MASTER THESIS

Master's Program in Learning in Complex Systems

November 2018

A Behavior Analytic Account of Social Categorization

Rebekka Cecilie Wood Strand

OsloMet – Oslo Metropolitan University

Faculty of Health Sciences

Department of Behavioral Science

Acknowledgements

I am very grateful for all the participants who came all the way to the lab and made this paper possible. Special thanks to Professor Erik Arntzen for the support and guidance and to Jon Magnus Eilertzen for helping me in structuring my writing. Professor Arntzen research group should also be recognized as a creative space, with great constructive feedback and quality discussions.

Most importantly, to my loved ones who never stopped believing in me, thank you. Thank you, Mum, for always encouraging me to push my boundaries. Thanks to my fiancé Jeromy for always making sure I keep my deadlines and to motivate me when things gets though. Thanks to my dad for always making me question the world around me.

Contents

Article 1

Article 2

Overview of Tables and Figures

Article 1:

No Tables and Figures

Article 2:

Abstract

In the current article, social categorization was analyzed within the behavior analytic framework. Behavior analysis has been deemed unable to account for more complex behavior that seems to emerge without direct training. Furthermore, mentalistic explanations have been common in explaining this phenomenon. Conditional discrimination, motivational operations, verbal behavior, stimulus equivalence and the transfer of function are presented as behavior analytic approaches that can explain the complexity of social behavior in the real world. Some studies have shown that studying the process of stimulus equivalence in relevance to social categorization is more precise than studies that motivate mentalistic explanations. However, methodological errors were found in the study deemed as the foundation for behavior analytic research on social categorization. A systematic replication of Watt, Keenan, Barnes, and Cairns (1991) was conducted to investigate these methodological errors. The results of the experiment showed that social categorization interferes on the emergence of emergent relations. However, the results of interference were present in other units of measurement then of those in the study that were replicated. The study by Watt et al (1991) was therefore not found to be a valid test for social categorization, however the adaptions made to the study showed some promising results that should be investigated further.

Keywords: social categorization, behavior analysis, stimulus equivalence, pre-experimental history, reaction time, questionnaire, prior learning

Social Categorization in a Behavior Analytic Framework

Rebekka C. W. Strand

Oslo Metropolitan University

MALK5000

Abstract

Social categorization has a long history within social psychology, but not much effort has been done to explain it through the behavior analytic framework. Behavior analysis has been deemed unable to account for more complex behavior, especially when behavior seems to emerge without no direct training. Furthermore, mentalistic explanations have been common in explaining these untrained behaviors. Conditional discrimination, motivational operations, verbal behavior, stimulus equivalence and the transfer of function are presented as behavior analytic approaches that explain the complexity of social behavior in the real world. First, a short analysis of some of the studies on social categorization in social psychology found some discrepancies in their studies. The functional relations studied did not appear to be similar across the studies when analyzed within behavior analytic principles. Secondly, some of the studies conducted on social categorization in behavior analysis was then investigated in comparison to the studies in social psychology. The studies in behavior analysis were found to be more functionally similar in focus of stimulus equivalence, however some of the studies were found to show methodological errors. Replications of previous studies on social categorization within behavior analysis are therefore necessary to further our understanding of social categorization.

Keywords: social categorization, stimulus equivalence, behavior analysis, pre-experimental history, prior learning

Sammendrag

Sosial kategorisering har en lang historie innen sosial psykologi, men det har ikke blitt gjort mye innenfor atferdsanalyse i å forklare sosial kategorisering i det atferdsanalytiske rammeverket. Atferdsanalyse har blitt ansett som ute av stand til å redegjøre for mer kompleks atferd, spesielt når atferden synes å dukke opp uten direkte trening. Videre så har mentalistiske forklaringer vært vanlige i å forklare denne type atferd. Betinget diskriminasjon, motiverende operasjoner, verbal atferd, stimulus ekvivalens og transformasjon av funksjoner ble presentert som atferdsanalytiske konsepter som forklarer kompleksitet i sosial atferd i den virkelige verden. Først, ble noen uoverensstemmelser funnet i en kort analyse av noen av studiene om sosial kategorisering fra sosial psykologi. Funksjonene studert syntes ikke å være like på tvers av studiene når analysert innen atferdsanalytiske prinsipper. Videre ble noen av studiene innenfor atferdsanalyse på sosial kategorisering undersøkt i forhold til studiene i sosial psykologi. Studiene innen atferdsanalyse ble funnet i å være mer funksjonelt like i fokus på stimulus ekvivalens, men noen av studiene ble funnet i å vise metodologiske feil. Replikasjoner av tidligere studier om sosial kategorisering innenfor atferdsanalyse er derfor nødvendige for å fremme vår forståelse av sosial kategorisering.

Nøkkelord: sosial kategorisering, stimulus ekvivalens, atferdsanalyse, pre-eksperimentell historie, tidligere læring.

A Behavior Analytic Account of Social Categorization

Watt et al. (1991) conducted a study on *Social Categorization and Stimulus Equivalence*, that was inspired by The Conflict of Northern Ireland in 1968 to 1998 (BBC, 2018) and the Social Identity Theory (SIT) (McLeod, 2008; Tajfel, Billig, & Bundy, 1973), to conduct a behavior analytic study on Social Categorization (Tajfel, 1982; Van Knippenberg & Dijksterhuis, 2000). They chose to test for the process of stimulus equivalence, which has a long history in behavior analysis (Sidman, 2000; Sidman & Tailby, 1982). Through this study they showed that social categorization could be studied through a behavior analytic method. Although, they never defined the principles they were studying within the behavior analytic framework. Furthermore, the few studies that have been conducted in the years after are hard to find and with different definitions on the same phenomena e.g. (Adcock et al., 2010; de Carvalho & de Rose, 2014; Dixon, Rehfeldt, Zlomke, & Robinson, 2006; Haydu, Camargo, & Bayer, 2015; Kohlenberg, Hayes, & Hayes, 1991; Roche & Barnes, 1996).

The concept of social categorization comes from social psychology. Some of the more famous studies in social psychology are the basis for some of the theories in behavior analysis, like the study on imitation by Albert Bandura (Bandura, Ross, & Ross, 1963). Others are interesting in other ways like the study on conformity by Asch (1956) and the Robbers Cave study by Sherif (1954), they inspired the social identity theory and the studies by Tajfel et al. (1973). Social psychology has widened our understanding about group behavior and how we perceive ourselves in these groups.

The common understanding is that social psychology is the study on how people's thoughts, feelings, and behaviors are influenced by the actual, imagined, or implied presence of others (Argyle, 1973). The social identity theory is defined as; how group behaviors are based on the way the group perceives itself in relation to those outside of the group. (McLeod, 2008; Tajfel et al., 1973). Most will know the studies from this theory as studies on racism, bias, social attitudes, stereotyping, prejudice and social discrimination. An overall label of

these studies is social categorization, emerging in the 1970s (Tajfel, 1982; Van Knippenberg & Dijksterhuis, 2000). Social categorization is defined as; a process whereby we make sense of the world around us by separating things into different classes or groups (McGarty, 1999, pp. 186–191; Tajfel & Forgas, 2000).

A problem in modern times about replicating studies on social categorization is the ethical complications that it entails. Other ways of analyzing social categorization are necessary, where less intrusive methods are used. Also, the studies conducted are more focused on the general behavior of the group, which might lose information about the individual (Barlow & Hayes, 1979; Neuman & McCormick, 1995). Furthermore, the focus on mentalism can cause problems in the experimental analysis of behavior (Skinner, 1974). While social psychology relies on predicting behavior and explaining the predicted behavior as part of a mentalistic structure, behavior analysis explains changes in behavior, through prediction and control, in the relation to changes in the environment (Skinner, 1953). The explanations of behavior are fundamentally different. From a behavior analytic stand point the methods of investigation, definitions and focus in social psychology are questionable. An effort should be exerted to expand the level of analysis within the behavior analytic framework. One focus of expansion should be to explain social categorization through the relations between the individual and its environment.

Behavior Analysis

Complex human behavior can be explained through the variables in relation to that behavior in the behavior analytic framework. Through the focus on social behavior, its evolutionary functions, the events and objects that elicit and signal this behavior, and the environment that maintains it, we can study social categorization in more functional terms (Holth, 2016; Keller & Schoenfeld, 1950; Skinner, 1953, 1981). The principles of behavior, selection by consequences, verbal behavior, motivational operations, conditional discrimination and stimulus equivalence are procedures and processes within behavior

analysis that will help explain the complexity of human behavior (Catania, 2013; Laraway, Snycerski, Michael, & Poling, 2003; Michael, Palmer, & Sundberg, 2011; Sidman & Tailby, 1982; Skinner, 1957, 1969, 1981). These procedures and processes can in turn explain the evolutionary function of racism, bias, social attitudes, stereotyping, prejudice and social discrimination. A further discussion of how these labels will be presented further into the article.

Behavior analysis is the science of behavior and its philosophy is radical behaviorism (Moore, 2008; Skinner, 1964). Radical behaviorism is the fundamental belief that the unit of analysis is behavior and that it can only be studied through measures of observable events. These events are the relations between the environment and the behavior. The predictions of these relations are important in everyday life, but not precise enough in experimental procedures (Skinner, 1953, 1974). Only experiments that include prediction and control are considered as true experiments and its aim is to be a natural science with two assumptions. First, that behavior is a product of the organisms past and current environment. Secondly, the principles discovered by an experimental analysis applies to all animal life. This fundamental belief of scientific methods in behavior analysis differs from social psychology. In social psychology prediction is enough to validate a theory, the unit of analysis is mental constructs or structures in the brain or mind and they are the explanations of behavior (McLeod, 2008).

Mental structures as explanations for behavior is not seen as precise enough in behavior analysis. First, the explanation of behavior becomes circular, the explanation for A = B, but B can also = A. Secondly, manipulations of mental states are difficult if not impossible as they rarely describe the physical and observable phenomena which could be manipulated to test causal relations between events (Beach, 1955; Christoffer Eilifsen & Arntzen, 2014; Skinner, 1953). With this reasoning the mention of mental states add nothing to the analysis of behavior and should be avoided (Skinner, 1974). However, others argue that behavior analysis is missing out on information when excluding mental events, by declaring that only

explanations of observable events can be deemed as explanations of behavior (Svartdal, 2014).

Behavior analysis cannot add to the understanding of social categorization without an analysis of their mentalistic explanations. A discussion and constructive criticism should be made about the studies that explained them through mentalistic constructs. By understanding the behavior that is explained through mentalistic constructs, an explanation of that behavior can be found without depending on mentalistic explanations. The focus on observable events instead of mentalistic constructs adds a dimension to explanations of behavior where manipulations of those events are possible. Changing a mental construct is not as straightforward as exposing someone to changes in the environment (Baer, Wolf, & Risley, 1968; Moore, 2008; Skinner, 1974, 1981).

Another point is the lack of simple descriptions and how to measure them. Through the method of operationalizing phenomena, complex concepts of behavior become observable (Baer et al., 1968). This means that even emotions in some form can be explained through simpler measures of behavior. Such as, defining happiness in a certain condition: An individual orienting towards the experimenter, the edges of their mouth turning up showing teeth. Arguably, happiness is a complex behavior that for every condition should be operationalized differently. If the functions are found to be the same across conditions a common measure of happiness could be found. As a result, private behavior, behavior not accessible to others but the organism, can be measured by the behavioral outcome of these events. Similarly, social categorization can be explained through observable events and its relation to the environment without including mental constructs. However, a parsimonious definition of social behavior specific to social categorization has yet to be discussed in behavior analysis to this date. Instead it has been used as a concept on its own in behavior analysis (Kohlenberg et al., 1991; Watt et al., 1991). A discussion of studies on social

categorization in behavior analysis is presented after a further explanation of behavior analytic concepts.

Selection by Consequences

It is assumed in behavior analysis that all behavior is due to complex interactions between the environment and genetic influence. The foundation of behavior analysis is the theory of natural selection, on the behavioral level and cultural level (Beach, 1955; Charles, 1859; Holth, 2016; Skinner, 1969, 1981; Zagury Tourinho & Vichi, 2012). The behavior is selected by its consequences, as species are selected through survival. A consequence is any object or event that follows behavior, that increases or decreases the likelihood of a behavior or cultural practice produced by the organism. By the contact with consequences, behavior becomes increasingly complex through the process of conditional discrimination, where objects or events can signal multiple different opportunities for different consequences (DeLong & Wasserman, 1981; Sidman, 1980; Sidman & Tailby, 1982). These objects or events commonly precede behavior and may be private or public, they can also be a part of a chain of multiple steps of behavior where they both act as an antecedent and as a consequence. Antecedent refers to any event or object preceding behavior. The complexity of these relational conditions is based on the process of reinforcement and punishment under the temporary effect of motivational operations (Keller & Schoenfeld, 1950; Skinner, 1969, 1981). Motivational operation refers to changes that can alter the effectiveness of reinforcers or punishers (Cooper, Heron, & Heward, 2014; Laraway et al., 2003; Petursdottir, 2013). The principles of behavior are parsimonious and explain the relations between the organism and the environment without circular explanations. An increase in behavior is explained by its consequences in the environment, but the consequences cannot be in turn be explained by the behavior (Christoffer Eilifsen & Arntzen, 2014; Skinner, 1963).

The level of analysis on behavior can be as simple as reflexive behavior, where a stimulus elicits a certain behavior specific to the species, like; blinking when air hits the eye,

or complex as the cultural differences in social signals such as; political views. This example of a social signal is also an example where phenotype override genetic influence, through culturally reinforced behavior (Holth, 2016). Behavior in humans are commonly under the control of multiple stimuli, with multiple relations, and multiple consequences under the control of the current, past or future environmental contingencies (Michael et al., 2011; Petursdottir, 2013; Skinner, 1957). These classes of behavior and stimuli are called higher-order classes and operants as they are learned through pairings of conditioned events or behaviors (Catania, 1995).

However, these principles of behavior do not in themselves explain how an individual shouts ugly remarks at a man on the streets, compliment the same man in a work situation, votes against other people that are similar to this man and that his friends exhibit the same behaviors in similar situations. All of these situations are regarded as one type of behavior in social psychology, social categorization. In a behavior analytic perspective this is an example of complex human behavior. The understanding of complexity is needed to understand the concepts behind social categorization. To delve further into the behavior analytic explanation for complex human behavior some higher-order behavioral processes will be described.

Complex Human Behavior

Imitation, observational learning, rule-governed behavior, verbal behavior and stimulus equivalence are processes that increase the complexity of the basic environmental contingencies of reflexive behavior and the three-term contingency (Hayes & Hayes, 1992; Skinner, 1986). Imitation and observational learning refers to processes where behavior is learned by watching other organisms. The behavior that is learned through these processes are not learned through direct training and is only maintained if the behavior is followed by reinforcement in the individual's environment when it is emitted. This process accounts for how some behaviors are transferred from one individual to another (Catania, 1995). This also explains cultural practices and how they can spread from one organism to a whole group.

Examples can be found in most organisms and seems to be a process selected through the survival of the species (Bandura, 1978; Hayes & Hayes, 1992). However, imitation and observational learning does not account for the complexity of human language compared to other organisms and how humans can emit behavior without direct correspondence between the behavior, antecedents and consequences.

The first account of the complexity of human language was suggested by Skinner (1957) in the book *Verbal Behavior*. In this book he made a clear distinction between language and verbal behavior. Language is the topography of verbal behavior and it is learned through the lifetime of the organism. The complexity of the topography of verbal behavior can be explained through the selection of variation through generations. Verbal behavior is both vocal and non-vocal. Verbal behavior is behavior which is reinforced through the mediation of another organism (Skinner, 1957, p. 20). When talking about verbal behavior we are interested in only the behavior reinforced through the behavior of another organism (Skinner, 1957, p. 20). It is the function on other organisms that is the level of analysis not only the topography. Skinner also described six different verbal operants, a class of behavior that has different verbal functions. In this paper intraverbal behavior is the verbal operant that explains complex behavior in the most functional terms in regard to social categorization. An intraverbal operant is verbal behavior regulated by other organisms verbal discriminative stimuli. There is no point-to-point correspondence between the verbal behavior of the speaker and the listener as described by Skinner (1957). This can explain the contingencies that select a certain verbal repertoire deemed as racist or discriminative in the interaction between individuals. However, the account of verbal behavior does still not account for how humans can emit behavior without direct correspondence between the behavior, antecedents and consequences.

An explanation of how individuals can follow verbal instructions weeks after they were described with no direct contact with consequences is the process of rule-governed

behavior (Catania, Shimoff, & Matthews, 1989). Rule-governed behavior is behavior under the control of contingency specifying stimuli. The contingency specifying stimuli is verbal behavior that describe these contingencies. Rule-governed behavior makes it possible for humans to perform novel behavior without direct observation (Catania et al., 1989). Examples of this type of contingency specifying stimuli are government laws that usually specify punishing consequences. The complexity of verbal behavior and rule-governed behavior comes from the variation and the large number of higher-order conditional discriminations.

With the previous description of the principles of behavior, operant behavior is dependent on antecedent signals through the direct contact with these stimuli or contingency specifying stimuli. The number of stimuli a human has to discriminate in a lifetime can be in some sense unlimited. In history this has been a main argument against behavior analysis, however the process behind this higher-order stimulus discrimination and explanation of how a child can discriminate between the word cat and the vocal response “cat” without direct training was found by Sidman and Tailby (1982). The process behind this phenomena of the emergence of untrained relations, emergent relations, is called stimulus equivalence (Sidman, 1994, 2000).

Stimulus Equivalence

In the sixty’s behavior analysts tried to explain how some relations occurred without training (Rosenberger, Mohr, Stoddard, & Sidman, 1968; Sidman, 1960; Sidman, Stoddard, & Mohr, 1968; Sidman & Stoddard, 1967; Stoddard & Sidman, 1967). They found that if participants were trained arbitrary relations, that A = B and B = C they would also be able to match A to C, B to A, C to B and C to A, without explicit training. This is described as the process of stimulus equivalence (Sidman, 1971; Sidman & Cresson, 1973). The relations that emerges without training is labeled as equivalence relations or emergent relations. Equivalence relations are the relations between the stimuli in an equivalence class that were not directly trained. Sidman (2000, p. 128) defined equivalence relations as a direct outcome

of reinforcement contingencies. Although this statement has later been argued against, a better alternative has not yet been presented in other theories (Clayton & Hayes, 1999).

The directions of the emergent conditional discriminations are the behavioral definitions of reflexivity, symmetry and transitivity (Sidman & Tailby, 1982; Sidman, Wynne, Maguire, & Barnes, 1989, p. 261). Reflexivity refers to the organism's ability to match a stimulus to itself A = A. Symmetry refers to the organism's ability to respond to the reverse relation B = A if only the relation A = B was trained. Transitivity refers to the organism's ability to respond accurately to the relations A = C and C = A when only the relations A = B and B = C were directly trained (Sidman, 2000). The later, transitivity, requires training of three classes or more to be observed. All of these relations have to be observed to prove the emergence of stimulus equivalence (Sidman, 2009; Sidman & Tailby, 1982).

The stimulus classes defined within the stimulus equivalence paradigm are mostly defined by the experimenter. However, in the real world these stimulus classes can be complex combinations with no physical, visual or auditory similarities. The only thing similar to these stimuli are the correlations of these stimuli to reinforcement in their natural environment unique to the individual organism's environmental history (Sidman, 2009). Arguably the scope of these stimulus classes is unlimited except to its relations within the stimulus class. The complexity of these classes is simplified within the experiments on stimulus equivalence, but the results can be just as complex. An example of this can be seen in Sidman (1971) study where he taught a boy to read through a matching-to-sample (MTS) procedure with conditional discrimination with arbitrary matching criteria, as the stimuli that were used in training bore no physical resemblance. By training the boy in matching the auditory name to a picture and then the picture to the written name, the boy learned to read the written name out loud with no direct reinforcement (Sidman, 1971, 2009). By training

simple directional relations, the boy learned to read. Which is regarded as a complex behavior that includes auditory, visual and physical dimensions.

Sidman et al. (1989) writes that if the emergence of equivalence relations can be found in non-humans, we can argue that language is not a requirement. However, no such proof has been found to this day although efforts have been made (Devany, Hayes, & Nelson, 1986; Hayes, 1989; Rehfeldt & Hayes, 1998; Sidman et al., 1982). Arguably stimulus equivalence is a distinctly human process that accounts for the complexity of human language, symbolic stimuli and social behavior, although some argues otherwise (Dube, McIlvane, Callahan, & Stoddard, 1993; Schusterman & Kastak, 1998; Vaughan, 1988). It also accounts for variety within the human language that can be selected as a consequence of reinforcement contingencies without direct training. Verbal behavior and observational learning are processes that can be found in varied complexity in other organisms, stimulus equivalence is distinctly different as it has only been observed in human behavior. As a consequence, stimulus equivalence has been the preferred method within behavior analysis to study complex human behavior (Sidman, 2009).

Complex behavior is not lost in the analysis of behavior. Through the discovered processes of verbal behavior, observational learning and stimulus equivalence, behavior analysis presents a functional explanation of complex human behavior that is observable in the environment, without using mentalistic explanations (Sidman, 2000). Furthermore, the mathematical procedure of stimulus equivalence has shown a unique general capacity to study social behavior such as social categorization (Clayton & Hayes, 1999; Dixon et al., 2006; Grey & Barnes, 1996; Haydu et al., 2015; Roche & Barnes, 1996; Watt et al., 1991) Although, relational frame theory (RFT) and the naming theory are also methods to research complex human behavior within behavior analysis, within this article stimulus equivalence is deemed as the more conceptually clear theory of emergent relations to this date in the discussion on social categorization (Clayton & Hayes, 1999; Sidman, 2000).

The Importance of Transfer of Function

Vaughan (1988) added a new dimension to the stimulus equivalence paradigm, when he found that stimulus equivalence had emerged in five pigeons through the transformation of functional stimulus classes. A functional stimulus class refers to stimuli that share a specific behavioral function and where contingencies applied to one or more stimuli within the class also affects the remaining stimuli of the class (Dougher & Markham, 1994; Sidman, 1994; Tonneau, 2001). However, Vaughan (1988) did not test for reflexivity, symmetry and transitivity as defined by Sidman and Tailby (1982) as properties that are necessary for a stimulus equivalence class to emerge.

As a response to this Sidman et al. (1989) systematically replicated the study by Vaughan (1988) to also test for the properties of stimulus equivalence with three human participants. All three participants showed the transformation of functional classes, but when they tested for reflexivity, symmetry and transitivity one of the participants did not respond in accordance with stimulus equivalence. They argued that functional classes are related to equivalence classes as shown in the two participants that responded in accordance with stimulus equivalence after the transformation of functional classes. The participant that did not show the emergence of stimulus equivalence made Sidman et al. (1989) argue that the two classes were under the control of two different behavioral processes. In later findings Sidman (1994) did however explain that the reason the participant did not form equivalence classes were due to methodological errors.

In the years following the discovery of transfer of function many have investigated its relation to complex human behavior in relation to stimulus equivalence (Arntzen, Eilertsen, & Fagerstrøm, 2016; Arntzen, Fagerstrom, & Foxall, 2016; Barnes & Keenan, 1993; Dymond & Rehfeldt, 2000; Ferro & Valero, 2006; Kohlenberg et al., 1991; Markham & Markham, 2002; Perkins, Dougher, & Greenway, 2007). However, little research has been conducted in direct relation to common understanding of social contexts and complex behavior. Even after

Kohlenberg et al. (1991) argued that social categorization might be explained by transfer of function, little has been done.

The lack of experiments on social categorization might be due to the assumption that it is not possible to study due to its complexity. However, as described earlier that is no excuse. Stimulus equivalence has proven itself as a valid method to investigate complex human behavior (Sidman & Tailby, 1982) and through the addition of transfer of function a new dimension of complex behavior is accessible to the analysis of behavior (Sidman et al., 1989). The terminology behind social categorization within social psychology and its interpretation to behavior analytic concepts might be the reason behind the lack of studies. A short description of social categorization within social psychology might clarify the problems of interpreting its definition to the behavior analytic framework.

A Short Description of Social Categorization in Social Psychology

In the literature in social psychology the concept of social categorization has been evolved since the study on conformity by Asch (1956). In this study they wanted to test if individuals would conform, or respond similarly as other individuals, even if they know their response was wrong when other individuals would openly respond wrong. They presented the participant with a matching test, not to be confused with MTS, where they were to match the length of a black line between three options, while 7 to 9 other individuals were pretending to be participants in the same room and were instructed to respond wrong. They found that only one fourth of the participants in the test groups responded with no errors, while 99% in the control group responded with no errors. They explained this as a consequence of group pressure in the immediate environment. From this study the focus evolved from the individual to the attitudes within the group to another group, intergroup or out-group behavior, to explain how individuals can show negative behavior towards another group by establishing a belonging to their own group (Tajfel et al., 1973).

Tajfel et al. (1973) conducted an experiment where they tested for social categorization versus non-categorization with similarity and non-similarity. They varied the instructions prior to the test which the participants were to award other participants money through points, they could not award themselves. In the social categorization condition, they were told they belonged to a defined group, in the non-categorization no mentioning of such groups were presented. They found that participants in the social categorization group rewarded those that belonged to their group. No such difference was found in the non-categorization group. They explained this difference as a consequence of the individuals need to define and place themselves in their social world. A further explanation of the processes behind social categorization were investigated by Kurzban, Tooby, and Cosmides (2001).

Kurzban et al. (2001) claimed that the categorization of sex, race and age is activated in the mind which encodes it through primitive processes. They also argued that the categorization of race is not evolved directly, but a bi-product of the encoding activated by sex and gender. Wigboldus, Dijksterhuis, and van Knippenberg (2003) also explained stereotyping as a part of social categorization. They described stereotyping as an action people activated automatically through categorization of stereotypical traits. They presented sentences describing behavior conducted by certain stereotypes defined by the experimenter on a computer screen, and once the participants read them, they were presented with another verb. The participants had to answer yes or no if that verb described the sentence they had just read. Wigboldus et al. (2003) predicted that those that were presented with sentences with stereotypes would spend more time before answering yes or no.

The understanding within social psychology is that social categorization is defined as; a process whereby we make sense of the world around us by separating things into different classes or groups (McGarty, 1999, pp. 186–191; Tajfel & Forgas, 2000). By looking at the previous studies on social categorization and the definition of social categorization as an area to investigate the described human behavior is not clear. Neither is the thread of focus within

the studies over the years. The replication of mental processes is hard, but the manipulations in the experiments can be interpreted into the behavior analytic framework.

A point should be made that the studies picked out within this article in social psychology were picked out through certain criteria, so a full picture of all the studies on social categorization is not presented. However, the goal was to present studies that manipulated the environment, with different studies in different time periods. All studies done only through questionnaires were excluded. Although, questionnaires could be a valid way of adding on to a functional study (Critchfield & Perone, 1993; Critchfield, Tucker, & Vuchinich, 1998). These experiments should be a valid presentation of how the studies are different from social psychology to behavior analysis. However, a further analysis of these studies should be conducted. First, some relations and context within these studies should be described in behavior analytic concepts. Secondly, a discussion of social categorization within these terms is necessary as a fundamental and conceptual discussion to aim for experimental precision within behavior analysis.

Social Categorization in the Behavior Analytic Framework

The experiment on conformity by Asch (1956) reveled a behavioral process that may be one of the processes that are necessary for social categorization. The process they found, that they called conformity, could be explained as verbal contextual control based on the individual learning histories. The stimuli presented is visual stimuli where the context is only relying on the individual ability to discriminate between visual differences and tact these differences. However, Asch (1956) added the verbal behavior into the matching criteria and increased the complexity with adding competing contingencies within the experimental context. The verbal behavior by the other participants, who were told to respond wrong on the matching task, can be interpreted as stimuli that elicited one behavioral processes dependent on the individuals learning histories. The verbal behavior by the fake participants cannot be

interpreted as stimuli eliciting rule-governed behavior as it is not contingency specifying, however it is an example of the verbal behavior of a speaker and the participant is the listener.

This experiment proves that with a choice between possible social reinforcement and non-specific reinforcement, in being right on the question, most individuals will respond to the possibility of social reinforcement. This is an example of simple conditional discrimination on concurrent schedule, as the participants responds to two reinforcement possibilities, the stimuli that in their learning history would signal more potent reinforcement. If conformity is a process within the concept of social categorization, then it can be interpreted as behavioral processes in constant mediation of stimuli signaling competing reinforcement contingencies under control of motivational operations. However, all higher-order operants are under such control and would not add to the understanding of social categorization specifically, but it should not be forgotten as a basic principle of complex human behavior in the individuals contact with the environment. A further investigation of this should be tested by replicating the study in behavior analytic terms.

The explanation of the experiment conducted by Tajfel et al. (1973) cannot be interpreted in a similar way however because their manipulation, although adding or withdrawing of certain verbal stimuli, also have a type of sorting task which complicates the relations measured. As described quickly earlier Tajfel et al. (1973) told some of the groups that they were sorted into one of two groups depending on their opinion of certain artists, while the other groups were not told that they had been sorted, although they were. Their results showed that this change made the participants respond differently on the pamphlet. This manipulation comes down to adding a verbal behavior which is contingency specifying as well as a discriminative stimulus that signals certain consequences. What Tajfel et al. (1973) actually may have tested, in the answers on the pamphlet the participants were handed in the social categorization group, were transfer of function and stimulus classes formed through contingency specifying stimulus and their previous learning history by establishing

stimulus classes. However, it is very hard to tell which variables were controlling the behavior of the participants and no clear explanation can be found until the variables are studied further.

The most interesting of these studies however on social categorization was the study conducted by Kurzban et al. (2001). This is because the method used to test for social categorization, bares some similarities to tests on verbal behavior and stimulus equivalence within behavior analysis. Some distinctions should be made however, not only on the explanations of the result, but also on the variables tested. Kurzban et al. (2001) argues that a test on social categorization can be conducted through measuring the reaction time between the presentation of the assumed socially loaded sentence and the response on the following verb. This is an interesting aspect to his experiment as this implies that private behavior in these matching tests are more complex than autonomic responses. However, using a sentence as a sample complicates the precision of the analysis as it may contain unidentified variables. A further investigation of the reaction time between stimuli assumed related to contradicting or complimenting stimulus classes is needed.

Through this analysis two things becomes apparent. First, there is no functional similarity between the studies. Secondly, some form of verbal behavior seems to be the common focus. Also, the definition of social categorization would be, as a direct translation into behavior analysis, similar to the behavioral process of conditional discrimination. However, social categorization might be a combination of multiple behavioral processes within verbal behavior. It seems like an assumption is being made about what social categorization is without a functional explanation of the processes behind it. A clarification is necessary on which processes might explain social categorization, which methods should be used to study them, which methods that already has been used and if social categorization is a unique process in itself or a bi-product of other processes as suggested by Kurzban et al. (2001).

An Attempt at Behavior Analytic Precision

The results from the studies on stimulus equivalence by Sidman (1971) confirmed the possibility of studying complex behavior through stimulus equivalence and this in turn started a wave of studies on complex behavior within behavior analysis. The first studies conducted on social categorization in the behavior analytic framework was conducted by Watt et al. (1991) and Kohlenberg et al. (1991). These studies have been the foundation for all the studies on social categorization within behavior analysis. However, the study that gained the most focus was the study by Watt et al. (1991). Presumably this might be due to that the study by Kohlenberg et al. (1991) focused on transfer of function a relatively new theory, while the theory of stimulus equivalence relations already had many studies backing it up. Due to this the focus will be on Watt et al. (1991) and some mention of the studies that followed (Dixon et al., 2006; Grey & Barnes, 1996; Haydu et al., 2015; Mizael, de Almeida, Silveira, & de Rose, 2016). Although, a note should be made that some more current studies are following the example of (Kohlenberg et al., 1991) and is investigating transfer of function in relation to social categorization.

Watt et al. (1991) presented six different names, three catholic, three protestant and three neutral. They were presented on a screen in a one-sample to three-comparison MTS procedure in a linear training structure. In comparison to the studies described earlier in social psychology, the focus is not a difference in responses as a result of mental processes, but measurement of the relations that emerge between the names presented based on the principle of reinforcement. They found twenty-three participants that came from three different social groups, North Ireland Catholics, North Ireland Protestants and English Protestants. They found that most of the participants that were from Ireland did not establish stimulus equivalence classes with the stimuli presented, while all in the control group did. They argued that previously trained social stimuli may have suppressed equivalence responding. This explanation describes social categorization as a bi-product of previously learned social

stimuli. This assumption might turn out to be true. However, such a definition has not been directly discussed and should not be assumed as true as discussed earlier: Social categorization might be a result of multiple behavioral processes.

Also, the method used by Watt et al. (1991) was not entirely within the scope of what is deemed as precise within studies on stimulus equivalence, they did not include a test for symmetry and equivalence. They claim to test for symmetry, however no reverse presentation of stimuli is presented in the test, namely the CB and BA relations. Also, they describe it themselves, that they did not include the CA relations in the test for transitivity. As described earlier, these are necessary relations to test for when testing for stimulus equivalence (Sidman, 2000). A skepticism must therefore be present when discussing social categorization as a bi-product of previously learned social stimuli.

Even with these methodical errors Watt et al. (1991) has been referred to as the foundation for the modern studies on social categorization within behavior analysis. Since then only a few studies have been conducted on social categorization, but they all carry their own weight on the future of understanding social categorization. The results in these studies also suggests, as Watt et al. (1991), that the emergence of new relations when using stimuli that are assumed to be a part of the individual's prior social learning history. Haydu et al. (2015) also used words assumed to be a part of the participants social learning history to test for the emergence of stimulus equivalence classes. They similarly found that the participants that had a social learning history with the stimuli did not establish stimulus equivalence classes.

Due to the current instability in the cultural environment, more of these studies have been conducted in recent years finding similar, but also different results (Adcock et al., 2010; Dixon et al., 2006). Furthermore, many of the studies, using similar methods, name them studies on, attitudes, racism, pre-experimental history, emotions and so forth. This is a practice that should race some skepticism and some further discussion, although this will be

left open for now. It is also interesting to note that the suggestion by Kurzban et al. (2001) of studying reaction times as a more valid measure of social categorization when using matching tests has yet to be tested in behavior analysis. This could be valid information we can use from social psychology and easily adapt to a stimulus equivalence test. The growing interest in such studies also requires a growing skepticism of the study by Watt et al. (1991) as the foundation for this research.

Process or Bi-Product

An account of social categorization is not complete without the discussion of how some behaviors in some instances are seen as practical and efficient, while the same behavior is deemed as racist in other instances. Arguably through the definition of principles of behavior, the fundamental processes effecting behavior in social categorization may also be the controlling variables of the cultural distinction that describes behavior as racist or efficient. A behavior is not deemed as racist unless another organism defines it as so, verbal behavior (Skinner, 1957). Social categorization could be a description of behavior under the control of specific stimuli related to respondent behavior, that has through the individuals contact with the environment transferred to a bigger stimulus class. In this description, social categorization may not be a process. It might be a bi-product of higher-order behavioral processes defined by its environment and cultural contingencies, in relation to stimulus classes. It is still unclear if social categorization is a behavioral process in itself. However, there is data that seem to point us in the direction of social categorization being a bi-product of other behavioral processes (de Carvalho & de Rose, 2014; Watt et al., 1991).

An interesting point made by Sidman (2009) about the possibilities of stimulus equivalence is, “Much of what we do is determined by things and events that we cannot possibly have experienced directly. For example, we can know historical events through words” (Sidman, 2009, p. 16). Most of these historical events are in turn events that can teach a group of individuals to not replicate those events, as the current environment does not

sanction such behavior. Through the interaction of the contingencies of described past events and the contingencies of the current environment some behaviors are deemed as social categorization by the individuals who observes them. Suggesting that behavior described in common language as; racism, bias, social attitudes, stereotyping, prejudice and social discrimination, is the same as social categorization. In summary it seems that social categorization may be the social behavior previously learned by the individual, as suggested by Watt et al. (1991) and the studies that followed. If so a better adjective than social categorization is available to use.

To describe social behavior previously learned resulting from complex behavioral processes another term is more precise, pre-experimental history (Wanchisen & Tatham, 1991). Pre-experimental history refers to all behavior that was learned in the individual's lifetime prior to the experiment. An argument against such a definition is that it becomes an umbrella term over all social behaviors that are under control of complex interactions between the environment and other organisms. However, by using specific stimuli in a controlled setting, these relations can be studied. Until some behavioral patterns in a certain complex behavioral interaction is distinctly different to others. The term of pre-experimental history is the best we currently have to describe these behaviors. Pre-experimental history is therefore deemed good enough of a definition with the current information available. However, more discussion on this should be done in relevance to new information that might be found in future studies.

Future Directions and Concluding Comments

When humans discriminate between their parent and a stranger, discrimination is useful for survival. However, when a race of humans is deemed as dangerous and results in violence, it is what we call social categorization. Discrimination between groups has been the largest cause of violence in the history of humans. Some examples of this are: Nazi-Germany 1939-1945, The Rwanda genocide in 1994 and Norway terror attack in 2011, but the list goes

on. An understanding of the processes behind this counter evolutionary behavior may give us some clues to prevent further acts of violence based on social categorization. In the current political environment, understanding social categorization has become more important than ever and a necessary step for improving the life of humans (Bear et al. 1968).

In the current paper, stimulus equivalence and transfer of function has been highlighted as methods to study social categorization in the behavior analytic framework. Social categorization was described as pre-existing verbal classes and operants, referred to as pre-experimental history, that may interfere with the individual's current responses to the environment and the emergence of relations. The study by Watt et al. (1991) in particular was also highlighted. Some interesting questions about their method and results were put in perspective to the modern criteria for studies on stimulus equivalence. Although many points have been made about why social categorization should be studied within behavior analysis, we stand with some questions left unanswered.

First, social categorization seems to be a bi-product of other behavioral processes related to verbal behavior and not a process in itself. Secondly, stimulus equivalence seems to be the most functional method to study social categorization to this date. Transfer of function may also contribute to the analysis. Third, conceptual differences in the descriptions of social categorization, and labels referring to similar behavior, should be investigated within behavior analysis. Fourth, pre-experimental history should be discussed as a potential description of social categorization. Finally, earlier studies on social categorization with stimulus equivalence as their method should be replicated with the current understanding of stimulus equivalence.

In summary, the study by Watt et al. (1991) is important to replicate as the foundation of studies on social categorization within behavior analysis. No clear criticism has been presented in regard to their study, although clear methodological errors were found. To further our understanding of social categorization, replication is necessary to either validate

earlier claims, or disprove them. If disproven, some fundamental changes need to be made on how we speak about social categorization within behavior analysis, and even how it is measured. Science is built on small steps toward a consensus on the theories found through multiple measures of the same phenomenon. Replications are necessary to test the foundations we are building our theories on. Although, we still have a long way to go in order to understand how social categorization effects human lives, a replication of Watt et al (1991) will contribute as a small step in the right direction.

References

Adcock, A. C., Merwin, R. M., Wilson, K. G., Drake, C. E., Tucker, C. I., & Elliott, C. (2010). The problem is not learning: Facilitated acquisition of stimulus equivalence classes among low-achieving college students. *The Psychological Record, 60*(1), 43–55. doi:https://doi.org/10.1007/BF03395693

Argyle, M. (1973). *Social interaction* (Vol. 103). New Brunswick: Transaction Publishers.

Arntzen, E., Eilertsen, J. M., & Fagerstrøm, A. (2016). Preferences in equivalence classes by low potency benign valenced stimuli. *European Journal of Behavior Analysis, 17*(2), 142-153. doi:https://doi.org/10.1080/15021149.2016.1247637

Arntzen, E., Fagerstrom, A., & Foxall, G. (2016). *Equivalence classes and preferences in consumer choice*.

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *sychological monographs: General applied, 70*(9), 1. doi:http://dx.doi.org/10.1037/h0093718

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of applied behavior analysis, 1*(1), 91-97. doi:https://doi.org/10.1901/jaba.1968.1-91

Bandura, A. (1978). Social learning theory of aggression. *Journal of communication, 28*(3), 12-29. doi:https://doi.org/10.1111/j.1460-2466.1978.tb01621.x

Bandura, A., Ross, D., & Ross, S. A. (1963). Imitation of film-mediated aggressive models. *The Journal of Abnormal Social Psychology, 66*(1), 3. doi:http://dx.doi.org/10.1037/h0048687

Barlow, D. H., & Hayes, S. C. (1979). Alternating treatments design: One strategy for comparing the effects of two treatments in a single subject. *Journal of applied behavior analysis, 12*(2), 199-210. doi:https://doi.org/10.1901/jaba.1979.12-199

Barnes, D., & Keenan, M. (1993). A transfer of functions through derived arbitrary and nonarbitrary stimulus relations. *Journal of the Experimental Analysis of Behavior, 59*(1), 61-81. doi:https://doi.org/10.1901/jeab.1993.59-61

Beach, F. A. (1955). The descent of instinct. *Psychological Review, 62*(6), 401. doi:http://dx.doi.org/10.1037/h0049138

Catania, A. C. (1995). Higher-order behavior classes: Contingencies, beliefs, and verbal behavior. *Journal of Behavior Therapy Experimental Psychiatry, 26*(3), 191-200. doi:https://doi.org/10.1016/0005-7916(95)00033-V

Catania, A. C. (2013). *Learning* (5th ed.). Cornwall-on-Hudson: Sloan.

Catania, A. C., Shimoff, E., & Matthews, B. A. (1989). An experimental analysis of rule-governed behavior. In *Rule-governed behavior* (pp. 119-150): Springer.

Charles, D. (1859). *On the origin of species by means of natural selection*. London: Murray.

Clayton, M. C., & Hayes, L. J. (1999). Conceptual differences in the analysis of stimulus equivalence. *The Psychological Record, 49*(1), 145-161. doi:https://doi.org/10.1007/BF03395312

Cooper, J. O., Heron, T. E., & Heward, W. L. (2014). *Applied behavior analysis* (2nd ed.). England: Pearson.

Critchfield, T. S., & Perone, M. (1993). Verbal self-reports about matching to sample: Effects of the number of elements in a compound sample stimulus. *Journal of the Experimental Analysis of Behavior, 59*(1), 193-214. doi:https://doi.org/10.1901/jeab.1993.59-193

Critchfield, T. S., Tucker, J. A., & Vuchinich, R. E. (1998). Self-report methods. In *Handbook of research methods in human operant behavior* (pp. 435-470): Springer.

de Carvalho, M. P., & de Rose, J. C. (2014). Understanding racial attitudes through the stimulus equivalence paradigm. *The Psychological Record, 64*(3), 527–536. doi:https://doi.org/10.1007/s40732-014-0049-4

de Carvalho, M. P., & de Rose, J. C. (2014). Understanding racial attitudes through the stimulus equivalence paradigm. *The Psychological Record, 64*(3), 527–536. doi:https://doi.org/10.1007/s40732-014-0049-4

DeLong, R. E., & Wasserman, E. A. (1981). Effects of differential reinforcement expectancies on successful matching-to-sample performance in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 7*(4), 394. doi:http://dx.doi.org/10.1037/0097-7403.7.4.394

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*(3), 243-257. doi:https://doi.org/10.1901/jeab.1986.46-243

Dixon, M. R., Rehfeldt, R. A., Zlomke, K. R., & Robinson, A. (2006). Exploring the development and dismantling of equivalence classes involving terrorist stimuli. *The Psychological Record, 56*(1), 83–103. doi:https://doi.org/10.1007/BF03395539

Dougher, M. J., & Markham, M. R. (1994). Stimulus equivalence, functional equivalence and the transfer of function. In *Behavior analysis of language cognition* (pp. 71-90).

Dube, W. V., McIlvane, W. J., Callahan, T. D., & Stoddard, L. T. (1993). The search for stimulus equivalence in nonverbal organisms. *The Psychological Record, 43*(4), 761-778. doi:https://doi.org/10.1007/BF03395911

Dymond, S., & Rehfeldt, R. A. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst, 23*(2), 239-254. doi:https://doi.org/10.1007/BF03392013

Eilifsen, C., & Arntzen, E. (2014). Er alle abstraksjoner hypotetiske konstrukter? *Norsk Tidskrift for Atferdsanalyse*. doi:http://hdl.handle.net/10642/2562

Ferro, R., & Valero, L. (2006). Transfer of function of visual stimuli through equivalence relations with verbal stimuli. *European Journal of Behavior Analysis, 7*(1), 5-14. doi:https://doi.org/10.1080/15021149.2006.11434260

Grey, I. M., & Barnes, D. (1996). Stimulus equivalence and attitudes. *The Psychological Record, 46*(2), 243.

Haydu, V. B., Camargo, J., & Bayer, H. (2015). Effects of preexperimental history on the formation of stimulus equivalence classes: A study with supporters of Brazilian soccer clubs. *Psychology Neuroscience, 8*(3), 385. doi:http://dx.doi.org/10.1037/h0101276

Hayes, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 51*(3), 385-392. doi:https://doi.org/10.1901/jeab.1989.51-385

Hayes, S. C., & Hayes, L. J. (1992). Verbal relations and the evolution of behavior analysis. *American Psychologist, 47*(11), 1383. doi:http://dx.doi.org/10.1037/0003-066X.47.11.1383

Holth, P. (2016). Levels of Selection: A Place for Cultural Selection. *Norsk Tidsskrift for Atferdsvitenskap*. doi:http://hdl.handle.net/10642/4882

Keller, F., & Schoenfeld. (1950). *Principles of Psychology: A Systematic Text in the Science of Behavior*. New York: Appleton-Century-Croft.

Kohlenberg, B. S., Hayes, S. C., & Hayes, L. J. (1991). The transfer of contextual control over equivalence classes through equivalence classes: A possible model of social stereotyping. *Journal of the Experimental Analysis of Behavior, 56*(3), 505-518. doi:https://doi.org/10.1901/jeab.1991.56-505

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences, 98*(26), 15387-15392. doi:https://doi.org/10.1073/pnas.251541498

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of applied behavior analysis, 36*(3), 407-414. doi:https://doi.org/10.1901/jaba.2003.36-407

Markham, R. G., & Markham, M. R. (2002). On the role of covarying functions in stimulus class formation and transfer of function. *Journal of the Experimental Analysis of Behavior, 78*(3), 509-525. doi:https://doi.org/10.1901/jeab.2002.78-509

McGarty, C. (1999). *Categorization in social psychology*: Sage.

McLeod, S. (2008). Social identity theory. *Simply Psychology, Retrieved from https://scholar.google.com/scholar?start=0&q=Social+identity+theory&hl=en&as_sdt=0,5*.

Michael, J., Palmer, D. C., & Sundberg, M. L. (2011). The multiple control of verbal behavior. *The Analysis of verbal behavior, 27*(1), 3-22. doi:https://doi.org/10.1007/BF03393089

Mizael, T. M., de Almeida, J. H., Silveira, C. C., & de Rose, J. C. (2016). Changing racial bias by transfer of functions in equivalence classes. *The Psychological Record, 66*(3), 451-462. doi:https://doi.org/10.1007/s40732-016-0185-0

Moore, J. (2008). *Conceptual foundations of radical behaviorism*. Cornwall-on-Hudson, N.Y.: Sloan Pub.

Neuman, S. B., & McCormick, S. (1995). *Single-Subject Experimental Research: Applications for Literacy*: ERIC.

Perkins, D. R., Dougher, M. J., & Greenway, D. E. (2007). Contextual control by function and form of transfer of functions. *Journal of the Experimental Analysis of Behavior, 88*(1), 87-102. doi:https://doi.org/10.1901/jeab.2007.36-04

Petursdottir, A. I. (2013). Verbal Behavior and Motivating Operations. *The Analysis of verbal behavior, 29*(1), 1.

Rehfeldt, R. A., & Hayes, L. J. (1998). The operant-respondent distinction revisited: Toward an understanding of stimulus equivalence. *The Psychological Record, 48*(2), 187-210. doi:https://doi.org/10.1007/BF03395266

Roche, B., & Barnes, D. (1996). Arbitrarily applicable relational responding and sexual categorization: A critical test of the derived difference relation. *The Psychological Record, 46*(3), 451-475. doi:https://doi.org/10.1007/BF03395177

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

Schusterman, R. J., & Kastak, D. (1998). Functional equivalence in a California sea lion: Relevance to animal social and communicative interactions. *Animal Behaviour, 55*(5), 1087-1095. doi:10.1006/anbe.1997.0654

Sherif, M. (1954). *Experimental study of positive and negative intergroup attitudes between experimentally produced groups: Robbers cave study*. University of Oklahoma: Intergroup Relations Project.

Sidman, M. (1960). *Tactics of scientific research* (Vol. 16.2). University of Michigan: Basic Books.

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

Sidman, M. (1980). A note on the measurement of conditional discrimination. *Journal of the Experimental Analysis of Behavior, 33*(2), 285-289. doi:https://doi.org/10.1901/jeab.1980.33-285

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

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*(1), 127–146. doi:https://doi.org/10.1901/jeab.2000.74-127

Sidman, M. (2009). Equivalence relations and behavior: An introductory tutorial. *The Analysis of verbal behavior, 25*(1), 5-17. doi:https://doi.org/10.1007/BF03393066

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

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior, 37*(1), 23-44. doi:https://doi.org/10.1901/jeab.1982.37-23

Sidman, M., Stoddard, L., & Mohr, J. (1968). Some additional quantitative observations of immediate memory in a patient with bilateral hippocampal lesions. *Neuropsychologia, 6*(3), 245-254. doi:https://doi.org/10.1016/0028-3932(68)90023-7

Sidman, M., & Stoddard, L. T. (1967). The effectiveness of fadin in programming a simultaneous form discrimination for retarded children. *Journal of the Experimental Analysis of Behavior, 10*(1), 3-15. doi:https://doi.org/10.1901/jeab.1967.10-3

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*(1), 5–22. doi:https://doi.org/10.1901/jeab.1982.37-5

Sidman, M., Wynne, C. K., Maguire, R. W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the Experimental Analysis of Behavior, 52*(3), 261-274. doi:https://doi.org/10.1901/jeab.1989.52-261

Skinner, B. F. (1953). *Science and human behavior*: Simon and Schuster.

Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, N.J.: Prentice-Hall.

Skinner, B. F. (1963). Operant behavior. *American Psychologist, 18*(8), 503. doi:http://dx.doi.org/10.1037/h0045185

Skinner, B. F. (1964). Behaviorism at fifty. *Nursing Research, 13*(1), 94.

Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1974). *About behaviorism*. New York: Macmillan: Vintage Books.

Skinner, B. F. (1981). Selection by consequences. *Science, 213*(4507), 501-504. doi:10.1126/science.7244649

Skinner, B. F. (1986). The evolution of verbal behavior. *Journal of the Experimental Analysis of Behavior, 45*(1), 115-122. doi:https://doi.org/10.1901/jeab.1986.45-115

Stoddard, L. T., & Sidman, M. (1967). The effects of errors on children's performance on a circle-ellipse discrimination. *Journal of the Experimental Analysis of Behavior, 10*(3), 261-270. doi:https://doi.org/10.1901/jeab.1967.10-261

Svartdal, F. (2014). Hypotetiske konstrukter innenfor atferdsanalyse: Finnes de? ja, men der lever de dessverre ikke i beste velgående. *Norsk Tidsskrift for Atferdsanalyse*(41).

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual review of psychology, 33*(1), 1-39. doi:https://doi.org/10.1146/annurev.ps.33.020182.000245

Tajfel, H., Billig, M., & Bundy, R. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology, 3*(1), 27–52. doi:https://doi.org/10.1002/ejsp.2420030103

Tajfel, H., & Forgas, J. P. (2000). Social categorization: Cognitions, values and groups.

Tonneau, F. (2001). Equivalence relations: A critical analysis. *European Journal of Behavior Analysis, 2*(1), 1-33. doi:https://doi.org/10.1080/15021149.2001.11434165

Van Knippenberg, A., & Dijksterhuis, A. (2000). Social categorization and stereotyping: A functional perspective. *European review of social psychology, 11*(1), 105-144. doi:https://doi.org/10.1080/14792772043000013

Vaughan, W. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 14*(1), 36. doi:http://dx.doi.org/10.1037/0097-7403.14.1.36

Wanchisen, B. A., & Tatham, T. A. (1991). Behavioral history: A promising challenge in explaining and controlling human operant behavior. *The Behavior Analyst, 14*(2), 139-144. doi:https://doi.org/10.1007/BF03392564

Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41*(1), 33–50. doi:https://doi.org/10.1007/BF03395092

Wigboldus, D. H., Dijksterhuis, A., & van Knippenberg, A. (2003). When stereotypes get in the way: Stereotypes obstruct stereotype-inconsistent trait inferences. *Journal of Personality Social Psychology, 84*(3), 470.

Zagury Tourinho, E., & Vichi, C. (2012). Behavioral-analytic research of cultural selection and the complexity of cultural phenomena. *Revista Latinoamericana de Psicologia, 44*(1), 169-179.

Social Categorization and Stimulus Equivalence: A Systematic Replication

Rebekka C. W. Strand

Oslo Metropolitan University

MALK5000

Abstract

A systematic replication of Watt, Keenan, Barnes, and Cairns study, *social categorization and stimulus equivalence* in The Psychology Record (1991), was conducted on two groups of Norwegian soccer supporters. The twenty-four participants were trained conditional discriminations for the emergence of three 3-member equivalence classes, when members of two of the classes were assumed as part of the participants pre-experimental history. The stimuli used in these classes were pictures or names of soccer players relevant to their own team or the rivaling team. The test was split into three test blocks. Test Block 1, a replication of Watt et al. (1991) equivalence test, Test Block 2 an adapted generalization test and Test Block 3 an equivalence test. The results matched the original article, in Test Block 1, but these results did not match the result found in Test Block 3. In Test Block 2 participants scored as expected and the response patterns were distinctly different between the test groups and the Control Group. Also, the time used on the full experiment by the soccer team supporters were significantly higher than by the participants who had no interest in soccer, this was also reflected in the reaction times the participant showed on the emergent relations in Test Block 1 and Test Block 3. A correlation was found in the number of expected scores on the questionnaire, but no correlation was found in the number of participants that passed in Test Block 3. The study by Watt et al. (1991) was not found to be a valid test for social categorization, however the extensions made to the study showed some promising results that should be investigated further.

Keywords: social categorization, stimulus equivalence, pre-experimental history, reaction time, questionnaire, prior learning

Sammendrag

En systematisk replikasjon av Watt, Keenen, Barnes og Cairns studie, *sosial kategorisering og stimulus ekvivalens* publisert i The Psychological Record (1991), ble gjennomført på to grupper av norske fotball fans. De tjuefire deltakerne ble trent i seks betingede diskriminasjoner for fremveksten av tre 3-medlems ekvivalens klasser, da medlemmer av to av klassene ble antatt å tilhøre deltakernes pre-eksperimentelle historie. Stimuliene som ble brukt i disse klassene var bilder eller navn på fotballspillere som tilhørte deres eget lag eller motstanderlaget. Testen ble delt inn i tre test blokker. Test Blokk 1, en replikasjon av Watt et al. (1991) ekvivalens test, Test Blokk 2, en tilpasset generaliseringstest og Test Blokk 3, en ekvivalens test. Resultatene stemte overens med den opprinnelige artikkelen, i Test Blokk 1, men de stemte ikke overens med resultatene funnet i Test Blokk 3. I Test Blokk 2 responderte deltakerne som forventet, og responsmønstrene var tydelig forskjellig mellom testgruppene og kontrollgruppen. I tillegg var det tydelige forskjeller mellom test gruppene og kontrollgruppen i hvor mange minutter ble brukt på hele eksperimentet per deltager, som var betydelig høyere i testgruppene. Disse forskjellene ble også reflektert i reaksjonstidene deltagerne viste på transitive relasjoner i Test Blokk 1 og Test Blokk 3. En korrelasjon ble funnet mellom antall deltagere som etablerte ekvivalens relasjoner i Test Blokk 1 og forventede skåre på spørreskjema, men denne korrelasjonen var ikke tilstede med Test Blokk 3. Metoden i Watt et al. (1991) ble ikke funnet å være en gyldig test for sosial kategorisering, men interessante resultater ble funnet i forlengelsen av studiet som burde undersøkes nærmere.

Nøkkelord: social categorization, stimulus equivalence, pre-experimental history, reaksjonstid, spørreskjema, prior learning

Social Categorization and Stimulus Equivalence: A Systematic Replication

Stimulus equivalence or the derived relations paradigm (Sidman, 2000; Sidman & Tailby, 1982) is a process which results are the emergence of untrained relations, also known as derived relations or emergent relations. A stimulus equivalence test has been proven as a valid procedure, as a systematic, operational and empirically verifiable way of testing for the emergence of emergent relations in humans (Horne, Pauline J, & Lowe, 1996; Johnson & Sidman, 1993; Sidman, 1971). It allows researchers to directly study each specific relation and control for other variables that might affect the emergence of the emergent relations. Including a matching-to-sample (MTS) procedure to train conditional discriminations between stimuli, with at least two conditional relations (e.g. AB and BC) with a common member. The test phase in turn evaluates the emergence of relations that are symmetrical (BA and CB), transitive (AC) and equivalence (CA). Reflexivity might also be evaluated (AA, BB and CC) but this property is generally not necessary to evaluate the emergence of other relations (de Carvalho & de Rose, 2014; Sidman & Tailby, 1982).

As a behavior analyst there is a purpose to look into the studies on social categorization, not only to improve methods, but also to explore other areas of research. One good example is the study by Tajfel et al. (1973) in Social Psychology. They found that no social group or social category can be conceptualized outside of their relation to other groups or categories (Tajfel et al., 1973, p. 23). This concept can be interpreted into behavior analysis (Keller & Schoenfeld, 1950) and expanded upon through the process of stimulus equivalence. It makes it possible to look into the effects of pre-experimental history on the emergence of emergent relations. By testing for the number of correct responses on emergent relations a unit of measurement of complex human behavior becomes observable. If the participants respond differently when the stimuli are assumed to be a part of their pre-experimental history an inference could be explained. However, there is a lack of studies on Social Categorization within the derived relations paradigm.

A possible reason for this may be that there is not any clear definition for this area of research in the behavioral analytic literature. There are many labels in literature as; social categorization, pre-experimental history, bias, racial attitudes, stereotyping, social contexts and prior learning (Adcock et al., 2010; de Carvalho & de Rose, 2014; Dixon et al., 2006; Haydu et al., 2015; Peoples, Tierney, Bracken, & McKay, 1998; Watt et al., 1991). These labels are not precise or observable and a common behavioral definition should be established to solve this problem. It would make literature more accessible for replication and studies more precise. In this article pre-experimental history was chosen to describe the suggested interference on emergent relations. As social categorization; thinking about others in terms of their group membership, a process whereby we make sense of the world around us by separating things into different classes or groups (Tajfel et al., 1973). This definition is hard to operationalize. Pre-experimental history seems to be a more correct term, since it does not refer to anything mentalistic and refers to the general history of learning for each participant before the experiment (Haydu et al., 2015).

Behavior analysts have much to learn about experiments on social categorization and should improve the methods used to measure it. With a focus on developing a more systematic method of investigation, that is more efficient in the behavior analytic paradigm. There are some studies that have done this (Adcock et al., 2010; Haydu et al., 2015; Kohlenberg et al., 1991; Watt et al., 1991). The upside is that the derived relations paradigm makes it possible to study the effects of pre-experimental history. In a way that, although the social contexts are different, they are comparable to a certain point. This is what Watt et al. (1991) realized and conducted the first experiment on pre-experimental effects on the emergence of derived relations.

Watt et al. (1991) found that participants did not form equivalence classes when the experimenter defined classes consisted of stimuli related to their pre-experimental history. They asked 23 students to participate. They belonged to three different social groups in

Northern Ireland. One of the groups consisted of 12 northern Irish Catholics, the second of 6 Northern Irish Protestants and the last group were six English Protestants. The stimuli used were catholic names, protestant names and nonsense syllables in three classes with four to five members. The experimental method consisted of a pre-training phase, a training phase, and two test categories.

The pre-training phase in Watt et al. (1991) was introduced to make sure the participants understood the concept of the program. Once they were through pre-training the training phase began. The training structure they used was linear (A to B and B to C) (E Arntzen, T Grondahl, & C Eilifsen, 2010; Holth & Arntzen, 1998). The training phase consisted of three stages. First, they presented the protestant names one by one as the sample stimuli (A1A2A3) and used three random nonsense syllables as comparison (B1B2B3) with continuous feedback. The reinforcement, or feedback, consisted of messages on the screen after each correct response: Correct, with a matching high pitch sound, and after each incorrect response, wrong, with a matching low pitch sound. In the second stage they used the nonsense syllables as comparison and the catholic names (C1C2C3) as comparison. When the participants had a 100% correct response rate in two successive blocks per stage, AB and BC, they would move on to stage three. In this stage the relations trained in the previous stages were mixed randomly, each stimulus combination was presented at least two times and feedback was only given to 50% of the possible trials. Once the participants reached the criteria for mastery of 100% correct responses for all stimulus combinations, they initiated the test phase.

Watt et al. (1991) split the test into two different categories according to different arrangements of novel stimuli. In the first category they presented the protestant names (A1A2A3) as sample and the catholic names (C1C2) as comparison. In this category they also included a new protestant name as comparison (C3') (Fields & Reeve, 2001; Reeve & Fields, 2001). In the second category they did a generalization test by including three new stimuli

from each class (A3'tB3'tC3't) as comparison to match with the protestant names as sample. They set a mastery criterion at 100% in both tests to suggest formation of equivalence classes, those who didn't reach the mastery criterion was set as fail.

Watt et al. (1991) found that all of the six Northern Irish Protestants failed to respond in accordance with stimulus equivalence. Five of the Northern Irish Catholics also failed the equivalence test, but the other seven scored in accordance with the experimenter defined classes. The 11 participants that did not pass the equivalence test showed a tendency to pick protestant names in the presence of protestant symbols in both tests. All of the English Participants scored in accordance with stimulus equivalence, suggesting formations of equivalence classes.

An issue worth mentioning in the Watt et al. (1991) study is that they did not include a symmetry test in accordance with a linear training structure (BA and CB) and a full equivalence test (CA). Sidman (2000) would therefore argue that the claim by Watt et al. (1991) is false. The participants might be able to establish equivalence classes, but test would not qualify as a test for equivalence classes. Because of this, an extension should be included to test for equivalence classes with the modern criteria.

One of the more recent studies conducted by Haydu et al. (2015) is a good example of how to use a full stimulus equivalence procedure to research the effects of pre-experimental history on the emergence of emergent relations. They used the MTS procedure on twenty-eight men from three different soccer clubs in Brazil. They tested for all the properties of the equivalence classes, with training (AB and BC), a symmetry test (BA and CB), a transitivity test (AC) and an equivalence test (CA). They were trained to match club emblems (A) to abstract paintings (B), and then to match the same abstract paintings with the words "Good", "Poor" or "Regular" (C). The words "Good", "Poor" or "Regular" were presented systematically so that the main rival would be matched with "Good" and their own club matched with "Poor" for each of the three groups. The way they presented these words in the

conditional discrimination procedure were to test for how pre-experimental history with the soccer emblems and the words would interfere with the emergence of emergent relations. They found that none of the participants passed the test for equivalence and transitive relations between the club emblems and words. Corresponding with the results of Watt et al. (1991). However, they may not be comparable, since the method used in the original study were questionable at best, a replication of Watt et al. (1991) is therefore necessary for scientific precision.

Replications are important to any area within the experimental method (Pierce & Cheney, 2013, pp. 29–51). However, a direct replication of Watt et al. (1991) is not possible as social contexts change. As Haydu et al. (2015) also points out, every participant will have different pre-experimental histories and would therefore not be comparable. However, it is possible to replicate the method of investigation; by using three different social groups, stimuli relevant to the two of the groups and adapt it to modern methods of investigation. To do a replication some adaptations and extension needs to be incorporated to the original method of Watt et al. (1991).

By incorporating a pre- screening, to look into what social group the participants identify themselves as, one can make some assumptions (Slaton, Hanley, & Raftery, 2017). By also including a post-screening, a bipolar questionnaire using a Likert scale, one can aim to validate those assumptions. Furthermore, make it possible to look into the results in the stimulus equivalence test and how they may correlate with their results in the bipolar questionnaire. This might be a way of solving the problem of uncertainty in if the equivalence test results are due to methodic errors or to the participants pre-experimental history (Critchfield & Perone, 1993; Critchfield, Tucker, & Vuchinich, 1998).

Another area to focus on may be the concept of time. Although some of the participants showed the emergence of equivalence classes in the test groups, a difference might have been found in the amount of time they spent on the study compared to the English

protestants (Kurzban, 2001). Although it is discussed if the reaction times is relevant to the total results per participant in traditional stimulus equivalence tests, the variable of reaction times in the study on pre-experimental history might be more relevant than first assumed (Erik. Arntzen, Terje. Grondahl, & Christoffer. Eilifsen, 2010; Catania, 2013; C Eilifsen & Arntzen, 2009; Fields et al., 2002; Leppänen & Hietanen, 2004; Vaidya, Hudgins, & Ortu, 2015)

Ethical aspects of these types of experiments should be considered before beginning. First, the test in itself can elicit certain responses that can be very uncomfortable to participants. Some of these responses might be excessive and too disruptive to produce any valid results. The participants should be told and allowed to quit at any time. Secondly, when the participant is told what the study is about it might elicit social disruption (Pierce & Cheney, 2013, pp. 187–188). Therefore, it is important to make sure all information is clear and informative both before and after the experiment. Thirdly, the social groups picked for a study should be evaluated thoroughly, especially with groups that have long histories of aggressive conflicts.

Systematic replications are important, because they can validate or disclaim previous studies (Sidman, 1960). The Watt et al. (1991) study has been cited many times, but research methods have evolved and no replications has been conducted since. The purpose of this study was to systematically replicate Watt et al. (1991) and compare the results on a different social group. Next, adapting the method to see if an equivalence test would show different results, to see if reaction time as a variable might give a better picture of the processes of how the pre-experimental history interfere with the formation of equivalence classes and to see if questionnaires could add to studies on pre-experimental history.

Method

Participants

Twenty-six participants were assigned into three groups, one Control Group and two test Groups. Twenty-four participants were included in the final results. Participants were recruited through phone calls and were asked to fill out a survey either on their e-mail, at the university laboratory or at the soccer stadium. If they responded according to set criteria, they were asked to pick a date and time to meet. The minimum age for participation was set to 18. The age range of the Control Group was relatively young (M = 26,75) while the Test Group 1 (M = 41,7) and Test Group 2 (M = 35) was generally older. Age ranged from 19 to 62 years old with a mean age of 35. Both test groups consisted relatively of more male (85%) participants and the Control Group had a more even mix of gender (50%). To increase the probability in the recruiting process, the participants received a gift card with the value of 100 kr.

The participants were allocated to the different groups based on the survey. If they answered that they had little or no interest in soccer they were assigned to the control Group. If they responded that they supported one of the teams and disliked the other that were in focus for the study, and also responded that they either; had a tattoo of their team, attended all home games and some away games or more, or if they had been in a physical altercation regarding their team, then they were included in test Group 1 or 2. If they supported VIF they were placed in Test Group 1. If they supported LSK they were placed in Test Group 2. A criterion was also set on the length the participants could be exposed to the experiment, as it might promote strong reactions. This criterion meant that if the participants had not reached the test block by the end of the 60th minute they would be excluded from the data and also asked if they wanted to quit the experiment. The study was assessed and accepted by the ethical center for research in Norway before it was conducted ("Norsk Senter for Forskningsdata," 2018,January).

One of the participants had prior experience of the current type of experiment in stimulus equivalence and was therefore excluded from the final data set: The other participant

that was excluded from the data set did not reach the test block within the 60th minute and did not reach the criteria for inclusion. None of the others had any prior experience.

Instructions were provided on how to use the computer. They were informed that they could end the experiment at any time with no negative consequences and that the instructor would come in after an hour to check in on them. Before beginning the test, they were asked to read and sign a document of agreement, this also informed them that they would stay anonymous. When they were finished or ended the experiment the participants were given a second document of agreement in accordance with the Norwegian center for research (NSD) requirements after being fully informed of what the study entailed ("Norsk Senter for Forskningsdata," 2018,January).

Apparatus and Setting

Setting. The experimental sessions were conducted in two locations. In a quiet room with a table and a chair in a cubicle at the laboratory at Oslo Metropolitan University in Oslo, and in a meeting room with a round table and few pictures on the walls at the soccer stadium for LSK in Lillestrøm. The participants were left alone while the experimenter waited outside. The participants were allowed to drink coffee during the experiment, but they were asked to put away their phone and bag. The length of experimental sessions varied from 15 minutes to 69 minutes, with a mean length of 33,5 minutes.

Apparatus. The experiments were conducted on HP Elitebook laptop computers running Windows 7 operating system. The computers had 17-inch screens and external Dell computer mice were used to control the mouse cursor. All aspects of the training and testing used to establish equivalence relations and to establish discriminative functions were controlled by parameters in a custom-made software program made in collaboration with Professor Erik Arntzen. The software program conducted automatic data recording on the number of trials, the stimulus relations that were trained, the responses to the sample stimuli, reaction time, the correct/incorrect comparison choice, and the provision of feedback on each

trial. Finally, the software counted the number of trials for both training and testing. Verbal reports were recorded through a prepared questionnaire on the computer. The questionnaire contained information about what stimuli might be seen as aversive, neutral or positive, and on how they experienced the experiment.

Stimuli. Figure 1 shows the stimuli used. The stimuli were chosen after a pilot was conducted with three participants prior to the current study. The stimuli were defined as three classes with four to five members. The defined classes were pictures and names from the two opposing soccer teams, as well as some abstract pictures and one abstract name. Class A were stimuli related to LSK, assumed positive for test group 2. The name chosen in this class was Frode Kippe. Class B were neutral symbols as well and the neutral name Devon Larsen. Class C were stimuli related to VIF, assumed positive for test group 1. There were two names in this class, John Carew and Freddy Dos Santos. For simplification the names in the different classes were written as LSK_name_1, Novel_name_1, VIF_name_1 and VIF_name_2 in the article.

Design

A group design was used with one control group and two test groups.

Data Collection

The computer program collected all data in the training and testing condition. Data was also collected through indirect analysis in two questionnaires. The first questionnaire was used as a tool for recruitment. Some of these questions were such as “which soccer team is the best in Norway?” and “do you have a tattoo of your favorite team?”. The second questionnaire was based on a Likert scale and was meant to test for discomfort or comfort of the participant’s reactions to the different stimuli. This included 7 questions, with pictures of each stimuli from class A and C, where they could respond on a scale with 5 points from uncomfortable to comfortable. An added question was also included where they were asked to give an added feedback if there was anything that they wanted to elaborate on.

Procedure

The conditional discriminations were administered, by the MTS program, in a linear training structure using simultaneous matching. The programmed consequences were thinned in three steps. The programmed consequences were at 100% in the first step, in the second step the programmed consequences were gradually thinned from 100% to 0%, and in the third step, the test condition, no consequences were presented. The presentation of consequences, depending on the experiment stage, would pop up on the screen after a response. The consequences were present for 1000ms with each presentation. Following the consequences an inter-trial interval (ITI) lasted for 1000ms.

In the three test conditions ten novel arrangements of stimuli were presented in the two first test blocks and in the third test block nine additional novel arrangements were presented. If the participants responded according to stimulus equivalence, they were written up as PASS, the participants that did not establish stimulus classes according to criteria, was scored as FAIL. The criteria for PASS was set as 100% correct score on Test Block 1 and Test Block 2. Before starting the experiment, an instruction was presented on the screen in Norwegian.

Stage 1. Training with continuous programmed consequences was conducted in two phases. The first phase was the A to B training, A as sample to B as comparison. The A class was defined as visual stimuli related to the soccer team LSK and the B class were arbitrary stimuli, as unknown figures. More precisely three A stimuli and three B stimuli were presented randomly until all stimulus combinations were successfully completed in two successive cycles. The second phase of stage 1 was then introduced where the training order, B to C, was presented. Where the sample B was an abstract symbol, and the comparison C was the stimuli related to the VIF team. The criterion for mastery was set to 30 trials correct with 15 trials in each block for completion, a mastery criterion of 100% correct before next stage was presented, see Table 1.

Stage 2. Training was then introduced between AB and BC relations with intermittent consequences in random order. All previously trained combinations in stage 1 were presented at least two times while the programmed consequences were thinned from 100% to 0%, from 100% to 75% to 25% and then to 0%. The criterion for mastery was set to a minimum of 150 trials correct in succession, equaling 5 blocks with 30 trials per block for completion. If the participant successfully matched all relations two times with a 100% score rate within the 0% feedback condition, they would advance to stage 3, the test.

Stage 3. The test was then introduced and included three different test blocks that were presented in succession. The first block included baseline relations AB and BC, transitivity relations, AC, and a novel stimuli C3'. The second test block was a generalization test. The third test block was an equivalence test.

Test Block 1. The first test block included presentation of all baseline relations AB and BC, to test for transitivity AC and generalization C3'. At least 10 presentations of each baseline relation were mixed in with five presentations of transitivity and generalization trials. Test block 1 was set to 60 possible trials where 20 trials were baseline trials, 20 trials were transitivity trials and 10 trials were generalization trials, in a mixed order. The criterion for mastery was defined as 100% correct, 60/60 responses.

Test Block 2. This part of the test included a generalization test. The comparison stimuli in this test block were three novel names A3't, B3't and C3't, that had not been presented in any other part of the training or test block 1. C3't was the only comparison stimulus defined as a correct response in this test block. Only the stimuli from class A was used as sample stimuli. Each relation was presented ten times before the next test block was introduced. The number of possible trials in this test block was 30. The criterion for mastery was set as 33% correct responses with a leniency of 25%–50%.

Test Block 3. This test block included baseline relations, AB and BC, symmetry relations, BA and CB, transitivity relations AC, and equivalence relations CA. This test block

did not include the generalization stimuli. The number of possible trials in this test block was set to 90, 30 trials were baseline trials, 30 trials were symmetry trials, 15 trials were transitivity trials and 15 trials were transitivity trials. The criterion for mastery was defined as 100% correct, 90/90 responses. See Figure 3 for visual examples.

Results

Trials to Criterion

See Figure 5 for average number of trials during training and thinning for all participants across Groups. The number of average trials during training was generally less in the Control Group (M = 66.75, range 45–120), than in Test Group 1 (M = 165, range 60–300) and Test Group 2 (M = 122, range 90–375). The significance criterion was set as $p < .05$. There were a significance effect between the Control Group and Test Group 1 $t(54) = 3.39$ p, $< .001$. However between the Control Group and Test Group 2 there were no significant effect $t(54) = -1.35$, $p < .09$. There was also no significance effect between Test Group 1 and Test Group 2 with a $t(36) = 0.81$, $p < .21$.

The number of average trials during the thinning blocks were also generally less in the Control Group (M = 189, range 150–270) than in Test Group 1 (M = 270, range 180–540) and in Test Group 2 (M = 282, range 150–480). Between the Control Group and Test Group 1 there were a significant effect $t(54) = 1.92$, $p < .03$, also there were also a significant effect between the control Group and test Group 2 $t(54) = 1.92$, $p < 0.03$. There were no significant effect between the two test groups $t(36) = 0.19$, $p < .42$.

Test Blocks 1 and 3

See the bottom half of Figure 4 for a visual presentation of the following results. The percentage of participants that passed with a hundred percent test score in Test Block 1 was nearly three fourth (70%) in the Control Group, almost one fifth (14%) in Test Group 1 and almost half (42%) in Test Group 2. The Fisher exact test was conducted on the passes versus fails and shows no significant difference between the groups in Test Block 1. The Control

Group and Test Group 1 showed no significant effect $p < .15$, neither did it show any significant results between the Control Group and Test Group 2 $p < .64$, or between Test Group 1 and Test Group 2 $p < 1$.

In test block 3 the percentage of participants that passed with a hundred percent test score was more than four fifths (83,333%) in the Control Group, four fifths (80%) in Test Group 1 and more than half (57%) in Test Group 2. The Fisher exact test shows no significant difference between the groups in Test Block 3. The Control Group and Test Group 1 had no effect $p <1$, between the Control Group and Test Group 2 there was little significant effect $p < .56$, and there were similar results between the two test groups $p < .58$.

See the top half of Figure 4 for the average percentage correct scored per group in Test Block 1. The average percentage correct per group was high in the Control Group (M = 92,5%, range 78–100%), but was not as high in Test Group 1 (M = 76,665, range 66–100%) and Test Group 2 (M = 88%, range 66–100%). A t-test was conducted on the percentage correct results and showed no statistical significance between the Control Group and Test Group 1, $t(53) = 0.26$, $p < .39$, between the Control Group and Test Group 2 $t(53) = -1.04$, $p < .15$, or between the two Test Groups $t(36) = 0.69$, $p < .24$.

The average percentage correct scored per participant, on the test on emergent relations in test block 3, was very high in all groups. The percentage correct scored was very similar in the Control Group (M = 98,889%, range 93–100%), Test Group 1 (M = 98,66%, range 93–100%) and Test Group 2 (M = 97,59%, range 93–100%). A t-test showed that there was no significant effect between the different groups. There was no significant effect between the Control Group and Test Group 1 $t(20) = 0.12$, $p < .45$, the Control Group and Test Group 2 $t(30) = 0.6$, $p < .27$, and the two test groups $t(24) = 0.42$, $p < .33$.

Test Block 2

The results from Test Block 2 are presented in Figure 6 as a visual presentation of response patterns. A visual analysis of the response matric in figure 6 shows a distinct pattern

in the two test groups, where they choose A3't more than any other comparison, while the distribution is more random in the Control Group. Though a few of the participants in all groups deviate from the pattern, participant 18301 in the Control Group consistently responded to a different comparison per different sample, A1 to A3't, A2 to B3't and A3 to C3't. Participant 18307 in the Control Group shows the same pattern as in the Test Groups with all responses on A3't, this participant also responded to the questionnaire with a preference to the A stimuli, see figure 8. Participant 18315 in Test Group 1 distributed responses equally between A3't and B3't but showed no responding to C3't. Participant 18320 in Test Group 2 did not respond consistently on A3't as the other participants in the Control Group but responded more distributed on all three comparisons. A statistical analysis of the number of correct responses of 30% (25%–50%), defined as responding in accordance with experimenter-defined classes, were set against the number of deviating results 0% (0%–25%), defined as not responding in accordance with experimenter-defined classes, between groups. The patterns of responses were found, using a Fisher-exact test, to have a statistically significant effect between the Control Group and Test Group 1 $p < 0.01$. However, the comparison between the Control Group and Test Group 2 showed no significant effect $p < .13$ and similar was the result between the two test groups $p < 1$.

Duration and Reaction Times

See Figure 5 for the average duration spent by participants for the whole experiment. In the Control Group the participants spent less minutes in average then the two Test Groups (M = 25,8, range 15–32). Test Group 1 spent a few more minutes (M = 39.57, range 25–69) than test Group 2 (M = 38,7, range 26–50). The t-test between the control group and the test groups 1 showed a significant effect. The biggest effect was seen between the Control Group and Test Group 2 $t(53) = -3.68$, $p < .001$, and a very similar result was seen between the Control Group and Test Group 1 $t(53) = -2.6$, $p < .009$. There was no statistical significance between test group 1 and 2 $t(36) = 0.13$, $p > .45$.

See Table 3 for the average Reaction Time (RT) per participant on Transitive and Equivalence relation in Test Block 1 and Test Block 3, across participants who established equivalence classes and those who failed to do so. In the Control Group eight participants were quicker in RT in Test Block 3 compared to Test Block 1. Two participants did not increase their RT between test blocks. Participant 18309, which did not pass the test, slowed down on the equivalence relations. Although participant 18309 slowed down, the RT was respectively quicker then participant 18308 who did pass the test. Participant 18307 on the other hand had consistently a short RT and passed both test blocks. Three participants in Test Group 1 reduced their RT in Test Block 3 compared to Test Block 1. These participants also went from failing Test Block 1 to passing Test Block 3. However, participant 18316 showed an increase in RT during Test Block 3 and did not pass in any of the test blocks. Six out of the seven participants in Test Group 1 had generally slow RT. The only participant, 18313, that consistently showed quick RT was also the only participant to pass both test blocks in this group. As in Test Group 1, the participants In Test Group 2 that increased in RT in Test Block 3 failed Test Block 1 but passed test Block 3. Similarly, the two participants that had consistent RT were generally quicker than the other participants in the group and passed both test blocks. Also, the two participants that failed both test blocks, similarly, showed a slower RT in Test Block 3 then in Test Block 1. Participant 18322 was one of these participants and was also the participant that had the slowest RT of all participants in all groups. Over half of all participants in all groups had a generally slower RT to the generalization relation, A3C4.

See Figure 7 for the average RT of all groups to the transitive and the equivalence relations in Test Block 1 and Test Block 3. A significant difference in RT was found between groups in the transitive relations in Test Block 3 ($M = 2.2$ $SD = 0.8$), $F(5.5) = 1.57$, $p < .04$. However, no differences were found in the RT on transitive relations in Test Block 1 ($M = 2.6$ $SD = 0.5$), $F(0.01) = 0.003$, $p > .9$, or on RT on the equivalence relations in Test Block 3 ($M = 2$ $SD = 0.6$), $F(2.9) = 0.68$, $p > .12$. Test Group 1 ($M = 2.1$ $SD = 0.7$), $F(1.7) = 0.88$, $p > .25$,

and Test Group 2 (M = 2.7 *SD* = 0.57), F(0.98) = 0.32, p > .37, had a more consistent RT across Test Block 1 and Test Block 3, while The Control Group (M = 2.1 *SD* = 0.55), F(14.2) = 1.2, p < .01 varied more between the two Test Blocks.

Questionnaire

A visual presentation and a Chi Square test were conducted on the results of the bipolar questionnaire. See Figure 8 for a visual presentation of the results. The visual presentation show that the test groups have distinct response patterns that were opposite to each other, while the control group show a less distinct pattern. This corresponded with the hypothesis, that the test groups would respond negatively to their rivaling team stimuli and positively to their own team stimuli. However, some outliers can be seen in the stimuli that were written names, C2, C3't, A3't and B3't. The Chi Square test, as seen in Table 2, confirms that the visual pattern was statistically significant. The significance criterion was set as p < .01. The pre-experimental history seemed to affect the participants responses on the questionnaire in Test Group 1 and Test Group 2. X^2 (5, N = 6) = 0.872 p <.001. The responses were not scored as significant to the stimuli; C2 in both groups X^2 (2, N = 6) = 2.204 p > .683 and C3't in Test Group 2 X^2 (5, N = 6) = 2.204 p > .446. The pre-experimental history did not have any effect on the control group responses, 2 X^2 (5, N = 9) = 4.168 p > .135. However, the Control Group responses to B3't did show an outlier 2 X^2 (5, N = 9) = 2.088 p < .001.

Discussion

The purpose of this study was to systematically replicate Watt et al. (1991) and investigate how pre-experimental history effects the formation of equivalence classes. Furthermore, extend the study to (1) if an equivalence test would show different results, (2) to see if reaction time is a variable that should be included in future studies on pre-experimental history and (3) to see if a bipolar questionnaire would add to studies on pre-experimental history.

Twenty-four participants were training conditional discriminations and testing for the emergence of three 3-member equivalence classes in a linear structure. Fourteen of those participants were taught to match stimuli that were part of their pre-experimental history. Testing for emergent relations in three blocks; the first block consisted of a transitivity test and generalization test, the second block consisted of an arbitrary generalization test, and the third block consisted of an equivalence test. All participants were presented with the same stimuli, training, test, questionnaire, and information. The current experiment found that the participants in the two test groups responded more correctly and quicker than the Control Group, suggesting that pre-experimental history interferes with the formation of equivalence classes. However, the results did not correspond with the results in Watt et al. (1991) and some interesting aspects on studying pre-experimental history was found in relation to RT and trials to criterion.

Trials to Criterion

In Stage 1. The number of trials used per participant to reach the set criterion of the conditional discrimination training varied between the different groups. The participants in the Control Group and test Group 2 had consistently lower numbers of trials to criterion then the participants in Test Group 1. This suggest that the participants in Test Group 1 had a harder time matching then the other groups. Since all the participants were exposed to the same stimuli and training structure, these results suggest that the pre-experimental history in Test Group 1 may have interfered in the conditioning of new relations. However, since Test Group 2 did not show similar results one can argue that the order of which stimuli were presented first in training may be a contributing factor to this difference between test groups. This also corresponds with the results in Watt et al. (1991) where Test Group 1 also consistently scored lower then Test Group 2. Arguably there were no such differences in Haydu et al. (2015) which used the OTM training structure. This needs to be investigated

further as the difference might be a result of the linear training structure and not because of pre-experimental history (Erik. Arntzen et al., 2010).

In Stage 2. While the Control Group used almost the minimal trials needed to reach criterion, the participants in Test Group 1 and Test Group 2 used a considerably higher number of trials to reach the criterion during Stage 2. This suggests that the participants in the two test groups did not maintain the relations they learned in stage 1 as well as the Control Group. Furthermore, it suggests that pre-experimental history also interfered in how they maintained the conditioned relations. Arguably, the fact that the two test groups had different scores in trials to criterion in Stage 1 is another confirmation of this. As seen in previous studies, different training structures effects the number of repetitions in training and the likelihood the participants establish equivalence classes (E Arntzen et al., 2010; Holth & Arntzen, 1998). The results in the two test groups does not correspond with the expected number of trials as seen in the theory of linear training structures. However, the results in the Control Group corresponds with this theory, which might also explain why some participants in the Control Group did not establish equivalence classes.

Equivalence Formation

Equivalence test as in Watt et al. (1991). Three out of seven participants passed in Test Group 1, only one out of seven participants passed in Test Group 2, and seven out of ten passed In the Control Group. This corresponds with Watt et al. (1991) results. However, when applying a statistical test there are no significant differences between the groups, in how many passed in numbers or percentages. A visual analysis, as in Figure 4, seems to show very different results between the Control Group and the test groups, but when compared statistically this difference is not as clear. This suggests that the results in Watt et al. (1991) might not be as clear as they make them out to be. Although these results does not correspond with what they claim to have found in Watt et al. (1991), the number of participants that passed in Test Block 1 and Test Block 2 compared to the Control group are almost identical

to their results. Another point about the results in Test Block 1, is that the test does not test for symmetry or equivalence relations (CA). So, at best, one can argue that pre-experimental history interferes with transitive relations (AC). If a fuller understanding of how pre-experimental history effects how humans learn, then a fuller test should be conducted.

Equivalence test. The results of Test Block 3 did not correspond with the assumption that pre-experimental history interfere with the emergence of stimulus classes. Almost all participants in all groups responded in accordance with stimulus equivalence. Five out of six participants in the Control Group passed, four out of five passed in Test Group 1 and four out of seven passed in Test Group 2. These results paints an interesting picture in opposition of the results that was found by Watt et al. (1991). With these results there are no differences between the groups, which suggests that pre-experimental history did not interfere with the emergence of stimulus classes.

Another interesting part of these results is that most participants that did not pass in Test Block 1 passed in Test Block 3 with no extra training. Studies on pre-experimental histories, with an equivalence test as in Test Block 3, has shown that participants do not establish equivalence relations (Haydu et al., 2015; Peoples et al., 1998). This suggests that the delay between the two test blocks may be why more participants show the emergence of stimulus classes in test Block 3, as a delayed emergence of equivalence classes (Arntzen & Nartey, 2018; Holth & Arntzen, 1998). However, a comparison is not possible as these studies used an OTM training structure, so more studies should be conducted to investigate this.

Generalization test. The generalization test in Test Block 2 was changed and is an extension of the test in Watt et al. (1991). Only colored pictures were originally chosen as stimuli for the generalization test. However, the participants in the pilot responded to the colors of the stimuli instead of the trained relations. One name was therefore intermixed in the baseline test and three names were chosen for the generalization test.

Very distinct response patterns were found in Test Block 2 to the names presented as comparison in the generalization test. The patterns were distinctly different between the test groups and the Control Group. All participants failed to established experimenter defined classes. The response pattern, if a participant had no relation to the names presented was expected to be random. However, the test groups consistently responded incorrectly by clicking on A3't in response to every sample, which suggests that their pre-experimental history interfered with their responses. Also, it suggests that although they were trained to match A to B and B to C they did not retain this training when the comparison stimuli were different in the generalization test, although they scored correct on the generalization stimuli in Test Block 1. Similarly, the lack of pre-experimental history in the Control Group resulted random responses to the comparisons stimuli. This can be explained due to their inability to distinguish between the comparison stimuli, since the participants did not have any prior history with them. In contrast the participants in the test groups should have been able to recognize which stimulus fit into which class, but the participants still only responded to the A stimuli despite training. This suggests that pre-experimental history interfere with the emergence of new relations. More studies need to be conducted to investigate if such a generalization test could be used to test for prior histories to stimuli.

Duration and Reaction Time. The number of minutes spent on the experiment per participant was substantially different between the Control Group and the Test Groups. This corresponds with the difference of number of trials used per participants during training. The test groups used almost twice the number of minutes in average to complete the experiment then the Control Group. This suggest that the participants in the test groups may be producing more private behavior then the participants in the Control Group. This could be interpreted as a consequence of pre-experimental history or "remembering" (Catania, 2013, pp. 128–135). No assumptions should be made about this however as there are no previous studies with this particular test to compare these results.

RT was also investigated to see if pre-experimental history might interfere with the amount of time the participants used between clicking on the sample to clicking on one of the three comparison on transitive and equivalence relations in the test blocks. Interesting patterns were found in all group's dependent on; if they passed the test, failed the test and had pre-experimental history with the stimuli. In all groups, those participants that either had, a consistent RT over trials and test blocks, or, reduced RT from Test Block 1 to Test Block 3, passed Test Block 3. Out of these patterns, four participants in Test Group 1, one in the Control Group and two participants in Test Group 2, passed Test Block 1. Those participants that passed Test Block 1 and test Block 3 had consistent RT across all relations, except for the generalization relation (A3C3') which had slower RT across almost all participants that passed both test blocks. Another find was that all participants that had slower RT in Test Block 3 compared to Test Block 1 failed both test blocks. Furthermore, almost all participants in the test groups had slower RT then the participants in the Control Group.

Some clear patterns were found in the comparison of RT versus the emergence of stimulus classes. These patterns may be a distinct measure for pre-experimental history within stimulus equivalence studies. Arguably, private behavior that consists of longer chains of relations, as pre-experimental history, could be what interferes with establishing equivalence classes. This might explain why the participants in the test groups, had slower RT to the comparison stimuli then the participants from the Control Group, even when they establish equivalence classes (Catania, 2013, pp. 376-389; Leppänen & Hietanen, 2004; Vaidya et al., 2015). RT might be a way to study private events indirectly and can shed more light on how pre-experimental history interfere with emergent relations.

Self-Report and Comparison. In the current experiment the participants mostly reported as expected on the bipolar questionnaire. These reporting's correlated with the number of participants that passed and failed Test Block 1 in each group. Even though more participants in Test Group 2 passed Test Block 1 it still fit with the results from the

questionnaire. As the self-reports from Test Group 2 were not as clear as the reports from Test Group 1. However, the questionnaire results did not correspond with the results in Test Block 3. Although the participants reported themselves as strong opposers of the opposite team they still managed to establish equivalence classes with the same stimuli. Which leads to question how valid questionnaires are, as they are based on self-reports and not observable behavior (Domeniconi, de Rose, & Perez, 2014; Lane & Critchfield, 1996). Although, self-reports are not seen as valid measurements on their own in behavior analysis, they can be valuable in giving general information about the participants expectation of themselves and how it deviates from their behavior. Also, it can be a valuable tool in collecting information about stimuli and which one should be picked out for stimulus equivalence experiments. This is confirmed by the results in the current experiment.

Participants in the test groups responded on the questionnaire that the soccer players that were used as names during the test, especially VIF_name_1 was generally liked by both soccer teams. In general, all soccer player names that were used were not scored as 5 or 1 as expected on the scale of the questionnaire. Interestingly, the participants in the test groups had the most difficulty with matching VIF_name_1 to the sample stimuli. Suggesting that although questionnaires should not be considered as a valid way of studying human behavior, they can be sources of information that can guide experiments to some understanding of how some stimuli might affect the participants responses.

Limitations

More than one setting was used to conduct the test due to the difficulty of getting enough participants for each test group. Although the rooms used were very similar the second setting was at a soccer stadium and this in itself might affect the results. Similarly, a difference in the age and gender between the test groups and the control group might be affecting the RT of the participants, since younger participants are more used to computers

(Cooper et al., 2014). However, this might also be a natural consequence of doing studies on pre-experimental history as the number of participants are limited.

Another limitation as mentioned earlier is the use of a linear training structure, as older studies have proved them to be less efficient then OTM and the Many to One training structures (E Arntzen et al., 2010; Sidman, 2000, p. 144). Also, the use of a group design in behavior analysis might be argued to be less valid then some n=1 experiment designs, but the current design was chosen despite this because it was more relevant in a systematic replication (Barlow, Nock, & Hersen, 2009; Cooper et al., 2014). Furthermore, this made more sense since a study on pre-experimental history are comparisons of social contexts in groups and should be arguably studied as groups. Although a group design was used, participants individual responses and results were included in the tables so that less information was lost.

Other limitations can be found in the stimuli used in the experimenter defined classes. The stimuli used might not be stimuli associated to the participants pre-experimental histories, as seen in the results on the name VIF_name_1. However, the questionnaire might be a way to counter this problem. Another problem with the stimuli was that there were similar shapes and colors across classes, which means that participants might have been responding on shapes or color instead of training. If that was the case however, then most participants would have failed on the generalization stimuli I Test Block 1, which was not the case.

Some aspects of the original study were not included in this study. Such as, the low-pitch and high-pitch sounds paired with the written feedback, similarly the pre-training stage was also not included. This might be an argument against the validity of this study as a systematic replication. However, there are studies that show that a paired sound to the written feedback is not necessary for the participants to discriminate between the correct responses and wrong responses (Sidman & Tailby, 1982). The results also confirm that most participants did establish the relations trained, although some did not show the emergence of

emergent relations, without the pairing of the sounds. Also, studies have shown that pre-training in simple stimulus equivalence experiments do not show any significant difference compared to participants that did not receive pre-training (LeBlanc, Miguel, Cummings, Goldsmith, & Carr, 2003). Furthermore, when pre-experimental history is the focus of the experiment a limitation of pre-exposure to the experimental variables is always preferable.

Ethical considerations.

Studies on pre-experimental history should always consider ethical issues before starting any experiments. This is due to the narrow line between acceptable comparisons to guiding comparisons. By guiding comparisons, I refer to studies that guide the participant in a more extreme direction then when they started. In the current study this is handled by being as open as possible about what the study is about to all participants. Playing on the comradery the participants feel to their own group instead of comparing them to the other group. Instructions needs to be formulated thoroughly with ethical issues in mind. Although, even with preparing for this, one can never know the consequences of such studies on the participants after the experiment. That is why studies on more extreme groups, where a long history of violence is present, should be avoided. However, there are groups that aren't violent, and these groups can still shed light on the more extreme groups, as pre-experimental history in stimulus equivalence should in theory have the same function across groups. This however can only be confirmed by more studies on pre-experimental history.

Summary

Pre-experimental history did interfere with the emergence of equivalence classes although almost all participants passed the equivalence test. Also, using a linear training structure makes it unclear if the participants that failed, failed due to their pre-experimental history or due to the training structure. The participants in the different groups differed in reaction times during test although they had similar results in establishing equivalence classes. Reaction time should therefore be included in future studies on pre-experimental

history. Participants in the test groups had distinct response patterns that differed from the Control Group during the generalization test. The adapted generalization test should be replicated further as a method of testing for pre-experimental histories. The questionnaire did not correlate with the results found in the equivalence test, but it gave information that could guide an assumption of certain response patterns to certain stimuli and the generalization test. The results found in this experiment confirms that more studies are needed in this area to challenge the current experiment, previous research and methods.

References

Adcock, A. C., Merwin, R. M., Wilson, K. G., Drake, C. E., Tucker, C. I., & Elliott, C. (2010). The problem is not learning: Facilitated acquisition of stimulus equivalence classes among low-achieving college students. *The Psychological Record, 60*(1), 43–55. doi:https://doi.org/10.1007/BF03395693

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*(3), 437-461.

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*(3), 437–461. doi:https://doi.org/10.1007/BF03395720

Arntzen, E., & Nartey, R. K. (2018). Equivalence class formation as a function of preliminary training with pictorial stimuli. *Journal of the Experimental Analysis of Behavior, 110*(2), 275–291. doi:https://doi.org/10.1002/jeab.466

Barlow, D. H., Nock, M., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior for change*.

Catania, A. C. (2013). *Learning* (5th ed.). Cornwall-on-Hudson: Sloan.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2014). *Applied behavior analysis* (2nd ed.). England: Pearson.

Critchfield, T. S., & Perone, M. (1993). Verbal self-reports about matching to sample: Effects of the number of elements in a compound sample stimulus. *Journal of the Experimental Analysis of Behavior, 59*(1), 193-214. doi:https://doi.org/10.1901/jeab.1993.59-193

Critchfield, T. S., Tucker, J. A., & Vuchinich, R. E. (1998). Self-report methods. In *Handbook of research methods in human operant behavior* (pp. 435-470): Springer.

de Carvalho, M. P., & de Rose, J. C. (2014). Understanding racial attitudes through the stimulus equivalence paradigm. *The Psychological Record, 64*(3), 527–536. doi:https://doi.org/10.1007/s40732-014-0049-4

Dixon, M. R., Rehfeldt, R. A., Zlomke, K. R., & Robinson, A. (2006). Exploring the development and dismantling of equivalence classes involving terrorist stimuli. *The Psychological Record, 56*(1), 83–103. doi:https://doi.org/10.1007/BF03395539

Domeniconi, C., de Rose, J. C., & Perez, W. F. (2014). Effects of correspondence training on self-reports of errors during a reading task. *The Psychological Record, 64*(3), 381–391. doi:https://doi.org/10.1007/s40732-014-0009-z

Eilifsen, C., & Arntzen, E. (2009). On the role of trial types in tests for stimulus equivalence. *European Journal of Behavior Analysis, 10*(2), 187–202. doi:https://doi.org/10.1080/15021149.2009.11434318

Fields, L., & Reeve, K. F. (2001). A methodological integration of generalized equivalence classes, natural categories, and cross-modal perception. *The Psychological Record, 51*(1), 67-87. doi:https://doi.org/10.1007/BF03395387

Fields, L., Reeve, K. F., Matneja, P., Varelas, A., Belanich, J., Fitzer, A., & Shamoun, K. (2002). The formation of a generalized categorization repertoire: Effect of training with multiple domains, samples, and comparisons. *Journal of the Experimental Analysis of Behavior, 78*(3), 291–313. doi:https://doi.org/10.1901/jeab.2002.78-291

Haydu, V. B., Camargo, J., & Bayer, H. (2015). Effects of preexperimental history on the formation of stimulus equivalence classes: A study with supporters of Brazilian soccer clubs. *Psychology Neuroscience, 8*(3), 385. doi:http://dx.doi.org/10.1037/h0101276

Holth, P., & Arntzen, E. (1998). Stimulus familiarity and the delayed emergence of stimulus equivalence or consistent nonequivalence. *The Psychological Record, 48*(1), 81–110. doi:https://doi.org/10.1007/BF03395260

Horne, Pauline J, & Lowe, F. C. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*(1), 185–241. doi:https://doi.org/10.1901/jeab.1996.65-185

Johnson, C., & Sidman, M. (1993). Conditional discrimination and equivalence relations: Control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 59*(2), 333–347. doi:https://doi.org/10.1901/jeab.1993.59-333

Keller, F., & Schoenfeld. (1950). *Principles of Psychology: A Systematic Text in the Science of Behavior*. New York: Appleton-Century-Croft.

Kohlenberg, B. S., Hayes, S. C., & Hayes, L. J. (1991). The transfer of contextual control over equivalence classes through equivalence classes: A possible model of social stereotyping. *Journal of the Experimental Analysis of Behavior, 56*(3), 505-518. doi:https://doi.org/10.1901/jeab.1991.56-505

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences, 98*(26), 15387-15392. doi:https://doi.org/10.1073/pnas.251541498

Lane, S. D., & Critchfield, T. S. (1996). Verbal self-reports of emergent relations in a stimulus equivalence procedure. *Journal of the Experimental Analysis of Behavior, 65*(2), 355–374. doi:https://doi.org/10.1901/jeab.1996.65-355

LeBlanc, L. A., Miguel, C. F., Cummings, A. R., Goldsmith, T. R., & Carr, J. E. (2003). The effects of three stimulus-equivalence testing conditions on emergent US geography relations of children diagnosed with autism. *Behavioral Interventions: Theory Practice in Residential Community-Based Clinical Programs, 18*(4), 279-289. doi:https://doi.org/10.1002/bin.144

Leppänen, J. M., & Hietanen, J. K. (2004). Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychological research, 69*(1–2), 22–29. doi:https://doi.org/10.1007/s00426-003-0157-2

Norsk Senter for Forskningsdata. (2018,January). Retrieved from http://www.nsd.uib.no

Peoples, M., Tierney, K. J., Bracken, M., & McKay, C. (1998). Prior learning and equivalence class formation. *The Psychological Record, 48*(1), 111–120. doi:https://doi.org/10.1007/BF03395261

Pierce, W. D., & Cheney, C. D. (2013). *Behavior analysis and learning*. New York: Psychology Press.

Reeve, K. F., & Fields, L. (2001). Perceptual classes established with forced-choice primary generalization tests and transfer of function. *Journal of the Experimental Analysis of Behavior, 76*(1), 95-114. doi:https://doi.org/10.1901/jeab.2001.76-95

Sidman, M. (1960). *Tactics of scientific research* (Vol. 16.2). University of Michigan: Basic Books.

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

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*(1), 127–146. doi:https://doi.org/10.1901/jeab.2000.74-127

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*(1), 5–22. doi:https://doi.org/10.1901/jeab.1982.37-5

Slaton, J. D., Hanley, G. P., & Raftery, K. J. (2017). Interview-informed functional analyses: A comparison of synthesized and isolated components. *Journal of applied behavior analysis, 50*(2), 252-277. doi:https://doi.org/10.1002/jaba.384

Tajfel, H., Billig, M., & Bundy, R. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology, 3*(1), 27–52. doi:https://doi.org/10.1002/ejsp.2420030103

Vaidya, M., Hudgins, C. D., & Ortu, D. (2015). Conditional discriminations, symmetry, and semantic priming. *Behavioural processes, 118*, 90–97. doi:https://doi.org/10.1016/j.beproc.2015.05.012

Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41*(1), 33–50. doi:https://doi.org/10.1007/BF03395092

Figure 1. Visual examples of the stimuli used in the experiment. The vertical letters show the number of classes, the horizontal numbers show the number of members. The scarf represents the generalization stimuli in Test Block 1. The three members A3't, B3't and C3't represents the comparison used in test block 2.

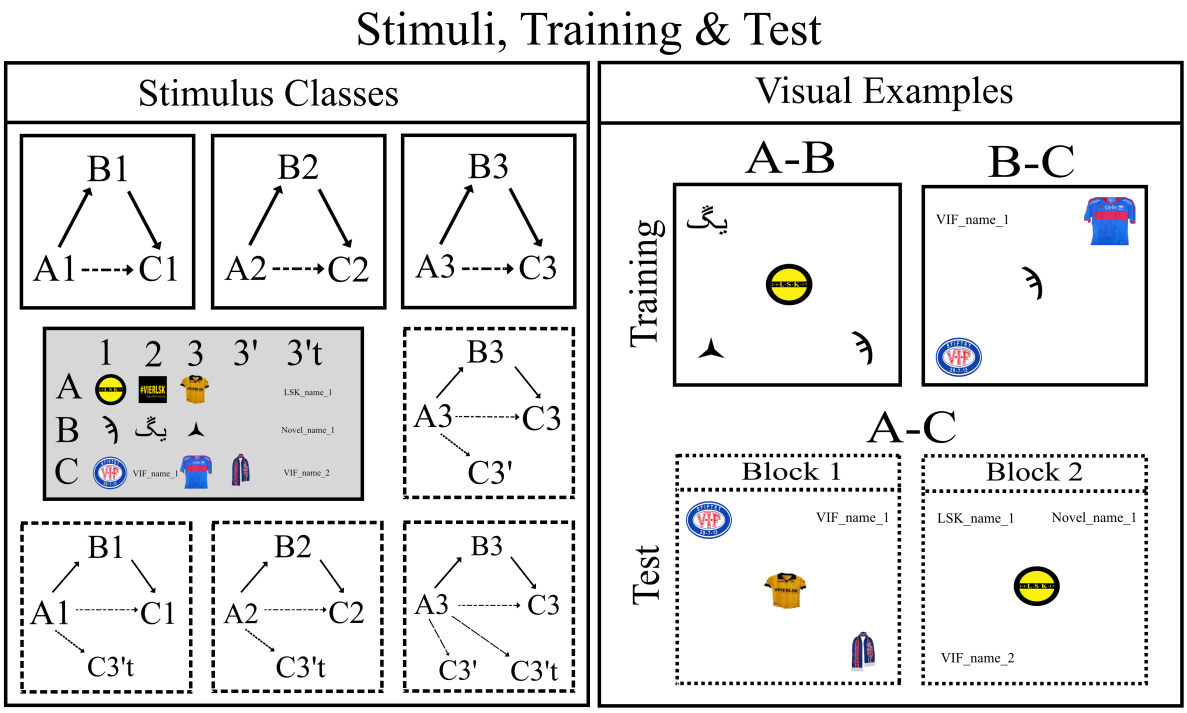

Figure 2. In the right square there are visual examples of the training and Test Block 1 and 2 in accordance with Watt et al. (1991). In the left square there are visual examples of the stimulus class relations. The trained relations are represented with full lined arrows, and the tested relations with dotted lined arrows. The full lined squares within the bigger boxes represent the three equivalence classes. The dotted lined boxes represent the classes with the extra test stimuli.

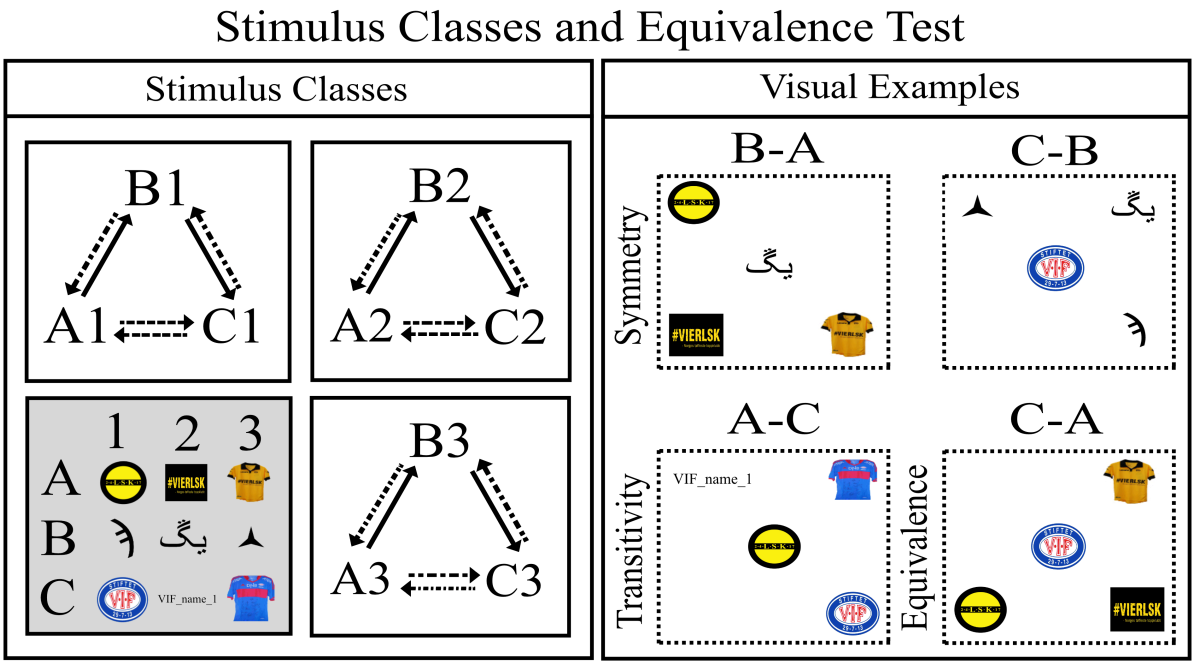

Figure 3. The right square is a visual example of the different emergent relations included in Test Block 3. The box to the left is visual examples of all relations trained and tested including all test blocks. The full lined squares within the left box represent the three full equivalence classes. The full lines represent the trained relations and the dotted lines represent the emergent relations that were tested.

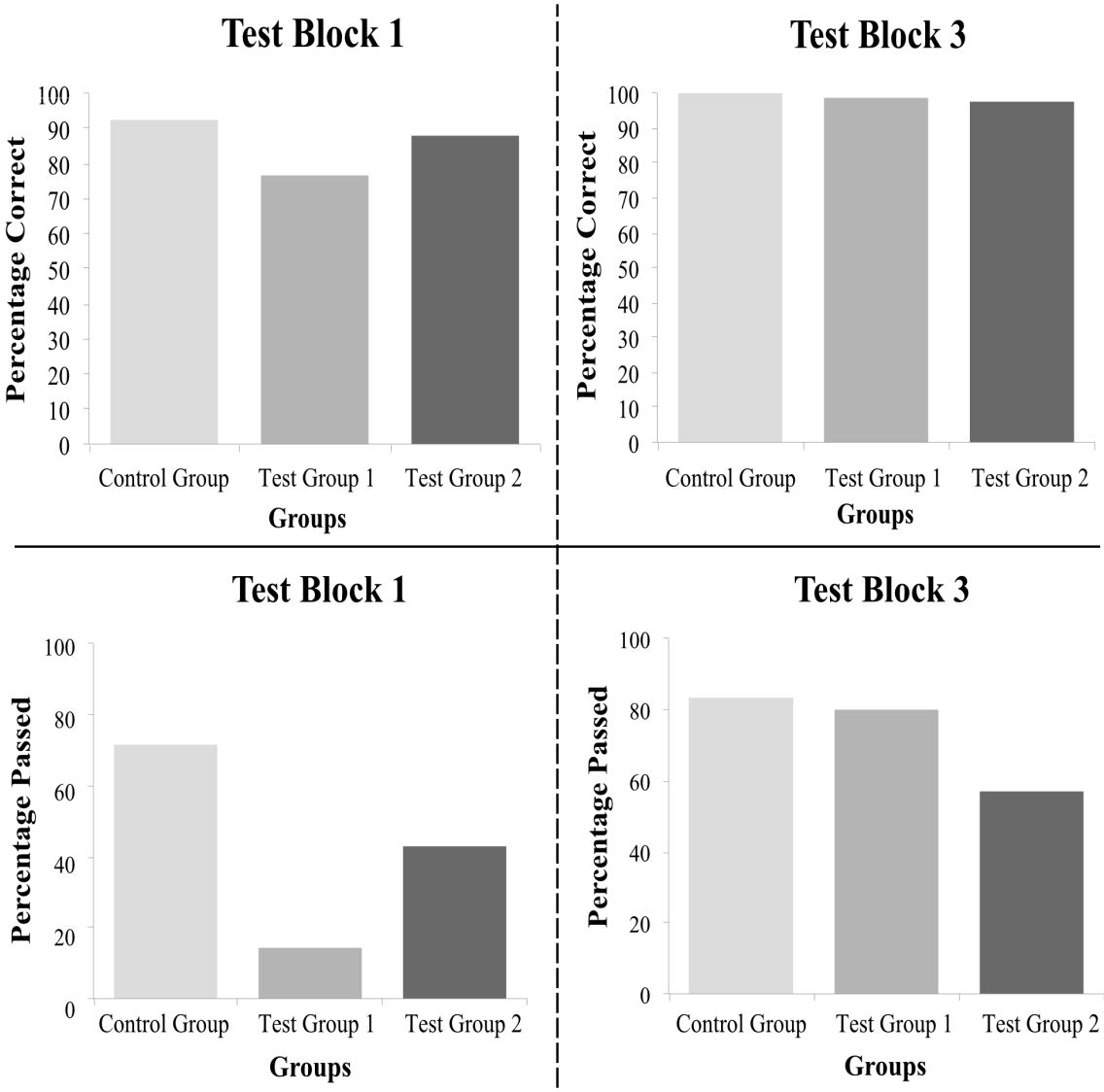

Figure 4. The top two figures show the average percentage correct across all groups. The left figure shows results from test block 1 and the right from Test Block 3. The two figures below the full line shows the average percentage of participants per group that established stimulus equivalence classes. The one to the left below the line shows the percentage passed in Test Block 1. The one to the right below the line shows the percentage passed in Test Block 3.

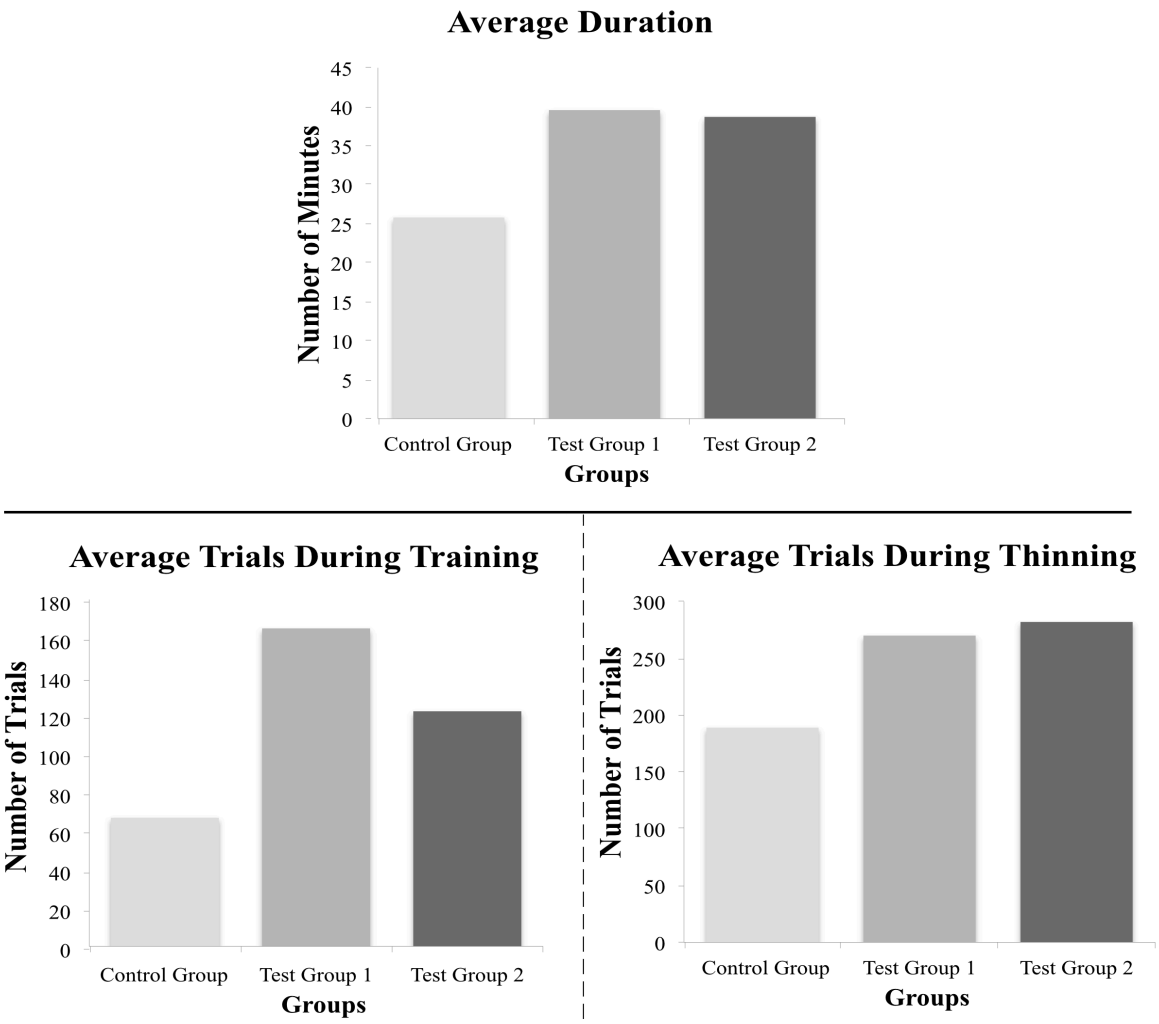

Figure 5. The figure over the full line shows the average minutes, duration, used for the full experiment in each group. The y-axis shows number of minutes and the x-axis shows the different groups. The figure on the bottom left shows the average number of trials used in training across all groups. The bottom right shows the average number of trials used in thinning across all groups. The y-axis represents the number of trials and x-axis represents the different groups in both of the bottom figures.

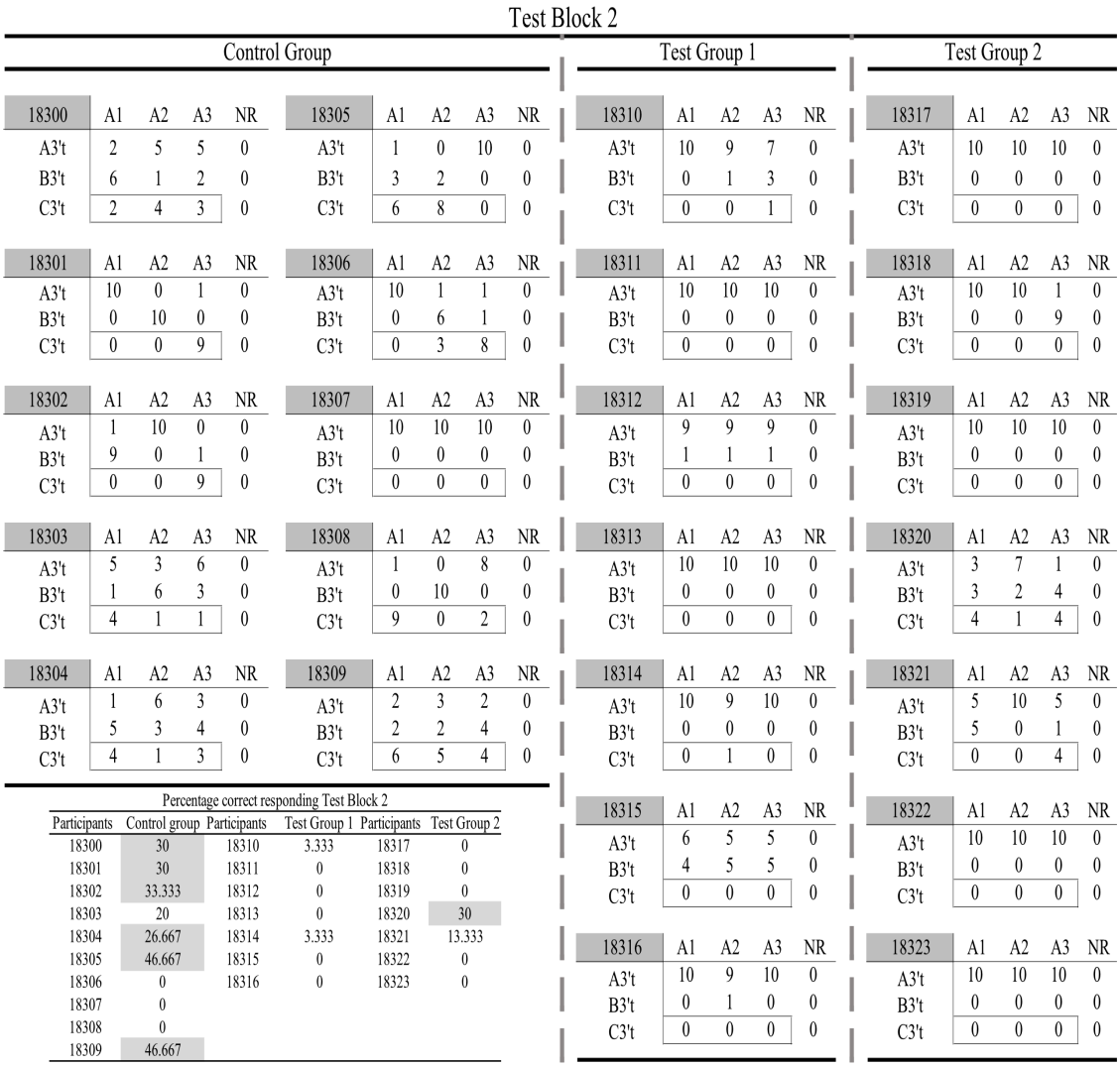

Figure 6. This is a visual presentation of responses made by participants during Test Block 2. The numbers in grey squares represent the participants per group. The two bars to the left represent the control group, the two longer bars represent the two test groups. NR stands for none responses during tests. A1, A2 and A3 are the sample used, A3't, B3't and C3't was the stimuli used as comparison. The three numbers that are inside squares within the matrices next to C3't represents the responses that were defined as correct. The smaller figure to the bottom left shows the percentage correct responding per participants in all groups. The grey squares in this smaller graph show the participants that reached the defined criteria for mastery.

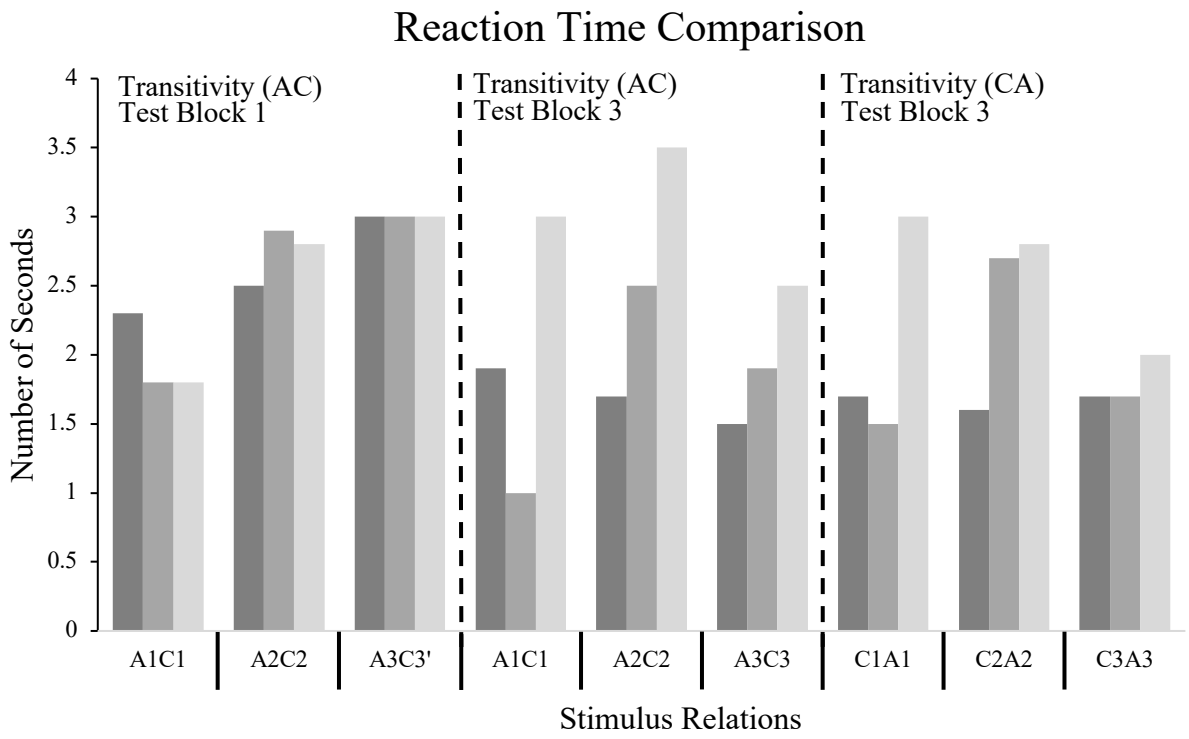

Figure 7. The figure is a presentation of the average reaction time of all participants responses within each group on transitive relations. The reaction time on transitive relations were calculated for Test Block 1 and Test Block 3. The darkest bar represents the Control Group, the middle bar represents Test Group 1 and the lightest bar represents Test Group 2. The y-axis represents number of seconds and the x-axis represents each transitive relation tested.

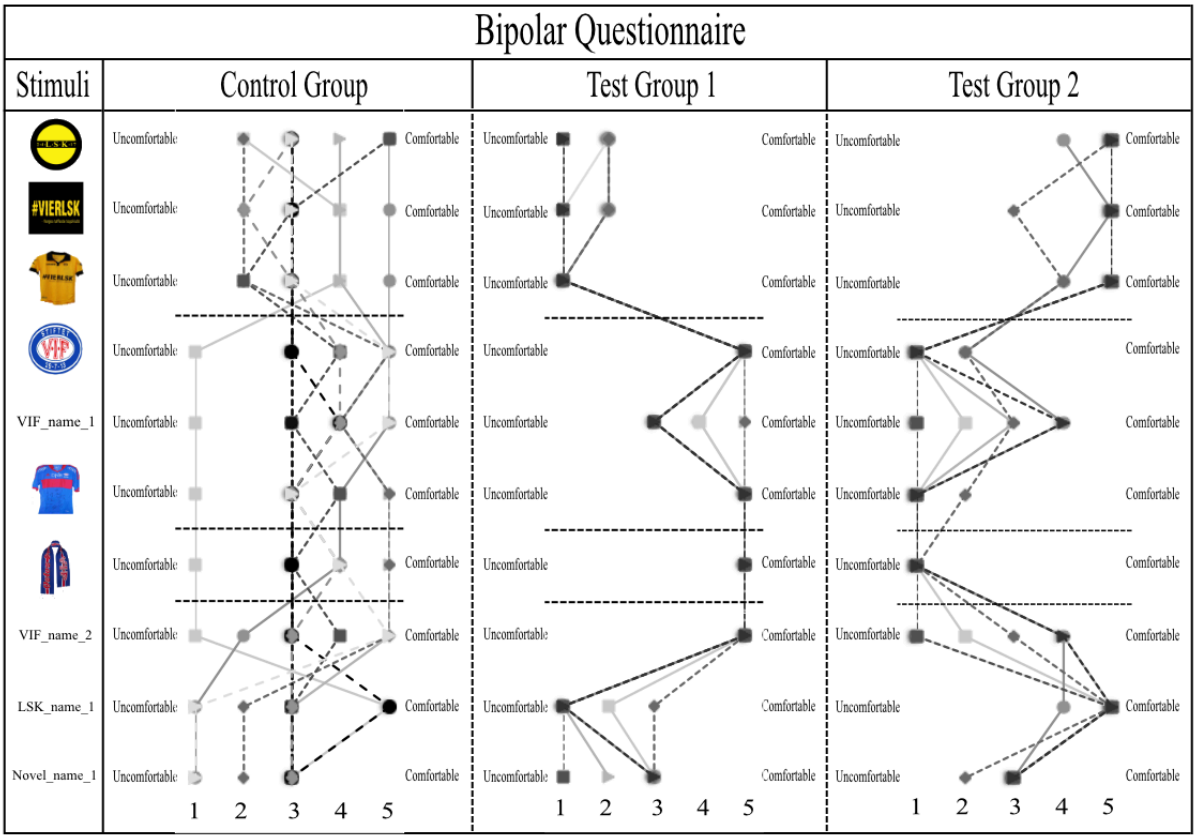

Figure 8. The figure represents the responses the participants made on the bipolar questionnaire after the experiment. The left row of pictures represents the stimuli presented per question, and the row of points to the right of the picture represents the responses made in each group. The left figure shows the control group, the middle test group 1 and the right test group 2. The word uncomfortable to the left and comfortable to the right of each figure represents the adjectives, where the dimension between them, was scored as a Likert scale of 1 to 5 as shown at the bottom of each figure on the x-axis.

Table 1. *Number of correct responses in all test blocks per participant. In test block 1 there were 60 mixed trials, Test Block 2 had 30 mixed trials and test block 3 had 90 mixed trials. Percentage correct on responses are also shown. The numbers to the left of the columns represent the participants. The left number per participant across relations, show correct responding while the number to the right show the correct responses possible. The abbreviations stand for the relations tested per test block.*

	Test Block 1				Test Block 2	Test Block 3				
Participants	BL	Tran. (AC)	Gen. 1	(% cor. responses)	Gen. 2	BL	Sym.	Tran. (AC)	Tran. (CA)	(% cor. responses)
Control Group										
18300	30 of 30	20 of 20	10 of 10	100	9 of 30	N/A	N/A	N/A	N/A	N/A
18301	30 of 30	20 of 20	10 of 10	100	9 of 30	N/A	N/A	N/A	N/A	N/A
18302	30 of 30	20 of 20	10 of 10	100	10 of 30	N/A	N/A	N/A	N/A	N/A
18303	30 of 30	12 of 20	10 of 10	86.6	6 of 30	N/A	N/A	N/A	N/A	N/A
18304	30 of 30	20 of 20	10 of 10	100	8 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18305	30 of 30	10 of 20	10 of 10	83.3	14 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18306	30 of 30	20 of 20	10 of 10	100	0 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18307	30 of 30	20 of 20	10 of 10	100	0 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18308	30 of 30	20 of 20	10 of 10	100	0 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18309	28 of 30	11 of 20	8 of 10	78.3	14 of 30	30 of 30	28 of 30	11 of 15	12 of 15	93.3
Test Group 1										
18310	29 of 30	20 of 20	10 of 10	88.3	0 of 30	N/A	N/A	N/A	N/A	N/A
18311	30 of 30	13 of 20	10 of 10	98.3	1 of 30	N/A	N/A	N/A	N/A	N/A
18312	30 of 30	19 of 20	10 of 10	98.3	0 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18313	30 of 30	20 of 20	10 of 10	100	0 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18314	25 of 30	19 of 20	8 of 10	86.6	1 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18315	30 of 30	17 of 20	10 of 10	95	0 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18316	30 of 30	10 of 20	0 of 10	66.6	0 of 30	30 of 30	30 of 30	12 of 15	12 of 15	93.3
Test Group 2										
18317	30 of 30	0 of 20	10 of 10	66.6	0 0f 30	30 of 30	28 of 30	8 of 15	11 of 15	93.3
18318	30 of 30	20 of 20	10 of 10	100	0 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18319	30 of 30	18 of 20	10 of 10	96.6	9 of 30	30 of 30	30 of 30	14 of 15	11 of 15	100
18320	30 of 30	20 of 20	10 of 10	100	0 of 30	30 of 30	30 of 30	15 of 15	15 of 15	94.4
18321	29 of 30	20 of 20	10 of 10	98.3	4 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100
18322	30 of 30	10 of 20	0 of 10	66.6	0 of 30	29 of 30	29 of 30	15 of 15	15 of 15	97.7
18323	30 of 30	20 of 20	10 of 10	100	0 of 30	30 of 30	30 of 30	15 of 15	15 of 15	100

Note. Some of the participants were not able to do test block 3 and are shown as N/A.

Table 2. *Pearson Chi Square test results on the bipolar questionnaire across all stimuli and all groups.*

Groups		Stimuli									
		Round Yellow	Black Square	Yellow shirt	Round Blue	VIF_name_1	Blue Shirt	Blue Scarf	VIF_name_2	LSK_name_1	Novel_name_1
Control Group	Pearson Chi Square	4.168	4.168	4.168	4.168	4.168	4.168	4.168	4.168	4.168	2.088
	df	9	9	9	9	9	9	9	9	9	9
	p	.135	.135	.135	.287	.287	.135	.287	.406	.135	.001
Test Group 1	Pearson Chi Square	1.653	0.872	0.872	0.872	2.204	0.872	0.872	0.872	0.872	0.872
	df	6	6	6	6	6	6	6	6	6	6
	p	.028	.008	.00001	.00001	.272	.00001	.00001	.00001	.015	.015
Test Group 2	Pearson Chi Square	0.872	0.872	0.872	0.872	2.204	0.872	0.872	2.204	0.872	0.872
	df	6	6	6	6	6	6	6	6	6	6
	p	.0006	.0006	.008	.008	.683	0.0006	.00001	0.446	.0006	.0006

Note. Both test groups did not show expected results when presented stimuli C2.

Table 3. *The average Reaction Time, per participant on the Transitive relations during Test Block 1 and Test Block 3. The darker fields represent the part of the test the participants did not establish equivalence classes.*

		Reaction Time								
		Test Block 1			Test Block 3					
	Participants	A1C1	A2C2	A3C3'	A1C1	A2C2	A3C3	C1A1	C2A2	C3A3
Control Group	18300	2.8	1.9	3.6	N/A	N/A	N/A	N/A	N/A	N/A
	18301	1.5	4.2	2.3	N/A	N/A	N/A	N/A	N/A	N/A
	18302	1.3	1.5	2	N/A	N/A	N/A	N/A	N/A	N/A
	18303	2.3	2.8	2.1	N/A	N/A	N/A	N/A	N/A	N/A
	18304	1.5	2.7	2.5	1.2	1	1	1	1.8	1.4
	18305	5.1	3.3	2.9	3.8	2.2	1	2.4	1.4	1.2
	18306	1.8	1.8	6.5	1.4	2.2	1.4	1	1.5	1.4
	18307	1.1	2	1.5	1.4	1.4	1	1.2	2	1.2
	18308	4.5	2.6	4.6	2.4	1.8	3	2.2	1.4	2.8
	18309	1.3	1.8	2.6	1.2	1.8	1.8	2.6	1.4	2.4
Test Group 1	18310	2.9	2.3	3	N/A	N/A	N/A	N/A	N/A	N/A
	18311	2.5	3.4	3.5	N/A	N/A	N/A	N/A	N/A	N/A
	18312	1.4	3.9	2.6	1	3.2	1.5	1	2	1.6
	18313	1.4	2.2	2.4	1.2	1.6	1.6	1.2	2	1.2
	18314	1.7	2.6	4.3	1.6	1.6	2.4	1.4	1.8	2
	18315	1.5	3.5	3.4	1	2	1.8	1.25	4.5	1.6
	18316	1.6	2.7	2.4	1.4	4	2.25	2.7	3.4	2.2
Test Group 2	18317	1.7	2	2	4	4.8	1.6	1.8	4.75	1.4
	18318	1.5	2	2.5	1	2.2	1.4	1.6	1.5	1.8
	18319	2.7	4.8	2.5	1.4	2.2	2.7	2.2	2	1.6
	18320	1.6	2.5	4.9	1.8	2.4	2.8	1.4	3.6	2
	18321	2.2	2.7	4.5	1.4	2.4	1.6	1.4	1.8	1.4
	18322	1.6	3.7	2.1	12.5	9	5.5	11.8	4.2	5
	18323	1.6	1.9	4.1	1.4	1.6	1.8	1.4	2	1.8

Note. Some of the participants were not able to do Test Block 3 and are shown as N/A.