# MASTER THESIS

# Learning in Complex Systems

# November 2015

Approaches to Understanding of Remembering

Tonje Gellein Åstad

**Faculty of Health Sciences**

**Department of Behavior Science**

OSLO AND AKERSHUS
UNIVERSITY COLLEGE
OF APPLIED SCIENCES

**Acknowledgement**

First, I would like to thank my supervisor Professor Erik Arntzen for all the excellent guidance through this process of conducting the experiment, and the constructive feedback during my writing. I am also very grateful for all the good feedback from the Lab, Experimental Studies of Complex Human Behavior, at Oslo University College of Applied Science, and the access to the Matching-to-Sample program. I will also thank all the participants who took their time and conducted the experiment. Finally, great thanks to my employer, for giving me the opportunity to study this master program beside my work. They all made this master thesis possible.

**Table of Contents**

Article 2: Effects of Including a Separate Conditional Discrimination Training or Not before Implementing a Titrating Delayed Matching-to-Sample Procedure

**List of Tables & Figures**

## Abstract

Article 1 gives an account of how cognitive psychology and behavior analysis understand and explain memory or remembering. Within behavior analysis, the most common procedure for studying remembering is matching-to-sample procedures. Cognitive psychology employs such procedure also, but the Article 1 gives examples of other procedures used to either recall or recognize earlier presented stimuli as words and items. Both cognitive psychology and behavior analysis focus on for how long a past event can control correct responding, and how fast participants respond immediately after the presentation of stimuli, or after a delay. The procedures may not seem to be so different. However, the main difference is related to how the cognitive psychology and behavior analysis interpret the outcome of these behavioral events. Cognitive psychology explains remembering or memory in models and structures, e.g., short-term memory, long-term memory, different systems of working memory, and the transfer of information from one system to another. On the other hand, behavior analysis describes the functional relation between the environmental conditions and the observed behavior, and argues that remembering is something we do, and not any hypothetical constructs. Furthermore, there is a distinction between remembering and reminding, where there are some stimuli guiding correct responding, and just remember without any present stimuli to evoke responses.

Article 2 presents an experiment using a titrating DMTS (TDMTS) procedure. Thirty participants are allocated to three different groups, 10 in each. For two groups, a baseline training is conducted prior the TDMTS condition and for one group, TDMTS condition only is investigated. The main results from this experiment show no differences in the accuracy of responding between the groups. However, the response patterns are more stable in the TDMTS procedure for those who had some training of conditional discrimination on beforehand. Finally, the reaction time data show a typical pattern of an increase in reaction time from training to test, and decrease during the test conditions.

*Keywords:* remembering, cognitive psychology, behavior analysis, adult human participants.

# MASTER THESIS

## Learning in Complex Systems

## November 2015

Different Approaches in the Study of Remembering

Tonje Gellein Åstad

**Faculty of Health Sciences**

**Department of Behavior Science**

OSLO AND AKERSHUS
UNIVERSITY COLLEGE
OF APPLIED SCIENCES

**Abstract**

In the study of memory or remembering, cognitive psychology and behavior analysis have in common some of the same variables that influence performance. Both traditions are recording accuracy of responding and reaction time under different environmental conditions. However, some of the differences are related to how they interpret the outcome from observed environment-behavior relations. Cognitive psychology focuses on explaining memory by different systems and models that is not directly measureable, but their data is very often aggregated observations of many participants. On the other hand, behavior analysis focuses on descriptions of the functional relationship between the manipulative variables and the behavior within participants, and argues that remembering is behavior of interest in its own sake. Another distinction is that cognitive psychology talks about transferring information from the short-term system to a long-term system of memory, while behavior analysis refers to this as some sort of a temporal variable between environmental events and behavior. In addition, it is important to notice that neuropsychology and neuroscience are focusing how memory is related to the activity in the brain under the different conditions. For example, some studies have correlated the measured brain activity with the observed reaction time and response accuracy. This article presents how the different traditions are presenting and studying remembering or memory, and discusses the interpretations of remembering within cognitive psychology and behavior analysis.

*Keywords:* remembering, cognitive psychology, behavior analysis, neuropsychology, recall, recognition, delayed matching-to-sample

## Different Approaches in the Study of Remembering

Remembering refers to behavior under the control of selection that has occurred between the environment-behavior or environment-environment relations in the past (Donahoe & Palmer, 2004). This is behavior observed in the present of stimulus similar to past stimulus when selection occurred like recognitions, or in the absence of stimuli like in recalling past event or stimulus.

Remembering or memory has been studied for many years in cognitive psychology and neuropsychology, and in the last 40 years within behavior analysis tradition. These traditions have in common that they are concerned about factors that influence and affect our ability to remember events and make us forget. Delay and other variables are studied to determine how long, and how human remember events and stimuli that no longer are observable, or recognize a similar stimulus that have been present in a temporal time back. The main differences between these two traditions (cognitive psychology and behavior analysis) may not be how remembering is studied, but in how they interpret the outcome of the study, where response accuracy and reaction time are of interest. Both traditions have some type of behavior as the datum of their experiments. Cognitive psychology is concerned with describing and explaining memory through theoretical structures or models that are derived from observations between behavior and environmental conditions of many participants, and uses metaphors as storage, encoding, retrieval, short-term and long-term storage to explain memory. Behavior analysis uses no such metaphors, and is concerned of describing the functional relation between the behavior under study and its controlling variables. Further, a neuropsychology tradition can supplement and make a connection between the two traditions by referring to activity and locations within the brain under different conditions.

Richard Atkinson and Richard Shiffrin made in the late 1960 a description of a modal model of memory by distinguish between a sensory register of input of information, and a system of short-term memory (STM) and a long-term memory (LTM) (as referred in Atkinson & Shiffrin, 1969). The difference is that limited information can only be stored in the STM a very short period of time before the information is lost to the human or transferred into the LTM system. In 1974, a new model of the memory system was described as the multicomponent model of working memory (WM) by Baddeley and Hitch (Cowan, 2008). From the behavior analytic perspective, the difference between STM and LTM are considered as types of different behavior patterns under delayed stimulus control, and is mainly separated from each other compared to the time between a stimulus was last presented to relevant behavior derived from this stimulus or another stimulus related to the past stimulus is shown. As Palmer (1991) has argued, about the difference; “it is not the memory of the earlier experience but the stimulus control over the behavior that has endured” (p. 265).

In the following discussion section, the paper will (1) present how cognitive psychology and behavior analysis are presenting and studying remembering, especially with focus on explicit memory and WM, (2) briefly describe the role of neuroscience, and (3) discuss the different procedures and the interpretation of memory and remembering. The term remembering will be used of both memory and remembering, except when presenting the cognitive memory models and procedures only, then the term memory is used.

## Learning and Remembering Defined

When studying remembering, it is the learned behavior that has been selected in the past environment-behavior or environment-environment relation that is studied under the present contingencies (Donahoe & Palmer, 2004). Cognitive psychology and behavior analysis define learning and remembering quite differently. Cognitive psychology defines memory as different systems were acquired behavior is stored and at a later time retrieved, and learning is about a process of how the events or stimuli are encoded, stored and afterward retrieved (Baddeley, 1997). Behavior analysis defines learning and remembering as descriptions of behavior in functional relation to the environment. From this point, learning is a process of change in behavior that will maintain over time, and remembering is about the endurance of the selected environment-behavior relations (Catania, 2007). Palmer (1991) defines remembering in two categories: (1) Behavior that "is brought under control of a stimulus at one time" (p. 265) and is still under the control of that stimulus, or stimulus of the same class, when "presented at a later time" (p. 265), and (2) "behavior is brought under control of a stimulus, and reinforcement is later made contingent on appropriate behavior in the absence of the stimulus" (p. 265). The differences are whether the behavior is shown in the presence or in the absence of the stimuli that is conditioned to the response. Donahoe and Palmer (2004) distinguish between reminding and remembering. Reminding refers to the first definition, where stimuli presented after a delay, remind participants of the correct responses, and the second definition for remembering were there are no such stimuli to remind the participants. White (2013) describes remembering as "a discrimination between concurrently available response alternatives" (p. 412). Further, White (2013) refers to forgetting when a stimulus no longer exerts control over the response. In this term of forgetting, the participants do not respond, or make an incorrect response to stimuli that in the past evoked correct responses.

## Cognitive Psychology

Cognitive psychology talks about memory as keeping information over a shorter or longer period of time in a STM and LTM storage system, which is located at different places in the brain (Teigen, 2011). This stored information must first through a process of being encoded, and then later retrieved in environmental conditions. The difference between short and long-term memory is the temporal time from an event or stimulus is presented to remembering responses are emitted.

**Different Memory Systems**

The two most known models or structures regarding memory systems are the modal model of Richard Shiffrin and Richard Atkinson, and the multicomponent model of Working Memory (WM) of Alan Baddeley and Graham Hitch. These will be described in the following. Shiffrin and Atkinson distinguished between a sensory system, a STM and a LTM system. The distinction is that the sensory system holds information while it is transferred to the STM system. And further, that information in STM is then stored for a short period of time before being transferred to a more or less permanently LTM store (Shiffrin & Atkinson, 1969).

**Long-term Memory System.** Cognitive psychology separates LTM between explicit and implicit or procedural memory. Explicit memory is further divided into episodic and semantic memory. Karlsen (2008) explains the differences of these three structures. (1) Explicit memory is about to remember, recall and recognize, events or stimuli from the past. (2) Episodic memory is about what we remember can be referred to an episode where a specific stimulus from an event can be identified. And (3) semantic memory is about our understanding of words and concepts and is referred to their meanings. This is information we have been selected and learned over time in several different conditions or situation. The Implicit (procedural) memory is about the skills we under various conditions perform, and apply to more complex relations and response patterns. This kind of memory is about perceptual identification and can be difficult to refer to specific events or situations where a pattern of behavior is learned as the skills can be acquired over a longer period of time.

**Short-term Memory System.** Both STM and working memory are common used in the daily speech which means that the acquired information or knowledge are limited in time, and depended on the increased difficulties in tests. STM is often used in studies that involve immediate recall or recognitions of presented stimuli. In Atkinson and Shiffrin's modal model, learning was about transferring the information from one system to the other, and that learning in the LTM system depends in time spent on the information in the STM system (Baddeley, 2007). Shiffrin & Atkinson (1969) argued that the information will proceed from the sensory register to STM system and held there for about 30 seconds. One problem of distinguish STM and LTM, is that studies have shown that patients with impaired STM performance are capable of learning in the LTM system. For example, in the study by Shallice & Warrington (1970), a patient with reduced STM in free recall procedure, performed well on an LTM recognition task.

**Working Memory.** Baddeley and Hitch (1974) suggested the multicomponent model of working memory as an alternative to Shiffrin and Atkinson's memory system. Baddeley (1992) explained the WM as “a brain system that provides temporary storage and manipulations of the information necessary for such complex cognitive task as comprehension, learning, and reasoning” (p. 556). Further, Baddeley (2003) wrote: “The theoretical concept of working memory assumes that a limited capacity system, which temporarily maintains and stores information, supports human thought processes by providing an interface between perception, long-term memory and action” (p. 829). As the STM system is widely used as a structure for many different span tasks, Baddeley and Hitch (1974) divided WM into three subsystems; the phonological loop, visuospatial sketchpad and central executive. In 2000, Baddeley supplemented WM with the episodic buffer system (Baddeley, 2000). These four subsystems are working together and are concerned about the process of encoding, storage and retrieval. The phonological loop is about storage of verbal and auditory stimuli for a few seconds, which is possible to increase with a rehearsal process, and play a role in language acquisition (Baddeley, 1992; 2003; 2007). The visuospatial sketchpad is about storage of visual information that is coded from speech or stories (Baddeley 2007). The central executive is not a storage system, but a control system that coordinates and manipulates information between the loop and the sketchpad system. Finally, the episodic buffer is about a limited storage system that integrates information from the phonological loop and the visuospatial sketchpad (Baddeley, 2000).

**Cognitive Memory Tests**

Cognitive psychology uses different span task to measure how many stimuli a person can remember. Recognition, free recall and cued recall, are typical memory test within the explicit memory. It is all about remembering events or stimuli presented a defined time back. Participants are in a recognition procedure about to remember a stimulus presented at one time and select the identic stimulus presented at a later time, where performance is depended on the last presented stimuli (Baddeley, 1997) (e.g., reminding). In recalling procedure, the stimuli have to be remembered in the absence of any stimuli guiding the behavior (e.g., remembering). Into account of episodic memory, a memory test is to recognize or recall words or items one has been presented in advance.

Furthermore, cognitive psychology distinguishes between simple and complex span task when measuring the limitations or capacity of memory. Simple span task is used to measure storage of a single subsystem. As in reading span task participants is, for example, about to recall the final word from a sentence. Complex span task is applied to study interference effect where participants have to perform an additional task in the delay prior to recall or recognition. Testing regarding the phonological loop is about presenting a list of words, that immediately or after a short break should be recalled. Introducing an irrelevant sound, simultaneous or right after the words to be remembered is presented, are used to study the effect of inference (Baddeley, 2003). Four procedures of interest are presented next (see Table 1 for an overview).

The first to mention is the *running memory span task*. This is a free recall procedure that measures the capacity of a number of stimuli to be remembered. In these tasks, a long list of stimuli is presented. The participants have, for example, to recall a defined number of them from the end of the list (Conway 2008). The second is a *serial recognition procedure*, where participants have to recognize probe stimuli, in which is presented after a defined delay, and previously has been presented in series of other stimuli (Baddeley, 1997). The third procedure is a *paired associate procedure*. After presenting pairs of words or other stimuli, one of the words or stimuli from the pair are presented and participants are to choose the other stimuli from the pair (Donahoe & Palmer, 2004). The final procedure is the *cued recall procedure*, which is similar to the paired associate procedure. After presenting numbers of words in pairs, and after a delay, the first word in the pair is presented and the participants are about to recall the second word of the pair (Thomson & Tulving, 1970).

To sum up, cognitive psychology is concerned about the numbers of stimuli human is capable of recognize or recall immediate or after a delay. Studies with use of such tasks have suggested that the STM system was limited to a short period of time and have limited capacity. This will be further outlined later in this article. Behavior analysis describes remembering and use different procedures to study remembering than the cognitive psychology does. This will be illuminated in the following.

## Behavior Analysis

Branch (1977) argued that there is no use of explaining STM and LTM through distinctive storage devices in the brain. Instead, behavior analysis explains remembering as behavior under different stimulus control, and it is just the time between the stimuli presentations and emitted response that distinguish those two systems. From this, remembering is operant behavior emitted in the environmental context. When a correct response evokes in time after a presentation of a stimulus, we say that the stimulus is remembered and that the response is under the control of the absent stimulus. To investigating complex human behavior such as learning and remembering, and to determine variables that influence remembering performance, the most used procedure is Matching-to-Sample (MTS) procedures. There are several ways of implementing this type of procedure. Hence, a distinction between the main procedures will be outlined below, and are listed in Table 1.

## Delayed Matching-to-Sample Procedure in Studying Remembering

In MTS procedure, a sample stimulus is a conditional stimulus that will functions as a controlling stimulus for the selection response of a comparison. In general, a sample is presented on the screen. An observing response or after a limited time, the comparisons will be present randomly in the corners of the screen. Making a choice response to the experimenter-defined correct comparison stimuli produces reinforcement only if one specific sample is presented or has been presented. This continues trial by trial until the conditional discriminations between all the sample-comparison pairs are established. For example, if sample A1, selecting B1 or C1 leads to reinforcement, and not selecting comparison B2 or C2. However, if sample A2 is shown, selecting comparison B2 or C2 will be reinforced and not selecting B1 or C1. In the beginning of an experiment session, the probability of selecting B1 or C1 in the presence or A1 is at chance level (50/50 when using two comparisons). When selection of B1 or C1 is reinforced, this selection is under stimulus control of A1.

**Delayed Matching-to-Sample.** Delayed Matching-to-Sample (DMTS) is an important procedure to study remembering and forgetting behavior (e.g.,Blough 1959). In Simultaneous MTS procedure, the sample is present at the same time as a choice response to comparison is emitted. In the DMTS procedure, the sample disappears for a defined period of time before the onsets of the comparisons. In this way, the use of delay in MTS procedure will contribute to identify when the sample no longer exerts control over the selection response (Sidman, 1969). The standard procedure in DMTS is to use fixed delay, where the delay is stable throughout the conditional discrimination training. An alternative procedure is titrating or adjusted DMTS (TDMTS) where the length of delay changes as a function of the participants' performance (Cumming and Barryman, 1965; Eilifsen & Arntzen, 2011; Kangas, Vaidya & Branch, 2010; Lian & Arntzen 2011; Rosenberger, Mohr, Stoddard, & Sidman, 1968; Sidman, 2013). The delay between the sample offset and the comparison onset will increase only as a result of correct responding, and it will be maintained at the same delay or have a decrease in delay after incorrect responses. DMTS procedure has been applied for children (e.g., Lian & Arntzen, 2011; Torgrud & Holborn, 1989), normal functioning adults (e.g., Arntzen, 2006; Vaydia & Smith, 2006), with older adults with and without dementia (Plaza, López-Crespo, Antúnez, Fuentes, & Estévez, 2012; Sidman, 1969; Stengrimsdottir & Arntzen, 2011a; 2011b; 2014a; 2014b), and people with mental retardation (e.g., Williams, Dube, Johnston, & Saunders, 1998; Williams, Johnston, & Saunders, 2006) in a variety of ways to identify both what humans remember, and at which delay humans remember or forget. In these procedures, there are both being used identity and arbitrary matching, which are explained below.

**Identity and Arbitrary MTS**. Sidman (1994) argued that identity matching means sameness and is “a prerequisite for the emergence of simple meanings, vocabularies, or semantic correspondence” (p. 340). Identity matching has been used in DMTS procedure with participants diagnosed with dementia, and is seen as the simplest form of MTS procedure where the correct defined comparison is identical to the sample stimulus. For example, if sample A1 is presented, then choosing comparison A1, and not any of the other comparisons, is right. In some cases, learning identity matching is a prerequisite for, and can enhance the matching performance of arbitrary stimuli, but it can also weaken the arbitrary matching as the participant starts looking for similarities between the stimuli (Sidman, 1994).

Therefore, Sidman (1994) suggested a test for generalized identity matching to determine if the matching performance is based on similarities and no other aspect of the stimulus. Here, a new set of stimuli other than from training will serve as sample and comparison, and correct choices of comparison can be called true matching (Iversen, 1997).

Instead of matching identical stimuli participants can conduct an arbitrary MTS procedure. In these procedures, participants match comparison to sample, which do not share any physical properties. For example, visual – visual matching (written word – picture) and auditory – visual (spoken word – printed word or picture). For example, a picture of a *cat* means the same as the written word *cat* and may lead to the same response saying *"cat".*

**Stimulus Control in DMTS Procedure**

In operant behavior, when a response is only reinforced in the presence of a particular stimulus, the response will most likely be repeated in the presence of that stimulus on a later occasion. The response strength from this selection is measured by accuracy of responding and reaction time. High accuracy and low reaction time indicate strong response strength, while the opposite indicates weak response strength. The main question in DMTS procedures, are how long a sample stimulus will maintain control of a selection response. Accurate responding in DMTS task shows that sample exerts a delayed stimulus control of the selection. As Palmer argued; "Current behavior is under control of current variables" (Palmer, 1991, p. 264). If participants have emitted correct responding to comparison, but with longer delay no longer do so, the conditional stimulus does not function as a controlling stimulus any more, and the incorrect response may be termed as forgetting (White, 2013). White (2013) argues with this that the discrimination occurs at the time of the comparison presentation and not at the time of sample onset.

## Variables Influencing Remembering Performance

In experimental studies of remembering in both traditions, many of the same variables function to increase the evocative effect or to suppress responses. Effects of manipulate the independent variable such as delay, rehearsal, and other supplemental stimulus controlling responses and stimuli, will show the changes in the dependent variable, the response rate of remembering, and in reaction time, from a remembering response is supposed to be emitted. Presenting more complex stimuli or distracting, remembering will often decline. Some of the independent variable that influences the remembering performance in human are listed in Table 2, and will be outlined in the following.

### Delay

The duration of delay is critical when studying remembering. Cognitive psychology talks about memory that do not pass from the STM system to the LTM system are lost from the storage system, while behavior analysis is concerned about the decline of stimulus control as the time pass. Typical results are that the accuracy of responding will decrease with increase delay. However, some studies have shown that longer delay than shorter, results in fewer trials to establish new conditional relations, and more correct responding during tests. One reason for this may be that the use of delay, in both traditions, has shown that there is time for supplemental controlling stimuli or precurrent behavior to occur (Donahoe & Palmer, 2004). This has to be studied within each participant.

By increasing the delay, may both decrease and increase the remembering responses. In a study of using both associative word and non-associative word, Glanzer and Schwartz (1971) found better performance in recalling non-associative words immediate after presentations than after a delay, and that there were no clear differences between delay and no delay for the associative words. In MTS procedure, longer delay may facilitate the conditional discrimination learning using higher rather than shorter delay in a fixed DMTS procedure (Arntzen, 2006; Lian & Arntzen, 2013). Stengrimsdottir and Arntzen (2011b) found that increasing delay for a person with Alzheimer, gave less accuracy in identity matching in a DMTS procedure. The use of DMTS can study correct responding as a function of increasing delay, in fixed and titrated delays. Titrated instead of fixed delay can (1) enhance the remembering function as a function of the gradual increase in delay, (2) give us more precise information for how long a stimulus controls a correct responding if the titration steps are relative small, and (3) it can give better performance on higher delay compared to fixed delay (Arntzen, Steingrimsdottir, & Brogård-Antonsen, 2013).

**Precurrent Responses**

By giving name or a manual sign to stimuli, participants can rehears (overt or covert). These are mediating responses that function as supplemental controlling stimulus, and can enhance the rate of correct responding in test. In MTS procedure with arbitrary visual stimuli, participants often give names to stimuli. When rehearsing a stimulus during the delay, it involves responses to the prior stimuli, that are repeated until the remembering response is about to be emitted. Such responses can evoke an intraverbal response to the next, arbitrary stimulus is presented (Donahoe & Palmer, 2004). Rehearsal will in this way build a bridge between the presented stimuli to be remembered, to a response is supposed to be emitted after a delay. However, rehearsal is not always possible, as in for example serial recognition procedure. In these procedures the participants are prevented from rehearsing because of the multiple sample presentations prior the recognition test (e.g., Sands & Wright, 1980). The positive effect of rehearsal has been demonstrated in several studies. In a study with pigeons, Blough (1959) found that stereotypic responses during the delay lead to more correct responding as the delay increased. Likewise, repeated rehearsal in the delay for human can enhance remembering responses. In an experiment where children learned to emit sample specific response, had higher accuracy of responding after delay, than the children who did not emit such responses (Parson, Taylor, & Joyce, 1981; Torgrud & Holborn, 1989). Name giving to pictures in a DMTS task has also shown to improve the matching performance (Carrigan and Sidman, 1975; Gutowski & Stormer, 2003). Finally, Horne, Lowe and Harris (2007) showed that name giving, and manual sign to sample gave better matching performance.

**Distractor**

In the daily life, many variables can distract us from remembering. With increased delay, the extent of interfering variables may increase. Experimentally, using distractor in the delay between the presentations of the stimulus to be remembered or before the presentation of the stimuli that remind the participant of the correct target response may decrease the accuracy of responding. This has been studied by implement a distractor task in the delay after the relations between the stimuli in DMTS task were established (Arntzen, 2006; Arntzen & Vie, 2013; Torgrud & Holborn, 1989). This kind of procedure will prevent participants to rehears or repeat (overt or covert) the sample stimulus in the delay. Another variable that may distract correct responding is that later learning may affect the earlier learning. For example, if a participant initial has learned one set of paired words, for example, the stimulus word PEM and the response words big, and afterward learn PEM-bright relations in a paired associative procedure, this second learned list will infer with the first (PEM-big), and will typically respond to the second learned stimulus word (Donahoe & Palmer, 2004).

**Complexity of the Stimuli**

The complexity of the stimuli like un-familiar stimuli, un-associative stimuli, increase the word length and so on may affect the performance of remembering the stimuli presented in advance.

**Arbitrary and identity stimuli.** The use of identity stimuli in MTS performance can give more accurate responding than with arbitrary stimuli. In the study of Steingrimsdottir &

Arntzen (2011b), one man with Alzheimer responded incorrect on arbitrary matching, but did match on the identity matching.

**Familiar and unfamiliar stimuli, intraverbal responding and similarity.** In MTS procedure, the use of familiar stimuli may yield better performance in establishment of conditional discriminations (Arntzen & Lian, 2010) than using abstract and unfamiliar stimuli alone. Furthermore, it can be easier to remember the second word in a pair of words if they are intraverbal related like “bread” and “butter” (Palmer, 2010), and the recall performance decrease when they are supposed to recall unrelated words (Baddeley, 2000). In contrast, in according to similarity**,** the words used in recall test, has proven that it can be more difficult in remembering a sequence of words if they are similar to each other, rather than dissimilar.

**Cue word and category-name.** In cognitive memory test, cue word and category-name can improve remembering responses. Comparing recall and recognition, recalling words in a cue recall procedure can give better performance than recognize one of the words from a list among new words (Tulving & Thomson, 1973). Further, using a cue word in the training or encoding phase has shown to give higher response accuracy rather than if the cue word only was presented when the words were about to be recalled (e.g., Thomson and Tulving, 1970). In the light of distinguishing reminding and remembering, the participants can be given a category name as a cue word for the words from the list in a cued recall procedure (Palmer, 2010), and not be presented a category word. Result from the study of Tulving & Psotka, (1971), showed higher accuracy of responding when the category name was presented rather than during free recall procedure.

**Word length.** Studies have also shown that immediate recalling performance is better when words have short temporal duration of pronunciation (Baddeley, 2003; 2007; Baddeley, Thomson, & Buchanan, 1975; Cowan, Day, Saults, Keller, Johnson, & Flores, 1992). This is known as the word length effect**,** and is explained in the sense that long words are more complex than shorter (Baddeley, 2003). On the other hand, in the study by Tehan and Tolan (2007), there were demonstrated that long word gives better performance in recognition of items and cued recall procedure, in contrast to recall procedures.

Those variables mentioned here, should not, for behavior analysis, serve as explanation, but rather as a statement of results derived from performance during the test.

**Neuropsychology**

The history of patient HM, helped neuropsychology to correlate long-term memory and short-term memory to different parts of the brain. After removing the hippocampus, amygdala and part of the medial temporal Corti, HM lost his ability to store new memories, but was capable of recalling figures after a delay of 30 s. (Karlsen, 2008; Smith, Kosslyn, & Barsalou, 2007). Today, the neuropsychology uses distinctive measurement to determine activity in the different area within the brain during different test conditions. The fMRI (functional magnetic resonance imaging), and EEG (electroencephalography), can determine changes in brain activity and the locations of the activity under different environmental conditions (Ortu, 2010). The use of fMRI can be used to correlate the results with, for example reaction time and response rate during testing. Dickins (2005) sums up some studies from early in the 2000 century, where there are correlations between brain activity and the dependent measured response rate and reaction time under different conditions of more or less related and known stimulus relations. For example, Dickins, Singh, Roberts, Burns, Downes, Jimmieson et al. (2001), found in their study, correlations between the brain activity in the frontal cortex and the response accuracy during the test for the trained relations in a MTS procedure, and between the measured activity and reaction time between not directly trained relations.

One technique of EEG is the event related potential (ERP), that measure the electrical activity in the brain during test conditions. ERP can describe the correlations between the response strength and the electrical brain activity (Ortu, 2010). ERP has shown faster reaction time when the words are related to each other (e.g., intraverbal responses). A N400 effect is typically measured using ERP. For example, will the N400 effect measure lower activity at 400 ms after stimulus presentation when a stimulus is correlated to a known or evocative stimulus. And N400 effect will be typical shown by a negative peak (lower electrical brain activity), as when presenting unknown or no evocative stimulus. This means, as the decrease in observed reaction time indicates greater response strength with lower activity in the N400 effect does so too.

## Comparison of the Traditions

### The Procedural Differences

The tradition of the cognitive psychology and the behavior analysis has well-described procedures in which is concerned of studying the effect of manipulate with the independent variables, and observe the effect of remembering responses (see Tables 1 and 2 for an overview). In that, the different procedure supplements each other. When studying remembering in a behavior analytic way, DMTS procedures are most often conducted. In this procedure the first presented stimulus (sample) control the effect of the second presented stimulus (comparison). This is different from cognitive testing (described in this article) where participants have to memories several stimuli at the same time for then to remember them after a delay. As a recognition test differs from arbitrary matching, recognition is more similar to identity matching. In a recognition procedure, participants are supposed to select an identical stimulus (for example a word) to what was presented prior to the delay. Furthermore, in identity matching, identic stimuli are supposed to match together. Further, this differs from free recall procedures where the participants are to recall presented stimuli after a delay. Comparing recall and recognition, Baddeley (1997) claims that recognition procedure may be easier than recall because of the absence of the stimulus when recalling performance is about to take place.

In a cued recall procedure, a cue word is presented after a delay. This is much like arbitrary matching were the comparison, which is paired with the sample, is presented. However, there is a distinction in the explanations of the cue word between the two traditions. Cognitive psychology refers to this as an association to the target response, while the behavior analysis refers to this word as a stimulus that increases the probability to evoke a correct response to the second word, like for example an intraverbal response. Likewise, there is a distinction in the way they interpret the effect of a paired associate procedure. Cognitive psychology refers to this as a mediating process of association, while the behavior analysis describes the functional relation between the controlling stimuli and the responses.

Both traditions have investigated the effect of verbal name giving to the stimuli. By give distinctive names to the different stimuli in arbitrary matching and cued recall procedure, this functions as intraverbal responding in which can increase the probability of correct responding. The same results are shown when a participant rehears the name, or other responses are made for the sample during the delay. This way, the time between the stimuli presentations will become shorter in the sense that it functions as a supplemental controlling variable for the evocative effect of the response. It has been demonstrated that after naming sample pictures prior to comparison selection have improved the matching performance in DMTS procedure (Carrigan and Sidman, 1975; Gutowski & Stormer, 2003). In this case, it functions as a mediating response that facilitates remembering. Whether this is tact of the stimulus to be remembered, or function as a supplemental controlling stimulus for intraverbal responding to comparison, Sidman (1994) argues that such explanation is nothing more than behavior-behavior explanations. The effect of name giving and rehearsal can be studied if the participant talk aloud, or use distractor to prevent such responding in the delay. Without talking aloud, rehearsal will only be a speculation of what enhance the probability of remembering. Even though there are disagreements about the importance of the effect of rehearsing (Baddeley, 2007), it is important from the behavior analysis not to use this as an explanation of the remembering effect, in which applies to all manipulated variables. Even though delays facilitate rehearsal and remembering responses, it has also been shown that prior stimuli may lose its controlling function of selection as the delay increase (Sidman, 1969). In this way, the use of different delay in fixed or in titrated DMTS procedure, testing can provide a better knowledge of in which delay, stimuli loose its controlling functions of remembering.

Accuracy of remembering responses are measured in both cognitive and behavior analysis. Dymond and Rehfeldt (2001) and Palmer (2010) suggested that other dependent variables can contribute the study of complex human behavior, and see the importance of indirect measure, for example, reaction time because it is difficult to measure covert responses. For behavior analysis, high accuracy and low reaction time are demonstrated as a function of high degree of response strength. For example, measuring of neurological activity can contribute to the understanding of reaction time and response strength, like in the study of Polish & Kok, (1995) and Dickins et al. (2001). The use fMRI and ERP are ways to describe the correlations with the degree of relatedness of stimuli between measurable behavior and brain activity. As measured reaction time between stimulus presentations and response, an ERP can contribute to the fact that there are some overt responding going on (Ortu, 2010). As neuropsychology studies refer to brain activity, combining these traditions may give a wider and more concrete knowledge of complex behavior such as remembering. Remembering is explained in a more pragmatic way than cognitive psychology since it refers to the brain activity, and not just referring to collective metaphors as encoded, stored, and retrieved. However, quite different from what is common within behavior analysis, the neuroscience sum up the results from many experiments and do statistical tests based on the average performance of many participants.

**Interpretations of Remembering**

To draw the line of difference between the traditions of cognitive psychology and behavior analysis, we need to see how the terms of memory and remembering are used. Both memory and remembering are terms bases on some sort of behavior observed in the environment or experimental setting. It can be argued that the terms remembering and operant are abstraction and summary labels of operations that are measurable, and that memory is a hypothetical construct. MacCorquodale and Meehl (1948) distinguishes between intervening variables as an abstract variable summed up of observable events, and hypothetical constructs in which involve theoretical variable with meaning of something more than just the observed relations. Operant is a fundamental term that sums up repeated observations of environment-behavior relation. From this distinction, the term operant is an intervening variable. There are some disagreements in this perspective. While Svartdal (2014) argued that operant behavior goes under the term hypothetical construct as well as memory, Eilifsen and Arntzen (2014) argued that an operant fits in MacCorquodale and Meehl definition of intervening variables.

The WM model of its four subsystems (phonological loop, the visuospatial sketchpad, central executive and episodic buffer) seems to go under the term hypothetical construct. Neither of these names of the systems, nor the interaction between them, is observable. These are more like words that function as theoretical explanation of unobservable events. Instead of explaining what is going on between the presentations of a stimulus and the emitted behavior with hypothetical constructs as an unobservable mediating process, behavior analysis observe the behavior of interest and describe the environment-behavior relations. With the use of metaphors and models as an explanation for an environment-behavior relation, this relation is transferred to an unobservable mental thing that is going on in the delay between the presentations of stimuli and the emitted behavior. In this case, behavior analysis refers to behavior evoked by controlling stimuli from the past, rather than retrieved in which can associate memory as something that exists without being observed (Donahue & Palmer, 2004). From the cognitive tradition, what is to be remembered is transferred from one storage system to another, and some hypothetical process between the systems will search for the behavior needed to be remembered. Roediger (1980) summed up the use of metaphors of memory research during the history. He argued that metaphors like encoding, storage and retrieval do nothing more than explaining human memory as something that exists in a multidimensional space, and find it more functional referring to environment-behavior relations. With this, remembering responses are observable, memory is not.

From a practical point of view, the use of metaphors and models can convey the phenome to the public in the sense that they understand the metaphors as coding, storing and retrieval from the daily speech. In that, theoretical explanation of something unobservable is accepted as an explanation. Talking to the public about stimulus control may be more distances or complicated to understand. Even so, when explaining memory in light of the storage metaphor, we will not determine the role of the controlling stimulus-response relations in science. Even when neuro psychologists observe activations in the brain during different memory tests, will not do so. Moore (2015) brought up that science starts with observations of single cases and ends up with general rules and scientific laws. Cognitive psychology most often compares the performance of bigger groups and outlines some general rules, and explains the outcome with models and structures. From the behavior analysis point, this excludes other individual performances and capabilities.

## Summary and Conclusion

In the current paper, overviews over some of the experimental variables that can influence the remembering responses are presented for both traditions (see Table 2). Even though, there are similarities between the procedures and the variables in the studies within these traditions, the observed effects of remembering responses are often interpreted differently. The analysis of remembering within behavior analysis is defined functionally, and the performances are described in terms of what the participants do during the different conditions. The cognitive psychology rather use models, metaphors and hypothetical construct derived from the observation of behavior through many participants under the same conditions to explain remembering. Furthermore, while cognitive psychology refers to a distinction between STM system and LTM system in the brain, and that the information is transferred from one system to the other, the behavior analysis describes this as a temporal variable between the controlling stimulus and the emitted remembering responses. It has further been outlined the role of neuropsychology, that refers to the activity in the brain during test conditions. For example, how the N400 effect in ERP correlated to the results of the observed accuracy and reaction time during testing. From this point, the neuropsychology gives another way to describe the relation between stimulus presentations and the variable that affect remembering responses.

Regardless of which tradition that investigates the variables of remembering and how they interpret the outcome, Skinner argued in 1989 that:

> “The analysis of behavior need not wait until brain scientists have done their part. The behavioral facts will not be changed, and they suffice for both a science and a technology. Brain scientists may discover other kinds of variables affecting behavior,

but they will turn to a behavioral analysis for the account of the effects of these variables" (p. 18).

**References**

Arntzen, E. (2006). Delayed matching to sample: Probability of responding in accord with equivalence as a function of different delays. *The Psychological Record, 56*(1), 135–167. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html.

Arntzen, E., & Lian, T. (2010). Trained and derived relations with pictures versus abstract stimuli as nodes. *The Psychological Record, 60*(4), 659–678. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol60/iss4/8.

Arntzen, E., Steingrimsdottir, H. S., & Brogård-Antonsen, A. (2013). Atferdsmessige studier av demens: Effekten av ulike varianter av matching-to-sample prosedyrer. *Norsk Tidsskrift for Atferdsanalyse, 40*(1), 17–29. Retrieved from www.atferd.no.

Arntzen, E., & Vie, A. (2013). The Expression of Equivalence Classes Influenced by Distractors During DMTS Test Trials. *European Journal of Behavior Analysis, 14*(1), 151–164. doi: 10.1080/15021149.2013.11434453

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556–559. doi: 10.1126/science.1736359

Baddeley, A. D. (1997). *Human memory: Theory and practice*. Bristol, UK: Psychology Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423. doi: http://dx.doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839. doi: 10.1038/nrn1201

Baddeley, A. (2007). *Working memory, thought, and action*. New York, US: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. doi: 10.1016/S0022-5371(75)80045-4

Blough, D. S. (1959). Delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior, 2*(2), 151–160. doi: 10.1901/jeab.1959.2-151

Branch, M. N. (1977). On the role of "memory" in the analysis of behavior. *Journal of the Experimental Analysis of Behavior, 28*(2), 171–179. doi: 10.1901/jeab.1977.28-171

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*(1), 183–204. doi: 10.1901/jeab.1992.58-183

Catania, A. C. (2007). *Learning* (Interim, 4th ed.). New York: Sloan Publishing.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research, 169*, 323–338. doi: 10.1016/S0079-6123(07)00020-9

Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language*, *31*(1), 1–17. doi: http://dx.doi.org/10.1016/0749-596X(92)90002-F

Cumming, W., & Berryman, R. (1965). The complex discriminated operant: Studies of matching-to-sample and related problems. In D. I. Mostofsky (Ed.), *Stimulus Generalization* (pp. 284–330). Standford, CA: Standford University Press.

Dickins, D. W. (2005). On Aims and Methods in the Neuroimaging of Derived Relations. *Journal of the Experimental Analysis of Behavior, 84*(3), 453–483. doi: 10.1901/jeab.2005.92-04

Dickins, D. W., Singh, K. D., Roberts, N., Burns, P., Downes, J. J., Jimmieson, P. et al. (2001). An fMRI study of stimulus equivalence. *NeuroReport*, *12*(2), 405–411. doi: 10.1097/00001756-200102120-00043

Donahoe, J. W., & Palmer, D. C. (2004). *Learning and complex behavior*. Needham Heights, MA, US: Allyn & Bacon.

Dymond, S., & Rehfeldt, R. A. (2001). Supplemental measures of derived stimulus relations. *Experimental Analysis of Human Behavior Bulletin, 19,* 8–12.

Eilifsen, C., & Arntzen, E. (2011). Single-subject withdrawal designs in delayed matching-to-sample procedures. *European Journal of Behavior Analysis, 12*(1), 157–172. doi 10.1080/15021149.2011.11434361

Eilifsen, C. & Arntzen, E. (2014). Er alle abstraksjoner hypotetiske konstrukter? *Norsk Tidsskrift for Atferdsanalyse*, 41(2), 133–138. Retrieved from www.nta.atferd.no.

Glanzer, M., & Schwartz, A. (1971). Mnemonic structure in free recall: Differential effects on STS and LTS. *Journal of Verbal Learning and Verbal Behavior, 10*(2), 194–198. doi: http://dx.doi.org/10.1016/S0022-5371(71)80013-0

Gutowski, S. J., & Stromer, R. (2003). Delayed matching to two-picture samples by individuals with and without disabilities: an analysis of the role of naming. *Journal of Applied Behavior Analysis*, *36*(4), 487–505. doi: 10.1901/jaba.2003.36-487

Horne, P. J., Lowe, C. F., & Harris, F. D. A. (2007). Naming and categorization in young children: v. Manual sign training. *Journal of the Experimental Analysis of Behavior, 87*(3), 367–381. doi: 10.1901/jeab.2007.52-06

Iversen, I. H. (1997). Matching-to-sample performance in rats: a case of mistaken identity? *Journal of the Experimental Analysis of Behavior*, *68*(1), 27–45. doi: 10.1901/jeab.1997.68-27

Kangas, B. D., Vaidya, M., & Branch, M. N. (2010). Titrating-delay matching-to-sample in the pigeon. *Journal of the Experimental Analysis of Behavior*, *94*(1), 69–81. doi: 10.1901/jeab.2010.94-69

Karlsen, P. J. (2008). *Hva er hukommelse*: Oslo, Norway: Universitetsforlaget.

Lian, T., & Arntzen, E. (2011). Training conditional discriminations with fixed and titrated delayed matching to sample in children. *European Journal of Behavior Analysis, 12*(1), 173–193. doi: 10.1080/15021149.2011.11434362

Lian, T., & Arntzen, E. (2013). Delayed matching-to-sample and linear series training structures. *The Psychological Record*, *63*(3), 545–561. Retrieved from http://search.proquest.com/docview/1442999079?accountid=26439.

MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review, 55*(2), 95–107. doi: 10.1037/h0056029

Moore, J. (2015).Pragmatism, mathematical models, and the scientific ideal of prediction and control. *BehaviouralProcesses,114*,2–13.doi:10.1016/j.beproc.2015.01.007.

Ortu, D. (2012). Neuroscientific Measures of Covert Behavior. *The Behavior Analyst*, *35*(1), 75–87. Retrived from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3359857/.

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

Palmer, D. C. (2010). Behavior Under the Microscope: Increasing the Resolution of Our Experimental Procedures. *The Behavior Analyst, 33*(1), 37–45. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2867504/.

Parsons, J. A., Taylor, D. C., & Joyce, T. M. (1981). Precurrent self-prompting operants in children: “remembering”. *Journal of the Experimental Analysis of Behavior*, *36*(2), 253–266. doi: 10.1901/jeab.1981.36-253

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*(4), 483–489. doi: 10.1037/a0028485

Roediger, H. (1980). Memory metaphors in cognitive psychology. *Memory & Cognition, 8*(3), 231–246. doi: 10.3758/BF03197611

Rosenberger, P. B., Mohr, J. P., Stoddard, L. T., & Sidman, M. (1968). Inter- and intramodality matching deficits in a dysphasic youth. *Archives of Neurology, 18*(5), 549–562. doi: 10.1001/archneur.1968.00470350107010

Sands, S. F., & Wright, A. A. (1980). Serial probe recognition performance by a rhesus monkey and a human with 10- and 20-item lists. *Journal of Experimental Psychology: Animal Behavior Processes, 6*(4), 386–396. doi: 10.1037/0097-7403.6.4.386

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *The Quarterly Journal of Experimental Psychology, 22*(2), 261–273. doi: 10.1080/00335557043000203

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review, 76*(2), 179–193. doi: 10.1037/h0027277

Sidman, M. (1969). Generalization gradients and stimulus control in delayed matching-to-sample. *Journal of the Experimental Analysis of Behavior, 12*(5), 745–757. doi: 10.1901/jeab.1969.12-745

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*, 149–149. doi: 10.1080/15021149.2013.11434452

Skinner, B. F. (1989). The origins of cognitive thought. *American Psychologist, 44*(1), 13–18. doi: 10.1037/0003-066X.44.1.13

Smith, E. E., Kosslyn, S. M., & Barsalou, L. W. (2007). *Cognitive psychology: Mind and brain*: Upper Saddle River, NJ: Pearson Prentice Hall.

Steingrimsdottir, H. S., & Arntzen, E. (2011a). Identity matching in a patient with Alzheimer's disease. *American Journal of Alzheimer's Disease and Other Dementias, 26*(3), 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*(3), 179–192. doi: 10.1002/bin.334

Steingrimsdottir, H. S., & Arntzen, E. (2014a). Discrimination learning in adults with neurocognitive disorders. *Behavioral Interventions, 29*(3), 241–252. doi: 10.1002/bin.1389

Steingrimsdottir, H., & Arntzen, E. (2014b). Performance by Older Adults on Identity and Arbitrary Matching-to-Sample Tasks. *The Psychological Record, 64*(4), 827–839. doi: 10.1007/s40732-014-0053-8

Svartdal, F. (2014). Hypotetiske konstrukter innenfor atferdsanalyse: Finnes de? Ja, men der lever de dessverre ikke i beste velgående. *Norsk Tidsskrift for Atferdsanalyse*, 41(2), 119–131. Retrieved from www.nta.atferd.no.

Tehan, G., & Tolan, G. A. (2007). Word length effects in long-term memory. *Journal of Memory and Language, 56*(1), 35-48. doi: http://dx.doi.org/10.1016/j.jml.2006.08.015

Teigen. K. H (2011) Hukommelse og glemsel. In F. Svartdal (Eds.). *Psykologi: en introduksjon* (pp.107–132). Oslo, Norway: Gyldendal Akademisk.

Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology, 86*(2), 255–262. doi: 10.1037/h0029997

Torgrud, L. J., & Holborn, S. W. (1989). Effectiveness And Persistence Of Precurrent Mediating Behavior In Delayed Matching To Sample And Oddity Matching With Children. *Journal of the Experimental Analysis of Behavior, 52*(2), 181–191. doi: 10.1901/jeab.1989.52-181

Tulving, E., & Psotka, J. (1971). Retroactive inhibition in free recall: Inaccessibility of information available in the memory store. *Journal of Experimental Psychology, 87*(1), 1–8. doi: 10.1037/h0030185

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. doi: 10.1037/h0020071

Vaidya, M., & Smith, K. N. (2006). Delayed matching-to-sample training facilitates derived relational responding. *Experimental Analysis of Human Behavior Bulletin, 24*, 9-16. Retrieved from http://www.eahb.org/eahb-volume-2-2006/.

White, K. G. (2013). Remembering and forgetting. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley & K. A. Lattal (Eds.), *APA handbook of behavior analysis, Vol. 1: Methods and principles* (pp. 411–437). Washington, DC, US: American Psychological Association.

Williams, D. C., Dube, W. V., Johnston, M. D., & Saunders, K. J. (1998). Conditional Versus Trial-Unique Delayed Matching-to-Sample. *American Journal on Mental*

*Retardation, 103*(2), 186–192. doi:10.1352/0895-8017(1998)103<0186:CVTDM>2.0.CO;2

Williams, D. C., Johnston, M. D., & Saunders, K. J. (2006). Intertrial sources of stimulus control and delayed matching-to-sample performance in humans. *Journal of the Experimental Analysis of Behavior, 86*(2), 253–267. doi: 10.1901/jeab.2006.67-01

Table 1

*Overview of Some Procedures Used in Studying Human Remembering*

| Procedure | Cognitive Psychology | Behavior Analysis |
|---|---|---|
| DMTS | X | X |
| - Identity matching | X | X |
| - Arbitrary matching | X | X |
| Running memory span task | X | |
| Serial recognition | X | |
| Paired associative | X | |
| Cued recall procedure | X | |

Table 2

*Overview of Some Independent Variables of Interest in Human Remembering Procedures*

| Variable | Cognitive Psychology | Behavior Analysis |
|---|---|---|
| **Delay** | X | X |
| fixed | X | X |
| titrated | | X |
| **Precurrent responses** | X | X |
| Rehearsal | X | X |
| Name giving | X | X |
| **Distractor** | | |
| **Stimuli Complexity** | X | X |
| Arbitrary and identity | | X |
| Familiar and unfamiliar | | X |
| Cue word – category name | X | |
| Word length | X | |
| Intraverbal related | X | |
| Similar and dissimilar word | X | |

# MASTER THESIS

## Learning in Complex Systems

## November 2015

Effects of Including a Separate Conditional Discrimination Training or Not before Implementing a Titrating Delayed Matching-to-Sample Procedure

Tonje Gellein Åstad

**Faculty of Health Sciences**

**Department of Behavior Science**

OSLO AND AKERSHUS
UNIVERSITY COLLEGE
OF APPLIED SCIENCES

## Abstract

Delayed Matching-to-Sample (DMTS) procedures are used in several studies to study variables that influence “remembering” and how long a stimulus can exert control of a response. The use of titrating DMTS (TDMTS) may facilitate higher delay in the establishing the conditional discriminations. A quite small number studies have investigated the effects of TDMTS in humans. The purpose of the present experiment was to investigate the effect of with and without baseline training of the conditional discriminations prior to implementing the DMTS procedure with a titrating up to 12 seconds delay. Thirty adults were allocated into three groups. In Group 1, ten participants started with a Simultaneous Matching-to-Sample, while ten participants in Group 2 had a 0 s DMTS training of the baseline relations before starting TDMTS procedure. In Group 3, ten participants started directly with TDMTS procedure without any training as described for Groups 1 and 2. The findings show no significant differences in accuracy of responding among the groups. However, the response patterns during TDMTS procedure were more stable for Group 1 and 2. All three groups showed also high accuracy in the test conditions for equivalence relations. Finally, reaction time data show slower responding to comparison for equivalence trials, than for baseline trials, and also a decrease during test conditions in baseline, symmetry, and equivalence trials.

*Keywords:* Stimulus equivalence, titrating delayed matching-to-sample, adults human participants, response strength.

**Effects of Including a Separate Conditional Discrimination Training or Not before Implementing a Titrating Delayed Matching-to-Sample Procedure**

Matching-to-Sample (MTS) procedures are used to study both learning and remembering, and also to establish necessary conditional discriminations to test for responding in accordance with stimulus equivalence. Learning and remembering are descriptions of different behavioral patterns in a functional relation with the environment. Catania (2007) refers to learning as a process of change in behavior that will maintain over time, while remembering is behavior emitted in the absence of the controlling stimulus (Palmer, 1991).

When describing the training and testing arrangement in an MTS procedure the nomenclature used is letters and numbers like A1B1C1. Letters refer to sets of stimuli, while numbers refer to classes. These nomenclatures represent arbitrary signs, words, pictures, sound and so on, in which can elicit the same bodily reactions, or evoke the same operant behavior (Sidman, 1994). For example, if we are going to establish three 3-member classes (C←A→B) in an MTS format, a choice response to the comparison B1 and not comparisons B2 and B3 produce reinforcement only if a specific sample, A1, are presented. If sample A2 is presented selecting comparison B2 is correct and not B1and B3. If sample A3 is presented selecting B3 is correct and not B1 and B2. Likewise, when training AC relations. Following establishment of AB and AC relations tests for equivalence class formations can be introduced, were the untrained relations BA, CA, BC and CB may emerge. To sum up, if relations between stimuli in a class that is not directly trained to each other become related, they have to be observed in test for equivalent relations after establishing baseline relations. Murray Sidman and colleagues demonstrated the occurrence of relations between not directly trained stimuli, after using MTS procedure with people with developmentally disabilities (Sidman, 1971; Sidman & Cresson; 1973, Sidman, Cresson, & Willison-Morris, 1974;

Sidman & Tailby; 1982). Later, several studies have been done with MTS procedure to investigate conceptualization, emergent responses and remembering.

After establishing baseline relations in, for example, three 3-member classes, responding in according to equivalence relations are tested. Equivalence responding must fulfill the requirements of (1) reflexivity, (2) symmetry and (3) transitivity (Sidman & Tailby, 1982). (1) Reflexivity means that each stimulus must have a conditional relation to itself (A1A1, B1B1, etc. matching). (2) The relations must be symmetrical where the stimuli must be interchangeable. After learning the AB and AC relations, testing BA and CA relation will confirm symmetrical relation. (3) Transitivity can only be tested directly by using a linear training structure (LS), training AB and BC, and test for AC. Testing CA is symmetrical transitivity relation and is called global equivalence test. In other training structures, training BA and CA in a Many-to-One (MTO) procedure, and AB and AC and in a One-to-Many (OTM) procedure, there are no direct transitivity tests. Therefore, in these training structures, testing for BC and CB is usually called combined transitivity and equivalence test.

Training structure is an important determinant for the result of both numbers of trials to criterion and accuracy of responses during testing. Some studies have shown best result using MTO and OTM structure, while LS is the structure that generates fewest emergent relations (Arntzen & Hansen, 2011; Arntzen, Grondahl, & Eilifsen, 2010). OTM has shown to be slightly more effective with better results in tests, and fewer numbers of training trials to establish the conditional discriminations than MTO using three 3-member classes (Arntzen; 2006, Arntzen & Hansen; 2011; Arntzen & Holth, 1997; 2000; Arntzen & Nikolaisen; 2011; Fields, Hobbie-Reeve, Adams, & Reeve, 1999; Arntzen, Grondahl, & Eilifsen, 2010), and fewer trials in training with two 7-members classes (Fields et al. 1999). In the present experiment, OTM training structure is used because it benefits responding in according to stimulus equivalence.

Emergence of equivalent relations does not always occur in the beginning of testing, but can be acquired through repeated testing, even without any programmed consequences (Fields, Arntzen, Nartey, & Eilifsen, 2012; Saunders & Green, 1992; Sidman, Kirk, Willson-Morris, 1985; Travis, Fields, & Arntzen, 2014; Tyndall, Roche & James, 2004). This is called a delayed emergence (e.g., Sidman, 1994) and suggests that training necessary baseline relations cannot alone generate equivalence class formation, but can only serve as a prerequisite for the emergence of the new derived relations as symmetry and combined transitivity and equivalence. In the following study, the effect of maintenance and delayed emergence is examined by running two test blocks without any baseline training between the tests.

In most research, a simultaneous matching-to-sample task (SMTS) is conducted where the sample and comparisons are available at the same time when selection responses are required. However, in a delayed matching-to-sample (DMTS) procedure, the sample disappears prior to the presentation of the comparisons. This has to be referred to as a case of remembering. Thus, the participants have to discriminate and emit a correct comparison selection when there is a delay between the sample offset and the onset of the comparisons. This delay is an independent variable that can influence the response rate and reaction time, and, therefore, gives an opportunity to investigate variables that influence remembering (Palmer, 1991).

DMTS procedures are used to determine how long a stimulus can exert control of a response in both nonhumans (Blough, 1959; Kangas, Berry & Branch, 2011; Kangas, Vaidya, & Branch, 2010), and in humans (Arntzen; 2006; Eilifsen, & Arntzen, 2011; Lian & Arntzen, 2011; Sidman, 1969; Vaydia & Smith, 2006). Typical outcome is that sample stimulus may lose its conditional control as the delay increase (Constantine & Sidman, 1975; Sidman, 1969) and an increase in training trials to reach the baseline criterion with increasing delay in pigeons (Kangas, Berry & Branch, 2011; Sargisson & White 2001; 2007) and with human (Arntzen, Steingrimsdottir, & Antonsen-Brogård, 2013). Furthermore, responding in according to stimulus equivalence can also be enhanced by using a delay between sample and comparison, and result in a lower number of trials to establish conditional discriminations (Arntzen, Nartey, & Fields; 2015, Saunders, Chaney & Marquis; 2005). Arntzen (2006), Arntzen, Galaen and Halvorsen (2007), Vaydia and Smith (2006), and others have shown high accuracy in test conditions using relative long delays.

The delay can be fixed from 0 s too *n*s where the delay stays unchanged throughout all training trials. Arntzen (2006) and Vaydia and Smith (2006) studied the effect of different fixed delay. The main findings of the three first experiment in Arntzen (2006), using SMTS, 0 s, 2 s and 4 s DMTS, is that starting with longer delay gave more correct responses in testing for emergent relations in MTO and OTM training structure. Furthermore, 0 s after SMTS gave a higher number of training trials than both SMTS and longer delay. In Experiment 1, starting with SMTS, then 0 s, 2 s and 4 s DMTS, gave less correct responding in test after SMTS but correct responding increased as the delay increased up to 4 s, and 0 s delay required the highest number in training trials. When starting with 4 s delay, the participants had more correct responding in test than staring with SMTS, and the training trials decreased as a function of decreasing the delay. In Experiment 2, 0 s delay resulted in more training trials in OTM training structure, than SMTS, 2 s and 4 s delay when starting with SMTS. All participants responded in accordance with stimulus equivalence. In Experiment 3, the delay increased from SMTS, 0 s, 3 s and 9 s delay. The results replicated the findings in Experiment 2 were 0s produced more training trials, and all participants responded in accordance with stimulus equivalence. Arntzen, Galaen and Halvorsen (2007), investigated the effects of increasing or decreasing fixed delay (0 s, 6 s, and 12 s), one group started with 0 s, and another started with 12 s. The results showed no difference between the groups in responding in the test conditions. The study from Vaidya and Smith (2006) showed that exposure to longer delay in training trials in the conditional discriminations, increased the likelihood of responding correct in test for symmetry when using 0s, 2s and 8s delay, but no significant differences in number of training blocks to establish the conditional discriminations between the groups at the different delays.

Another variant of DMTS is a titrating DMTS (TDMTS) procedure where the delays increase as a result of correct responding or decrease as a result of incorrect responding (e.g., Cumming & Barryman, 1965). This way, the participants will not necessary reach a stable responding at one titration step before increasing the delay. As far as the author knows there are not that many published studies with the TDMTS procedure (Arntzen et al., 2013; Arntzen & Steingrimsdottir, 2014; Ferraro, Francis, & Perkins, 1971; Eilifsen, & Arntzen, 2011; Lian & Arntzen, 2011) and adjusting delay (Rosenberger, Mohr, Stoddard, & Sidman, 1968). Ferraro et al. (1971) examined the effects of different delays in a TDMTS procedure in children with delays from 0 ms–12000 ms. The results showed increased accuracy and delay as a function of increased age. Eilifsen and Arntzen (2011) studied the effect of using different starting points in TDMTS procedures. The result of starting with 0 ms delay up to 3000 ms is for some participant less effective in establish both the necessary prerequisite for, and the formation of stimulus equivalence classes than start at 5000 ms up to 8000 ms. These result supports earlier finding from Arntzen (2006) that using higher fixed delay in baseline training can give a better outcome in test for emergent relations.

TDMTS training procedure can be used to investigate whether a gradual increase of the delay would enhance the effect of DMTS procedure. In the study by Arntzen et al. (2013), one man with dementia emitted correct responding at higher delay in a TDMTS procedure compared to DMTS with a fixed delay. Furthermore, Arntzen and Steingrimsdottir (2014) found more stable response patterns with titration step of 100 ms compared to 500 ms. These results showed that number of correct responses change as a function of titration steps. Furthermore, Lian and Arntzen (2011) investigated the differences in using fixed and titrated DMTS procedure and found no significant differences between using 3 and 6 second fixed or titrated delays in children. The main findings were arranging experiments with different delays seem to be relevant in studying remembering and especially short-term memory. Hence, to see how long the sample continues to have control to comparison selection, and to determine what can help to continue such control of a sample stimulus. One of the purposes of using TDMTS, is to see if gradually increasing the delay, can improve the establishment of conditional discriminations.

A research question which is not answered yet, is the importance of starting the training by establish the baseline relations before introducing titration. One of the purposes of the present experiment was to investigate if baseline training before titrating could make more stable responding in the titrations steps, and if there was any difference in the number of training trials and equivalence outcome by starting training the baseline relations with SMTS, 0 s DMTS or without baseline training.

Dymond and Rehfeldt (2001) suggest different additional measurements for research on derived relations. One such measurement is reaction time, the time from the comparison onset to a selection response is emitted. Results from studies by Bentall, Dickins and Fox (1993) and by Arntzen and Lian (2010), shows an increase in reaction time when using arbitrary stimuli rather than with namable pictograms or pictures. The increase of reaction time from baseline training to equivalence test is demonstrated in several studies (Arntzen, 2010; Arntzen & Hansen, 2011; Arntzen & Holth, 1997, 2000; Arntzen & Lian; 2010; Eilifsen & Arntzen; 2009, Holth & Arntzen, 1998; 2000) and an increase in reaction time from test of symmetry to combined transitivity and equivalence (Arntzen, Galaen, & Halvorsen, 2007). These findings may indicate that there are some mediating behaviors going on. However, to measure the mean reaction time may be problematic because the reaction time may vary for the same participants during the same experimental conditions (Whelan, 2008). In that case, the mean of median reaction time from each individual is investigated in this study, to determine the change of reaction time from baseline training to the different conditions of equivalent relations in test, and during the test condition.

The purpose of this experiment was to investigate effects of titrated delayed matching-to-sample (TDMTS) with and without baseline training in a SMTS or a 0 s DMTS procedure before testing for emergent relation of symmetry and equivalence (combined transitivity and equivalence) using three 3-member classes in a OTM training structure. A group of participants had a baseline training of conditional discrimination (CD) with SMTS procedure and a second group with 0 s DMTS procedure before the TDMTS procedure were initiated, and a third group of participants started directly on the titrating procedure without such baseline training of the conditional discriminations. Titration steps in the TDMTS procedure are the same for all groups. The important factor is the role of the pretraining of CD. The present experiment investigates the differences between groups with the number of trials to master criteria, response pattern in the titration conditions, reaction time, and accuracy of responding in accord to the properties of stimulus equivalence.

## Method

### Participants

Total 30 experimental naive adults, 22 female and 8 male between 18 and 45 years old were allocated into three groups. The average time spent on the experiment was 1 hour and 55minutes. No participants were trained in matching performance prior to the experimental conditions.

When participants were recruited, they were informed that the experiment was not paid, and it came just to conduct an experiment for research purposes to give a better insight into what influences how we learn and not an IQ or personality test. Participants were also informed that it was one experiment session only, and that it will last about 2.5 to 3 hours. Moreover, they were informed they could withdraw from the experimental session at any time. In addition, that it is anonymous, and finally, that they did not need any prior knowledge about computers to carry out the experiment. When the participants came to the experimental session, they signed a consent form. They were here informed that it was an experiment about stimulus equivalence and why it is important. In this form they were presented with the word stimulus equivalence for the first time and they were asked if they knew anything about it to ensure that they were naive. If they needed a break, they had to contact the experimenter. Furthermore, that they could not get more information in advance, but would be debriefed after the experimental sessions. For any questions during the experiment, standard answers as "I can’t answer," "keep going" and "do your best" were given.

**Groups and Design**

One group started with SMTS (SBT = simultaneous baseline training group) before TDMTS procedure, a second group started with a 0 s DMTS procedure (DBT = delayed baseline training group) before TDMTS procedure, and the third group started direct on the TDMTS procedure (Without group). Ten participants (15000-15010) between 19 and 41 year, 4 male and 6 female was allocated to the SBT group, 10 participants (15011-15020) between 18 and 45 year, 2 male and 8 female in DBT group, and 10 participants (15021-15032) between 21 and 40 year, 2 male and 8 female in the Without group. Two participants in the Without group who did not finish the experiment are excluded and replaced by two others. One participant resigned during the experiment, and for the other participant the software program stopped during the session.

**Stimuli, Setting and Apparatus**

The stimuli used in this experiment were 9 visual abstract stimuli, in three 3-members classes (see Figure 1). All relations between the stimuli were initially arbitrary for all participants. The same stimulus set was used for all three groups (SBT, DBT and Without). Prior to the experiment, the participants were asked to presort the stimuli used in the experiment. This was to ensure that they had no knowledge of the stimuli sets before starting. They were also asked if they knew anything about the experiment to catch up if they had received explanations of other participants in advance.

All participants completed the experiment in a quiet room. In the room there was a desk with a computer, a mouse, and an office chair. The desk was located towards a blank wall. Some other furniture in the room was placed behind the participants.

The experiments were carried out with matching-to-sample software on a HP EliteBook 8730w with Windows 7 professional 32-bit operating system and the following specification; Intel Core ™ 2 Duo CPU 2,67GHz with 4GB of RAM, screen 17 ".

**Procedure**

**Instructions**. When the participants sat in front of the computer, they pressed the start button with a mouse click. The following instruction was presented on the screen prior to the first training and test block:

> A stimulus will appear in the middle of the screen. Clicking on the sign will make it disappear and three other signs will appear in the corners of the screen. Choose one of these by clicking it. If you choose the one we have defined as correct, the word "correct," "good," "well done" and the like will be displayed on the screen. If you choose an incorrect one, "incorrect" will be displayed on the screen. During some stages of the experiment, the computer will not tell whether your choices are correct or incorrect. But from what you have learned, you can get all the tasks correctly.

> Please do your best to get everything right. Good luck. Press start to begin the experiment.

The following instruction was presented on the screen after pressing the start button prior to the second test block:

> A stimulus will appear in the middle of the screen. Clicking on the sign will make it disappear and three other signs will appear in the corners of the screen. Choose one of these by clicking it. You will not receive any feedback on whether your choices are correct or incorrect but from what you have learned, you choose the stimulus you mean is the correct one. Please do your best to get everything right. Good luck. Press start to begin the experiment.

The first instruction was read out loud to the participants to ensure that the information was known. The experimenter was in the room for the first two trials to ensure that the participant followed the instruction. If they did not, they were told to click on the stimulus. Questions about the experiment were answered by referring to the instruction text.

**Structures and parameters.** All participants in the three groups were trained to establish 6 conditional discriminations, AB and AC relations (**A1**B1B2B3, **A2**B2B1B3, **A3**B3B1B2, **A1**C1C2C3, **A2**C2C1C3 and **A3**C3C1C2). The training was arranged concurrently in a simultaneous protocol with a One-To-Many (OTM) training structure. There were used 3 classes with 3 members in each class (see Figure 1) in all conditions (see Table 1).

The sample was presented in the center of the screen and an observing response, a mouse click, to the sample, was followed by presentation of the comparisons randomly located in different corners on the screen, always leaving a blank corner. After a choice response to comparison was made, there were 1500 ms inter trial interval (ITI) including 1000 ms with programmed consequences and blanking of the screen for 500 ms. The consequence of making a correct choice was a presentation of the words “great”, “nice”, “excellent” and so on. If the choice was incorrect the word “wrong” was displayed. After making a choice, and the programmed consequences, the mouse was reset to the middle of the screen. There was always a break for about 2-5 minutes after baseline training conditions of SMTS or DMTS before starting the TDMTS procedure, and between Test 1 and Test 2. Regarding the training and test phases, see Table 1. These conditions and phases are described below.

**Training phases.** The SBT group started the AB and AC conditioned discrimination training with a SMTS procedure. All 6 relations of AB and AC were trained 5 times within one training block . After an observing response to the sample, the comparisons came up and all four stimuli were present on the screen until a choice response was made. After 90% correct responding in one training block, the consequences were gradually thinned to a probability of 75%, 50% and 0% for the programmed consequences. Each block consisted of 30 trials with master criteria of 90%. If a block was failed, the same block was presented until the mastery criterion was reach.

The DBT group completed a 0-s DMTS in all training blocks prior to the TDMTS procedure. An observing response to the sample made it disappear, and the comparisons were onset at once. Otherwise, it was similar to the procedure for the SBT group.

The Without group started directly on the titrating procedure without having baseline training of the conditional discriminations on beforehand.

***TDMTS procedure.*** The participants started with 0-s delay, increasing with 1000 ms after responding 90% (6 out of 6) correct in each block. If one or more incorrect response was made, the delay decreased with 1000 ms delay, or continued on 0-s delay. The titrating steps were arranged from 0-s and up to 12000 ms. After having reached the criterion of 6 correct responses with 12000ms delay, thinning of consequences started. The thinning phases were arranged 1200 ms delay in 30 trials blocks with 100%, 75%, 25%, and 0% probability of programmed consequences. Master criterion was 95% correct. If the participants did not reach the criteria, the same training block was repeated.

***Testing phases for emergent relations.*** After mastering 95% in a block with 0% probability of programmed consequences**,** baseline relations (A1B1, A1C1, A2B2, A2C2, A3B3, A3C3, A3C3), relations of symmetry (B1A1, C1A1, B2A2, C2A2, B3A3, C3A3) and combined transitivity and equivalence (B1C1, B2C2, B3C3, C1B1, C2B2, C3B3) were tested in two mixed blocks of 90 trials. All test trials were conducted with a 0-s DMTS without any programmed consequences. As in training, the participants had to make an observing response to the sample. All participants were exposed to two test blocks to study maintenance or delayed emergence. The second test block started 2-5 minutes after finishing the first. The master criterion for responding in accordance with the relations of equivalence responding was 95%.

**Validity and data analysis**. The matching-to-sample program used, automatically recorded all the data of correct and incorrect choice responses to sample and comparison in training and test conditions. Furthermore, the reaction time from the onset of the comparison to a choice response was made, was also recorded. The data were then visually inspected by the experimenter to compare the results between the different groups.

Likewise, different statistical analyses are conducted to compare the effect of trials to master criterion, response pattern, and reaction time. To compare the mean accuracy of responding during training of the conditional discrimination between the different conditions (SBT and Without, DBT and Without, and SBT and DBT), an Independent Sample t-Test was conducted. Regarding the response pattern, a Kruskal-Wallis test was conducted to find out if there were any different in the mean rank among the three groups. Furthermore, a

Paired Sample t-Test compared the means of median reaction time from training trials to the test trials and the change in reaction time during Test 1.

## Results

The purpose of this experiment was to determine the effect of starting with conditioning discrimination with SMTS or 0 s DMTS before TDMTS, and start direct to TDMTS. All participants completed baseline training and both test phases.

### Trials to criterion

Table 2 shows the individual scores of training trials to establish the conditional discriminations at a level of 12000 ms without the training blocks with thinning of consequences. When counting the number of trials to criterion, 30 trials are subtracted from the SBT and DBT group because of the differences in the procedure from starting right on the TDMTS training (see Table 1). An independent sample test was conducted to compare the number of trials to master criteria between the groups. We cannot find any significant differences between starting with conditional training prior titration, or start direct on the TDMTS condition. Looking at the individual scores, two participants showed extreme scores above the mean level for their respective groups. In the SBT group, participant 15001 had the highest number from the mean with 1088 more trials, and the lowest number was for participant 15004 and 15009 with 208 trials below the mean. In the DBT group, it was 181 trials above the mean for participant 15014 and 101 trials below for participant 15012 and 15016. In the WITHOUT group, participant 15022 had the highest number from the mean with 1004 above, and the lowest was 238 trials below the mean for participant 15021.

### Emergent Responding and Delayed Emergent Responding

Two tests were conducted to determine whether there was a change in responding from Test 1 to Test 2. When baseline relations were successfully established, the participants was presented for test for emergent relations for symmetry and equivalence (combined transitivity and equivalence) intermixed with test for the direct trained baseline relations. As shown in Table 2, there were no significant differences between the groups. Totally seventeen participant had all responses correct in the tests, 6 in the SBT group, 6 in the DBT group, and 5 participants in the WITHOUT group. Five Participants had 3 or more wrong responses to comparison in Test 1 and did not meet the criteria of 95% correct. Ten participants in the SBT group and 8 participants in both DBT and Without group reached the criterion in Test 2. Three participants in SBT group, 3 participants in DBT group, and 2 participants in the Without group, did better on the second test. When we look at the individual scores in Table 2, two participants (15013 and 15022) did not maintain the score from Test 1 to Test 2. Participant 15022 had lower score on equivalence relation in Test 2. One participant (15013) had lower scores on baseline relation in Test 2, but maintained the results for symmetry and equivalence. Participant 15006, 15020 and 15030 did not respond in according to equivalence in Test 1, but showed delayed emergence of the untrained relations, in Test 2. Table 3 shows the response matrix for 15024 and 15030. Participant 15024 had no correct responding in Test 1 for the B1C1 relation, but 4 out of 5 correct responses in the following second test. Participant 15030 showed less consistent responding for equivalent relations in Test 1 than in the second test. Only one participant (15024) did not reach the goal in either tests, but performed better on Test 2.

**Response Pattern**

Another research question was how the groups responded in the different delays during the TDMTS procedure, and how stable the responding was as the delay increased. The delay increased with 1000 ms after 6 out of 6 correct responses in a titration block, and decreased with 1000 ms if one or more responses were incorrect. A Kruskal-Wallis test was conducted to score the significance level among the three groups. The scores showed that there are more likely to have a more stable responding through the titration steps after having a conditional discrimination training before starting with the TDMTS procedure ($H=7.155$, $p<.028$) with a mean rank of 12.65 for the SBT group, 12.45 for the DBT group and 21.40 for the Without group.

**Reaction Time**

Reaction time is defined as time from the onset of the comparison to the choice response to a comparison is made. Data selected is the mean score across all participants, regardless which group and how they formed equivalent classes. The data collected to run a Paired Sample t-Test, was the mean of the median reaction time on the last 5 baseline trials, and in tests the first and last 5 baseline relations, symmetry and equivalence for all participants across groups in Test 1 (see Figure 2). When conducting a Paired Sample t-Test (see Table 5) there are significant chances in reaction time between the last 5 baseline training trials and the first 5 equivalence test trials (BLT – EQ1), first 5 and the last 5 in baseline relations during testing (BL1 – BL2). There are also significant changes from the first 5 to the last 5 testing trials for the baseline, symmetry and equivalence relations. The greatest difference in reaction time was between the first 5 (which had the highest reaction time) and the last 5 test trials of the equivalence relations. When adding the first 5 and last 5 test trials in test for baseline, symmetry and equivalence, there are significant differences in reaction time between test for symmetry and equivalence, and between baseline relations and equivalence relations.

## Discussion

The main purpose of the current experiment was to investigate the effects of conducting a conditional discrimination training with SMTS procedure (the SBT group) or 0 s DMTS procedure (the DBT group) prior a TDMTS procedure, or conducting the TDMTS procedure only before testing for properties of equivalent relations. Furthermore, the purpose to include two separate test blocks was to observe the immediate, delayed and maintenance effect of the emergent relations. The reaction time was recorded as a supplement measurement for the response rate to describe the response strength. There were no significant differences among the groups in response rate during training or in test conditions. However, there were significant reductions in reaction time from training trials to the equivalent relations.

**Number of Trials in Training**

In the study by Lian and Arntzen (2011) and Eilifsen and Arntzen (2011), the participants started the TDMS procedure directly without any separate conditional discrimination. The current experiment shows no significant differences among the groups in the number of trials to master criterion at 12000 ms delay. Hence, there are no significant differences between the SBT group and the DBT group, and it differs from the results in both Arntzen (2006) where 0 s delays required more trials than SMTS, and Saunders, Chaney, and Marquis (2005) where 0 s delays gave fewer trials in training than SMTS.

Ten participants in each group are a quite small number in a between-group design, and can then weaken the result of not showing significant differences. It could be that a larger group may have shown differences in the average number of trials to criterion. Therefore, further research should arrange the same conditions with and without baseline training with a larger number of participants or using different stimulus set for each condition in a within-subject design.

**Stimulus Control and Accuracy in Training Trials**

The accuracy of responding may be depended of how the participants attend to the stimulus presentations. Low accuracy in training trials in the beginning of the experiment may be described by the stimulus control topography (SCT) coherence theory, were other than the relevant experimenter SCT was controlling the selection responses (McIlvane, & Dube, 1992; 2003; Stoddard & Sidman, 1971). One possibility is that the participants select the stimulus in the same corner as the last reinforced trial, such that the position could control the selection responses. However, in the present experiment, this could not be the case since the comparisons appeared randomly in one of the four corners on the screen. Another possible explanation is that comparison selection was based on the previously reinforced comparison regardless of the sample. Obviously, the increased accuracy may be a result of differential reinforcement in where selection of the experimenter-defined incorrect comparison resulted in less reinforcement than selecting the experimenter-defined correct comparison. Hence, a future experiment should include pretraining of matching performance since it will enhance the effect of established the conditional discriminations. Finally, longer inter trial interval (ITI) than shorter can give a greater degree of stimulus control in the way that the participants are more attentive to the sample, and do not intermix the sample with the previous stimuli presented on the screen (Williams, Johnston & Saunders, 2006). Therefore, 1500 ms as used in the present experiment gave relative long ITI that can have enhanced better conditional stimulus control.

It could be argued that SMTS may have resulted in fewer trials to criterion because of the increased probability of sample control, and that this control will decline at longer delay (Sidman, 1969). In SMTS procedure, participants can look back and forth between the sample and comparison. Whether this is a case, may be studied by using eye-tracking technology (Palmer, 2010). However, as the delay above 0 s has been favored, there can be argued that the absence of the sample evokes some kind of supplemental controlling stimuli such as rehearsal, echoic or intraverbal responding, in the delay that controls the choice of comparison, or facilitates the remembering of the absent sample (e.g., Blough, 1959; Torgrud & Holborn, 1989). The use of a distractor task in the delay prior comparison onset can result in fewer correct responses (e.g. Arntzen, 2006; Arntzen &Vie, 2013). This can strengthen the assumption that there is some mediating behavior that seems to control the selection of the comparison, and that distraction during the delay may overshadow this kind of behavior. Even though this is not investigated in the present study, it is possible that some kind of supplemental controlling stimuli maintained the accuracy as the delay increased, and still maintained at the delay of 12000 ms. A future experiment with and without a comparison distractor task during the delay can give some better insight into this assumption.

**Response Pattern in Different Titration Steps**

Conducting baseline training with SMTS or DMTS prior to TDMTS, gives more stable responding during the different titrating steps (see Table 4). One argument could be that both SBT and DBT groups had stable responding before gradually increasing the delay. In the titration conditions the interval increased without a very strict criterion of stability of the conditional relations. Therefore, the stimulus control of selecting the correct comparison when the delay was above 0 s might have been weaker than in simultaneous or 0 s delayed MTS procedure. Thus, resulted in less stable responding.

Response accuracy is one measure of the degree of response strength. In TDMTS procedure, the relations of sample-comparison relations are not conditioned before the delay increase, and then the sample could loose its controlling function. The present results indicate that accuracy responding at 12 s delay is a function of high degree of stimulus-stimulus strength. This means; with high sample control, the response accuracies are high. If the relation between the sample and experimenter-defined comparison were low, the accuracy of responding would have been lower too.

In future research, a condition with 12000 ms fixed delay should be included. Such an experiment could extend the information of the effect of using TDMTS, as it has been argued that TDMTS procedure will contribute to a more precisely description of when forgetting occurs. Another alternative is to start with higher delay as in Eilifsen and Arntzen (2011) that showed better results of starting with delay above 0 s.

**Immediate, Delayed and Maintenance of Emergent Relations**

Overall, 29 out of 30 participants reached the 95% criterion of correct responding in test. These high scores can be a result of (1) using three 3-member classes, that reduce the possibility of choosing the correct comparison as a result of rejecting the wrong (Carrigan, & Sidman, 1992; Sidman, 1987), and (2) that it is easier to remember the relations, when only a small number of conditional discriminations are established during baseline training (e.g., R. R. Saunders & Green, 1999). This second argument is not investigated in the current experiment, but supported in the study by Arntzen and Hansen (2011). Further it suggest that it is highly probably that the participant had established the experimenter defined SCT coherence (McIlvane, & Dube; 1992; 2003) at 12000 ms delay, and the result may have been affected by the thinning of consequences. If the testing had started directly after a training block with 100% programmed consequences, it is possible that the test relations would be under extinction.

The high accuracy in the present experiment supports that (1) OTM training structure generates a high number of equivalent responding in adults (Arntzen, 2006; Arntzen et al., 2010; Arntzen & Holth, 1997), (2) exposure to relative long delay may enhance the responding in accordance with the properties of stimulus equivalence (Arntzen, 2006; Arntzen et al., 2007, Arntzen et al.; 2015), or symmetry (Vaydia & Smith, 2006). While for children, Lian and Arntzen (2011) showed the opposite effect. This may be a result of that adults have more experience of problem solving and generating more supplemental controlling stimuli during the experimental session. Such responses could have occurred during the delay of sample offset and the comparison onset. However, there has been demonstrated that responses as name-giving, are not necessary for new equivalent relations to occur (Sidman, 1994; Sidman, Willson-Morris, & Kirk, 1986**).** The result of testing can also depend on which training and testing protocol being used (Adams, Fields, & Verhave 1993;

Imam 2006; Imam & Warner, 2014). A simultaneous protocol was used in the present experiment with high accuracy in test, even though a study by Imam and Warner (2014) showed that this protocol was the least effective.

Arntzen and Hansen (2011) found that participants, who formed equivalence classes, acquired the baseline relation in fewer trials to criterion. This result is not replicated in the present experiment. Probably is the difference in the results because Hansen & Arntzen (2010) used SMTS. As Table 2 shows, the participants with the extreme score (15001 and 15022), and the participants with a relative high and low number of trials, all except one responded in accordance with the master criteria in test conditions. It can be argued that the large number of minimum training trials of the baseline increases the response strength during training that again results in high accuracy during test conditions.

Several studies have shown that participants that do not show immediate emergence of equivalent classes do so with continued exposure to test probes (Arntzen, 2006; Arntzen et al., 2015; Doran & Fields, 2012; Dube, Green, & Serna, 1993; Fields, et al., 2012; Travis et al., 2014; Tyndall, et al., 2004). Three participants in the present study showed delayed emergence of the novel stimuli relations (see Table 2). This result supports the finding in Arntzen (2006), Doran & Fields (2012), Dube et al., (1993) Fields et al., (2012), Travis et al., (2014), Tyndall, et al. (2004), that used more than one test block and seem to be important to include in future experiments.

Sidman, Willson-Morris, and Kirk (1986) argue that symmetry responding is necessary for the correct responding of combined transitivity and equivalence test. The result of participant 15024 and 15030, who both showed high score on symmetry but not on the equivalence trials, indicate that participants can respond in according to symmetry but not transitivity or equivalence (e.g. Sidman et al., 1986), and that symmetry relations can be established before the transitivity or equivalence relations (e.g. Bush, Sidman, & Rose, 1989;

Dube, et al., 1993). An assumption can be the novelty of the stimulus relations presented in the test probes for transitivity and equivalence (BC and CB). For example, in a OTM training structure, the B and C stimuli have never been presented together before.

The present study found both improvements and high degree of maintenance from Test 1 to Test 2 (see Table 2). Participant 15006 improved the symmetry responding, and participant 15020 improved both symmetry and equivalence responding. Participant 15030 did not reach the master criterion in Test 1, but did so in Test 2. Participant 15024 improved in both symmetry and equivalence responding, but did two errors on equivalence trials and therefore do not meet the criterion of responding in accordance with equivalence. Many other studies have used criteria of 90%, and only three trials of each trial type (e.g., Arntzen, 2004; Lian & Arntzen, 2011). If a criterion of 90% had been employed in this study, all participants would have responded in accordance with stimulus equivalence in one of the test blocks.

In the present study, the use of many test trials and more than one test block, gave the possibility of studying the effect of both gradual establishments of the untrained relations and maintenance of responding. The delayed emergence supports the assumption that the conditional discrimination training only serves as a prerequisite for emergent relations, and that equivalent classes do not exist before they are observed in test conditions (Sidman, 1994).

**Reaction Time**

Behavior analysis will have some benefit of measurement of other variables as for example reaction time and not only response accuracy to measure the response strength (Dymond & Rehfeldt, 2001; Palmer, 2010). This will give better information about complex behavior such as, for example, remembering. Reaction time may function as an indirect measure of covert behavior as precurrent behavior that is difficult to observe. Long reaction time will in the same way as low accuracy refers to a lower degree of response strength between the antecedent stimulus and the incorrect or absence response (forgetting). Furthermore, long reaction time will support that there is some kind of precurrent responding going on before a response is emitted. The results from Test 2 are not included because all participants had a short break between the test blocks. Comparing Test 1 and Test 2 will then not show any valid result.

The reaction time data in Test block 1 (see Table 5) shows that the participants in general had a decline from (1) baseline training (AB and AC) to equivalence test (BC and CB) and (2) responded faster to the last 5 trials in all test conditions than in the first 5 trials. The reaction time increased most from baseline to the equivalence relations (BC and CB), and the greatest decline in reaction time is observed for the equivalent relations. These results support the finding in several other studies (Arntzen & Hansen, 2011; Arntzen, 2004; Arntzen et al; 2010, Arntzen & Holt, 1997: 2000; Arntzen & Lian, 2010; Eilifsen & Arntzen; 2009, Holth & Arntzen, 1998; 2000; R. R. Saunders et al., 2005). Slower responding to comparison at the beginning of testing may support the argument that there is some precurrent behavior going on (Arntzen & Holt, 1997), and support that precurrent responding takes longer time (Horne & Lowe, 1996), than if the selection is controlled by the sample only. If so, the emergence of responding in accordance with stimulus equivalence could emerge gradually.

There was no significant increase in reaction time from the last 5 baseline trials to the first 5 baseline relations in test. Twelve seconds delay between the sample offsets, and the appearance of the comparisons may be the reason for long reaction time because of precurrent behavior—the participant had to remember the sample for 12 seconds. When exposure to test, there was only 0 s delay, making it easier to remember the sample. Anyway, the results show a decrease from the first 5 to the last 5 test for baseline trials. In fact, the reaction for first 5 baseline trials could be so high because it will be the first time the participant is exposed to novel trials, as symmetry and equivalence trials in the mixed test block.

The third experiment in Holth and Arntzen (2000) showed that more participants responded in accordance with stimulus equivalence when the time of the onset of comparison increased. In the current experiment, there was no restricted time to respond to the comparison. However, limiting the onset of the comparison to the reaction time in training would have shown less accuracy in responding to the emergent relations.

### Implications for Studies with TDMTS

Based on results in the present experiment, it seems unnecessary to conduct a baseline training prior to TDMTS. However, the number of participant is quite small. Therefore, it could be difficult to make a valid conclusion. Thus, replications are needed either with a larger number of participants in group design or single-subject design. For example, to use single-subject design with different stimulus set.

## Summary and Conclusion

In the current experiment, groups of participants conducted either SMTS or 0 s DMTS prior to TDMTS, or a TDMTS procedure only. The overall results show (1) no significant difference in number of trials in establishing the conditional discriminations at 12000 ms delay among the groups, (2) significant difference in the response pattern among the groups, (3) high accuracy of responding in test conditions for all groups, and (4) significant reduction of the reaction time from baseline training to the equivalence trials, and from the beginning to the last part of the test conditions. It has been argued that the individual accuracy can be a result of different degree of stimulus control, and regarding the high accuracy in test and the reduction of reaction time, that there could be some kind of precurrent behavior going on in the delay and prior to a selection response, and the use of delay may enhance the probability of responding in accordance to stimulus equivalence. The present study has supported the notion that fixed DMTS and TDMTS procedures are applicable in investigate the accuracy of remembering responses using arbitrary stimuli.

In future research, a condition with 12000 ms fixed delay should be included. Such an experiment could extend the information of the effect of using TDMTS, as it has been argued that TDMTS procedure will contribute to a more precisely description of when forgetting occurs. Another alternative is to start with higher delay as in Eilifsen and Arntzen (2011) that showed better results of starting with delay above 0 s.

**References**

Adams, B. J., Fields, L., & Verhave, T. (1993). Effects of test order on intersubject variability during equivalence class formation. *The Psychological Record, 43*(1), 133–152. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html.

Arntzen, E. (2006). Delayed matching to sample: Probability of responding in accord with equivalence as a function of different delays. *The Psychological Record, 56*(1), 135–167. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html.

Arntzen, E. (2004). Probability of equivalence formation: Familiar stimuli and training sequence. *The Psychological Record, 54*(2), 275–291. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html.

Arntzen, E., Galaen, T., & Halvorsen, L. R. (2007). Different retention intervals in delayed matching-to-sample: Effects of responding in accord with equivalence. *European Journal of Behavior Analysis, 8*(2), 177–191. doi: 10.1080/15021149.2007.11434281

Arntzen, E., Grondahl, T., & Eilifsen, C. (2011). The effects of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *The Psychological Record, 60*(3), 437–462. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol60/iss3/4.

Antzen, E., & Hansen, S. (2011). Training structures and the formation of equivalence classes. *European Journal of Behavior Analysis, 12*(2), 483–503. doi: 10.1080/15021149.2011.11434397

Arntzen, E., & Lian, T. (2010). Trained and derived relations with pictures versus abstract stimuli as nodes. *The Psychological Record, 60*(4), 659–678. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol60/iss4/8.

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47*, 309–320. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol47/iss2/9.

Arntzen, E., & Holth, P. (2000). Differential probabilities of equivalence outcome in individual subjects as a function of training structure. *The Psychological Record, 50*, 603–628. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html.

Arntzen, E., Nartey, R. K., & Fields, L. (2015). Enhanced equivalence class formation by the delay and relational functions of meaningful stimuli. *Journal of the Experimental analysis of behavior, 103*(3), 524–541. doi: 10.1002/jeab.152

Arntzen, E., & Nikolaisen, S. L. (2011). Establishing equivalence classes in children using familiar and abstract stimuli and many-to-one and one-to-many training structures. *European Journal of Behavior Analysis, 12*, 105–120. doi: 10.1080/15021149.2011.11434358

Arntzen, E., & Steingrimsdottir, S. H. (2014). On the use of variation in delayed matching-to sample procedure in a patient with neurocognitive disorder. In iConcept Press (Ed.), *Mental disorder* (pp. 123–137): iConcept Press Ltd.

Arntzen, E., Steingrimsdottir, H. S., & Brogård-Antonsen, A. (2013). Atferdsmessige studier av demens: Effekten av ulike varianter av matching-to-sample prosedyrer. *Norsk Tidsskrift for Atferdsanalyse, 40*(1) 17–29. Retrieved from www.atferd.no.

Arntzen, E., & Vie, A. (2013). The expression of equivalence classes influenced by distractors during DMTS test trials. *European Journal of Behavior Analysis, 14*(1), 151–164. doi: 10.1080/15021149.2013.11434453

Arntzen, E., & Vaidya, M. (2008). The effect of baseline training structure on equivalence class formation in children. *Experimental Analysis of Human Behavior Bulletin, 29*, 1-8. Retrieved from: http://eahb.org/bulletin/bulletin-archives.

Bentall, R. P., Dickins, D. W., & Fox, S. R. (1993). Naming and equivalence: Response latencies for emergent relations. *The Quarterly Journal of Experimental Psychology, 46*(2), 187–214. doi: 10.1080/14640749308401085

Blough, D. S. (1959). Delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior*, *2*(2), 151–160. doi: 10.1901/jeab.1959.2-151

Bush, K. M., Sidman, M., & Rose, T. d. (1989). Contextual control of emergent equivalence relations. *Journal of the experimental analysis of behavior, 51*(1), 29–45. doi: 10.1901/jeab.1989.51-29

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*(1), 183–204. doi: 10.1901/jeab.1992.58-183

Constantine, B., & Sidman, M. (1975). Role of naming in delayed matching-to-sample. *American Journal of Mental Deficiency, 79*(6), 680–689. Retrieved from http://europepmc.org/abstract/med/1146860.

Cumming, W., & Berryman, R. (1965). The complex discriminated operant: Studies of matching-to-sample and related problems. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 284–330). Stanford, CA: Stanford University Press.

Doran, E., & Fields, L. (2012). All stimuli are equal, but some are more equal than others: measuring relational preferences within an equivalence class. *Journal of the Experimental Analysis of Behavior, 98*(3), 243–256. doi: 10.1901/jeab.2012.98-243

Dube, W. V., Green, G., & Serna, R. W. (1993). Auditory successive conditional discrimination and auditory stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 59*(1), 103–114. doi: 10.1901/jeab.1993.59-103

Dymond, S., & Rehfeldt, R. A. (2001). Supplemental measures of derived stimulus relations. *Experimental Analysis of Human Behavior Bulletin, 19*, 8–12. Retrieved from http://www.eahb.org/NewSitePages/BulletinHomepage.htm.

Eilifsen, C., & Arntzen, E. (2009). On the role of trial types in tests for stimulus equivalence. *European Journal of Behavior Analysis, 10*, 187–202. doi: 10.1080/15021149.2009.11434318

Eilifsen, C., & Arntzen, E. (2011). Single-subject withdrawal designs in delayed matching-to-sample procedures. *European Journal of Behavior Analysis*, *12*(1), 157–172. doi 10.1080/15021149.2011.11434361

Ferraro, D. P., Francis, E. W., & Perkins, J. J. (1971). Titrating delayed matching to sample in children. *Developmental Psychology, 5*(3), 488–493. doi: 10.1037/h0031598

Fields, L., Arntzen, E., Nartey, R. K., & Eilifsen, C. (2012). Effects of a Meaningful, a discriminative, and a meaningless stimulus on equivalence class formation. *Journal of the Experimental Analysis of Behavior, 97*(2), 163–181. doi: 10.1901/jeab.2012.97-163

Fields, L., Hobbie-Reeve, S. A., Adams, B. J., & Reeve, K. F. (1999). Effects of training directionality and class size on equivalence class formation by adults. *The Psychological Record, 49*(4), 703-724. Retrieved from http://thepsychological-record.siuc.edu.

Fields, L., Reeve, K. F., Rosen, D., Varelas, A., Adams, B. J., Belanich, J., & Hobbie, S. A. (1997). Using the simultaneous protocol to study equivalence class formation: the facilitating effects of nodal number and size of previously established equivalence classes. *Journal of the Experimental Analysis of Behavior, 67*(3), 367–389. doi: 10.1901/jeab.1997.67-367

Holth, P., & Arntzen, E. (1998). Stimulus familiarity and the delayed emergence of stimulus equivalence or consistent nonequivalence. *Psychological Record, 48*(1), 81–110. Retrieved from http://search.proquest.com.ezproxy.hioa.no/docview/212672443?accountid=26439.

Holth, P., & Arntzen, E. (2000). Reaction times and the emergence of class consistent responding: A case for precurrent responding? *The Psychological Record, 50*(2), 305-337. Retrieved from http://search.proquest.com.ezproxy.hioa.no/docview/212764237?accountid=26439.

Imam, A. A. (2006). Experimental control of nodality via equal presentations of conditional discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of the Experimental Analysis of Behavior, 85*(1), 107–124. doi: 10.1901/jeab.2006.58-04

Imam, A., & Warner, T. (2014). Test order effects in simultaneous protocols. *Learning & Behavior, 42*(1), 93–103. doi: 10.3758/s13420-013-0128-4

Kangas, B. D., Berry, M. S., & Branch, M. N. (2011). On the development and mechanics of delayed matching-to-sample performance. *Journal of the Experimental Analysis of Behavior, 95*(2), 221–236. doi: 10.1901/jeab.2011.95-221

Kangas, B. D., Vaidya, M., & Branch, M. N. (2010). Titrating-delay matching-to-sample in the pigeon. *Journal of the Experimental Analysis of Behavior, 94*(1), 69–81. doi: 10.1901/jeab.2010.94-69

Lian, T., & Arntzen, E. (2011). Training conditional discriminations with fixed and titrated delayed matching to sample in children. *European Journal of Behavior Analysis, 12*(1), 173–193. doi: 10.1080/15021149.2011.11434362

McIlvane, W. J., & Dube, W. V. (1992). Stimulus control shaping and stimulus control topographies. *The Behavior Analyst, 15*(1), 89–94. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2733401.

McIlvane, W. J., & Dube, W. V. (2003). Stimulus control topography coherence theory: Foundations and extensions. *The Behavior Analyst, 26*(2), 195–213 retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2731455/.

Palmer, D. C. (2010). Behavior Under the Microscope: Increasing the Resolution of Our Experimental Procedures. *The Behavior Analyst*, *33*(1), 37–45. Retrived from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2867504/.

Rosenberger, P. B., Mohr, J. P., Stoddard, L. T., & Sidman, M. (1968). Inter- and intramodality matching deficits in a dysphasic youth. *Archives of Neurology*, *18*(5), 549–562. doi: 10.1001/archneur.1968.00470350107010

Sargisson, R. J., & White, K. G. (2001). Generalization of delayed matching to sample following training at different delays. *Journal of the Experimental Analysis of Behavior, 75*(1), 1–14. doi: 10.1901/jeab.2001.75-1

Sargisson, R., & White, K. G. (2007). Remembering as discrimination in delayed matching to sample: Discriminability and bias. *Learning & Behavior, 35*(3), 177–183. doi: 10.3758/BF03193053

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*(4), 539–559. Retrieved from http://search.proquest.com.ezproxy.hioa.no/docview/212766566?accountid=26439.

Saunders, R. R., & Green, G. (1992). The nonequivalence of behavioral and mathematical equivalence. *Journal of the Experimental Analysis of Behavior*, *57*(2), 227–241. doi: 10.1901/jeab.1992.57-227

Saunders, R. R., & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior, 72*(1), 117–137. doi: 10.1901/jeab.1999.72-117

Sidman, M. (1969). Generalization gradients and stimulus control in delayed matching-to-sample. *Journal of the Experimental Analysis of Behavior*, *12*(5), 745–757. doi: 10.1901/jeab.1969.12-745

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech & Hearing Research, 14*(1), 5–13. doi: 10.1044/jshr.1401.05

Sidman, M. (1987). Two choices are not enough. *Behavior Analysis, 22*(1), 11–18.

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

Sidman, M., & Cresson Jr, O. (1973). Reading and crossmodal transfer of stimulus equivalences in severe retardation. *American Journal of Mental Deficiency, 77*(5), 515–523.

Sidman, M., Cresson, O., & Willson-Morris, M. (1974). Acquisition of matching to sample via mediated transfer1. *Journal of the Experimental Analysis of Behavior, 22*(2), 261–273. doi: 10.1901/jeab.1974.22-261

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior, 43*(1), 21–42. doi: 10.1901/jeab.1985.43-21

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*(1), 5–22. doi: 10.1901/jeab.1982.37-5

Sidman, M., Willson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities, 6*(1–2), 1–19. doi: 10.1016/0270-4684(86)90003-0

Sidman, M., Willson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities, 6*(1), 1–19. doi: 10.1016/0270-4684(86)90003-0

Stoddard, L. T., & Sidman, M. (1971). The removal and restoration of stimulus control. *Journal of the Experimental Analysis of Behavior, 16*(2), 143–154. doi: 10.1901/jeab.1971.16-143

Travis, R. W., Fields, L., & Arntzen, E. (2014). Discriminative functions and over-training as class-enhancing determinants of meaningful stimuli. *Journal of the Experimental Analysis of Behavior, 102*(1), 47–65. doi: 10.1002/jeab.91

Tyndall, I. T., Roche, B., & James, J. E. (2004). The relation between stimulus function and equivalence class formation. *Journal of the Experimental Analysis of Behavior, 81*(3), 257–266. doi: 10.1901/jeab.2004.81-257

Vaidya, M., & Smith, K. N. (2006). Delayed matching-to-sample training facilitates derived relational responding. *Experimental Analysis of Human Behavior Bulletin, 24*, 9-16. Retrieved from http://www.eahb.org/eahb-volume-2-2006/.

Whelan, R. (2010). Effective analysis of reaction time data. *The Psychological Record, 58*(3), 475–482. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol58/iss3/9.

Williams, D. C., Johnston, M. D., & Saunders, K. J. (2006). Intertrial sources of stimulus control and delayed matching-to-sample performance in humans. *Journal of the Experimental Analysis of Behavior, 86*(2), 253–267. doi: 10.1901/jeab.2006.67-01

Table 1

*Experimental Conditions and Phases for All Three Groups*

| Experimental Phases | Group | Trial Type | Condition | Master Criterion | No of Trial in Blocks |
|---|---|---|---|---|---|
| Baseline training | | | | | |
| Mixed trials | SBT<br>DBT | A1B1B2B3, A2B1B2B3, A3B1B2B3,<br>A1C1C2C3, A2C1C2C3, A3C1C2C3 | SMTS or<br>0s delay | 90% | 30 |
| Consequence thinning<br>75%, 50% & 0% | | A1B1B2B3, A2B1B2B3, A3B1B2B3,<br>A1C1C2C3, A2C1C2C3, A3C1C2C3 | SMTS or<br>0s delay | 90% | 30 |
| TDMTS | | | | | |
| Mixed trials | SBT<br>DBT<br>Without | A1B1B2B3, A2B1B2B3, A3B1B2B3,<br>A1C1C2C3, A2C1C2C3, A3C1C2C3 | Min 0ms<br>Max 12000ms<br>Step: 1000ms | 90% | 6 |
| Consequence thinning<br>100% , 75%, 50% & 0% | SBT<br>DBT<br>Without | A1B1B2B3, A2B1B2B3, A3B1B2B3,<br>A1C1C2C3, A2C1C2C3, A3C1C2C3 | 12000ms delay | 95% | 30 |
| Test 1 & Test 2 | | | | | |
| Mixed trials of: | SBT<br>DBT | | 0ms delay | 95% | 90 |
| Baseline | Without | A1B1B2B3 A1C1C2C3<br>A2B1B2B3 A2C1C2C3<br>A3B1B2B3 A3C1C2C3 | | | |
| Symmetry | | B1A1A2A3, B2A1A2A3, B3A1A2A3,<br>C1A1A2A3, C2A1A2A3, C3A1A2A3 | | | |
| Equivalence | | B1C1C2C3, B2C1C2C3, B3C1C2C3, C1B1B2B3,<br>C2B1B2B3, C3B1B2B3 | | | |

*Note*. In the third column, the first letter and number in each set describes the sample (ex. A1 in A1B1B2B3), and the underlined letters and numbers (ex. B1) describes the correct comparison given the described and defined sample.

Table 2

*Number of Trials to Criterion in Training and Correct Test Trials*

| No | Participant Gender | Age | Training Trials | Test 1 BL | Test 1 SYM | Test 1 EQ | Test 2 BL | Test 2 SYM | Test 2 EQ |
|---|---|---|---|---|---|---|---|---|---|
| | | | SBT | | | | | | |
| 15004 | M | 27 | 162 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15005 | F | 28 | 318 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15008 | M | 41 | 180 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15009 | F | 23 | 162 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15010 | F | 34 | 360 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15002 | F | 23 | 348 | 30 | 30 | 30 | 30 | 30 | 29 |
| 15001 | M | 26 | 1458 | 30 | 29 | 30 | 30 | 30 | 30 |
| 15003 | F | 24 | 288 | 30 | 29 | 30 | 30 | 30 | 30 |
| 15007 | F | 41 | 192 | 29 | 30 | 30 | 29 | 30 | 29 |
| 15006 | M | 19 | 228 | 29 | 28 | 30 | 30 | 30 | 30 |
| *Mean* | | *28* | *370* | | | | | | |
| | | | DBT | | | | | | |
| 15011 | F | 45 | 282 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15012 | F | 37 | 198 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15016 | M | 35 | 198 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15017 | F | 41 | 306 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15018 | F | 43 | 348 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15014 | F | 43 | 480 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15019 | F | 27 | 258 | 30 | 30 | 30 | 29 | 30 | 30 |
| 15015 | F | 18 | 228 | 30 | 29 | 30 | 30 | 30 | 30 |
| 15013 | M | 43 | 294 | 30 | 30 | 30 | 28 | 30 | 30 |
| 15020 | F | 35 | 402 | 30 | 27 | 27 | 30 | 30 | 29 |
| *Mean* | | *37* | *299* | | | | | | |
| | | | Without | | | | | | |
| 15021 | F | 29 | 174 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15027 | F | 25 | 216 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15028 | M | 30 | 264 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15031 | M | 33 | 192 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15032 | F | 33 | 270 | **30** | **30** | **30** | **30** | **30** | **30** |
| 15029 | F | 23 | 324 | 30 | 30 | 30 | 30 | 30 | 29 |
| 15026 | F | 29 | 270 | 30 | 30 | 29 | 30 | 30 | 29 |
| 15022 | F | 33 | 1416 | 29 | 29 | 29 | 30 | 30 | 27 |
| 15030 | F | 21 | 486 | 29 | 27 | 21 | 29 | 30 | 30 |
| 15024 | F | 40 | 504 | 30 | 29 | 24 | 30 | 30 | 28 |
| *Mean* | | *31* | *412* | | | | | | |

*Note:* Individual results in all three groups for number of trials, accuracy in responding in according to stimulus equivalence, and maintenance of the trained conditional relations in test 1 and test 2. Trials = number of trials to mastery criterion at 12 seconds delay. BL = test for baseline relations (AB and AC), SYM = test for symmetry (BA and CA), EQ = test for combined transitivity and equivalence (BC and CB). F = female, M = male. In each test, the maximum score is 30 (6 relations tested 5 times for BL, SYM and EQ).

Table 3

*Response Matrix*

| Test 1 | | | | Test 2 | | | |
|---|---|---|---|---|---|---|---|
| **Responding in Test for BC Relation for Participant 15024** | | | | | | | |
| EQ | C1 | C2 | C3 | EQ | C1 | C2 | C3 |
| B1 | 0 | | 5 | B1 | 4 | | 1 |
| B2 | | 5 | | B2 | | 5 | |
| B3 | | | 5 | B3 | | | 5 |
| **Responding in Test for Symmetry and Equivalence for Participant 1530** | | | | | | | |
| Symmetry | A1 | A2 | A3 | Symmetry | A1 | A2 | A3 |
| B1 | 5 | | | B1 | 5 | | |
| B2 | | 5 | | B2 | | 5 | |
| B3 | 1 | | 4 | B3 | | | 5 |
| Symmetry | A1 | A2 | A3 | Symmetry | A1 | A2 | A3 |
| C1 | 5 | | | C1 | 5 | | |
| C2 | 1 | 3 | 1 | C2 | | 5 | |
| C3 | | | 5 | C3 | | | 5 |
| EQ | C1 | C2 | C3 | EQ | C1 | C2 | C3 |
| B1 | 4 | | 1 | B1 | 5 | | |
| B2 | | 4 | 1 | B2 | | 5 | |
| B3 | 1 | 1 | 3 | B3 | | | 5 |
| EQ | B1 | B2 | B3 | EQ | B1 | B2 | B3 |
| C1 | 3 | 1 | 1 | C1 | 5 | | |
| C2 | | 4 | 1 | C2 | | 5 | |
| C3 | 1 | 1 | 3 | C3 | | | 5 |

*Note:* Shows number of responses to comparison. The numbers in the boxes are the number of correct responses. The number outside the boxes shows the number of incorrect responses. The vertical letters and numbers are the samples, and horizontal letters and numbers are the comparisons.

Table 4

*Number of Times the Delay was Reduced in the TDMTS Procedure*

| Group | SBT | | DBT | | WITHOUT | |
|---|---|---|---|---|---|---|
| | Participant | No | Participant | No | Participant | No |
| | 15001 | 0 | 15011 | 2 | 15021 | 2 |
| | 15002 | 0 | 15012 | 0 | 15022 | 18 |
| | 15003 | 0 | 15013 | 3 | 15024 | 5 |
| | 15004 | 2 | 15014 | 1 | 15026 | 2 |
| | 15005 | 0 | 15015 | 0 | 15027 | 1 |
| | 15006 | 5 | 15016 | 0 | 15028 | 2 |
| | 15007 | 1 | 15017 | 1 | 15029 | 2 |
| | 15008 | 1 | 15018 | 0 | 15030 | 26 |
| | 15009 | 2 | 15019 | 0 | 15031 | 2 |
| | 15010 | 1 | 15020 | 7 | 15032 | 1 |
| *Mean* | | *1,2* | | *1,4* | | *6,1* |
| *Median* | | *1* | | *0,5* | | *2* |

*Note:* Number of times the delay was reduced from the respective participant in the three groups.

Table 5

*Statistical Significance in Reaction Time in Test 1*

| | | | 95% CI | | | | |
|---|---|---|---|---|---|---|---|
| Pair | Mean | St.d. | CL | CU | t | df | Sig. (2-tailed) |
| **BLT – EQ1** | -891.87 | 1619.17 | -1496.47 | -287.26 | -3.017 | 29 | **.005** |
| **BL1 – BL2** | 512.60 | 823.51 | 205.10 | 820.10 | 3.409 | 29 | **.002** |
| **SYM1 – SYM2** | 801.07 | 1504.95 | 239.11 | 1363.03 | 2.915 | 29 | **.007** |
| **EQ1 – EQ2** | 1140.37 | 1498.71 | 580.74 | 1699.99 | 4.168 | 29 | **.000** |
| **SYM3 – EQ3** | -534.12 | 1174.65 | -837.56 | -230.67 | -3.522 | 59 | **.001** |
| **BL3 – EQ3** | -538.98 | 1281.01 | -869.90 | -208.06 | -3.259 | 59 | **.002** |

*Note:* The table shows the significant differences in reaction time in Test 1 between the different trial types. BLT = Baseline training, BL = Baseline test, SYM = Symmetry test, EQ = Equivalence test. The number behind the letters; 1 = the first 5 in the test block, 2 = the last 5 in the test block, 3 = the first 5 and the last 5 added together.

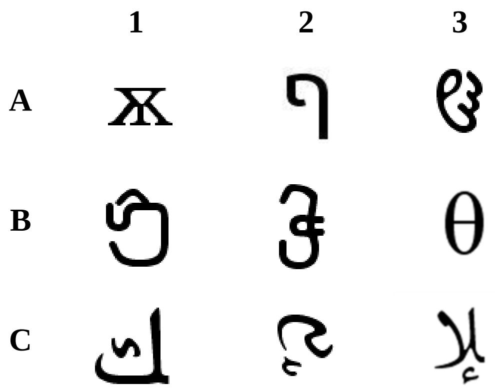


*Figure1:* Stimulus set used in the experiment for all three groups in all training and testing conditions. The number above describes the class, and the letter to the left describes the member of each class, A1B1CI is one class, A2B2C2 and A3B3C3.

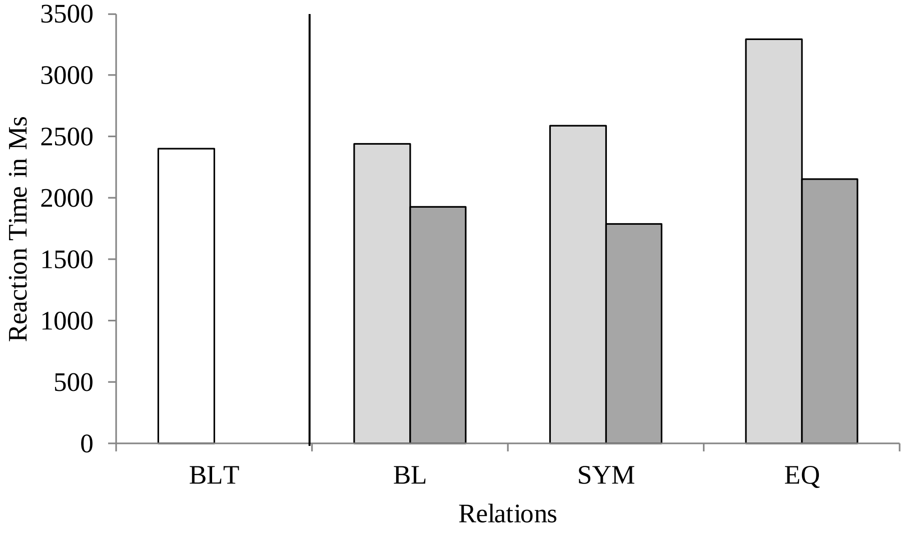


*Figure 2.* Reaction Time in Test 1: Showing means of the median reaction time for all three groups together in Test 1. The white bar = the last 5 trials in baseline training, the light grey bars = the 5 first in the test conditions, and the dark grey bars = the last 5 trials in the test conditions. BL = Baseline trials, SYM = Symmetry trials, EQ = Equivalence trials.