thousand Oaks, CA: SAGE Publications, Inc.

Nunley, K. A., Rosano, C., Ryan, C. M., Jennings, J. R., Aizenstein, H. J., Zgibor, J. C., &…Saxton, J. A. (2015, September). Clinically relevant cognitive impairment in middle-aged adults with childhood onset type 1 diabetes. *Diabetes Care, 38*, 1768 – 1776. doi: 10.2337/dc15-0041

Ortman, J. M., Velkoff, V. A., & Hogan, H. (2014, May). An aging nation: The older population in the United States. *Current Population Reports*. United States Census Bureau.

Palmer, D. C. (1991). A Behavioral Interpretation of Memory (p. 261 – 279). In L. J. Hayes & P. N. Chase. *Dialogues on verbal behavior: The first international institute on verbal relations.* Reno, NV: Context Press

Pérez-González, L. A., & Moreno-Sierra, V. (1999). Equivalence class formation in elderly persons. *Psicothema, 11*, 325 – 336. Retrieved from: www.ebscohost.com

Plaza, V., López-Crespo, G., Antúnez, C., Fuentes, L. J., & Estévez, A. F. (2012). Improving delayed face recognition in Alzheimer's disease by differential outcomes. *Neuropsychology, 26*, 483 – 489. doi: 10.1037/a0028485

Potkin, S. G. (2002). The ABC of Alzheimer's disease: ADL and improving day-to-day functioning of patients. International *Psychogeriatrics, 14*, 7 – 26. Retrieved from: www.ebsohost.com

Rehfeldt, R. A., & Hayes, L. J. (2000). The long-term retention of generalized equivalence classes. *The Psychological Record, 50*, 405 – 428. Retrieved from: www.ebscohost.com

Rehfeldt, R. A. & Root, S. (2004). The generalization and retention of equivalence relations in adults with mental retardation. *The Psychological Record, 54*, 173 – 186. Retrieved from: www.ebscohost.com

Roberts, R. O., Knopman, D. S., Pryzbelski, S. A., Mielke, M. M., Kantarci, K., Preboske, G. M., &…Jack, C. R. (2014). Association of type 2 diabetes with brain atrophy and cognitive impairment. *Neurology, 82*, 1132 – 1141. doi: http://dx.doi.org.proxy.lib.siu.edu/10.1212/WNL.0000000000000269

Saunders, A. M., Strittmatter, W. J., Schmechel, D., St. George-Hyslop, P. H., Pericak-Vance, M. A., Joo, S. H., & …Alberts, M. J. (1993). Association of apolipoprotein E allele e4 with late-onset familial and sporadic Alzheimer's disease. *Neurology, 43*, 1467 – 1472. Retrieved from: www.ebscohost.com

Saunders, R. R., Chaney, L., & Marquis, J. G. (2005). Equivalence class establishment with two-, three-, and four-choice matching to sample by senior citizens. *The Psychological Record, 55*, 539 – 559. Retrieved from: www.ebscohost.com

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14*, 5 – 13.

Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.

Sidman, M. (2013). Techniques for describing and measuring behavioral changes in Alzheimer's patients. *European Journal of Behavior Analysis, 14*, 141 – 149. Retrieved from: http://www.ejoba.org/

Sidman, M. & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*, 5 – 22. Retrieved from: http://www.ncbi.nlm.nih.gov.proxy.lib.siu.edu/pmc/articles/

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. Action, MA: Copley Publishing Group.

Skinner, B. F. (1953). *Science and human behavior*. NY: The Free Press.

Skinner, B. F. (1957). *Verbal Behavior*. NY: Appleton-Century-Crofts.

Steingrimsdottir, H. S., & Arntzen, E. (2011a). Identity matching in a patient with Alzheimer's disease. *American Journal of Alzheimer's Disease & Other Dementias, 26*, 247 – 253. doi: 10.1177/1533317511402816

Steingrimsdottir, H. S., & Arntzen, E. (2011b). Using conditional discrimination procedures to study remembering in an Alzheimer's patient. *Behavioral Interventions, 26*, 179 – 192. doi: 10.1002/bin.334

Steingrimsdottir, H. S., & Arntzen, E. (2014). Performance by older adults on identity and arbitrary matching-to-sample tasks. *The Psychological Record, 64*, 827 – 839. doi: 10.1007/s40732-014-0053-8

Stewart, S., O'Riley, A., Edelstein, B., & Gould, C. (2012). Preliminary comparison of three cognitive screening instruments in long term care: The MMSE, SLUMS, and MoCA. *The Gerontologist, 35*, 57 – 75. doi: 10.1080/07317115.2011.626515

Tariq, S. H., Tumosa, N., Chibnall, J. T., Perry III, H. M., & Morley, J. E. (2006). The Saint Louis University Mental Status (SLUMS) Examination for detecting mild cognitive impairment and dementia is more sensitive than the Mini-Mental Status Examination (MMSE) - A pilot study. *American Journal of Geriatric Psychology, 14*, 900-910. Retrieved from: www.ebscohost.com

Thompson, R. F., & Woodruff-Pak, D. S. (1987). A model system approach to age and the neuronal bases of learning and memory. In M. W. Riley, J. D., Matarazzo, & A. Baum (Eds.), *The aging dimension* (pp. 49 – 76). Hillsdale, NJ: Erlbaum.

Tombaugh, T. N., McDowell, I., Kristjansson, B., & Hubley, A. M. (1996). Mini-mental state examination (MMSE) and the modified MMSE (3MS): A psychometric comparison and normative data. *Psychological Assessment, 8*, 48 – 59. doi: l040-3S90/9ti/S3.00

Toussaint, K. A., & Tiger, J. H. (2010). Teaching early Braille literacy skills within a stimulus equivalence paradigm to children with degenerative visual impairments. *Journal of Applied Behavior Analysis, 43,* 181 – 194. doi: 10.1901/jaba.2010.43-181

Urcuioli, P. J. Stimulus control and stimulus class formation. (2013). In G. J. Madden (Ed.), *APA Handbook of behavior analysis, Vol. 1: Methods and principles* (pp. 361 – 386). Washington, DC: APA.

Woodruff-Pak, D. S., Ewers, M., & Vogel, R. W. III. (2001). An associative learning measure sensitive to Alzheimer's disease early in disease progression. *Fifth International Meeting on Alzheimer's and Parkinson's Disease Abstracts*, Kyoto, Japan.

Woodruff-Pak, D. S., & Thompson, R. F. (1988). Classical conditioning of the eyeblink response in the delay paradigm in adults aged 18 – 83 years. *Psychology and Aging, 3*, 219 - 229. doi: 10.1037/0882-7974.3.3.219

APPENDICES

APPENDIX A

Saint Louis University Mental State Exam

VAMC
SLUMS EXAMINATION

Questions about this assessment tool? E-mail aging@slu.edu

Name________________________________ Age______________

Is the patient alert?______________ Level of education________________________

__/1 — 1 **1. What day of the week is it?**

__/1 — 1 **2. What is the year?**

__/1 — 1 **3. What state are we in?**

4. Please remember these five objects. I will ask you what they are later.

Apple Pen Tie House Car

5. You have $100 and you go to the store and buy a dozen apples for $3 and a tricycle for $20.

1 **How much did you spend?**

__/3 — 2 **How much do you have left?**

6. Please name as many animals as you can in one minute.

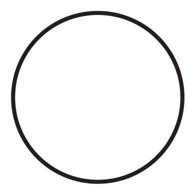
__/3 — 0 0-4 animals 1 5-9 animals 2 10-14 animals 3 15+ animals

__/5 — **7. What were the five objects I asked you to remember? 1 point for each one correct.**

8. I am going to give you a series of numbers and I would like you to give them to me backwards. For example, if I say 42, you would say 24.

__/2 — 0 87 1 648 1 8537

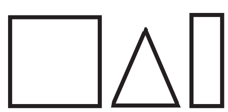
9. This is a clock face. Please put in the hour markers and the time at ten minutes to eleven o'clock.



2 Hour markers okay

__/4 — 2 Time correct

1 **10. Please place an X in the triangle.**



__/2 — 1 **Which of the above figures is largest?**

11. I am going to tell you a story. Please listen carefully because afterwards, I'm going to ask you some questions about it.

Jill was a very successful stockbroker. She made a lot of money on the stock market. She then met Jack, a devastatingly handsome man. She married him and had three children. They lived in Chicago. She then stopped work and stayed at home to bring up her children. When they were teenagers, she went back to work. She and Jack lived happily ever after.

2 **What was the female's name?** 2 **What work did she do?**

__/8 — 2 **When did she go back to work?** 2 **What state did she live in?**

_______ **TOTAL SCORE**

SCORING

High School Education		Less than High School Education
27-30	Normal	25-30
21-26	Mild Neurocognitive Disorder	20-24
1-20	Dementia	1-19

CLINICIAN'S SIGNATURE ________ DATE ________ TIME ________

SH Tariq, N Tumosa, JT Chibnall, HM Perry III, and JE Morley. The Saint Louis University Mental Status (SLUMS) Examination for detecting mild cognitive impairment and dementia is more sensitive than the Mini-Mental Status Examination (MMSE) - A pilot study. *Am J Geriatr Psych* 14:900-10, 2006.

APPENDIX B

Barthel ADL Index

Barthel Index of Activities of Daily Living

Instructions: Choose the scoring point for the statement that most closely corresponds to the patient's current level of ability for each of the following 10 items. Record actual, not potential, functioning. Information can be obtained from the patient's self-report, from a separate party who is familiar with the patient's abilities (such as a relative), or from observation. Refer to the Guidelines section on the following page for detailed information on scoring and interpretation.

The Barthel Index

Bowels
0 = incontinent (or needs to be given enemata)
1 = occasional accident (once/week)
2 = continent
Patient's Score: ____________

Bladder
0 = incontinent, or catheterized and unable to manage
1 = occasional accident (max. once per 24 hours)
2 = continent (for over 7 days)
Patient's Score: ____________

Grooming
0 = needs help with personal care
1 = independent face/hair/teeth/shaving (implements provided)
Patient's Score: ____________

Toilet use
0 = dependent
1 = needs some help, but can do something alone
2 = independent (on and off, dressing, wiping)
Patient's Score: ____________

Feeding
0 = unable
1 = needs help cutting, spreading butter, etc.
2 = independent (food provided within reach)
Patient's Score: ____________

Transfer
0 = unable – no sitting balance
1 = major help (one or two people, physical), can sit
2 = minor help (verbal or physical)
3 = independent
Patient's Score: ____________

Mobility
0 = immobile
1 = wheelchair independent, including corners, etc.
2 = walks with help of one person (verbal or physical)
3 = independent (but may use any aid, e.g., stick)
Patient's Score: ____________

Dressing
0 = dependent
1 = needs help, but can do about half unaided
2 = independent (including buttons, zips, laces, etc.)
Patient's Score: ____________

Stairs
0 = unable
1 = needs help (verbal, physical, carrying aid)
2 = independent up and down
Patient's Score: ____________

Bathing
0 = dependent
1 = independent (or in shower)
Patient's Score: ____________

Total Score: ______________

(Collin et al., 1988)

Scoring:
Sum the patient's scores for each item. Total possible scores range from 0 – 20, with lower scores indicating increased disability. If used to measure improvement after rehabilitation, changes of more than two points in the total score reflect a probable genuine change, and change on one item from fully dependent to independent is also likely to be reliable.

Sources:

- Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. *Int Disabil Stud*. 1988;10(2):61-63.
- Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J*. 1965;14:61-65.
- Wade DT, Collin C. The Barthel ADL Index: a standard measure of physical disability? *Int Disabil Stud*. 1988;10(2):64-67.

APPENDIX C

Procedural Integrity Evaluation

Date: _____________ Phase: ________________ Session #: _________

Participant #: _______ Researcher Initials: ______ Primary/Secondary (Circle)

Directions: Complete the following 7 yes or no questions based on the primary data collector's behavior during the session.

1. Is the research session occurring in a quiet, private environment (door closed, television and music muted, no other residents or staff in room)?

YES NO

2. Was the participant presented with the correct instructions?

YES NO

3. Was the participant exposed to the correct training or testing phase?

YES NO

4. Was the participant exposed to the correct stimuli during the session?

YES NO

5. Did the experimenter refrain from providing differential consequences for participant responses (did not tell the participant "Good job", "That's correct", "That's wrong)?

YES NO

6. Did the experimenter respond with, "We can talk at the end" if asked a question by the participant?

YES NO

7. Was the session terminated if the participant indicated he/she no longer wanted to participate?

YES NO

VITA

Rehabilitation Institute
Southern Illinois University

Dawn A. Seefeldt

dawn.seefeldt@gmail.com

University of South Dakota
Bachelor of Arts, Psychology; Bachelor of Arts, Criminal Justice, May 2009

Minnesota State University, Mankato: Mankato, MN
Master of Arts, Clinical Psychology, July 2011

Special Honors and Awards

Guy A. Renzaglia Award	April 2014
Dean's List	December 2005 – May 2009
Chapter Service Award, Pi Beta Phi,	April 2009
Pi Beta Phi: Undergraduate Scholarship	September 2008
Philip T. Roche Service to Others Award	September 2008
Greg Nothdurft Memorial Award	April 2008
Pi Beta Phi: Ethel Gunderson Scholarship	April 2007
Promise Scholarship	September 2005

Dissertation Title:
Evaluation of equivalence relations: Models of assessment and best practice.

Major Professor: Jonathan C. Baker, Ph.D., BCBA-D

Publications:

Baker, J. C., Seefeldt, D. A., Fairchild, K. M., Ringdahl, J., & Mattingly, J. (2015). The effects of reinforcement and extinction on responding in older adults with cognitive impairment. (In preparation).

Hartman, M., Seefeldt, D. A., Dysterheft, J., Bushey, J., & Lewinski, W. (2015). Recall accuracy of law enforcement officers involved in an analogue traffic stop scenario. (In preparation).

Lewinski, W., Everett, W., Seefeldt, D., & Azar-Dickens, J., O'Neill, J. (2015). The differences between law, logic, and the real way officers make force decisions. (In preparation).

Lewinski, W., Seefeldt, D., Redmann, C., Gronin, M., Sargent, S., & Thiem, P. (2015). The speed of a prone subject. (In preparation).

O'Neill, J., Seefeldt, D. A., Dysterheft, J., & Lewinski, W. (2015). Factors of memory in law enforcement use-of-force. (In preparation).

Seefeldt, D. A., Lewinski, B., & Spence, B. (2015). Mistaken identity: A qualitative study on LEO experience, perception, and training for potentially lethal blue on blue confrontations. (In Preparation).

Baker, J. C., Fairchild, F., & Seefeldt, D. A. (2015). Behavioral gerontology: Research and clinical considerations. In H. S. Roane, J. E. Ringdahl, & T. S. Falcomata (Eds.). *Clinical and Organizational Applications of Applied Behavior Analysis*. Academic Press.

Dysterheft, J. L., Lewinski, W. J., Seefeldt, D. A., & Pettitt, R. W. (2013). The influence of start position, initial step type, and usage of a focal point on sprinting performance. *International Journal of Exercise Science, 6* (4), 320 – 327.

Lewinski, W. J., Dysterheft, J. L., Seefeldt, D. A., & Pettitt, R. W. (2013, March). Influence of officer positioning on movement during a threatening traffic stop scenario. *Law Enforcement Executive Forum, 13*, 98 – 109.

Seefeldt, D. & Buchanan, J. (Writers). (Broadcast in 2011). Flying Saucers and Floating Spirits may be Figments of Happiness and Intuition [Radio Broadcast]. In D. Albertson, E. Stark, & D. Sachau (Producers), Psychological Frontiers: Exploring Human Behavior. Mankato, MN: KMSU.

Seefeldt, D. & Buchanan, J. (Writers). (Broadcast in 2010). Marriage, It's A "We" Thing [Radio Broadcast]. In D. Albertson, E. Stark, & D. Sachau (Producers), Psychological Frontiers: Exploring Human Behavior. Mankato, MN: KMSU.