MASTER THESIS
Study programme MALKA
Year 2011

Title

ON THE ROLE OF CONCEPT FORMATION

Name

CAMILLA HARANGEN

Faculty of Health Sciences

Department of BEHAVIORAL SCIENCES ….

OSLO AND AKERSHUS
UNIVERSITY COLLEGE
OF APPLIED SCIENCES

Acknowledgement

I would like to thank my supervisor Erik Arntzen for his guidance and supervision throughout this experiment. Thank you to the kindergarten staff for being very helpful throughout the experiment. I am also very grateful for the support and helpful guidance from the lab group.

Abstract

Within behavior analysis the main focus has been on the functional variables affecting concept formation, which have been widely studied within stimulus equivalence. Cognitive psychology has on the other hand been more interested in describing the structure of concepts, and the functional relations have received little attention. Article 1 involved a comparison between cognitive psychology and behavior analysis on concept formation. Laurence and Margolis (1999) presented five different theories about concepts. These included the classical theory, the neoclassical theory, the theory-theory, conceptual atomism and the prototype theory. This article focused on comparing these five theories against behavior analysis and stimulus equivalence on the role of concept formation. Article 2 involved the investigation of whether children in the age of 5–6 years old would be able to form three 6-member equivalence classes by using all visual stimuli with an MTO training structure by training one relation between two trained stimulus sets, and testing for all of the relations in the end. The results from this experiment demonstrated that both participants responded according to stimulus equivalence in tests for Stimulus Set 1 and 2, but not in the tests for the three 6-member classes. However, one of the participants sorted the stimuli according to the defined classes, and showed an increase in correct responding with extended testing.

Keywords: concept formation, stimulus equivalence, cognitive psychology, sorting, expansion.

A Comparison Between Cognitive Psychology and Behavior Analysis on Concept Formation

Camilla Harangen

Oslo and Akershus University College

Department of Behavioral Sciences

Abstract

The study of concepts has been a major field within cognitive psychology. Within the field of behavior analysis, research on concept formation has been one focus within the area of stimulus equivalence. However, cognitive psychologists have given little attention to studying the variables that affect concept formation. Research on stimulus equivalence has focused on behavior in terms of stimulus and response classes, while cognitive psychology has investigated the structural aspect of concepts. There have been several different theories, and the theories have disagreed as to what should be unit of analysis. Five theories from the cognitive perspective were compared to behavior analysis and stimulus equivalence, with a special focus on the prototype theory. It was suggested that a complete analysis of concepts should involve functional relations of conceptual behavior, and that a complete account of concepts seems impossible with only studying the structure.

Keywords: concept formation, stimulus equivalence, cognitive psychology, *stimulus classes, generalization, discrimination.*

A Comparison Between Cognitive Psychology and Behavior Analysis on Concept Formation

In the late 1950's and 1960's psychology was redefined, and many psychologists started practicing cognitive psychology. Others proposed that science had more to gain from studying behavior, and its consequences and antecedents, instead of studying unobservable mental phenomena. These proponents included behavior analysts, which pointed out the weaknesses in studying mental phenomena. Some of these weaknesses involved that the cognitive psychologists said that mental processes could be measured (Overskeid, 2008).

One of the most important arguments against cognitive psychology has been that thoughts have been used to explain behavior. A problem with this is that thoughts are derived from people's behavior (Pierce & Cheney, 2008). This problem has also been termed explanatory fictions. When cognitive psychologists have tried to explain behavior with referring to unobservable mental events, this is only new behavior. This kind of explanation involves that behavior is treated both as the cause, and effect. For instance, when someone says that I am sad because I am depressed. Depressed is only an umbrella term of different behavior, and thus behavior is explained with new behavior (Skinner, 1956). The problem with this is that it is impossible to initiate measures when the behavior has not been described in terms of its relation to the environmental contingencies.

Ryle (1949) presented the term category mistakes, which have been a problem within cognitive psychology. This involves that the human body and the mind are treated as two different units, and that the mind must be studied as something apart from the human body, in terms of its own causes and effects. Ryle called this the ghost in a machine, which involves that the body and mind are two different units to analyze. He stated that the human body cannot be divided into the mind and body, the body consists of the mind, and this becomes an unnecessary doubling of the world. An example is to say that a person is sad, and crying.

Another example could include that someone says that cognitive psychology study both mental phenomena and behavior. These are both unnecessary distinctions, because mental phenomena such as private events are also behavior.

In cognitive psychology, many psychologists have been interested in the different aspects of concepts. The word concept has been defined in many different ways (Laurence & Margolis, 1999). Rosch and colleagues (e.g., Mervis & Rosch, 1981; 1973, 1978; Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) used the word categories instead of concepts. Mervis and Rosch (1981) considered a category to exist when two or more objects or events were treated equivalently. This could involve applying the same name for two or more objects or events, or performing the same action across the different objects. Even though different objects and events are unique in different stimulus situations, living organisms respond according to categorization and past learning histories. However, the definition of concepts made by Mervis and Rosch seems to be a lot different from the other theories presented in this article. During earlier research within cognitive psychology there has been little focus on the empirical issues that concerns the study of concepts.

Within behavior analysis the focus has been on the formation of concepts in terms of stimulus and response classes. Keller and Schoenfeld (1950) proposed the importance of concepts being investigated and viewed in terms of behavior. They defined a concept as when a group of objects form a class, and all the objects are reacted to in a similar way. Research on concept formation within behavior analysis has been one focus within the field of stimulus equivalence. Stimulus equivalence has involved that stimuli within a class are mutually exchangeable, and responded to equivalently (e.g., Green & Saunders, 1998).

Concepts have been widely studied within the field of cognitive psychology, while within behavior analysis the research focus has mainly been on concept formation and the study of stimulus equivalence. Behavior analytic research has been empirical and investigated the functions of concepts, while in cognitive psychology the main focus has been on the definitions of a concept, and the structure. Laurence and Margolis (1999) presents five different theories about concepts. These include the classical theory, the neoclassical theory, the theory-theory, conceptual atomism and the prototype theory. This article will focus on comparing these five theories against behavior analysis and stimulus equivalence on the role of concept formation. A special focus will be on the prototype theory.

Cognitive Psychology On Concept Formation

According to Barsalou (1992) “categorization provides the gateway between perception and cognition. After a perceptual system acquires information about an entity in the environment, the cognitive system places the entity into a category” (p. 15). The categories are representations and structures in the cognitive system that refers to what is perceived in the environment. These categories may be stored in memory, or may be combined with other representations, and so on.

A common assumption among many cognitive psychologists has been that the fundamental units of cognitive processing are the representations that are assigned during categorization, and that a form of categorization is what starts most cognitive processes. Barsalou (1992) said that most cognitive psychologists have been more interested in the ability of humans to categorize, than to investigate how it happens. Psychologists have been good to describe the sounds of speech, but have been far from able to explain how categorization occurs. The difference between concept formation and categorization is that the study of concept formation implicates the study of how concepts are formed, while categorization involves investigating how people “put” objects into the different categories or concepts. Cognitive psychologists have disagreed on how categorization occurs, and there has been no complete account of it. One of the reasons may be that they have only studied the structure of the concepts, and not been concerned with the functions of them. Concepts will tend to vary with experience, and it may be impossible to construct a complete account of how categorization takes place while only studying the structure.

Psychology in general has received a lot of criticism. This criticism has concerned their methods, goals, and achievements of psychology (Lee, 1988). In general this means the concern of what psychology is trying to do, what it is about, how are their methods for doing it, and how far have they progressed in doing it? The lack of unity has been one of the biggest problems within psychology. There are several theories and models within each subfield, and the different subfields do not agree on the subject matter of psychology.

Sidman (1960) presented the importance of science being coherent. For a science of behavior to exist, the experimental observations have to be brought into some kind of order. He said that the greatest experiments were the ones where the experimenter had no hypothesis about the results, and where the procedure was designed to answer a question of "what would happen if..". This was because this would implicate that the experimenter would not have any objections against the results. This has been a problem within other fields of psychology, where researchers have had hypotheses that they would like to verify, and if they did not get the results they hoped for, they could just change their hypotheses in order to get the results they wanted. Sidman said that the results of an experiment should always be seen as positive results, because this means that science has progressed, and maybe these results would generate new research questions. Another important factor when building a science of behavior concerns that the conceptual framework has to be coherent (see for example Baer, Wolf, & Risley, 1968).

Another problem has according to Watkins (1990) involved that it has been very easy to formulate theories involving hypothetical constructs. The reason for this, he says, has been that other researchers have only focused on their own theories and have not bothered about others, and if the theory is criticized, this critique is almost always misunderstood.

Furthermore, if the theory has to be changed, it has been sufficient if the researcher makes some small adjustments. This implicates that as long as people nourish their own theories, the theories will stand, and this is probably why psychology has had so many different theories. Theories about concepts are no different. There have been many different theories about them, and none of them agree with another on what a concept consist of. However, this article will only focus on the five theories from the book written by Margolis and Laurence (1999).

The Classical Theory

The classical theory have this definition of concepts: “most concepts are structured mental representations that encode a set of necessary and sufficient conditions for their application, if possible, in sensory or perceptual terms” (Laurence & Margolis, 1999, p. 10). The environmental contingencies are reflected through the perception of the joined sensory properties. A complex concept is composed when a learner notice how these sensory properties correlate in the environment. An object would be included as a member of a concept when it satisfies the defined features of the existing concept. Categorization takes place when something falls under the features of the concept, for instance if something has a top, four legs etc., it falls under the concept table (Laurence & Margolis, 1999).

Within the classical theory of concepts, the exemplars have shared properties that are necessary in order to be a part of that particular concept. Thus, a concept involves a class of exemplars rather than looking at each exemplar exclusively. The classical theory psychologists have paid attention to the representations rather than the processes (Smith & Medin, 1981). Categorizing a lamp would for instance comprise checking if this object has a light bulb, a button to switch on the lamp, a wire connected to an outlet, and so on. One problem with this theory is according to Palmer (2002) that for an object to be part of a concept, it has to hold the properties that are defined as necessary for the concept. By using the definition they have about concepts, they suggest that they can sort the nonexamples from the examples. To differentiate between lamp and not lamp is according to Palmer meaningless, because the boundaries between these classes are the models we have created of nature, and not the feature of nature. Objects and events will vary along a dimension, and by only using a checklist to categorize an object may lead to objects not being categorized into the correct concept.

The Neoclassical Theory

"Most concepts are structured mental representations that encode partial definitions, i.e., necessary conditions for their application" (Laurence & Margolis, 1999, p. 54). This theory resembles the classical theory, but some theorists abandoned the classical theory because of its limitations and problems. The psychologists associated with the neoclassical theory of concepts have been investigating the meaning of words, with a special interest in verbs. Jackendoff (1999) was one of the psychologists within the neoclassical theory. He suggested that when someone was categorizing a new object to for instance the concept dog, this happened by the use of a finite schema in the brain, and comparing these objects with mental representations that people had about arbitrary new objects. Jackendoff said that for example examples of causatives may be the verbs kill, give, persuade etc. and these verbs indicate that one variable will cause another one to change state. For example a concept in the category of "path" involves a relation to a destination, and a place of origin (as cited in Palmer, 2002). One of the main things that distinguish this theory from the classical theory has been that the neoclassical theory suggested that concepts consisted of partial definitions. The extensions of a concept must satisfy a set of the necessary conditions that applies for the concept's structure (Laurence & Margolis, 1999). A problem with this theory according to Palmer (2002) is that it is not possible to sort members of a concept from nonmembers which was possible with the classical theory, and some concepts have special status, such as thing, cause and event. This theory is not a clearly articulated position, and seems incomplete. This theory also, does not talk about concepts in behavioral terms.

The Theory-Theory

"Concepts are representations whose structure consists in their relations to other concepts as specified by a mental theory" (Laurence & Margolis, 1999, p. 47). The reason that psychologists associated themselves with this theory were because they wanted to distance themselves from the theory that categorization only consisted of checking different instances against particular sensory properties. These psychologists viewed cognition as being assimilated to scientific reasoning. They wanted to get philosophical treatments of theoretical terms into psychology by comparing concepts to theoretical terms. Furthermore, they desired to provide an explanation of how concepts change in the same way as theories change in science. The theory-theory psychologists have had a tendency to talk about concepts as being both like theoretical terms and like theories, which would be structures at different levels. When two concepts participate in the same mental theory, the structure of each of them would include the other. However, if the first concept contained the second concept, it was not possible for the second concept to contain the first concept. The theory-theory psychologists also focused on the structure of the concept. The internal structure of the object was the important aspect when categorizing the different objects. These psychologists did not focus on the observable patterns of behavior that occurs during categorization (Laurence & Margolis, 1999).

Psychological essentialism has been important within the theory-theory view of concepts (Medin & Ortony, 1989). Medin and Ortony suggested that when something was to be categorized, the person had to think as well as look to be able to categorize the object or event correctly. An example from Medin and Ortony (1989) was that when someone was categorizing an object, for instance a whale to the concept mammals, they had to think about the fact that whales are also mammals in order to categorize correctly. It was not enough to look at the whale to know that it was a mammal. It was the people's representations of the thing that was of importance when categorizing. Psychological essentialism was described by Palmer (2002) as the view that class memberships is determined by natural phenomena's set of necessary and sufficient properties.

This theory of concepts seems to bump into problems when they talk about people having theories about the nature of organisms, when these theories are just a fiction inferred from people's behavior. Psychological essentialism has been a problem because the necessary features of a concept only comes from the theories we have about nature, and not the reality of nature (Palmer, 2002).

Conceptual Atomism

"Lexical concepts are primitive; they have no structure" (Laurence & Margolis, 1999, p. 62). All of the other theories about concepts share the notion that all concepts have structure. Conceptual atomism, however, does not agree with that. Fodor (1998) proposed that for instance the mental representation DOG does not have structure, but rather that such mental representations are atoms. The view of conceptual atomism that concepts do not have structure has created discussions, and according to Laurence and Margolis (1999), the theory can be considered as a negative view about what concepts do not have. According to Fodor (1999) it cannot be said that for instance smoke carry information about smoke, but rather that it represents the mental representation smoke. He implied that carrying information of something involves transitivity, while a representation of something does not. An example he proposed was that the relation between smoke and fire is not transitive in the sense that smoke means fire. Smoke does not necessarily mean fire; it only represents the concept smoke.

One of the most developed theories within conceptual atomism have been the Asymmetric Dependence Theory, and was proposed by Fodor (1999). Fodor said that the content of a primitive concept was decided by the causal relation it had to other things in the world. For instance when talking about the concept bird, the bird's content are not expressed by being related to an animal or having wings, but it is expressed by the property of being a bird. Expressing the property of being a bird has a causal relation to the concept bird (as cited in Laurence and Margolis, 1999). According to Palmer (2002) this theory has been one of the most carefully reasoned, but a problem with this theory has concerned that a concept is used as a mental symbol. This theory may only be used to describe such mental symbols, and is not suitable for describing behavior.

The Prototype Theory

"Most concepts (especially lexical concepts) are structured mental representations that encode the properties that objects in their extension tend to possess" (Laurence & Margolis, 1999, p. 31). This theory involved that when categorizing an object or event, this object had to satisfy a sufficient number of features, and some of these features seemed more important than others. When a new object is to be categorized this would involve that the person categorizing checks if the representation of the category and the representation of the new object are sufficiently similar.

Rosch has been one of the central psychologists associated with the prototype theory (e.g., Mervis & Rosch, 1981; Rosch, 1973, 1978; Rosch & Mervis, 1975; Rosch et al., 1976). According to Rosch (1978) there are two basic principles in the formation of concepts, which she calls the formation of categories. Firstly, there is the function of the category systems, which means that the category system's task is to use the least cognitive effort to provide the most possible information about the environment. Considering a stimulus to be a part of a category would involve considering whether the stimulus is equivalent to other stimuli in that category, and considering if the stimulus is different from stimuli not in that category.

Reducing the many differences between stimuli to cognitively and behaviorally proportion was one purpose of categorization.

The second basic principle concerns the structure of the provided information. It has been important that the categories provide information that gives a correct picture of how the world is perceived. These two principles have implications for how the categories are abstracted in a culture and for the internal structure of these categories. Rosch (1978) suggested to divide into a horizontal and vertical dimension for explicating the category systems. The vertical dimension involves the level of what is included in the category. For instance abyssinian, cat, mammal, animal etc. vary along a dimension. The basic level has seemed to be the most effective level, in order to show how the structure of the characteristics is perceived in the world. Not all levels of categorization will be useful. The horizontal dimension refers to the segmentation of categories where for instance the dimension of pig, cat, sofa, table etc. varies. Basic objects have the highest cue validity because the members of the category have many attributes in common. “Cue validity is a probabilistic concept; the validity of a given cue x is a predictor of a given category y” (Rosch et al., 1976, p. 384). Categories that are subordinate to the basic level share many characteristics with other subordinate categories, and thus have lower cue validity. Categories at higher levels of abstraction share few characteristics, and also have lower cue validity compared to the basic objects.

Rosch et al. (1976) implied that the basic object categories are the categories that are learned first and named by children, and are also the most necessary for language universally. Furthermore, they said that it is the principle of category formation that is universal. Categories on the most general level, thus the basic level, will form so that they are maximally differentiable from each other.

According to Rosch and Mervis (1975) the most prototypical members at the different levels of categories are the members that have the most family resemblance to other members of the same category, and the least family resemblance to other categories. For instance in a category of furniture, the most representative members may for example be chair or sofa, as opposed to for example footstool. These are called the prototypes in a category, the ones that people typically think of first when they are thinking about the concept furniture. Mervis and Rosch (1981) suggested that the reason that the basic level categories seemed to be the ones that people learn first, are probably because they are the ones that are most similar to the other objects in the same category, and the least similar to the ones not in that category. The exemplars that are most representative for the category may also be the best to compare with when generalizing to other exemplars within the same category.

According to Palmer (2002) this theory is the most promising theory of the five. Rosch and her colleagues have conducted a lot of experiments on concepts and categorization, where several of the other theories about concepts have only been theoretical and not empirical. The theory points towards stimulus and response classes in interactions with the environment, and away from the classical theory with its logical categories. Rosch and her colleagues have investigated several aspects of concepts and categorization, and they have focused on some of the same things as in stimulus equivalence, for instance measuring reaction time. However, the prototype theory never comes to speak of concept formation in terms of behavioral processes. Thus, he sees the theory as inadequate in giving a full account of concepts and its formation.

These five theories of concepts within cognitive psychology have some similarities. They talk about concepts as mental representations, relative to talking about them in behavioral terms, and as derived from direct observation. They all have the tendency to talk about concepts in an essentialistic way, that for an object to be a part of a concept it has to have certain necessary properties. The behavior analytic approach have been more interested in studying concepts by studying how they form, and their functions.

The Behavior Analytic Approach To Concept Formation

Behavior analysis has obtained an unfortunate rumor of being uncaring in the way they study behavior, and many have avoided behavior analysis just because of this. Sidman (2007) said that private events are important for the science of behavior analysis as well as cognitive psychology, for instance that private events may become reinforcers for our behavior. The study of complex human behavior, and private events has been an important field of study within stimulus equivalence. Behavior analysts have focused on the study of the environmental variables influencing behavior, and how they can use the results from science to change the world, in terms of changing behavior. This does not mean that behavior analysts do not care about people, but that they are interested in using the best methods in order to create behavior of importance for the person in question. On the other hand, in cognitive psychology where the main field of study have been unobservable mental phenomena, the behavior is never explained, which leads to a difficulty when forming interventions for behavior change.

This approach has been very different from cognitive psychology, and its focus on the structure of concepts. Behavior analysis has focused on the functions of behavior. Catania (2007) talked about the importance of focusing on both structure and function. An example he used is that when you are teaching a child to read, the child needs to be trained in the structure of words and sentences, and he/she needs to be trained in the functions of the words in order to understand the relations between the words in the text. Skinner (1956) talked about the variables that behavior is a function of. These variables have to be available for scientific analysis. This means that the variables have to be outside of the organism, they have to be in the environment of the organism, and that behavior is influenced by its environmental history.

The definition of verbal behavior implicates that the individual's behavior only achieves its effects through the behavior of another person, thus a listener. The listener must have been trained in repertoires that include reinforcing the speaker, because this facilitates social control (Skinner, 1957).

Within behavior analysis concepts have not been given a lot of focus. However, research on stimulus equivalence has been a field that can be seen as the behavior analytic approach to the study of concept formation (see for example Sidman, 1994). Keller and Schoenfeld (1950) said that conceptual behavior occurs when the same response is evoked by a group of objects. Calling something a concept would involve that the members of a stimulus class are reacted to similarly. An example may be the concept dog. This concept would involve a group of stimuli, for instance different dogs that have different shapes and colors. Concepts are developed by generalizing within classes of objects or events, and through discrimination between the different classes.

> When an organism is conditioned to respond to one stimulus, it will respond in the same way to certain others. We call this generalization… Behavior can show a specificity with respect to stimuli, and when this specificity is developed in the face of generalization, we speak of stimulus discrimination (Keller & Schoenfeld, 1950, p. 116).

Both generalization and discrimination has to be trained over several different objects before concepts are formed, especially when children are starting to develop words for the different objects or events. As children grow older, the concepts are altered. For instance, when children are taught that the different fish have different names, they fall into different categories, and the concept fish changes (Keller & Schoenfeld, 1950). Another example of discrimination is that a person will discriminate between a friend's face and another unknown person's face, and would probably talk to the friend instead of the unknown person. An example of stimulus generalization can be that children say fish to all fish. After a while the children will discriminate the fish, and say that this is an eel and this is a salmon. Thus, one stimulus class has become two different stimulus classes, where eel is one class and salmon is another.

Stimulus classes may be formed based on primary stimulus generalization or based on arbitrary matching. A match-to-sample (MTS) procedure is a procedure that arranges four-term contingencies in order to establish conditional discriminations. Identity MTS involves that the sample stimulus that appears first is physically similar to the correct stimulus comparison. This stimulus is defined as the discriminative for reinforcement if a certain response occurs, and the other comparisons are not (Green & Saunders, 1998). For example when an apple is the sample stimulus, and the comparisons includes an apple, a pear, and an orange, the correct comparison stimulus will be the apple.

Stimulus classes that do not share the same physical features are called abstract stimuli and the stimulus classes they form have been called arbitrary classes. Arbitrary stimulus classes implicate that there is no obvious relation between the stimuli in the class. The experimenter defines which stimuli belong together in a class. The relation between the stimuli has to be trained, which can be done by using a procedure called arbitrary MTS. The stimuli could for instance involve Greek letters for people that have had no previous training with these letters. For others the stimuli may include different kinds of abstract symbols. Arbitrary classes have been widely studied within stimulus equivalence (Green & Saunders, 1998).

Zentall, Galizio and Critchfield (2002) suggested to talk about concept learning instead of concepts solely. They proposed to distinguish between three types of relations that are united within a category. These are perceptual concepts, relational concepts and associative concepts. Perceptual concepts refer to stimuli within a group based on shared physical features. Similarity between the stimuli within a category has been defined by relatively little difference between the stimuli along a well-defined physical dimension, and that the participants respond similarly to these stimuli.

Within the research on Sidman's account of stimulus equivalence, they have studied identity MTS as described above. Perceptual concepts form based on identity MTS procedures, by physically similar objects being matched to each other. Relational concepts are stimuli that are grouped into a category by the relation between them. These stimuli have often been abstract and an example may be same versus different that Zentall and Hogan (1978) did an experiment on with pigeons. This can be done by using an MTS procedure, and have for instance been studied within relational frame theory (e.g., Steele & Hayes, 1991), which is one of the three approaches to stimulus equivalence (e.g., Arntzen, 2010). The other two consist of Sidman's equivalence, and the naming hypothesis (e.g., Horne & Lowe, 1996).

Associative concepts involve stimuli that are grouped because of a shared function, this may be that they are correlated with a common consequence or evoke a common response. Thus, they have shared functional properties (Zentall et al., 2002). Associative concepts have been studied a lot within Sidman's stimulus equivalence, as described below.

Adams, Fields, and Verhave (1993) talked about open-ended categories or concepts. Open-ended categories have involved categories that consist of an infinite number of stimuli. The stimuli within these categories are physically similar. Categories that only consist of a limited number of stimuli are called functional classes and equivalence classes, and they contain stimuli that are disparate. Through the processes of equivalence class formation and primary generalization a category can be formed, either with stimuli physically similar, or those that are physically disparate. These types of classes have been called generalized equivalence classes. Generalized functional classes involve that a response trained to one of the members of a class transfers to the other members of this class.

Griffee proposed that even if words do not resemble each other or the pictures they represent, under some conditions, all of the stimuli would function interchangeably. The same response would be evoked by all of the stimuli in this kind of class, even if it was only trained to only a few stimuli in this class (as cited in Fields & Reeve, 2000).

Stimulus Equivalence

Stimulus equivalence has been the study of complex human behavior within the field of behavior analysis. In addition to studying concept formation, research has also focused on the study of remembering, classification and categorization, and other aspects of learning (see Sidman, 1994).

Stimulus equivalence involves that stimuli within a class are mutually exchangeable (e.g., Green & Saunders, 1998). Responding according to stimulus equivalence involves that relations emerge without having been trained directly. These new relations emerge as a result of the trained relations and have been called derived relations (Sidman & Tailby, 1982). The definition of stimulus equivalence includes that the participant responds according to reflexivity, symmetry and transitivity. Reflexivity involves that the participant is able to respond according to identity, that A1 equals A1, B1 equals B1 and so on. Symmetry includes that the participant responds to the reversed relation, that if the participant was taught A1 equals B1, then B1 equals A1. Transitivity involves that if the participant has trained that A1 equals B1 and B1 equals C1, then A1 equals C1. It has to be at least two classes with three members in each class (e.g. A1, B1, C1 and A2, B2, and C2) to be able to test for these three relations. There have been discussions as to which variables influence stimulus equivalence, and the different approaches have different meanings about this. Researchers within the naming hypothesis have proposed that naming the different stimuli are necessary to respond according to stimulus equivalence (Horne & Lowe, 1996). Researchers within the relational frame theory talk about relational frames, and one of these frames are equivalence, or the same as (e.g., Hayes, Barnes-Holmes, & Roche, 2001).

Some examples from what they have studied within stimulus equivalence have included reaction time, speed, class size, nodal density, number of classes, and how using the same names on the stimuli in the same class influence stimulus equivalence formation. Research on stimulus equivalence has differed from other experimental procedures in that new relations emerge that have not been explicitly taught (Sidman, 1994). Most of the experiments on stimulus equivalence have been conducted in laboratory settings under controlled conditions. One of the goals with this research has been to find ways to apply findings from the laboratory and create procedures that may be used in training situations, such as in school. Sidman wrote that it has been difficult to apply stimulus equivalence procedures in education settings because many institutions have disregarded fundamental research. He noted that establishing teaching procedures based on the findings from the laboratory ought to be a primary focus in the future.

There have been several variables that influence stimulus equivalence responding. Fields and Verhave divided the variables into “class size, number of nodes, the distribution of “singles” among nodes, and directionality of training” (p. 317). Class size involves the number of members in an equivalence class. Research on class size has focused on the maximum number of stimuli that may participate in the same equivalence class. The number of nodes involves the number of stimuli connected to at least two other stimuli during training. For instance an A, B, C class would have one node, B, and two singles, A and C, because they are only connected to one stimulus each in the class. The distribution of singles among nodes concerns that if the classes contain more than six stimuli, it can be distributed in many ways among the nodes. Thus, the singles may be distributed differently to each node, for instance there may be three singles linked to one node in one end and two singles to another node (Fields & Verhave, 1987).

Directionality of training is the last variable that Fields and Verhave (1987) discussed. They differentiated between unidirectional and bidirectional training. Unidirectional training involves that one stimulus is only sample and the other is only comparison. For instance, A can be trained to B or B to A. In bidirectional training, the stimuli may function as both sample and comparison. For example; A and B are both sample and comparison to each other.

Stimulus equivalence has been studied with different groups of people, for example with adults, college students, people with developmental disabilities, and children. Studies have focused on which variables influence stimulus equivalence, and how young the participants can be and still respond according to stimulus equivalence. Sidman, Kirk, and Willson-Morris (1985) conducted a study including eleven children as participants. Seven were in the age of 5–6 years old. This study consisted of three experiments. The purpose was to investigate the possibility of training three 3-member classes in two different stimulus sets, and combin them into three 6-member classes. The stimuli consisted of three classes with the A, B, and C stimuli in Stimulus Set 1, and the D, E, and F stimuli in Stimulus Set 2. They used an OTM training structure. The three experiments differed in the way the different relations were trained and tested. In Experiment 1 they trained the two stimulus sets one by one, before training EC combined them, and then there was a big overall test for all relations. The relations were tested one at a time, for example BF, then EF etc. Experiment 2 trained all of the relations together in a mix. Subsequently, the participants were tested for the ABC classes, and DEF classes before they trained EC, and had a big overall test. In Experiment 3 they enlarged the baseline progressively by training AB, and AC, and then they were tested for the appropriate stimulus class before training the next relations. The next step was to train EC, test, train DE, test, train DF, test, and then the last test with all of the remaining relations.

Sidman et al. (1985) found that eight of the eleven participants managed to from the three 6-member equivalence classes. They suggested that training one relation at a time and testing for the appropriate stimulus classes was probably the most effective procedure in order to expand stimulus classes.

Discussion

Palmer (2002) did a review of the book by Margolis and Laurence, and said that the concepts they wrote about were mainly conceptual rather than empirical. Most of the theories about concepts within cognitive psychology have focused on the structure of the concepts. Little focus has been on how concepts form, and the variables that influence concept formation. Within behavior analysis the function of the concept has been the main focus. Within stimulus equivalence the goal has been to study how concepts form in behavioral terms, and which variables influence this formation.

One of the main problems with concepts within cognitive psychology in general has been that the psychologists do not agree on what a concept is, and the unit of analysis. There have been very many theories, and it seems that the psychologists within each theory criticize the other theories. According to Laurence and Margolis (1999) the classical theory was the first theory about concepts, and the other theories that have been discussed have been seen as reactions to this view.

According to Palmer (2002) there are two conceptual problems with most of the papers in the book by Margolis and Laurence. The first problem is relative to psychological essentialism, where objects are categorized to be part of a concept based on fulfilling a set of criteria or properties. Concepts are not presented as stimulus classes and response classes, rather as abstractions, which may be hard to define operationally. Furthermore, concepts are presented as structured mental representations, and not as derived from direct observation. An example that Palmer presents is that the distinction between classes of dog and not dog is meaningless. The boundary between classes is only a feature of our models of nature, and not of nature as it is. Natural classes produced by selection contingencies do not have essential properties in order to be part of that class. There are for instance several species where the members of the same species settled in different parts of the world a long time ago, and now these species have come to be different species because of natural selection, and different environmental conditions. Some species have even become so dissimilar that it is no longer possible for them to interbreed.

Zentall et al. (2002) chose to talk about concept learning instead of only concepts because they indicated that this would provide the ability to identify the functional relations between the environment and behavior. Furthermore, this would provide a basis for investigating behavior defined as conceptual. Cognitive psychologists have until now not given the functionality of concepts much consideration, and only focused on the structure of them. Palmer (2002) proposed three probable reasons that psychological essentialism overshadows most of the papers in Margolis and Laurence (1999). The first reason concerns that when contingencies are stable, natural variability is hard to detect. Secondly, the unit of analysis in behavior analysis is different from other psychological disciplines in that it is identified empirically, in oppose to a priori. Thirdly, different response classes are revealed when behavior is defined topographically, without looking at the controlling variables of the behavior. Thus, it may seem obvious that some stimuli are members of one stimulus class, and others are members of another stimulus class. However, all instances of behavior are considered as the same, instead of looking at the response as a member of a response class. This may lead to serious error in the case of verbal behavior where the same word may have different meaning in relation to the different environmental contingencies.

Variability has not been accounted for in many of the theories about concepts. However, it is fundamental in the selection of behavior, and also an important aspect in theories about concepts (Palmer, 2002). If there had not been any variation in behavior or the organisms, selection would not occur either. Selection does not specify the criteria objects have to meet in order to be part of a concept. Variability within the objects is necessary and important. There are many instances that would classify as members of a concept that would not necessarily fulfill a set of criteria. It is the reinforcement contingencies that select which instances that are considered as members of a particular concept. The selection contingencies will change and the lineages of behavior or organisms will change along the contingencies, by behavior being shaped, the organisms' evolution, and the specialization that emerge among the domestic plants and animals. Theories about concepts have to acknowledge that contingencies will change and variability within living things will have to adjust. This means that the theories about concepts must accommodate this fact, and respect the generic nature of response and stimulus classes where each instance of a response and a stimulus varies.

In the classical theory of concepts the objects were categorized by fulfilling certain necessary features, and a focus was on the joined sensory properties that reflected the environment (Laurence & Margolis, 1999). A problem with this definition is that these sensory properties may be hard to define. What are the joined sensory properties, and how is it possible to see these for other people? These joined sensory properties are unobservable, and thereby hard to measure. Furthermore, the psychologists within this theory do not talk about concepts in terms of behavior, and have not given any attention to the variables influencing concept formation (Palmer, 2002). They have rather been interested in what a concept consist of and not. It seems incomplete to only talk about what a concept consist of, and not how someone comes to acquire a concept. Also, this theory seems to only be a theory, not derived from direct observation and empirical research.

The neoclassical theory included the psychologists that were interested in the meaning of words, especially verbs, and that concepts consist of partial definitions (Laurence & Margolis, 1999). There seems to have been no focus on behavioral principles within this theory either (Palmer, 2002). Also, Jackendoff (1999) said that when people categorized objects they used finite schemas. He posted that conceptual development are ordered in structures in the mind. Furthermore, Jackendoff's theory suggested that there are some concepts that have special status and are part of the mind's architecture, and other concepts only become meaningful in relation to this architecture (as cited in Palmer, 2002). This would involve that we have some structures that exist in the mind, and that the brain are divided into special systems. This includes the fallacy of using hypothetical constructs to explain something that they do not know exist (Watkins, 1990). This would have to be investigated by neurologists, and it is not suitable to make theories without having any empirical background.

The theory-theory psychologists promoted that concepts may only be understood in relation to our theories about the world because they play a role in the theories we have about the world (Palmer, 2002). For instance when categorizing different objects or events into a special concept we use our theories about the world to decide which concept it may be part of. A problem with this may concern that most people have different theories about the different concepts, and thereby categorize differently. This leads to a problem with what should be parts of the different concepts. Some may also have insufficient theories about some concepts, and then not be able to categorize. According to Palmer (2002) people do not have theories about the nature of organisms, it is rather talk about the basic processes of discrimination and generalization. People are trained to discriminate between classes of stimuli, and between stimuli in a class/concept.

Conceptual atomism is according to Palmer (2002) the most abstract theory out of the five. Fodor (1998) has been one of the leading psychologists within this theory. In his book he indicated that the existing theories about concepts were useless, and that it was necessary to investigate concepts in another way. He did not agree that concepts had structure, but he still talked about concepts as mental representations. According to Palmer (2002) a concept is demonstrated here as a mental symbol and not a response class, and that Fodor's conclusions about concepts may only be related to these symbols. They are not relevant in the case of talking about behavior, and thus the theory cannot be used.

The psychologists within the prototype theory were interested in the members of a concept that were more typical than others, and called these the basic level objects or the prototypes (e.g., Rosch et al., 1976). Palmer said that the classical theory may be seen as a formal theory that focus on the nature of the world, while the prototype theory focus on our relationships to the world and may be seen as a psychological theory (2002). Furthermore, he implied that what is defined as a basic level object, or the central exemplar, has to depend on the person's experience. Within behavior analysis this would be called the person's learning history. The prototype theory may be translated into stimulus and response classes, and they define the basic level objects in terms of the processes of discrimination and generalization. Rosch (1978) presented two basic principles involved in how concepts are formed, but these two principles do not focus directly on what kind of behaviors that occur during the formation of the concepts, they have given more attention to the task of the category system. Palmer (2002) refers to the work by Rosch as promising, but because she never talks about concepts in terms of behavioral processes this theory also seems incomplete.

Skinner (1956) posted that the variables affecting behavior has to be outside of the organism, behavior is influenced by the environmental history, and the variables have to be in the organism's immediate environment. When it comes to these five theories about concepts as discussed above, they do not talk about concepts in behavioral terms, and they have not accounted for the environmental contingencies. However, the formation of concepts should be talked about in terms of behavior and their relation to the environment, because when we form concepts or what we call stimulus classes there are many different behaviors going on. When stimulus classes are being formed, the individual have to be presented to many exemplars of the class before the individual starts to generalize to other stimuli in the same class. Also, the individual has to be presented to other stimuli that do not belong in that class to discriminate between the classes. The stimulus classes or concepts will vary, and it depends on the environmental contingencies that the individual has been exposed to throughout its life span, also known as learning history.

The study by Sidman et al. (1985) was an example of how young children down to the age of 5–6 years old may expand stimulus classes by training a relation between two existing stimulus sets, and that many relations may emerge as a result of training. This master's thesis Article 2 consist of a systematic replication of the Sidman et al. study. In order to contribute to the progress of the science of behavior it is important to use empirical methods. Rosch and her colleagues conducted empirical research on concepts and categorization, but they focused on the structure of the concepts. There are similarities between the research done by Rosch and her colleagues, and research within stimulus equivalence. Rosch (1978) referred for instance to evidence based on previous experiments that prototypicality is related systematically to priming effects, rate of acquisition, response rate, and other response strength measures. These measures have also been studied within stimulus equivalence. The article of Rosch et al. (1976) consist of several experiments where they talk about different levels of categorization, the superordinate, basic, and subordinate level. The superordinate level involves categories such as musical instruments or fruit, while the basic level categories involves guitar, piano, and drum, and apple, peach, and grapes. This can be related to the formation of stimulus classes within stimulus equivalence, where different stimulus classes may be formed by the use of MTS procedures.

Rosch et al. (1976) also conducted experiments in a similar way to how typical conditional discrimination procedures have been conducted within stimulus equivalence research. However, in their experiments they focused mostly on the different levels of categorization. They investigated for instance the difference between reaction times to the superordinate level, basic level, and the subordinate level of categorization. They found that responding to objects classified as members of categories on the basic level had lower reaction times than the superordinate, and subordinate levels. Research on stimulus equivalence has also focused on reaction time. They have for example measured the different reaction times when comparing the use of pictures versus abstract stimuli in the formation of equivalence classes. Other similarities between the experiments done by Rosch et al. and stimulus equivalence experiments include that both have studied class size, in how many members that may be considered as a part of a concept. Rosch et al. (1976) found that the participants shared a higher agreement on the shared attributes for the basic level objects, than at the other levels of categorization. They also found that there was generally more correct categorization for the objects at the basic level, compared to the other levels. Findings from research on stimulus equivalence have for example included that accuracy may decrease as a result of increasing class sizes, and the number of nodes (Fields & Verhave, 1987).

Overskeid (2008) addressed some central critiques of cognitive psychology, seen from a behavior analytic perspective. He divided the critiques into three, consisting of the problem of mentalism, private causation, and functional relations. This critique involved that the if the mental is used as a name for something nonphysical cognitivists cannot be said to perform science. This leads to problems when having discussions in scientific terms, and since private events are unobservable it has to be measured in another way than observable behavior. The problem of private causation refers to that private events cannot be the only cause of behavior; it has to be something in the public environment. An important point to why behavior analysts say this is that public stimuli are the only ones that are available for behavior change (Pierce & Cheney, 2008). And, the last problem includes functional relations. Cognitive psychologists have not on focused the functions of behavior, which provides an insufficient account of the behavior. Catania (2007) suggested that it is necessary to study both structure and function, and that the study of behavior should involve both. Although behavior analysis have not given much attention to explaining behavior called cognition, cognitive psychologists has not focused on the functional relations, which are a part of their central field of study (Overskeid, 2008).

Conclusion

Concepts have been studied in different ways within cognitive psychology and behavior analysis. The five theories about concepts as discussed here have all studied concepts as mental representations in some or another way. None of the theories have the same definition of what a concept is, and they all have different definitions of how categorization occurs. This has been one of the bigger problems with cognitive psychology, and psychology, in general. There have been many different subfields within psychology, and they do not agree on what should be the unit of analysis, and how to study it. Behavior analysis on the other hand has united on the field of study, which is behavior. Behavior analysts studying concept formation have referred to it as stimulus equivalence within a class of stimuli. It has according to Watkins been fairly easy to come up with a theory. Maybe one of the reasons that there have been so many theories about concepts is that many of the theories are not based on direct observation, and not empirical. Thus, many have not even been tested. None of the theories have discussed concepts in relation to stimulus or response classes. However, Rosch and her colleagues have done a lot of research in many of the aspects also studied within stimulus equivalence. Concepts and especially concept formation should be studied in behavioral terms, in relation to the environment. The environment is in a constant change, and the study of concepts has to acknowledge that, and follow the change.

References

Adams, B. J., Fields, L., & Verhave, T. (1993). Formation of generalized equivalence classes. *The Psychological Record, 43*, 553–566. Retrieved from http://www.web.ebscohost.com.

Arntzen, E. (2010). Om stimulusekvivalens. In S. Eikeseth & F. Svartdal (Eds.), *Anvendt atferdsanalyse. Teori og praksis.* (2 ed., pp. 100–138). Oslo: Gyldendal Norsk Forlag AS.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of applied behavior analysis, 1*, 91–97. doi: 10.1901/jaba.1968.1-91 Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1310980/pdf/jaba00083-0089.pdf.

Barsalou, L. W. (1992). *Cognitive psychology: An overview for cognitive scientists*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Fields, L., & Reeve, K. F. (2000). Synthesizing equivalence classes and natural categories from perceptual and relational classes. In J. C. Leslie & D. E. Blackman (Eds.), *Experimental and applied analysis of human behavior* (pp. 59–82). Reno, NV: Context Press.

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior, 48*, 317–332. doi: 10.1901/jeab.1987.48-317

Fodor, J. (1998). *Concepts. Where cognitive science went wrong*. New York, NY: Oxford University Press Inc.

Fodor, J. (1999). Information and representation. In E. Margolis & S. Laurence (Eds.), *Concepts: core readings* (pp. 513–524). Cambridge, MA: A Bradford Book. The MIT Press.

Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of Reseach Methods in Human Operant Behavior* (pp. 229–262). New York: Plenum Press.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). Relational frame theory: A précis. In S. C. Hayes, D. Barnes-Holmes & B. Roche (Eds.), *Relational frame theory. A post Skinnerian account of human language and cognition* (pp. 141–155). New York: Kluwer Academic/Plenum.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*, 185–241. doi: 10.1901/jeab.1996.65-185

Jackendoff, R. (1999). What is a concept, that a person may grasp it? In E. Margolis & S. Laurence (Eds.), *Concepts: Core readings* (pp. 305–333). Cambridge, MA: MIT Press.

Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology*. USA: Appleton-Century-Crofts, Inc.

Laurence, S., & Margolis, E. (1999). Concepts and cognitive science. In E. Margolis & S. Laurence (Eds.), *Concepts* (pp. 3–81). Cambridge, MA.: A Bradford Book, The MIT Press.

Lee, V. (1988). *Beyond behaviorism*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Margolis, E., & Laurence, S. (1999). *Concepts: core readings*. Cambridge, MA: A Bradford Book. The MIT Press.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York, NY: Cambridge University Press.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 82–115. doi: 10.1146/annurev.ps.32.020181.000513

Overskeid, G. (2008). They should have though about the consequences: the crisis of cognitivism and a second chance for behavior analysis. *The Psychological Record, 58*, 131–151. Retrieved from http://opensiuc.lib.siu.edu/tpr/.

Palmer, D. C. (2002). Psychological essentialism: a review of E. Marcolis and S. Laurence (Eds.). Concepts: core readings. *Journal of the Experimental Analysis of Behavior, 78*, 597–607. doi: 10.1901/jeab.2002.78-597

Pierce, D. W., & Cheney, C. D. (2008). *Behavior Analysis and Learning* (4 ed.). New York, NY: Psychology Press.

Rosch, E. (1973). Natural categories. *Cognitive Psychology, 4*, 328–350. doi: 10.1016/0010-0285(73)90017-0

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573–605. doi: 10.1016/0010-0285(75)90024-9

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439. doi: 10.1016/0010-0285(76)90013-X

Ryle, G. (1949). *The concept of mind*. New York: Barnes.

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology.* USA: Basic Books.

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

Sidman, M. (2007). The analysis of behavior: What's in it for us? *Journal of the Experimental Analysis of Behvaior, 87*, 309–316. doi: 10.1901/jeab.2007.82-06

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

Skinner, B. F. (1956). *Science and human behavior*. New York, NY: The Macimillan Company.

Skinner, B. F. (1957). *Verbal behavior*. Massachussets, MA: Cambridge.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, Massachussets: Harvard University Press.

Steele, D., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior, 56*, 519–555. doi: 10.1901/jeab.1991.56-519

Watkins, M. J. (1990). Mediationism and the obfuscation of memory. *American Psychologist, 45*, 328–335. doi: 10.1037/0003-066X.45.3.328

Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: An introduction. *Journal of the Experimental Analysis of Behavior, 78*, 237–248. doi: 10.1901/jeab.2002.78-237

Zentall, T. R., & Hogan, E. (1978). Same/different concepts learning in the pigeon: The effects of negative instances and prior adaptation to transfer stimuli. *Journal of the Experimental Analysis of Behavior, 30*, 177–186. doi: 10.1901/jeab.1978.30-177

Equivalence Class Establishment and Expansion

Camilla Harangen

Oslo and Akershus University College

Department of Behavioral Sciences

Abstract

Responding according to stimulus equivalence includes responding correctly to relations without being directly trained. In the present experiment, the participants were two typically developing children at the age of 5 and 6 years old in their last year of kindergarten. The purpose of the present experiment was to investigate if it was possible to train two stimulus sets each consisting of three 3-member classes, and combine them into three 6-member classes, by training one relation between the stimulus sets. The training structure employed in the present experiment was a many-to-one (MTO) structure. The participants were first trained and tested with one stimulus set with the potentially three 3-member classes (ABC), thereafter trained and tested in another stimulus set with the potentially three 3-member classes (DEF). Subsequently, the participants experienced a re-training of the two sets before training the relation CF (the first node, –C and the second node, –F), in order to combine the sets. Finally, the participants were exposed to an overall test to examine whether they responded according to stimulus equivalence, and had formed the three 6-member classes. The results showed that both participants responded according to stimulus equivalence in the tests for Stimulus Set 1 and Stimulus Set 2. However, none of the participants responded according to stimulus equivalence in the overall test. Hence, one of the participants sorted the stimuli according to the defined stimulus classes, suggesting that she may have formed the classes if extended testing had been employed.

Keywords: expanding, stimulus classes, children, MTO training structure, sorting.

Equivalence Class Establishment and Expansion

Three properties have to be present in order to talk about someone responding according to stimulus equivalence. These properties include reflexivity, symmetry and transitivity. An example may include the potentially three 3-member classes, ABC. Training involves that the sample stimulus is presented successively in the middle of the screen. When the participant clicks on this stimulus, two or more comparison stimuli will appear on the screen simultaneously. Training may involve that the A stimuli is presented as sample to the B stimuli as comparisons, and the B stimuli as sample to the C stimuli. After training there would be a test for emerged relations where the three properties of stimulus equivalence are tested. To be able to test for the three defining properties of stimulus equivalence training has to include at least two 3-member classes (e.g. A1B1C1 and A2B2C2). The tests for potentially derived relations would involve the relations BA, CA, AC, and CB, in addition to the trained relations AB, and BC. Responding to reflexivity would involve that the participant is able to respond according to identity, that A equals A, B equals B, and C equals C. Symmetry includes that the participant responds to the reversed relation, that if the participant was taught A equals B, then B equals A. Transitivity involves that if the participant has trained that A equals B, and B equals C, then A equals C. Responding to stimulus equivalence would involve that the stimuli within each class are substitutable for each other (e.g., Sidman & Tailby, 1982).

Within the research area of stimulus equivalence it has been normal to denote the different sets by letters: A, B, C etc. Numbers, for example 1, 2 and 3, has defined the different classes. A1, B1, C1 defines the first class, while the second class would include A2, B2, C2 (Arntzen, 2010). A conditional discrimination procedure involves that when a conditional stimulus called A1 is presented, the discriminative stimulus is B1, and if A2 is presented, B2 is the discriminative stimulus. When A3 is presented, B3 is the discriminative stimulus, and so on. During the test for emerged relations all possible relations within each class are presented. The equivalence relations are inferred from the results of the test (e.g., Sidman, 1992; Sidman & Tailby, 1982). The conditional discrimination procedure may generate equivalence relations. If that happens, the participant's performance has been called matching-to-sample (MTS). The relationship between the stimuli within the experimenter-defined class is arbitrary (e.g., Sidman & Tailby, 1982). The MTS performance involves that the participant develops stimulus classes.

There have been three training structures used in equivalence research. In three 3-member classes, the one-to-many (OTM) training structure may involve training A→B and A→C relations, thus one sample stimulus to many comparison stimuli. The linear series training structure (LS) can involve training A→B and B→C, and the many-to-one (MTO) structure may involve training A→C and B→C, thus many sample stimuli to one comparison stimulus (e.g., K. J. Saunders, Saunders, Williams, & Spradlin, 1993). There have been several experiments investigating the different effects of the training structures. Some experimenters have found that the OTM structure provides a better outcome on tests for stimulus equivalence (e.g., Arntzen, Grondahl, & Eilifsen, 2010; Arntzen & Holth, 1997, 2000; Arntzen & Nikolaisen, 2011). Others have found results indicating that MTO provides a better outcome on equivalence tests (Arntzen & Vaidya, 2008; e.g., Hove, 2003; Saunders, Chaney, & Marquis, 2005; Saunders, Drake, & Spradlin, 1999).

Both Arntzen and Vaidya (2008), and Saunders et al. (1999) had children as participants and found that the MTO structure gave higher yields of equivalence responding than the OTM structure. This may suggest that the MTO structure may be the better training structure to use in order to facilitate responding according to stimulus equivalence with children. The LS training structure seems to have been the training structure that facilitates the poorest outcomes on tests for stimulus equivalence (e.g., Arntzen et al., 2010; Arntzen & Hansen, 2011; Arntzen & Holth, 1997, 2000).

According to Fields and Verhave (1987), there are four parameters that influence equivalence responding. These are class size, directionality of training, distribution of singles among nodes, and the number of nodes. This paper will not focus on the directionality of training and the distribution of singles among nodes. Singles and nodes may be described by how they are linked to the other stimuli in a class directly through training. Each stimulus in an equivalence class may function either as a single, or as a node. A single is a stimulus that is only linked to one other stimulus in a stimulus class, while a node is a stimulus that is linked to at least two other stimuli. For instance, in a class formed by using an LS training structure with A→B→C→D→E→F stimuli, the B, C, D and E stimuli are nodes, while the A and F stimuli are singles (Fields & Verhave, 1987).

According to Skinner (1935), every stimulus and response is a member of a larger class. In experimental studies that have focused on the area of stimulus equivalence the amount of stimulus classes and members per class has varied. Fields and Verhave (1987) said that one of the parameters that influenced equivalence responding was class size, in terms of the number of stimuli that may be included in a stimulus class.

Arntzen and Holth (2000) compared class size versus number of classes relative to results on equivalence tests. They had 66 participants in two experiments. In Experiment 1 they used an LS training structure, and in Experiment 2 they used an MTO training structure. They trained increasing class sizes, and the increasing number of classes. The results showed that the probability of equivalence class formation decreased as a function of increasing the number of classes, and of increasing the number of nodes. However, the decrease of equivalence outcome was much more sensational as a function of the increasing number of nodes. By comparing results from Experiment 1 and 2, the experimenters found that the decrease in equivalence responding was more influenced by the number of nodes, not as much by the increase in class size. In Experiment 2 the class size was increased in an MTO training structure. The results from Experiment 2 showed no difference in equivalence responding compared to an increase in the number of classes in Experiment 1 with the LS training structure.

According to Green and Saunders (1998), expanding a class to include more members involves training a conditional relation between the new stimulus and a stimulus in the already existing class. There have been extraneous variations within the research on expansion. Some experiments have focused on the expansion of stimulus classes with humans (e.g., Devany, Hayes, & Nelson, 1986; Dube, McIlvane, Maguire, Mackay, & Stoddard, 1989; Fields, Newman, Adams, & Verhave, 1992; Goyos, 2000; Lazar, Davis-Lang, & Sanchez, 1984; Saunders, Saunders, Kirby, & Spradlin, 1988; Schenk, 1994; Sidman, Kirk, & Willson-Morris, 1985; Sidman & Tailby, 1982). Some have tried to expand a class by training a new stimulus to an existing set (e.g., Fields et al., 1992; Goyos, 2000; Lazar et al., 1984), while others have trained two sets of stimuli and then tried to combine them to one (e.g., Sidman et al., 1985).

Experiments that have investigated whether expansion would occur by pairing different reinforcers to the different stimulus classes as proposed by Sidman (2000) have for example been conducted by Dube et al. (1989), Goyos (2000), and Schenk (1994), while other experiments have been conducted to explore stimulus-stimulus relations (e.g., Devany et al., 1986; Fields et al., 1992; Lazar et al., 1984; Saunders et al., 1988; Sidman et al., 1985; Sidman & Tailby, 1982).

Goyos (2000) conducted an experiment with four participants in the age of 4- and 5-years old. The procedure involved identity matching, before arbitrary matching was introduced with the stimuli A, B, C and D. The classes were trained by using stimulus-specific reinforcers. Training the AB and BC relations first led to the formation of the ABC classes for all participants when tested. The D stimuli were then trained by applying the stimulus-specific reinforcer for that particular stimulus class. The results from this experiment showed that all four participants responded according to stimulus equivalence after having trained AB and AC relations, while three of them showed the formation of the expanded classes ABCD after having applied the specific reinforcers for each class. The last participant was taught to name the D stimuli, and thereby demonstrated the expansion of classes to include the ABCD stimuli. Goyos found that it was possible to generate equivalence classes by using specific reinforcers for each class. However, he suggested that the effects of using stimulus specific reinforcers were relatively weak. He underlined this suggestion with the fact that many other experiments have demonstrated results of people forming equivalence classes by only using one single type of reinforcer for all of the stimulus classes.

Lazar, Davis-Lang and Sanchez (1984) conducted an experiment with four normal children where three of them were five years old, and one was 7 years old. They used the stimuli, A, B, C, D and E with three stimuli in each class (e.g., Class 1 consisted of A1B1C1D1E1). The stimuli in the A class were used as samples, and the D stimuli were used as comparisons. The participants in this experiment showed that after they had been trained in the AD and DC relations, the participants responded correct in a test for the AC and CA relations without having had the direct training in these relations. AE/EA and EC/CE relations emerged as a result after having trained the ED relations. When the CB relations were trained, AB/BA, EB/BE and DB/BD equivalence relations emerged. These findings demonstrate that many relations may emerge in addition to the relations that are directly trained. Adding other trained performances to the already existing matching baseline can be done to expand classes. Another finding from this experiment was that naming was not necessary for equivalence relations to emerge.

Sidman et al. (1985) conducted a study including 11 participants, where seven of them were 5–6 years old. The study consisted of three experiments in order to compare different procedures in the formation of expanded equivalence classes. The purpose of the experiment was to investigate the possibility of combining six 3-member stimulus classes and form three 6-member classes. The stimulus classes consisted of A, B, C, D, E and F with three stimuli in each class (e.g. class one consisted of A1B1C1D1E1F1). The general method included training the conditional relations AB, AC, DE, DF, and EC. Thus, the training structure was OTM. The A stimuli were auditory stimuli, which made the testing of global equivalence and symmetry with the A stimuli impossible. After training, there was an overall test to examine whether the participants had formed the three 6-member equivalence classes. The existence of 6-member equivalence classes would be demonstrated by the emergence of the conditional relations BF and FB (Sidman et al., 1985).

The sequence of the presented trial types differed between participants in the different experiments (Sidman et al., 1985). In Experiment 1 the conditional relations were trained separately, and the three relations within a class was taught in a prescribed sequence, thus they used a serialized training order. Experiment 1 involved training all of the conditional relations first as described above, before having an overall test for all of the relations involved in the three 6-member classes. One test consisted of 27 trials with each sample being presented nine times. For instance, all of the BF relations were presented in one test, and all of the FB relations were presented together in another test. These tests were inserted among 45 baseline trials, as probes.

Each scheduled test session for all of the experiments involved a review of the baseline relations where the relations were presented with reinforcement (Sidman et al., 1985). If the participants did not respond to the accuracy criterion in these relations, they would go back to an earlier teaching stage before moving forward again. In Experiment 2 there were none in the age of 5–6 years old. In Experiment 3 the conditional relations were presented progressively, thereby increasing the size of the baseline. The conditional relations were prerequisites for increasing the class size. With each new step the relations were tested. Training started with the AB and AC relations before the participants were tested for emergent relations appropriate for this stimulus class. The next step was to train the DE and DF relations, and then test. In the end they trained the EC relations, and tested if the classes had expanded to involve the three ABCDEF classes (Sidman et al., 1985).

Results from the study by Sidman et al. (1985) demonstrated that eight of the eleven participants responded according to equivalence. Four of the participants who managed to form the three 6-member equivalence classes were in the age of 5–6 years old. In Experiment 1, there were two out of four participants in the age of 5–6 years old that formed the classes. In Experiment 3, two of the three participants managed to form the three 6-member equivalence classes, while the last one did not complete all tests. She showed equivalence responding in several of the 3-, 4- and 5-stage relations, and may have showed equivalence formation in 6-stage equivalence relations if she had completed the tests.

One of the implications of the study by Sidman et al. (1985) was that when the 6-stage conditional discriminations did not emerge, the prerequisites were also deficient. This deficiency was also present during the 3-stage relations, which included the three 3-member classes of ABC or DEF. When extended testing was employed, the missing prerequisites started to emerge. The tests functioned as a necessity for class formations to emerge; all of the participants in the age of 5–6 years old required extended testing (Sidman et al., 1985). Lazar et al. (1984) supported these findings, and suggested that including more test trials would make it more likely for the relations to emerge. According to Sidman et al. (1985) it is a possibility that before larger stimulus classes can emerge, the baseline conditional discriminations have to be extendedly trained. Sometimes the baseline conditional discriminations are not established well enough, and this may be the reason that equivalence relations do not emerge during testing without reinforcement. They suggested that the equivalence classes did not exist before the conditional relations are tested, and that this was supported by the fact that many of the relations required to be tested repeatedly before the equivalence classes emerged.

Another observation from the study by Sidman et al. (1985) implicated that the higher stage relations did not emerge before the lower stage relations had been tested. The participants could have formed many kinds of classes. For instance, they may have formed classes based on physical features among the stimuli. The fact that they did not form other classes than what was experimentally defined; also support the suggestion that the equivalence classes formed during the testing of the derived relations. An analysis of the results made by Sidman and his colleagues suggested that enlarging classes progressively might be the most effective method in order to expand stimulus classes. They insinuated that the addition of three new members to a class in one time might break down or prevent 3-member classes from forming, as well as the formation of the three 6-member classes. Some findings that exist about expansion with children have been that Goyos (2000), Lazar et al. (2000) and Sidman et al. (1985) found that naming was not necessary for expansion to occur. It is possible for children to expand classes by adding new performances to the matching baseline (e.g., Goyos, 2000; Lazar et al., 1984; Sidman et al., 1985). It is possible to use stimulus specific reinforcers to expand stimulus classes (Goyos, 2000). Also, Lazar et al. and Sidman et al. employed different test procedures, and both studies suggested that extended testing was necessary before equivalence classes formed.

This current experiment was a systematic replication of the Sidman et al. (1985) study. Some of the differences between this experiment and theirs include that this experiment used only visual stimuli, thus making it possible to test all relations. This experiment excluded naming procedures, and used an MTO training structure as opposed to the OTM training structure that they used. Also, the testing procedures used in this current experiment differed from the testing procedures that were used by Sidman and his colleagues. The purpose in the present experiment was to investigate whether children in the age of 5−6 years old would be able to form three 6-member equivalence classes by using all visual stimuli with an MTO training structure by training one relation between two trained stimulus sets, and testing for all of the relations at the end.

Method

Participants

The participants consisted of two children in a kindergarten in Oslo. They were in their last semester of kindergarten. The participants were selected to this experiment if they managed to form the AC conditional discriminations within 500 trials. Out of two kindergartens with 17 children there were only four participants that managed to form the conditional discriminations, only two participants completed the experiment. In this first kindergarten, there was only one participant that managed to form the first conditional discriminations, but he had to withdraw from the experiment. He said he was bored, and wanted to quit. The other participant that had to withdraw from the experiment was due to that his correct responding decreased when there was no reinforcement during subsequent tests. At the beginning of the experiment Participant 1, named Emma, was 5 years and 8 months old, Participant 2, named Theresa, was 6 years and 2 months old. The duration of the experiment was 88 days, one session per day per participant, and they participated about four days a week. The participants had no previous experimental experience. The reference number for the approval of this experiment from the Norwegian Social Sciences Data Service Approval was 31034. The participants were asked every day if they wanted to participate.

Apparatus and setting

The apparatus was a computer with a MTS program developed by Cognitive Science Partners Match to Sample 3.12 in collaboration with professor Erik Arntzen. The computer that was used was an HP EliteBook 8740w with an Intel Core i5 2, 40 GHz processor. The stimuli consisted of twelve abstract symbols and six familiar stimuli. The stimuli were presented on a white background on the computer screen. The familiar stimuli were colored. The height of the arbitrary stimuli varied between 0,8 cm to 3 cm, and in width from 1,2 cm to 2,3 cm. The familiar stimuli varied in height from 4,7 cm to 4,9 cm and in width from 2,6 cm to 4,9 cm. The distance between comparison to comparison horizontally was measured to 31 cm, while it was measured to 17 cm vertically. The distance from sample to comparison was 17,5 cm. See Figure 1 and 2. for an overview of the stimuli.

The programmed consequences presented by the computer were written words in Norwegian. The room had a pitched roof, and was measured to 2x3 m. From the floor to the roof in the pitched part, the height was measured to 1 m. At the highest point, the height was measured to 2, 40 m. The participant sat behind a table and the researcher sat next to the participant.

Design

Within participant manipulation replicated over two participants.

Measurements. The formula used to calculate the percentage of correct responses during the tests were (number of corrects x 100)/the total number of trials. The amount of correct in the directly trained trials was divided by the total number of directly trained trials, as were the same for the symmetry trials, transitivity/equivalence trials.

Sorting instructions. The sorting instruction that was used in the pre- and postcategorization involved asking the participants “Can you sort these, and let me know when you are finished”. With Theresa two other variations of the instructions were applied in the post categorization after the last overall tests. These included “Put them the way you think they should be”, and the last one was “Can you put these in different groups”. This was done to see if she would sort the stimuli differently in accordance with changing the instructions.

Procedure

General Information. The kindergarten staff was informed about the experiment at an information meeting before the experiment started. This information included what the children would be doing, and some general information about the duration of the experiment, and the purpose of the experiment. The parents received information about the experiment with the phone numbers to the experiments in case they had questions about the experiment. The parents also had to sign a consent form before the experiment started. See Appendix number 1.

Before the experiment started, the experimenters used a day with the children in the kindergarten. This day was used to play with the children, and attend their daily activities. The first day of the experiment began by informing the participants about the experiment. The participants were asked if they wanted to be a part of the experiment, and that they could ask questions if they wanted to, but that the experimenters might not be able to answer. This would not mean that the experimenters were angry, but they could not answer all questions.

The participants were told that they were going to do some tasks on a computer. They were notified that the experimenters did not know the duration of the experiment, but that the computer would let them know when they were finished. A collection of stickers and homemade books called the researcher book were shown to the participants prior to the startup, and they were told that they would get one sticker for every completed token economy board with ten crosses. The participants wrote their name on the outside and were told that they would collect stickers in this book. They were instructed that this book would stay in the kindergarten until they had finished the experiment. The experimenters made a list that contained each day the participants would be in the kindergarten for the upcoming week.

Each day the participants participated in the experiment they could write a cross next to their name. The participants were notified that if they participated every day the experimenters were present during one week, they would be allowed to attend something fun after lunch on Fridays. This activity was decided either on the Friday before or on Mondays, and the participants were told what the activity was. This could for instance involve watching a movie, popcorn party, youtube party, and so on.

General Procedure. The training structure used in this experiment was the MTO structure. The pauses were programmed to every 25 trials, and more often during tests. The participants were told that they could ask for extra pauses. Each session varied in duration, from 15 minutes to one and a half hour. The duration increased during the tests. Normally, one training session lasted for half an hour, and a test session lasted for an hour.

The participants were guided one by one to the experimental room and seated in front of a computer. When the participants were ready to start, the experimenter read the instructions presented on the computer screen. These instructions involved: a picture will appear in the middle of the screen. You shall click on this with a computer mouse. Three other pictures will appear on the screen. Choose one of these by clicking on it. If you choose the correct one, it will say good, super, and fantastic on the screen, which I will read aloud. If you click on the wrong one, the screen will say wrong. At the bottom of the screen you will see the number of correct. After a while the computer will not tell you if you choose the right or wrong. Do as best as you can to get all correct. Good luck!

The instructions were repeated if the participants said that they did not know what to do when the stimuli appeared on the screen, and throughout the experiment if the participant asked questions related to the task. A token economy system was used in addition to the programmed consequences from the computer program. The relation between the stimuli was initially arbitrary for both Emma, and Theresa. One block consisted of each trial type being presented five times. This applied for all of the blocks, except for the last overall tests, where each trial type was presented one time. The intertrial interval was fixed and set to 2000 milliseconds. This interval involved the time between the time the participants pressed the comparison stimulus, and until the new sample appeared. The duration of the programmed consequences was set to 1500 milliseconds.

Before starting the experiment, both participants did a precategorization of the stimuli in Stimulus Set 1. The stimuli lay in a pile, and the experimenter said “can you sort these, and let me know when you are done”. The experimenter took a picture of this categorization. At the startup, the sample stimulus was presented first successively in the middle of the screen. The participants had to click on this stimulus for three comparison stimuli to appear in the corners simultaneously with the sample stimulus in the middle. They had to use the mouse to choose one of the comparison stimuli. Programmed consequences were presented on the screen and read aloud by the experimenter. If they chose the correct comparison stimulus, the programmed consequences were “super”, “good”, “fantastic”, “brilliant”, “great” or “correct”, and they received a cross on the token economy board. If the participants made the wrong choice, the programmed consequence presented on the screen was “wrong”, and they did not receive a cross on the token economy board.

Emma. The stimuli used with Emma included pictures as C-stimuli, and F-stimuli. The rest of the stimuli were abstract, and consisted of Greek letters (see Figure 1). Stimulus Set 1 consisted of the three 3-member classes, ABC. Class 1 consisted of the A1, B1, and C1 stimuli, Class 2 of A2, B2 and C3, and Class 3 of the A3, B3, and C3 stimuli (see Table 1 for an overview of the procedure). The AC relations were presented first. The correct choice was defined by the presentation of A1 – choose comparison C1, A2 – choose C2, A3 – choose C3 and so on. When Emma achieved a mastery criterion of 95 % correct during one block, Emma would move on to the next phase in the experiment. The next phase involved training the BC relations. The same procedure was implemented with these relations, as with the AC relations. When Emma mastered this phase she moved forward to the mix phase. This phase involved that the AC and BC trials were presented interlaced. One block with AC or BC included 15 trials, while a mixed phase with AC/BC involved 30 trials per block. The next phase included the AC/BC relations with 100 % programmed consequences, next the programmed consequences was faded to 75 %, 25 % and then 0 %. The probability of receiving the programmed consequences was calculated within each trial. The token economy was removed in the first break after starting with 0 % programmed consequences. Emma was instructed that she would collect stickers in each break. [1]

After the phases with faded programmed consequences there was a test for derived relations in addition to the directly trained relations. The tests involved no programmed consequences. The tested trial types were AB, AC, CB, CA, BA and BC in each of the three classes, thus a total of 18 relations were tested. If Emma did not respond in accordance with the 95 % accuracy criterion, Emma started cycle 2, which involved the mix phase AC/BC, 100 %, 75 %, 25 % and 0 % fading of programmed consequences, and then she had a new test. After this Emma had two more tests. After the fourth and last test for Stimulus Set 1, Emma was asked to do a post categorization of the stimuli and a picture was taken. Training and testing were not always conducted over consecutive days. The time between each session varied from a day and up to four days at the most.

The next phase was to train Stimulus Set 2 with different stimuli. This stimulus set involved the three 3-member classes DEF. Class 1 included D1, E1, and F1, Class 2 included D2, E2, and F2 and Class 3 involved D3, E3, and F3 stimuli. The same procedure was applied with this set as with Stimulus Set 1. Due to a mistake, there was no post categorization of

[1] Emma had four trials into 75 % fading during Stimulus Set 1 before easter. After easter she was retrained in 100 % fading before moving forward to the next phases in the experiment as Due to an experimental error she had eight trials into a new test after the third test in this set, thus a total of 98 test trials in test three.

Stimulus Set 2. The next phase involved connecting the two stimulus sets. This involved training the CF relations. This phase included training only, and Emma received 100 % programmed consequences. The next phase was to retrain Stimulus Set 1 and Stimulus Set 2. The sets were trained separately with 100 % programmed consequences. The last phase included three overall tests consisting of all the possible relations between all the A, B, C, D, E and F stimuli in each of the three classes. If Emma formed the three 6-member classes this would involve that Class 1 consisted of A1, B1, C1, D1, E1 and F1, Class 2 of A2, B2, C2, D2, E2, AND F2, and Class 3 consisted of A3, B3, C3, D3, E3, and F3. The tested directly trained relations included AC, BC, DF, EF and CF. The test of the possible emerged relations were AB, AD, AE, AF, BA, BD, BE, BF, CA, CB, CD, CE, DA, DB, DC, DE, EA, EB, EC, ED, FA, FB, FC, FD and FE. After the overall tests Emma was asked to post categorize the stimuli and a picture was taken. The post categorization was done four days after the final tests. The multiplier was set to one during the overall tests, thus the trial types were presented one time each in one test. This was done to ensure that the amount of trials presented in one day were manageable for the participants. Due to a mistake by the experimenter, the multiplier was set to five in the first overall test, and Emma completed 127 trials. During this test some of the relations were presented more and some none at all. Emma completed the second test during one day, but had to do the third test over two consecutive days. This was done because Emma was too tired to complete the first day. None of the tests included testing for reflexivity, only symmetry, directly trained relations, and equivalence relations.

Theresa. The stimuli used in this experiment were all abstract and the relations between the stimuli were arbitrary (see Figure 2). The procedure was the same as with Emma (see Table 1). Due to an error in the rollback function in the MTS program, Theresa was presented with the BC phase after having had one block with the mix phase AC/BC in training Stimulus Set 1. This happened after 255 trials, and between two consecutive days.

There were two blocks of BC before Theresa was back in the mix phase. Theresa did not do a pre categorization before starting the training in Stimulus Set 2. Another rollback happened during training in Stimulus Set 2. Theresa stepped back from 75 % fading of programmed consequences to 100 % fading. The trials during the rollbacks were not graphed. There was one and a half block of 100 % fading before returning to 75 % programmed consequences.

The experimenters decided to start errorless training after Theresa had consistently chosen the same comparison stimulus, C1, in the presence of the sample stimulus, B3 for several blocks. The errorless training involved first training the AC relations involving one comparison. Each of the A stimuli were presented successively, and only one of the C-stimuli was presented as comparison in one of the corners. The next phase was to train the BC relations in the same way. This procedure was repeated. All of the AC relations were presented one time each, with one comparison, then the BC relations with one comparison. The next phase was AC with two comparisons, then BC with two comparisons. Subsequently, the AC and BC relations were mixed, with two comparisons. The last two phases of the errorless training included AC, then BC, with three comparisons. After this the AC and BC relations were mixed with the multiplier set to five. After this Theresa went through the same phases as Emma.[2]

Results

Emma. As shown in Figure 3, Emma had 554 training trials before the test in Stimulus Set 1. The first test showed that the directly trained relations were 96,7 % correct, symmetry trials were 93,3 % correct, and equivalence relations included 100 % correct responses (see Table 2 for the results for Stimulus Set 1 and 2). In the second test Emma showed 100 % correct responding to the directly trained relations, 96,7 % to the symmetry relations, and equivalence relations were 90 % correct. The third test showed that the directly

[2] After the first test in Stimulus set 2, Theresa had one trial in errorless training. Some blocks involved an error in the multiplier, this applied for both participants.

trained relations were 94 % correct, symmetry relations were 96,7 % correct, and equivalence relations were 100 % correct. In the fourth test the directly trained relations implicated 100 % correct responses; symmetry relations were 96,7 % correct, and equivalence 96,7 % correct. Emma showed responding according to stimulus equivalence after the fourth test in Stimulus Set 1, and formed the three 3-member classes (ABC). She also sorted the stimuli according to the experimenter-defined classes after the tests.

Emma had 185 training trials before the test in Stimulus Set 2 (see Figure 3). The first test resulted in 93,3 % correct responses to the directly trained relations; symmetry relations were 90 % correct, and equivalence 86,7 % (see Table 2). The second test resulted in 100 % correct responding to the directly trained, symmetry, and equivalence relations. She responded according to stimulus equivalence after the second test in Stimulus Set 2, and formed these three 3-member classes (DEF). Due to an experimental error there was no post categorization after Stimulus Set 2.

As presented in Table 3, the amount of correct responses in the directly trained relations were 48 % in the first of the overall tests, 73,3 % in the second test, and 93,3 % correct in the last overall test. Correct responding in equivalence trials was 40 % during the first test, 61,7 % during the second, and 70 % in the last test. The amount of correct responses in symmetry trials was 31,8 % in the first test, 73,3 % in the second test, and 86,7 % in the last test. Emma showed 40 % correct responding according to stimulus equivalence. In the second test, she had 64,4 % correct responses according to stimulus equivalence. In test 3, she showed 76,7 % correct responding to stimulus equivalence.

During the first overall test, Emma responded more correctly in the first part of the test, and had a lot more wrong responses in the last part of the test, especially after 35 trials. In the first 35 trials she had 28 correct responses, while after 35 trials she had 69 wrong responses, and 23 correct. In the second test the allocation of correct and wrong responses were more equal throughout the test. In the third, and final overall test, Emma had more wrong responses after 65 trials. She had 15 correct responses and 10 wrong responses after 65 trials. Thus, the 12 remaining wrong responses were allocated to the first 65 trials.

In the first overall test Emma was not presented with all of the trial types, and some of the trial types were presented more than once. After these final three overall tests, she sorted the stimuli according to the defined three 6-member classes. Figure 5 displays the class formation for the three last overall tests. In the first overall test Emma showed 54 % formation of Class 1, 47 % of Class 2, and 25 % of Class 3. In the second test she showed 60 % class formation of Class 1, 73 % of Class 2, and 60 % of Class 3. In the third and final overall test she showed 93,3 % class formation of Class 1, 67 % of Class 2, and 70 % of Class 3.

Theresa. See Figure 4 for an overview of the responding during training. Theresa had 465 training trials before the first test in Stimulus Set 1. Results from the first test in Stimulus Set 1 showed that the percentage of correct for the directly trained relations were 73,3 %, for the symmetry relations it was 80 %, and 50 % for the equivalence relations (see Table 2 for the results for Stimulus Set 1 and 2). The second test showed that Theresa responded correct to 96,7 % of the directly trained trials, 100 % correct to the symmetry trials, and 96,7 % correct to the equivalence trials. Thus, Theresa responded according to stimulus equivalence after the second test in Stimulus Set 1.

Theresa had 642 training trials before the test in Stimulus Set 2 (see Figure 4). The dashed lines between block 46 and 53 indicate the blocks for the errorless training. These responses were not graphed. The first test in Stimulus Set 2 resulted in 100 % correct responding to the directly trained relations, symmetry relations were 93,3 % correct, and equivalence showed 93,3 % correct responding (see Table 2). In the second test she showed that the directly trained relations were 93,3 % correct, symmetry relations were 80 % correct, and equivalence relations were 90 %. The third and last test in Stimulus Set 2 resulted in 100 % correct responding to the directly trained relations, symmetry relations were 100 % correct, and equivalence relations were 96,7 % correct. The test was divided in two, and these results were from the first 45 trials. See Table 3 for an overview of the final results for the three last overall tests. These tests showed that Theresa responded correct to 84,6 % of the directly trained relations in the first test, 80 % in the second test, and 86,7 % in the last test. When the symmetry relations were presented, she responded correct to 84,6 % in the first test, 66,7 % in the second test, and 66,7 % in the last test. The equivalence relations showed that she responded correct to 32,8 % in the first test, 43,3 % in the second test, and 38,33 % correct in the last test. Responding according to stimulus equivalence involved 47,8 % correct in the first test, 54,4 % correct in the second test, and 50 % in the last test.

During the first overall test Theresa had more wrong responses in the last part of the test. She had 27 correct responses and 18 wrong responses during the first 45 trials, while during the last 45 trials she had 16 correct responses and 29 wrong responses. In the second test she had 25 correct responses and 20 wrong responses during the first 45 trials. The last 45 trials, she had 24 correct responses and 21 wrong responses. In the third test she had 25 correct responses and 20 wrong responses in the first 45 trials and 20 correct responses and 25 wrong responses in the last 45 trials.

During the first of the three overall tests Theresa showed 58 % class formation of Class 1, 43 % of Class 2, and 36 % in Class 3 (see Figure 5). The second test showed 57 % in Class 1, 50 % in Class 2, and 53 % in Class 3. The third and final test showed 43 % class formation in Class 1, 57 % in Class 2, and 50 % in Class 3.

Theresa did not sort the stimuli according to the defined stimulus classes after the tests in Stimulus Set 1, Stimulus Set 2, or after the overall tests. A variation of two other sorting instructions (as described in the method section) was used after the overall tests, but she still did not sort the stimuli according to the defined classes.

Discussion

The research question was whether children in the age of 5–6 years old would be able to form three 6-member equivalence classes by using all visual stimuli with an MTO training structure, and training one relation between two trained stimulus sets. The results showed that both of the participants formed the three 3-member classes after the tests in both Stimulus Set 1, and Stimulus Set 2. Emma had four tests in Stimulus Set 1 before she reached the 95 % mastery criterion, and Theresa had two tests. Emma had two tests in Stimulus Set 2 before she reached the 95 % criterion, and Theresa had three. In the final overall tests for the three 6-member classes, none of the participants responded according to stimulus equivalence. Thus, they did not form the three 6-member classes in these tests for emergent relations. However, Emma sorted the stimuli according to these defined classes, and demonstrated an increase in the number of correct responding with extended testing.

In the first of the overall tests Emma showed 40 % stimulus equivalence formation, while the third, and last test showed that it had increased to 76,7 % formation (see Table 3 for an overview). During the last three tests the amount of correct responses to the directly trained relations increased from 48 % to 93,3 %, symmetry relations increased from 31,8 % to 86,7 %, and the equivalence relations increased from 40 % to 70 %. Emma sorted the stimuli according to the defined three 6-member classes in the post categorization, which may suggest that she might have shown responding according to stimulus equivalence if extended testing had been employed.

During the first test after combining the two stimulus sets Emma had 48 % correct responses to the directly trained relations, thus only half of the relations were correct. Sidman et al. (1985) found that when the 6-stage relations did not emerge, the prerequisite relations were also deficient, including the lower stage relations. They found that the prerequisite relations and the derived relations emerged with repeated testing. The results from Emma replicated the findings from Sidman et al. The relations were deficient in the first test, but started to emerge with repeated testing. Furthermore, Sidman and his colleagues suggested that the classes did not exist until they were tested, and that training some of the relations only provides a potential for class formation. They suggested that this was one of the reasons that extended testing had a positive effect on equivalence class formation. This also seems consistent with the data from Emma. Class formation for Class 1 showed an increase from 54 % to 93,3 %. Class 2 increased from 47 % to 67 %, and Class 3 increased from 25 % to 70 %. See Figure 5 for an overview.

As displayed in Figure 3, Emma had 554 training trials in Stimulus Set 1 compared to only 185 training trials before the test in Stimulus Set 2. Thus, she acquired the baseline conditional discriminations quicker in Stimulus Set 2 than Stimulus Set 1. Arntzen et al. (2010) conducted an experiment on the effects of training with the different training structures. The results from their experiment demonstrated a lower number of responses to criterion when the participants trained with a new stimulus set, after having been exposed to another stimulus set first. These results suggest that pre-exposure to another stimulus set may result in quicker acquisition of a new stimulus set.

Fields et al. (2012) conducted an experiment where they investigated the use of both an emergent relations test, and a sorting test. They found that when the participants formed the stimulus classes during the emergent relations test, they also sorted the stimuli according to these defined classes. However, for some of the participants there was no correlation between the results from the sorting task and the emergent relations test. Some of the participants sorted some of the classes correctly, but did not respond correctly in the test for these emergent relations.

Fields, Arntzen and Moksness (In press) conducted an experiment to investigate how equivalence class formation may be displayed by using a sorting task and a regular MTS test of the derived relations. They found that the sorting task appeared to be more sensitive in tracking class formation than the MTS test. They suggested that the sorting test might be an alternative to use in tracking equivalence class formation. Their findings demonstrated that there was a correspondence between responding according to the defined classes in the MTS test and the sorting test. They also found that some participants did not respond according to the defined classes in the MTS tests, but that some of the classes were revealed during the sorting task. The results from Fields et al. (2012), Fields et al. (In press), and the results from Emma in the present experiment may suggest that the sorting task may be a more sensitive measure in tracking class formation.

A question is whether the sorting task shows something else than the emergent relations test. The Sidman equivalence has been defined as responding to the three properties of reflexivity, symmetry, and transitivity (e.g., Sidman & Tailby, 1982). The sorting task may be a measure of class formation on another level, thus these two tests for class formation may not be the same. A sorting task may be used as an extra measurement of class formation in addition to the emergent relations test. However, the sorting task is different from the emergent relations test/MTS test in that all of the stimuli are present at the same time in the sorting task, and it may be easier to discriminate between the stimuli when all of them are presented simultaneously.

Theresa did not form the three 6-member equivalence classes and the results from the tests suggest that she may not have formed the classes with extended testing. Theresa had 465 training trials in Stimulus Set 1 compared to 642 training trials in Stimulus Set 2. The reason for the increased number of trials was that she consistently chose the same wrong comparison, and had to go through errorless training before moving on in the experiment. As displayed in

Figure 3 and 4, Theresa used longer time than Emma, and had an unstable baseline performance. This may suggest that her responding took longer to come under the control of the stimuli in the experiment. This can also be seen in Figure 1 and 2. Emma's responding is more stable, and she acquires the relations in Stimulus Set 2 very quickly. Theresa on the other hand uses longer time than Emma in acquiring Stimulus Set 2.

Theresa showed 47,8 % formation of the three 6-member classes during the first of the overall tests, and 50 % formation in the last test (see Table 3). The directly trained relations started with 84,6 % correct, and in the last test she had 86,7 % correct. The symmetry relations involved 84,6 % correct responses in the first test, and 66,7 % in the last test. The equivalence relations were 32,8 % correct in the first test, and 38,3 % correct in the last test. Theresa did not show any increase in the formation of any of the classes either (as displayed in Figure 5). These results suggest that extended testing did not have a positive effect on equivalence class formation for Theresa. A possibility would have been to retrain the CF relations and then test again to see whether she would have responded differently. However, Theresa had all abstract stimuli in her two stimulus sets as opposed to Emma that had familiar c- and f-stimuli.

Some experimenters have investigated the use of pictures among abstract stimuli in the formation of stimulus classes (e.g., Arntzen & Lian, 2010; Arntzen & Nikolaisen, 2011; Fields et al., 2012). Arntzen and Lian (2010) had 16 children as participants and used an MTO training structure with three 3-member classes in two stimulus sets. The procedure involved that the first group received training with a stimulus set with three 3-member classes where all of the stimuli were abstract. Subsequently, this group trained a new stimulus set with three 3-member classes with abstract stimuli, and familiar pictures as nodes. The other group had these conditions reversed. The results demonstrated that the use of pictures among abstract stimuli were more effective in training, and gave higher yields of equivalence responding than using all abstract stimuli.

Fields, Arntzen, Nartey, and Eilifsen (2012) did a similar experiment with adults. The stimuli consisted of the potentially three 5-member classes with the stimuli A-E. The results indicated that using meaningful stimuli in a set with meaningless stimuli would influence how the stimulus set would convert to an equivalence class. By involving meaningful stimuli in a set of meaningless stimuli 8 out of 10 participants responded according to equivalence, against 10 participants with all abstract stimuli that did not respond according to equivalence.

Arntzen and Nikolaisen (2011) had participants in the age of 8–9 years old and found results that also replicated the findings from Fields et al. (2012). They found that the use of familiar stimuli among abstract stimuli were more effective in equivalence formation than when the stimulus set consisted of all abstract stimuli. In this present experiment Emma had familiar c- and f-stimuli, and also showed a higher percentage of correct responses to stimulus equivalence than Theresa that had all abstract stimuli. If Theresa would have had pictures as c- and f-stimuli this may have given another outcome on the tests for stimulus equivalence.

Theresa said she did not understand the original sorting instructions, so the instructions were changed after the overall tests. The instructions changed to “Put them the way you think they should be”, and the last one was “Can you put these in different groups”. These instructions did not seem to change how she categorized the stimuli. The fact that she did not categorize the stimuli after the overall tests according to the defined classes supported the results from the emergent relations test that she had not formed the three 6-member equivalence classes, and may not have with extended testing. However, she did respond according to equivalence in Stimulus set 1 and 2, but did not sort the stimuli into these defined classes either. It would have been interesting to see if she would have categorized the stimuli correctly if the other two sorting instructions had been applied for the first two stimulus sets.

In the study by Sidman et al. (1985) the number of tests varied from 27 to 42 in Experiment 1 and 3. Each test consisted of 27 trials in addition to the baseline trials, while this experiment had only three tests consisting of 90 trials a total per test. In the overall tests for the present experiment all of the relations were mixed, as opposed to being presented one by one as in the Sidman et al. study. The total amount of test trials in the Sidman et al. study was a lot more than the number of test trials employed in the present experiment. Furthermore, as described in the introduction, Sidman and his colleagues had all the participants review the baseline relations (with reinforcement) before every test session. If this had been employed before each of the three overall tests in the present experiment the results may have been different. The use of reinforcement before each test session may be one way to keep the participants at this age interested, because the participants in this experiment expressed that they thought it was fun when they received reinforcement on the trials, and boring when they did not. This may also have implications relative to the possibility to prolong the duration of the experiment when the participants are children.

There was no particular design used in this experiment, this was because the experiment does not per definition have a baseline phase before the intervention starts. However, the pre categorization before training, and post categorization after training could be seen as a control for class formation for Emma. Due to the fact that Theresa said that she did not understand the sorting instructions, these categorizations could not be used as a control for class formation. However, both Figure 3 and 4 demonstrate that the stimulus sets are not formed before the training started, and are acquired during training. Also, all of the stimulus sets were retested, and Table 2 and 3 show that the conditional relations increased with extended testing, and the acquired relations were maintained throughout the tests.

According to Saunders and Green (1999) the training structure MTO presents all of the simple discriminations involved in each trial type that involves equivalence and symmetry relations. However, in this current experiment the training structure changed after having trained the two stimulus sets. When the stimulus sets were combined before the overall tests, the only relations that combined them were the CF relations. This involves that many of the stimuli in Stimulus Set 1 were not trained directly to the stimuli in Stimulus Set 2. The training structure involved training of the AC, BC, CF, DF, and EF relations. This means that the relations AD, AE, BE, BD, EA, DA, DB and EB are combined by two-nodes, and that this form of training structure may not be the easiest when forming larger equivalence classes with children. Sidman et al. (1985) suggested that a parameter that would be of importance when enlarging classes embrace how many members of stimuli that can be added to a stimulus class at one time. The results from Goyos (2000), Lazar et al. (1984), and Sidman et al. suggest that adding one member at a time and testing for class formation after each new trained relation, may give higher yields of equivalence responding than by combining stimulus sets before testing for all of the relations at one time.

Lazar et al. (1984) proposed that the role of testing in teaching should be clarified. This present experiment should be replicated with a focus on extended testing, and with more participants. As described earlier, Sidman et al. (1985) employed many more test trials compared to in this present experiment, and because Emma demonstrated an increase in her correct responding with extended testing, this should be a focus in further research. Another possibility could involve as Sidman and his colleagues posted: to train and test one stimulus at a time to an existing set of stimuli. Such an experiment could be another systematic replication without the naming procedures, and with an MTO training structure. The procedure could have included training AC and BC before testing for these 3-member classes, and then train CF before testing for 4-member classes, train DF, test for 5-member classes, train EF, and then test for 6-member classes.

Further research should focus on comparing different testing procedures with children to see what would be the most effective method to expand stimulus classes. An experiment could have included comparing two groups of participants. The first group could have the same training and testing procedure as in the present experiment, and the other group could have had the same training procedure as the first group, and a similar testing procedure to the one employed in the study by Sidman et al. (1985). Such an experiment may have demonstrated interesting results on the role of testing in forming expanded stimulus classes with children.

Another important suggestion concerns the validity of using a sorting task as a second measure of stimulus equivalence with children. It is necessary to find out what the sorting instructions have to include for children to be able to categorize the stimuli correctly. This could have been conducted by employing different sorting instructions with a group of children, by using familiar pictures where the stimuli should be matched based on identity. Another suggestion involves examining whether the application of a sorting test between the consecutive tests would influence equivalence responding in the later tests. It would have been interesting to find out if the sorting tests could have caused the relations to emerge quicker compared to only applying extended tests.

In summary both of the participants responded according to stimulus equivalence in the tests for Stimulus set 1 and Stimulus Set 2, but none of them responded according to equivalence for the three 6-member classes. However, Emma may have continued to demonstrate an increase in correct responding if extended testing had been employed, and may have responded according to stimulus equivalence with the extended tests, which may be supported by the correct sorting test. Theresa on the other hand did not show an increase in correct responding, suggesting that she probably would not have responded according to the defined classes if extended testing had been employed. From the two stimulus sets, 12 of the relations had been directly trained, while 18 relations emerged. This created a total of 30 conditional relations. Expansion by only training one relation between two stimulus sets may not be the best way in order to create larger equivalence classes.

References

Arntzen, E. (2010). Om stimulusekvivalens. In S. Eikeseth & F. Svartdal (Eds.), *Anvendt atferdsanalyse. Teori og praksis.* (2 ed., pp. 100–138). Oslo: Gyldendal Norsk Forlag AS.

Arntzen, E., Grondahl, T., & Eilifsen, C. (2010). The effects of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *The Psychological Record, 60*, 437–462. Retrieved from http://opensiuc.lib.siu.edu.

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

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

Arntzen, E., & Holth, P. (2000). Probability of stimulus equivalence as a function of class size vs. number of classes. *The Psychological Record, 50*, 79–104. Retrieved from http://opensiuc.lib.siu.edu.

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

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. Retrieved from http://www.ejoba.org.

Arntzen, E., & Vaidya, M. (2008). The effect of baseline training structure on equivalence class formation in children. *Experimental Analysis of Human Behavior Bulletin, 26*, 1–8. Retrieved from http://www.eahb.org/.

Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 46*, 243–257. doi: 10.1901/jeab.1986.46-243

Dube, W. V., McIlvane, W. J., Maguire, R. W., Mackay, H. A., & Stoddard, L. T. (1989). Stimulus class formation and stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior, 51*, 65–76. doi: 10.1901/jeab.1989.51-65

Fields, L., Arntzen, E., & Moksness, M. (In press). Stimulus sorting: A quick and sensitive index of equivalence class formation. *The Psychological Record*. Retrieved from http://opensiuc.lib.siu.edu/.

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*, 163–181. doi: 10.1901/jeab.2012.97-163

Fields, L., Newman, S., Adams, B. J., & Verhave, T. (1992). The expansion of equivalence classes through simple discrimination training and fading. *The Psychological Record, 42*, 3–15. Retrieved from http://www.ebscohost.com/public/psychology-behavioral-sciences-collection.

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior, 48*, 317–332. doi: 10.1901/jeab.1987.48-317

Goyos, C. (2000). Equivalence class formation via common reinforcers among preschool children. *The Psychological Record, 50*, 629–654. Retrieved from http://opensiuc.lib.siu.edu/.

Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of Reseach Methods in Human Operant Behavior* (pp. 229–262). New York: Plenum Press.

Hove, O. (2003). Differential probability of equivalence class formation following a one-to-many versus a many-to-one training structure. *The Psychological Record, 53*, 617–634. Retrieved from http://opensiuc.lib.siu.edu/.

Lazar, R. M., Davis-Lang, D., & Sanchez, L. (1984). The formation of visual stimulus equivalences in children. *Journal of the Experimental Analysis of Behavior, 41*, 241–266. doi: 10.1901/jeab.1984.41-251

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

Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of Experimental Analysis of Behavior, 71*, 195 – 214. doi: 10.1901/jeab.1999.71-195

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

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior, 50*, 145–162. doi: 10.1901/jeab.1988.50-145

Schenk, J. J. (1994). Emergent relations of equivalence generated by outcome specific consequences in conditional discriminations. *The Psychological Record, 44*, 537–558. Retrieved from http://opensiuc.lib.siu.edu/.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*, 127 – 146. doi: 10.1901/jeab.2000.74-127

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

Skinner, B. F. (1935). The generic nature of the concepts of stimulus and response. *Journal of General Psychology, 12*, 40–65. doi: http://dx.doi.org/10.1080/00221309.1935.9920087

Table 1.

Overview of the Procedure

Phase	Emma	Theresa
Training Set 1		
AC, BC, AC/BC, TPC: 100 %, 75 %, 25 % and 0 %	yes	yes
Test: AC, BC, BA, AB, CA, and CB	yes	yes
Cycle 2: training AC/BC, 100 %, 75 %, 25 % and 0 %	yes	yes
Test: AC, BC, BA, AB, CA, and CB	yes	yes
Repeated tests	2	-
Training Set 2		
DF, EF, DF/EF, TPC: 100 %, 75 %, 25 % and 0 %	yes	yes
Errorless training	-	yes
Test: DF, EF, FE, FD, DE and ED	yes	yes
Cycle 2: training DF/EF, 100 %, 75 %, 25 % and 0 %	yes	yes
Test: DF, EF, FE, FD, DE and ED	yes	yes
Repeated tests	-	1
Training: CF	yes	yes
Retraining Set 1: AC, BC, AC/BC	yes	yes
Retraining Set 2: DF, EF, DF/EF	yes	yes
Final overall tests:	3	3
AB, AD, AE, AF, BA, BD, BE, BF, CA, CB, CD, CE, DA, DB, DC, DE, EA, EB, EC, ED, FA, FB, FC, FD and FE		

Note: TPC = thinning of programmed consequences.

Table 2

Results for Stimulus Set 1 and Stimulus Set 2

Test	Stimulus Set 1			Stimulus Set 2		
	DT	SYM	EQ	DT	SYM	EQ
1	96,7	93,3	100	93,3	90	86,7
2	100	96,7	90	100	100	100
3	94	96,7	100			
4	100	96,7	96,7			

Test	Stimulus Set 1			Stimulus Set 2		
	DT	SYM	EQ	DT	SYM	EQ
1	73,3	80	50	100	93,3	93,3
2	96,7	100	96,7	93,3	80	90
3				100	100	96,7

Note: All of the results are in percent. The upper panel shows Emma's results, and the lower panel shows Theresa's results. DT = directly trained, SYM = symmetry, EQ = equivalence.

Table 3

Results for the Three Overall Tests

Test	DT	SYM	EQ	SE
1	48	31,8	40	40
2	73,3	73,3	61,7	64,4
3	98,3	86,7	70	76,7

Test	DT	SYM	EQ	SE
1	84,6	84,6	32,8	47,8
2	80	66,7	43,3	54,4
3	86,7	66,7	38,3	50

Note. DT = directly trained. SYM = symmetry. EQ = equivalence. SE = stimulus equivalence. The results are in percent. The upper panel shows the results from Emma, and the lower panel shows the results for Theresa.

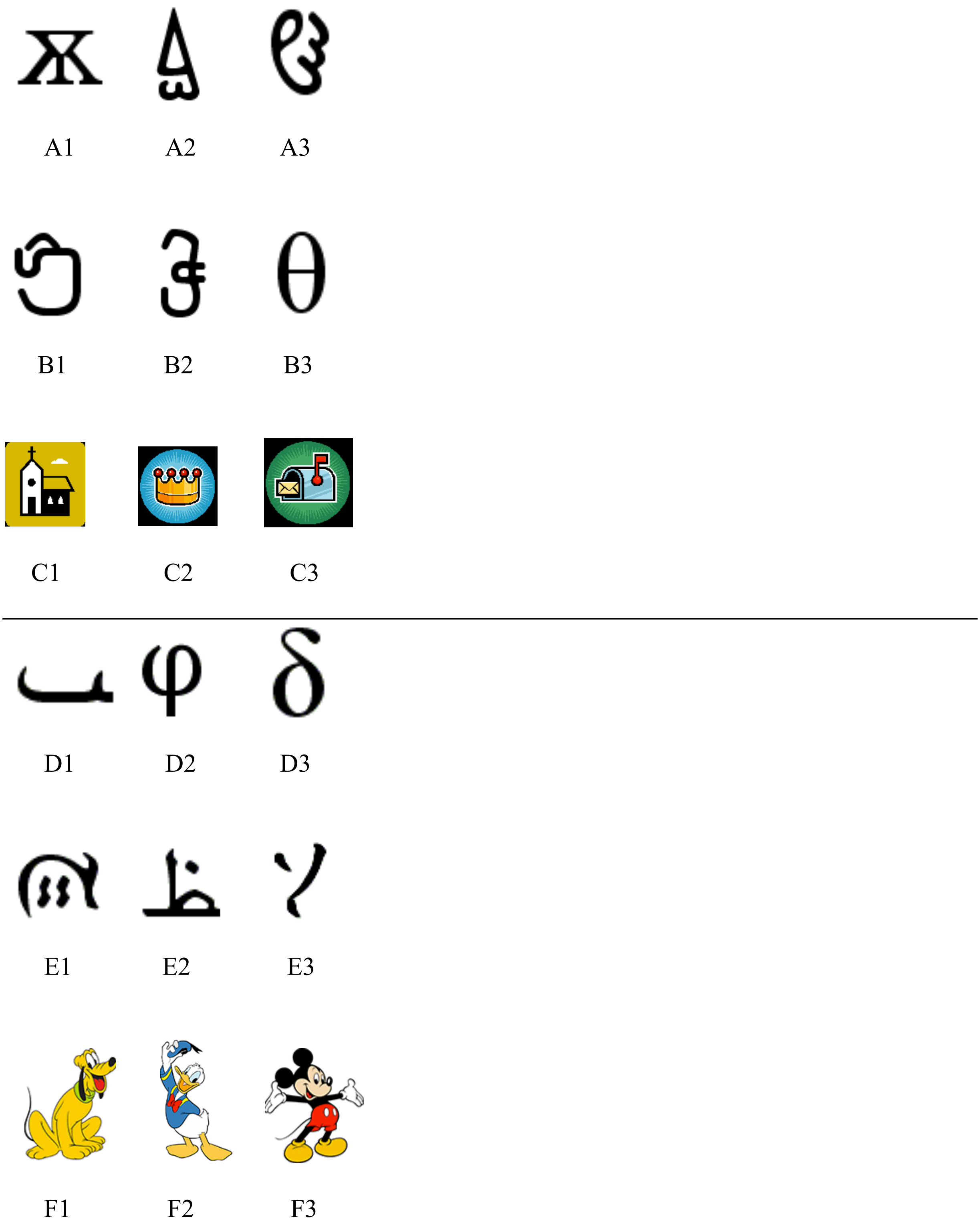

Figure 1. The upper panel shows Stimulus Set 1 for Emma. The lower panel shows Stimulus Set 2. The different classes are shown horizontally and the different members of each class are shown vertically. The letters indicate the different members and the numbers indicate the different classes.

Figure 2. The upper panel shows Stimulus Set 1 for Theresa. The lower panel shows Stimulus Set 2. The different classes are shown horizontally and the different members of each class are shown vertically. The letters indicate the different members and the numbers indicate the different classes.

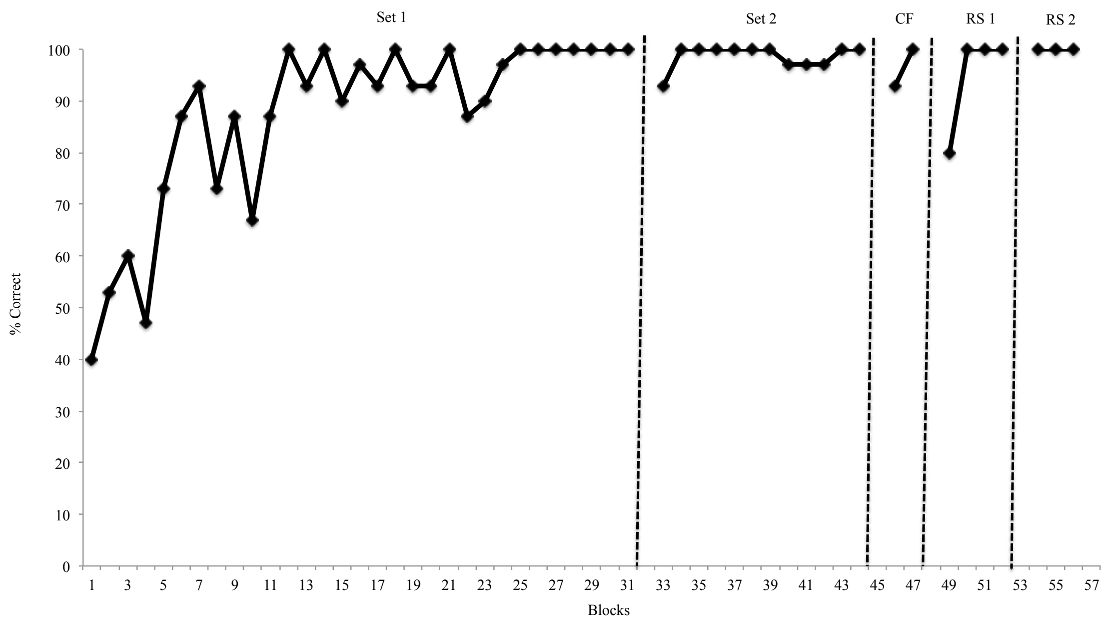

Figure 3. This figure shows the percentage of correct responses for Emma on the y-axis, and number of bloks on the x-axis. CF = training of CF relations. RS1 = retraining of Stimulus Set 1. RS2 = retraining of Stimulus Set 2.

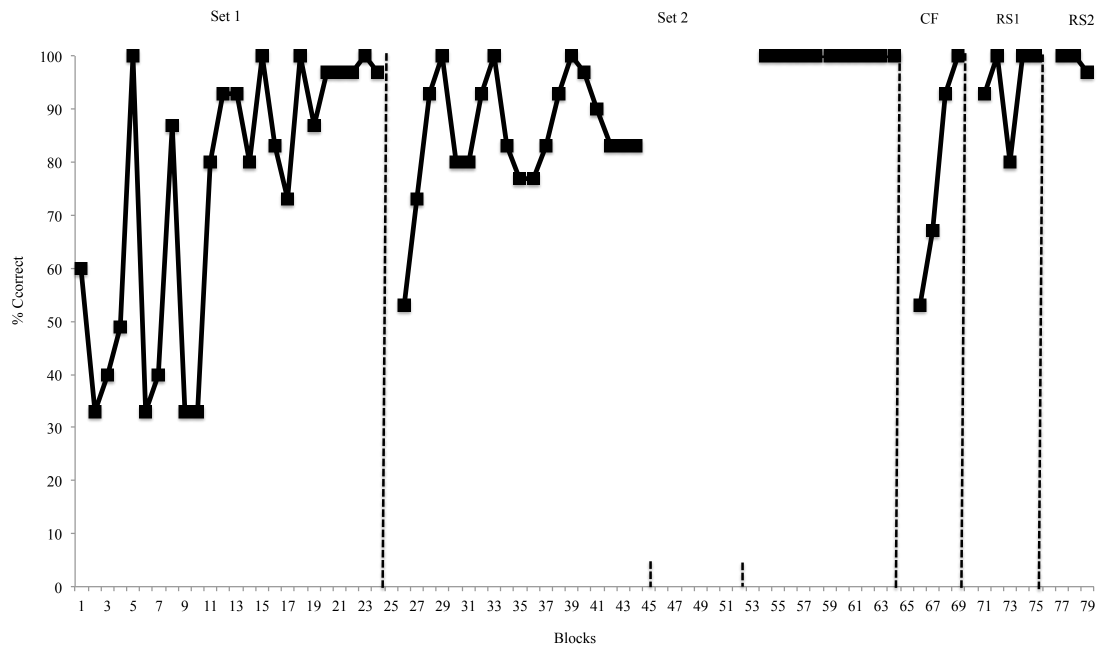

Figure 4. This figure shows the percentage of correct responses for Theresa on the y-axis, and number of bloks on the x-axis. CF = training of CF relations.

RS1 = retraining of Stimulus Set 1. RS2 = retraining of Stimulus Set 2. The dashed lines between blocks 46–53 indicate the errorless training blocks.

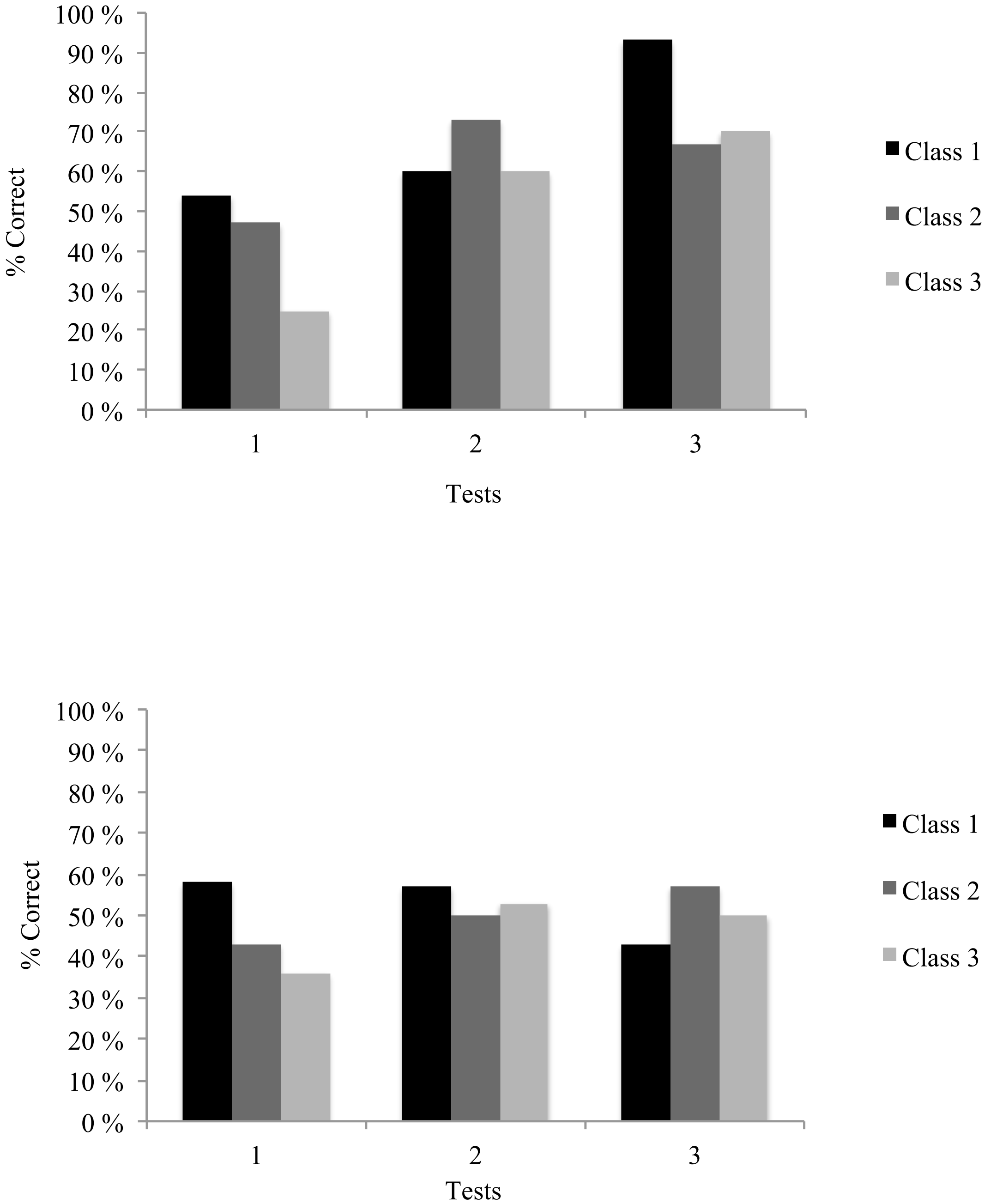

Figure 5. The upper panel shows class formation in percent for Emma. The lower panel shows class formation for Theresa.

Request about participation in the research project

Symbolic functions with children

Background and purpose

This is a request to ask whether your child may participate in a research project within studies on symbolic functions or what has been designated as stimulus equivalence. The project starts during the autumn of 2012, and the collection of data will end spring 2013. The project period will be over 4 years. The study will be based on existing research conducted in our research group and international research.

Stimulus equivalence is a research area within learning psychology, memory and problem solving. A big part of the research in this area has now been conducted with adults, but it is of considerable interest to uncover which variables influence these phenomena's with children. In this study we wish to investigate how learning processes may be influenced with children.

Implementation of the experiment and the person responsible

PhD- and graduate students from Oslo and Akershus University College (OAUC) by the institute of behavioral science will be collecting the data during the project period. Professor Erik Arntzen by OAUC will be responsible for the project that is also part of a bigger research project by the University College. If guardians have other questions considering this project they may contact professor Erik Arntzen (erik.artnxen@equivalence.net, PhD student Hanna Steinunn Steingrimsdottir on tlf. (47907208 and email hannasteinunn@simnet.is, or gradutate student Marie Moksness (41619278 and email mairemoksness@hotmail.com.

What does the study involve?

The experiments will be conducted in the kindergarten and involves that the child will solve tasks on a computer. There are no required prerequisites to use the computer in order to be able to solve the tasks. The experiment has a training phase and a test phase. During training there are different symbols/signs that will appear on the screen that the child should press. Mastery of the assignment does not require any pre knowledge of these signs. The same signs will be used during the test phase, but now in other locations on the screen. The leader of the research will not be present with the child all of the time to be of least disturbance, however he or she will be available instantly, and take regular trips to see if the child is ok and how the tasks are solved. The duration of the experiment will vary depending on how the child solves the tasks, but for about 3 hours. It will be given opportunities to take breaks and/or divide the experiments into shorter sessions over more days.

Possible advantages and disadvantages

The experiment will be conducted in a calm and safe environment in familiar surroundings, and there is no form of discomfort in the implementation of the experiment. Professor Erik Arntzen has long experience with implementation of similar studies and the people that conduct these experiments are specially trained to do them. By the end of the experiment the participant will have an oral review of the tasks and feedback on how he or she has solved them. The parents may have an oral review of the results for their child, and a Norwegian article on stimulus equivalence by contacting the PhD fellow Hanna Steinunn Steingrimsdottir.

What happens with the information about you?

The information registered about the child and their results should only be used as described in the purpose of the study. All information will be treated without names or other recognizable details. Consent forms will be deleted at the end of the project at the latest. It will not be possible to identify the child in the results of the study in the case of publication.

Voluntary participation

It is voluntary to participate in the study. It is possible to withdraw from the study at any time without any particular reason. This will not result in any consequences. It you agree that your child may participate in this project, you may sign to the consent form on the last page. If you say yes to participate, you may withdraw your consent later. If you would like to withdraw from the study at a later time or have any questions about the project you may contact.

Hanna Steinunn Steingrimsdottir
Høgskolen i Oslo og Akershus, institutt for atferdsvitenskap
Telefon 47 90 72 08 eller e-post adresse hannasteinunn@simnet.is

Policy

Information registered about the child is

- age
- sex
- how the child solves the task, for instance which signs the participant presses, how long time it takes before he or she presses the different signs, number of repetitions, and if the child has learned more than what was trained directly.

Oslo and Akershus University College by the principal is the computer controller.

Permission to viewing and deleting information about the child

If you say yes to participate in this study you have the right to get a view of what information that has been registered. Furthermore, you have the right to correct any wrong information in what we have registered. If you would withdraw from the study, you have the possibility to demand that all information are deleted, unless the information already has been used in analyzes or scientific publications.

Economy

The study is financed by Oslo and Akershus University College.

Information about the results of the study

Beyond the information about each participant's results, the parents may receive a copy of any future publications about the results by contacting Hanna Steinunn Steingrimsdottir.

Consent to participation in the study

I/we consent that our child ______________________________________,
(The child's name)

age ___ and _____, participate in a research project about stimulus equivalence that is described above
(Years) (Months)

I hereby confirm that i have received information about the study

----
(date) (Signature, role in the study)