Stimulus ekvivalens og temporale aspekter

Stimulus Equivalence and Temporal Aspects

Silje Haugland

Læring i komplekse systemer, MALKL 5000, Masteroppgave

Høsten 2009

Høgskolen i Akershus

Abstract

Article 1 is an outline of time as an independent and dependent variable in the study of stimulus equivalence (SE). A general introduction of SE and the different theoretical explanations is describes as a background. Which predictions the different views implicit and empirical findings related to these predictions are discussed. The temporal aspect within other disciplines such as cognitive psychology and neuroscience are discussed as a way of interpreting result from research in SE and methodological challenges in combining disciplines. The equivalent class is discussed in light of findings indicating unequal relatedness among members in an equivalent class, with a focus on Sidman`s account of SE. Article 2 is a systematic replication of Tomanari, Sidman, Rubio and Dube (2006). Restricted time to respond was introduced after training of relations of 3 classes with 3 members in a one-to-many (OTM) training structure. A titrating limited hold (LH) was introduced until the participants responded within 1200 ms to comparison. Test for emergent relations showed that 1 of the 5 participants responded in accordance with equivalence. There was no clear pattern in reaction time (RT) to different relational trial typed during testing.

Time Restrictions and Temporal Analysis –

Important Variables for the Understanding of Stimulus Equivalence

Silje Haugland

Akershus College

Abstract

The topic of the current article is time and stimulus equivalence (SE). An outline of SE is provided with a specification of influencing factors. Three major explanatory theories (Relational Frame Theory, naming and Sidman account) of SE are described and their predictions about time as independent and dependent variable are interpreted. Empirical studies related to naming as a theory and time is analyzed. The support is mostly of indicative nature due to the difficulties of falsifying this theory. The usefulness of findings from cognitive psychology with emphasize on chronometry is discussed. There are methodological difficulties making direct transfer of results from this chronometry to the study of SE. This interdisciplinary approach is supplying new perspectives, but not giving any conclusive support to naming theory. Then different results indicating unequal relatedness between stimuli in an equivalent class is outlined. These results are discussed on the basis of the defining characteristics of SE. A discussion of Sidman account and the finding that reaction time (RT) seems to vary as a function of different relational types are discussed. Future research suggestions are outlined with a focus on research within the inductive paradigm of science.

Time Restrictions and Temporal Analysis –

Important Variables for the Understanding of Stimulus Equivalence

The study of stimulus equivalence (SE) is not a new topic. An example is psychologist and behaviorist (Hull, 1939), who did include SE in his theories about learning. In this early phase, SE was connected theoretically and methodically to techniques of classical conditioning (Green & Saunders, 1998). But then SE was not focused much until Sidman and colleagues revived the field. The major new contribution was the fact that the formation of equivalent classes revealed that after relations between arbitrary stimuli, the participants generated new and untrained relations that could not be explained through the mere presence of reinforcement (Sidman, 1971; Sidman & Cresson, 1973; Sidman, Cresson, & Willson-Morris, 1974). This represented a need for new theory and started a new lineage of research within behavior analysis.

The study of stimulus classes usually involves teaching subjects a series of conditional discriminations and then testing to determine whether new conditional discriminations emerge without direct training in a matching-to-sample (MTS) format (R. R. Saunders, Saunders, Kirby, & Spradlin, 1988). The participant is typically presented a sample stimulus (A1) and then a display of comparison stimuli, where one is correct (B1), then a sample stimulus (A2) followed by comparisons where B2 is correct and then sample A3 with comparison B3 as correct. The stimuli labeled A, B and C are members of the respective classes 1, 2 and 3. Further the participant learns that given A1, A2 and A3 the stimuli C1, C2 and C3 respectively are correct. When these relations are trained, the participant is tested for the relations B1A1, B2A2, B3A3, C1A1, C2A2, C3A3, C1B1, C2B2, C3B3, B1C1, B2C2 and B3C3. These are the emergent relations between stimuli that have never been paired with directly related to each other (Green & Saunders, 1998).

Sidman and Tailby (1982) adapted the definition of equivalence from mathematical set theory in order to provide pivotal criteria for judging if a relation is equivalent. Equivalence requires that the stimuli exhibit the properties of reflexivity, symmetry and equivalence. Reflexivity is demonstrated by generalized identity matching. Symmetry is demonstrated through interchangeability of sample and comparison. Transitivity is demonstrated by responding to the indirect relation between stimuli through their direct relation to a common stimulus. All these relations need to be demonstrated before one conclude that the participants respond in accordance with equivalence and the stimuli constitute an equivalent class (Spencer & Chase, 1996).

This phenomenon of SE has been contributing in different areas. The formation of the classes results in an exponentially number of untrained relations in a generative fashion (Sidman, 1994). Since children master language despite its complexity, the principles of learning has been considered inadequate to account for language acquisition (Chomsky, 1965; Wulfert & Hayes, 1988). The generative nature of SE has been claimed to explain the rapidness of human acquisition of language (Horne & Lowe, 1996). The research linage has been criticized for lack of utility, but there are numerous examples like for instance teaching of people with language disabilities (e. g. Carr & Felce, 2000). The field of SE has also called for conceptual clarifications and a different view on the analytic units in behavior analysis (Sidman, 1994, 2000). And at last SE have been claimed to present a behavioral analytic approach to phenomena like symbolizing and meaning long disputed both within philosophy, behavior analysis and cognitive psychology. Sidman (1986) point is that if stimuli like spoken word dog, the picture of a dog and printed word dog are demonstrated to be equivalent, they will have the same meaning. The relation between a spoken word, a heard word and their common referent have been said to form an equivalent class. Findings indicate that stimulus equivalence can serve as a model for the experimental study of the acquisition of meaning (Wilkinson & McIlvane, 2001), its contextual dependence (Rose, 1996) and categorization (Lane, Clow, Innis, & Critchfield, 1998; Sidman, 1994).

Temporal aspects

Two of the lines of research, that these diverse topics within SE has lead to is the search for influencing factors on the formation of equivalent classes and a search for a theoretical explanation of the phenomena. Temporal aspects, meaning the abstract dimension of the world operationalized in chronometric variables like milliseconds and seconds etc., will inevitably be a part of all human behavior. The general topic of this article is to put time into the context of research on SE. Some basic terms and influencing factors on formation of equivalent classes will outlined first, since these are essential for the understanding of time within SE. The questions raised in this article are how time as a variable (independent or dependent) is predicted given different theoretical explanations of SE. The second question is to what extent predictions have been confirmed through experimental findings. Analogue findings from other psychological disciplines which can be of importance, will be presented. At last suggestions implications of the findings and suggestions for future research will be presented.

Basic terms

Time will always be linked with other variables, since it exists as a continuous dimension alongside the rest of the physical world. To analyze the temporal aspect, we first need a clarification of how these other variables influence the formations of equivalent classes. All variables are not put on experimental trial, but the following gives an outline of some of the findings. Training structure is the sequence of conditional discriminations and arrangements of common stimuli presented to subjects in baseline training (R.R. Saunders &

Green, 1999). Some experiments show that many-to-one (MTO) is superior to one-to-many (OTM) in producing positive equivalence outcome (K. J. Saunders, Saunders, Williams, & Spradlin, 1993; R. R. Saunders, Wachter, & Spradlin, 1988; Spradlin & Saunders, 1986), while other experiments have found the opposite pattern (Arntzen & Holth, 1997, 2000). Training protocols refer to the chronology of how training and testing of different relations are tested. Simple – to – complex (STC) refers to the order of first symmetry, transitivity, equivalence and at last mixed testing. Simultaneous testing refers to training of baseline relations first and then testing of all relations afterwards. Fields et al. (1997) found that the STC protocol was the most effective and that simultaneous protocol was the least effective procedure.

Equivalence have shown to be influenced by the familiarity of stimuli (Arntzen & Holth, 2000; Bentall, Dickins, & Fox, 1993; Holth & Arntzen, 1998; Mandell & Sheen, 1994), to be a direct function of the size and number of nodes in previously established equivalence classes (Buffington, Fields, & Adams, 1997; Fields, Hobbie-Reeve, Adams, & Reeve, 1999) and the number of stimuli in a class also has (Buffington et al., 1997).

Different units of time can be influencing just like the previously mentioned factors and variables. Temporal units will then be independent variables, systematically manipulated to reveal their effect on formation of equivalent classes. Time as an independent variable can be manipulated in the establishment of the prerequisites for SE, meaning the training phase, but also included as independent variable during testing for emergent relations. On the other side, time can be viewed as a dependent variable. Response latency (Bentall et al., 1993) or reaction time (RT) (Spencer and Chase, 1996) is the term for one possible segment of time, which is time measured from a stimulus appears until the participant is responding to it. This can be RT to a sample or a comparison stimulus and it can be measured both in training and testing (Tomanari, Sidman, Rubio & Dube, 2006).

Explanations of stimulus equivalence

Within the behavioral analytic paradigm, there are proposed different theoretical explanations for stimulus equivalence. The following is an attempt to synthesize what the different theories implicitly predicts about the temporal aspect of SE, even though this is not explicit within the different perspectives. Sidman (2000) argues that SE is a phenomenon that is explained through the reinforcing contingencies, probably due to phylogenese. This means that SE is a basic behavioral process just like reinforcement, generalization and discrimination.

In contrast, two other theoretical views consider SE as a product of and depending on learning. *Naming* is an explanation, which chore is that SE is depending on naming as an overarching operant, which is a constituted by listening behavior, echoic and tacting coming together in the same body. Every operant does not require to be reinforced or directly trained, but the behavior will emerge as a result of entering into this loop of behaviors. Success on MTS is theoretically a result of the stimuli entering into the naming relations and practically that the stimuli are labeled covertly or overtly. Naming accounts for SE, by changing arbitrary stimuli into semantic categories through either giving them common names or intraverbal naming (Horne & Lowe, 1996). One major controversy between Sidman`s theory and the naming theory is the question of *mediation*, which is a necessary behavioral processes between the observable responses when the subject is relating stimuli in an equivalent class. Mediation refers to an association of two stimuli taking place because of at third event, a mediator (Sidman, 1994). Note that mediation is not a cognitive process, but behavior that can be unobservable and covert. A third view is Relational Frame Theory (RFT) that considers SE to be one of many relational frames that human beings have the capacity to learn (Hayes, Fox, Gifford, Wilson, Barnes-Holmes and Healy, 2001).

Relational Frame Theory

In RFT the relational responding is a generalized operant caused by multiple exemplar training. The operant is called a relational frame and is characterized by mutual entailment which is similar to reflexivity, combinatorial entailment which is similar to symmetry and transformation of function which is similar to transitivity. The major issue is that equivalence is a characteristic of the relational frame coordination, but this is only one of many relational frames (Hayes, Barnes-Holmes & Roche, 2001). The view of RFT is that mutual entailment, combinatorial entailment and the transformation of stimulus functions are themselves learnt behavior (Hayes et al. 2001). This theory is less concerned with the issue of mechanisms behind the formation of the classes, and refers in general to a history of multiple exemplar training. The view consider derived stimulus relations to be acquired quite early and naturally, but concludes that a fully empirical analysis is difficult to achieve (Hayes, Gifford, & Wilson, 1996). Hayes (1996, p. 311) is also critical to this deduction of predictions from theory when he states: “It would be a very bad thing if the development of behavioral theories leads to traditional hypothesis-testing research. The goal is not to test theories. The goal is to predict and control behavior with precision and scope.”

This rejection of a more specified theory makes clear predictions difficult, but has also lead to a criticism of the theory as impossible to falsify, and thereby not fulfilling scientific criteria. Within the hypothetic deductive paradigm, a axiom is that it is impossible to prove a hypothesis, and a theory is only scientific if it is falsifiable and scientific knowledge is based on accumulation of hypotheses that have resisted many attempts of falsification (Popper, 1959). An implication of this is that it is only possible to draw hypotheses out of a theory which are falsifiable meaning have potential for experimental disconfirmation (Pilgrim, 1996).

Naming

The same critique based on incapability of disproof have been raised against naming (Pilgrim, 1996). Naming as a theory relies on naming as a necessary, mediating response for the formation of equivalent classes (Horne & Lowe, 1996). Several researchers have suggested a close relationship between stimulus equivalence and verbal relations (Barnes, McCullagh, & Keenan, 1990; Catania, 1992; Devany, Hayes, & Nelson, 1986). Responding in accordance with equivalence has been shown in language-able subjects, but not in language-disabled subjects (Barnes et al., 1990; Devany et al., 1986). The studies are based on small numbers of participants and repeated training and testing was not provided, like in some studies of language able grown ups not forming equivalent classes (Sidman, Kirk, & Willson-Morris, 1985; Wulfert & Hayes, 1988). The overall picture so far is that nameing, and the making of connections between stimuli through labeling, may help or may even be sufficient to the formation of equivalence relations, there is an ongoing debate concerning naming and SE (Randell & Remington, 2006).

There are two variants of how naming is involved in stimulus equivalence; common naming or intraverbal naming. The former version is based on a mediating response of naming all members and the latter involves naming each stimulus with an intraverbal as the mediating response in a MTS task (Horne & Lowe, 1996).

One prediction is that reaction time would be shorter using common naming compared to intraverbal naming, due to the length of time taken to pronounce one word compared to two. Meaning that time is influenced by the strategy the subject uses in responding to the relations. One early study examined time as a dependent variable measured as reaction time (Bentall et al., 1993). One group was taught class names and another group was taught individual names of abstract stimuli, prior to establishment of conditional discriminations in a LS training structure. During testing RT on transitivity test trials were longer for the group taught individual names compared to the group taught class names. This result was inconclusive about naming as a theory, but it was an early empirical support that covert naming in SE will require time.

A recent study by Arntzen and Lian (in press) used two conditions and two groups receiving the conditions in different order: nameable picture as sample versus abstract picture as sample, with abstract pictures as comparisons in both conditions in a MTO training structure. RT did increase from baseline to symmetry test trials and from symmetry to equivalences test trials for the group who received abstract sample first. For the group who started off with pictures as sample first, the difference in RT according to relational test type was much shorter. An explanation is that presenting a nameable sample at first, is like prompting a common name, which is similar to the condition in Bentall et al. (1993). Differences in RT based on relational type given common class names to the stimuli was not found. This recent results also just indicates that naming could occur, naming could be a mediating behavior and that mediating behaviors take time.

Naming and LH

This verification that naming requires time, is some of the basis for research on time restrictions. Fields (1996) suggests evidence for naming theory of SE will be emergence of classes without naming and no emergence of classes if naming is absent. Evidence against naming will be failure of class formation when naming is present and class formation when naming is absent. The theoretical logic is to manipulate time as an independent variable to such an extent that naming could not occur. The practical problem is how to experimentally arrange for these contingencies.

One attempt has been introducing the restrictions of LH. This is the time limit, within which the participant has to respond, which results in no reinforcer and sometimes adds feedback that specifies that time was out. In a MTS format this LH will be calculated from comparisons are present and until the participant chooses by responding to one of them (Dickins 2005). When computer software programs are used, there is also a possibility to have LH to sample, meaning that participants sample appears and the participant has to respond to it (e.g. by touching it) within a certain period of time, before the next step in the MTS format is introduced. Additionally LH can be a set point (Holth & Arntzen, 2000) or it can change according to certain criteria, like titration which is a reduction or increase in LH based on the participant's correctness or speed in responding (Tomanari et al. 2006)

Few studies have been conducted putting time restrictions on responding. In a study conducted by Holth and Arntzen (2000) a set LH of 2. 0 s was introduced, after training had reached mastery criteria without LH. This resulted in only 5 out of 10 participants completing training. Then a test of equivalence with a LH of 2.0 s was introduced, which resulted in none of the participants responded in accordance to equivalence. When a second testing without LH was introduced, 1 responded in accordance wiht equivalence in the first half of the testing while 2 others showed correct responding only in the second part of the testing. In general, this study showed, that time restrictions have disruptive effects on mastering training in the MTS format, even though OTM training strucuture was used, which has shown to be quite efficient in training (Arntzen & Holth, 1997, 2000). The authors discussed if this indicated that LH had inhibited some kind of precurrent or mediating behavior. The question if naming is a necessary prerequisite for formation of SE (Horne & Lowe, 1996) or as a contributor but not necessity is still not answered. Results left a question if participants formed equivalent classes, but due to the time restrictions during testing they did not have time to engage in the precurrent behaviors necessary to respond correctly. This was just an indication, since the result is threatened by order effect, since the testing without LH might have lead to learning through testing. In the future there is a need for further studies with LH in training and no LH in testing (Arntzen, 2004; Arntzen, Grohndal & Eilifsen, in press; Arntzen & Lian, in press; Holth & Arntzen, 2000).

Time was in a later study even more restricted. Tomanari et al. (2006) did introduce titrating LH both to comparison and sample during training. The LH was after a training phase with OTM training structure, titrated down in 100 ms steps based on the participants` performance, until an asymptotic level where no further reduction in RT was occurring. This resulted in 3 out of 5 participants responding in accordance to equivalence, based on the immediate transfer data (Dymond & Rehfeldt, 2000) This study is in line with Holth and Arntzen (2000) in the sense that LH seems to inhibit the formation of equivalent responding.

But the major underlying topic of these studies, is to narrow the time to engage in some kind of "precurrent" or mediating behavior. Tomanari et al. (2006) specifically addresses naming as this mediating behavior, but concludes that the study can not answer the question if there were time to name the stimuli.

Analogues in cognitive psychology

Although the experiment is inconclusive about naming as a theory, Tomanari et al. (2006) refers to studies within cognitive science that can give indications about this question. The logic is based on a finding from studies like Schatzman & Schiller (2004), which found that participants after extended practice reached a asymptotic level of 0.58 s in naming aloud of high frequency words in a picture labeling task of words occurring in high frequency. Given that the LH to sample was 0.4 s in Tomanari et al., there was not enough time to label samples. Their calculations regarding comparisons were that time left for observing response and manual response was a LH of 1.1 s minus the required time for labeling (0.58 s). This left 0.52 s to observe, select and perform response to a comparison out of a display of 4 stimuli. To clarify whether this was enough time Tomanari et al. draw an analogue to Dehaene et al. (2001) where participants needed 0.62 s to decide if a printed word was equivalent to natural or manufactured object, which is an indication of RT when more than one stimulus is involved. Adding the motor demand with two stimuli, the time estimate from Schatzman and Schiller (2004) and the fact that Tomanari et al. included 4 stimuli instead of two lead Tomanari et al. to conclude that it is more reasonable to believe that the reinforcing contingencies was a more probable explanation than naming. This line represents a linkage between disciplines representing different paradigms. The authors conclude that there seems to be little literature directly relevant to the MTS format (Tomanari et al., 2006), but there are findings that could indicate interpretations of time as an independent variable in the MTS format.

Chronometry

Is it possible to find indications that not enough time is given to name stimuli in MTS training and testing as a basis of an experimental design. One way cognitive psychology has approached this, is through *chronometry*, where RT has been the central variable. Mental chronometry seeks to measure the time course of mental operations in the human nervous system. In a seminal article, Donders (1869) put forward a subtraction method, based on the principle that behavior can be parted into steps or stages. The method was based on time from a single stimulus appeared until a single response was carried out. This was subtracted from the time it took to make the response to one of two stimuli, and this was assumed to represent the time required to discriminate. Subtraction of time to discriminate was then subtracted from a situation with two possible responses, which was the time estimate of making a choice between to responses. This method provided a way to investigate the cognitive processes underlying simple perceptual-motor tasks, and formed the basis of subsequent developments within time analysis in cognitive psychology. Later these methods have been used together with neuro imaging experiments, like functional Magnetic Resonance

Imaging (fMRI), and can give considerable insight into the nature of human thought (Posner, 2005).

Translated into the behavior analytic paradigm the question is if these methods can reveal characteristics of covert behavior and presumably mediating behaviors. How does the area of research in chronometry map on to the format of MTS used in most studies of SE? Dickins (2005) puts forward the hypothesis that when trained relations had been established there might be five steps in the participant's performance:

> 1) As soon as the sample appears, registration (perception and recognition) of the sample stimulus. 2) This would immediately evoke retrieval of the (by now well established) specific sample-comparison link, leading to the production of an anticipatory representation of the appropriate comparison which would be held thereafter in some kind of working memory. 3) When after the delay, the comparison appear, scanning and registration of the comparison stimuli. 4) Recognition of the comparison stimulus that matches the representation retrieved and held in working memory from stage 2. 5) Organization of corresponding motor response. (p.465).

The challenge is then how to study these steps in isolated variables in experiments. One way is the previous mentioned chronometry and additionally the use of fMRI technology present in the field for the last 20 years (Dickins, 2005).

Starting with RT experiments, these can be grouped into two different types; simple reaction time experiments and choice reaction time experiments. Simple reaction time experiments do not require participants to make any decisions between the presented stimulus, since there are one stimulus present (Ulrich, Mattes, & Miller, 1999). This simple RT will be relevant to response to sample in research on SE, since this response also involves just one stimulus. An analogue to response to sample can be recognition of objects. General findings from cognitive psychology and neuroscience are that recognition of objects is rapid in humans. The estimate is that detection of an object starts about 150 ms after image presentation (Thorpe, Fize, & Marlot, 1996). This latency seems to be a constant both for new novel and previously learnt images of objects, which his indicates that recognition of an object have an asymptotic level when it comes to RT (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001). This finding is relevant for presentation of both comparison and sample, since there seems to be a lower set point for how fast pace of presentation could be used. On the other hand, the as brief presentations as 40 ms have resulted in neural activation (Grill-Spector, Kushnir, Hendler, & Malach, 2000). This means that rationale for length of presentation of stimuli is difficult to find based on research on RT.

But recognition is not the same as naming or labeling an object. Schatzman and Schiller 2004) reported that latencies in labeling of pictures of high frequency reached an asymptote of 580 ms after repeated exposure. This is longer time than the mean RT to sample in Tomanari et al. (2006). But there is a major difference in the fact that the list of words was only 4 in Tomanari et al. meaning 1 sample belonging to 1 of 4 classes. In Schatzman and Schiller (2004) 90 different pictures were included. Retrieval RT of an object name has shown to be linked to how many stimuli in a list of words is to be memorized (Sternberg, 2004), which indicates that RT involving naming would be shorter in Tomanari et. al (2006) compared to Schatzman and Schiller (2004) due to the difference in stimuli numbers. There is therefore no clear connection between the time values of these two studies, but one can suspect that shorter RT is to be expected given only 4 classes.

Choice reaction time experiments involve multiple stimuli that must be discriminated from one another (Ulrich et al., 1999) This is relevant for RT to comparison in a MTS format, since there is a display of more than one comparison to choose from. Johnson and Olshausen (2003) did a study which could be an analogue to the process of making a decision if a stimulus (comparison) is a part of a class or a category (sample). They flashed words belonging to categories, and then presented images to participants, to decide if these were members or no-members of this category. Neurological activation started earliest at 152-300 ms after presentation. These findings are different though from MTS, since there is only one comparison present, but gives indications more than one response alternative increases RT.

Linking of these findings to the MTS format can be done by an additive approach. If there are 4 comparisons, the time it requires to label these will be four times as long as labeling 1. But there are findings complicating such an estimate. Some components may deal with simultaneous operations and may be limited only by a total capacity of central mechanisms. We know that many situations involve parallel processing and feedback loops at many levels (Posner, 2005). Rousselet, Fabre-Thorpe and Thorpe (2002) found that in a categorization task using one image versus two images, there was no difference in RT, indicating parallel processing. There is no conclusion though on tasks including more than two images, but the results supports skepticism against additive conclusions when it comes to RT involving choice between many stimuli.

Another issue that has implications for the findings from chronometry and research within SE, is *processing* speed, which refers to the finding that RT seems to be individual and stable across different tasks (Williams, Myerson, & Hale, 2008). Meaning that some people responds slower in general on a diversity of tasks and the individual differences increase with difficulty of the tasks (Myerson, Hale, Zheng, Jenkins, & Widman, 2003). This makes interpretations of averaged RT across participant somewhat more complicated. This will also touch the methodological element of introducing LH at set time values (e. g. Holth & Arntzen, 2000; Imam, 2001, 2006). This will based on the individuality represent different contingencies for the participants in the studies and complicate comparisons between them. Of course it also means that if RT is going to be a independent variable, this has to be controlled for by the experimental design or through statistical analysis based on high numbers of participants, since the RT does not follow a standard human norm.

This indicates that the symbiosis between chronometry and experimental behavior analysis can be helpful, but that direct transfer of knowledge is cumbersome. There is a need for thorough analysis of this broad field before analogues can be drawn. The new developments in neuroimaging, has lead to finer measure of emergent relations through measuring brain activity during testing. Sclund (2007) found RT shorter transitivity and equivalence than to symmetry relations in a LS training structure with 3 members in a class. He also found that symmetrical relations elicited activation in the nearby parahippocampus while transitive and equivalence relations elicited bilateral activation in the anterior hippocampus, which resulted in the author linking equivalent and transitive supporting hippocampus as important for maintaining relational structures and memory. There are difficulties in the fMRI studies, that correlates between a particiapant`s behavior and the activated area can be confounded with other variables and that the definitions of the areas are blurred (Dickins, 2005). But at least such results as Sclund (2007) indicates that responding to different relational types, does not follow a similar neurological trace. This combination of measures of RT, behavioral responding and neural activity, seems to be reality confirming Skinner (1989) predicting the future of behavior analysis:

> There are two unavoidable gaps in any behavioral account: one between the stimulating action of the environment and the response of the organism and one between consequences and the resulting changes in behavior. Only brain science can fill those gaps. In doing so, it completes the account; it does not give a different account for the same thing. (p.18).

This outline of the findings from cognitive science indicates that conclusions drawn from RT and simple reaction times interpreted additively is a simplification of an area full of open patches in the carpet of knowledge. There are methodological difficulties within each separate discipline as well as cautions to be taken in transferring knowledge from one discipline directly into interpretations of findings in another one. If we assume these interdisciplinary approaches to be useful in exploring the strategies proceeding equivalent responding, there is still a question if RT is an important feature in defining the chore of an equivalent class of stimuli.

The Sidman account

To elaborate the definitional side of an equivalent class an outline the account put forward by Sidman (1994, 2000) is necessary. Sidman (2000) does not see naming as a critical determiner of emergent relations. His point is that reinforcers, stimuli and responses can enter into an equivalent class. Sidman argues that equivalence classes that include all the elements involved in the contingency will automatically follow when a direct reinforcement contingency is put in place, this including the response contingency. This is connected to another feature emphasized by Sidman. Equivalent relations can only be deducted from responding and the criteria are not to stimuli, but to relation between them. Equivalence classes can not be observed, but must be inferred from tests of reflexivity, symmetry and equivalence (Sidman & Tailby, 1982). What bridges the gap between trained and emergent relations is the formation of equivalent classes. When a participant demonstrates all of these relations, the stimuli are said to be equivalent and can be used interchangeably. This interchangeability defines an equivalence class. This is specified in what Sidman (1994) refers to as the *bag analogy*:

> An equivalence relation can be thought of as a bag that contains ordered pairs of all events that the contingency specifies; the bag can be shaken and the elements mixed without regard to any spatial or temporal relations among them. To document the relation, all we have to do is reach into the bag and pull out its member pairs. And reinforcers and response components can be included in this “bag”. (p. 381).

But this interchangeability does not define the criteria for when reflexivity, symmetry and transitivity is concluded. The establishment of emergent stimulus relations is usually inferred when an individual responds correctly on a certain percentage of test trials assessing the emergence of untrained conditional discriminations. Usually this inference is based on percentage of test trials correct. Note that this criteria is just a measure for studying SE.

This has lead to a debate over the nature of SE, where other means of measuring responding in a accordance to equivalence to shed new light on the (Dymond & Rehfeldt, 2000). One of these suggested measures is RT, which could validate the standard measures (Dymond & Rehfeldt, 2001). Differences in RT of responding to stimuli in an equivalent class may be a more precise description of the relations and should be included and not just accuracy alone (Spencer & Chase, 1996).

Nodal distance

A series of studies studying establishing of MTS based on LS training, have incorporated RT as a measure during testing for emergent relations. One of the findings have been that accuracy and RT are inversely related to nodal number (Bentall, Jones, & Dickins, 1999; Imam, 2001; Spencer & Chase, 1996). A node is a stimulus that connects with two other stimuli during training. Another characteristic found is that RT is shortest for the relations involving the first and the last stimuli in a stimulus set, and RT is highest for the relations in between these, forming reaction time as a U-shaped function (Fields et al.,1995). This empirical finding has been explained by *associative distanc*, put forward by Fields et al (1993) and elaborated in later works as *nodal distance* (Fields, Adams, Verhave, & Newman, 1990; Fields et al., 1995). In a study of Bentall et al. (1993) the authors discuss a hypothesis that both naming and nodal distance effect can occur depending on the stimuli. One group were taught conditional discrimination in serial training structure using nameable stimuli and another group using abstract symbols. There was a clear tendency that test trials for transitivity lead to longer RT than for symmetry, and trained relations in the case of abstract symbols, but nameable pictures in testing resulted in equal RT to all test trial types. The critical factors are not the training and the reinforcement history, but number of nodes when it comes to formation of equivalent classes.

Imam (2001, 2003, 2006) did studies based on the assumption that unequality of reinforcement history could be the cause of the nodal distance effect. The studies showed that securing equal number of reinforcers and equal number of tests, diminished or eliminated the nodal distance effect. The results of no nodal distance effect held over two conditions in the experiment, one with and one without a LH of 2.0 s for 7 member classes in all three studies and held across different training protocols (Imam, 2006). These results are in line with reinforcer-contingency perspective on SE (Sidman, 2000).

Different RT to different relational types

Even though there is are empirical contradictions when it comes to RT in LS training structures, most studies seem to find a somewhat similar pattern of responding to different relational trial types. In general the tendency that RT is longest on transitivity and equivalence test trials, shorter on symmetry trials and shortest on trained trials have been supported by quite a few studies (Arntzen, 2004; Arntzen & Eilifsen, in press; Arntzen, Grondahl, & Eilifsen, in press; Arntzen & Lian, in press; Bentall et al., 1993; Holth &

Arntzen, 2000; Spencer & Chase, 1996). The latter study found that RT was lowest on baseline trials, then symmetry and then transitivity and shortest on combined testing, but found no difference between the two latter (Spencer & Chase, 1996). Other studies have only reported a difference between trained and derived relations (Wulfert & Hayes, 1988). This tendency seams to show across LS training structure (Eilifsen & Arntzen, in press, Holth &Arntzen, 2000) OTM training structure (Arntzen & Lian, in press), children as participants (Arntzen & Lian, in press) and having two comparisons (Wulfert & Hayes, 1988) versus three(Eilifsen &Arntzen, in press). One recent study compared MTO, OTM and LStraining structure and found same pattern as mentioned above (Arntzen et al., in press).

Imam (2001) found the same pattern of different RT according to relational types, with increasing RT across trained, symmetry and transitivity test trials and no difference between transitivity and equivalence. In his study the pattern was similar across a condition of LH of 2.0 s to comparison and one condition with no LH. The tendency that RT is shorter to trained than to emergent relations is quite consistent, even though the studies vary in training structure. In comparison, Tomanari et al. (2006) which had shorter LH to comparison and to sample as well, did not find this clear temporal pattern according to relational type. One explanation is that fast responding makes stimuli in an equivalent class more equally related. The lack of difference could be due to the short RT in the study, making the participants more vulnerable to confounding variables like sneeze or noise.

This latter explanation is in line with a parallel in cognitive sciences. Also methodological issues can be refined by a symbiosis of disciplines. Responding involving time is a variable considered a product of underlying behavioral processes will be threatened by the confounding variable fatigue. Examples are Welford (1980), who found that RT got slower when the participant was fatigued. McKeever (1986) notes that undertaking RT experiments, it is essential to motivate particiapants to be as quick as possible. Reaction tasks are usually repetitive and somewhat boring for the participants. He notes based on his experience that participants maintain good motivation for as many as 300 trials within 50 minutes if the trials are grouped into blocks of no more than 30 or 32 trials and brief breaks of 2 minutes or so taken at the end of each block. The actual duration of works sessions and block distributions are often not explicit in studies on SE within the topic of time, which makes comparisons more difficult (e.g. Spencer & Chase, 1996) especially considering that RT is a fragile variable easily confounded with other variables (Tomanari et al. 2006). These latter methodological aspects might be incorporated in future research on RT. The tendency to pattern in RT holds across quite a few studies, so this might not change the general finding.

Unequal relatedness

This pattern of different RT as a function of different relational trial types could reflect unequal relatedness. This is problematic for the reinforcement contingency or Sidman account of SE principally because reinforcement contingencies specify the properties that determine class membership, and to that extent that these properties do not change, the stimuli in the class are substitutable for one another (Imam, 2006). Differential relatedness contradicts that it is the mere reinforcing contingency which would have resulted in substitutable stimuli. (Fields et al. 1993) and is therefore in contrast to Sidman`s theory (R.R. Saunders & Green, 1999; Sidman, 1994).

In his theory, Sidman does not give any specific account of time as a dimension. Sidman (1994) himself calls it a basic component and refers to the concept of class. But others have characterized stimuli in an equivalent relation substitutable (Green and Saunders, 1998). Different reaction times are a result that is incompatible with pure substitution of stimuli in an equivalent class. Will the different RT to relational types found and neuroscience` segregation of different neurological paths for these disconfirm Sidman`s theory of equivalence? Sidman (1994) says that "…the notion that members of a class differ from each other with respect to the criterion for class membership contradicts the very concept of classes." (p. 543). His suggestion is to look at procedural variables and methodological features.

One characteristic of an equivalent class has been that the stimuli are interchangeable can substitute the other (Critchfield & Fienup, 2008). In one way, one could argue that this is a linguistic controversy around interpretations of words like substitutable, interchangeable, and equivalent. Findings of different RT to different relations, does contradict a total or pure understanding of these phrases. On the other side if these are interpreted as behaviorally substitutable to the organism, the time differences will just represent different points of a continuum that all components behavioral analysis try to segregate for the sake of functions in contingencies. The challenge is how to define the operant in equivalent responding. The operant is defined as a correlation between stimuli and responses. But also the orderliness of changes in the correlation is important, meaning that modifiability are an essential feature of the operant relation (Catania, 1973). The question is if time differences in responding to different relations an indication that the relations defining equivalence class membership may be construed as different operant units (Pilgrim, 1996) or if the correlations between the stimuli, reinforcers and responses in an equivalent class verifies the class concept.

Future research

One answer is to this is to search for the orderliness through inductively looking for patterns in responding. The controversies in the theories explaining SE, have resulted in deductive instead of focusing on inductive methods. One plausible question is at what point is a phenomenon like SE described thoroughly enough for a theory to be put forward. One line of research is based on testing of theories like naming and nodal distance effect. Another line of research is more inductive, manipulating variables systematically and describing results with fewer tendencies to interpret findings in light of the major theories. Since the theories are hard for theoretical and practical reasons to falsify, the real interest should be in finding the patterns, the orderliness of changes. It is when orderliness has been thoroughly demonstrated, than we can define a principle and these are the ones which are applicable. The theories will just remain theories. In this way behavioral analysis will remain applied, in the sense Baer, Wolf and Risley (1968) outlined it as useful for people, not just basic research within the field of experimental behavior analysis.

Findings from other psychological disciplines seems to reveal variables worth pursuing, like more fine grained analysis of RT, but caution should be taken in transferring results on an analogue basis from one discipline to the other. Within a field there are methodological difficulties, leading to a need for specialist within these fields to cooperate on research questions to make a synthesis of cognitive, behavioral and neuropsychological sciences possible. According to (Posner, 2005) mental chronometry and thereby studies of RT, is still a cornerstone that binds psychology to the techniques of neuroscience. The fact that time can be represented in a variety of different variables, like RT to sample or comparison, inter-trial-interval, LH to sample or comparison, titration of LH, length of breaks and length of sessions, calls for more research in the area of manipulating time.

References

Arntzen, E. (2004). Probability of Equivalence Formation : Familiar Stimuli and Training Sequence. *The Psychological Record, 54*(2), 275-291. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Arntzen, E., Grondahl, T., & Eilifsen, C. (in press). The effect of different training structures in the establishment of conditional discriminations and the subsequent performance on the test for stimulus equivalence. *The Psychological Record.*

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47*(2), 309-320. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

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

Arntzen, E., & Lian, T. (in press). Trained and derived relations with pictures as nodes. *The Psychological Record*.

Baer, D. B., Wolf, M. M. & T. R. Riesley. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91-97. doi:10.1901/jaba.1968.1-91

Barnes, D., McCullagh, P., & Keenan, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *The Analysis of Verbal behavior, 8*, 1-11. Retrieved from https://apps.abainternational.org/store/ViewJournals.asp?strJournalType=avb

Bentall, R. P., Dickins, D. W., & Fox, S. R. A. (1993). Naming and equivalence: Response latencies for emergent relations. *The Quarterly Journal of Experimental Psychology, 46*(2), 187-214. Retrieved from http://www.eps.ac.uk/index.php/journals

Bentall, R. P., Jones, R. M., & Dickins, D. W. (1999). Control over emergent relations during formation of equivalent classes: Response error and latency for 5-member classes. *The Psychological Record, 49*, 93-116. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Buffington, D. M., Fields, L., & Adams, B. J. (1997). Enhancing the formation of equivalence classes by pretraining of other equivalence classes. *The Psychological Record, 47*, 1-20. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Carr, D & Felce, D. (2000). Application of stimulus equivalence to language intervention for individuals with severe linguistic disabilities. *Journal of Intellectual and Developmental Disability, 25*(3), 181-205. Retrieved from http://informahealtcare.com/jid

Catania. (1973). The concept of the operant in the analysis of behavior. *Behaviorism, 1*(2), 103.

Catania, A. C. (1992). B. F. Skinner, Organism. *American Psychologist, 47*(11), 1521-1530. doi:1510.1037/0003-1066X.1547.1511.1521.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Critchfield, T. S., & Fienup, D. (2008). Stimulus equivalence. In S. F. Davis & W. Buskist (Eds.), *21st century psychology: a handbook* (pp. 360-372). Los Angeles, CA: Sage.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience, 4*(7), 752-758. doi:10.1038/89551

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:10.1901/jeab.1986.46-243

Dickins, D. W. (2005). On aims and methods in the neuroimaging of derived relations. *Journal of the Experimental Analysis of Behavior, 84*(3), 453-483. doi:10.1901/jeab.2005.92-04

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412-431. doi:10.1016/0001-6918(69)90065-1

Dymond, S., & Rehfeldt, R. (2000). Understanding complex behaviour: The transformation of stimulus functions. *The Behavior Analyst, 23*(2), 239-254. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/journals/557/

Dymond, S., & Rehfeldt, R. (2001). Supplemental measures of derived stimulus relations. *Experimental analysis of human behavior bulletin, 19*, 8-12. Retrieved from http://www.eahb.org/NewSitePages/BulletinHomepage.htm

Eilifsen, C., & Arntzen, E. (in press). On the role of trial types in tests for stimulus equivalence. *European Journal of Behavior Analysis*.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience, 13*(2), 1-10. doi:10.1162/089892901564234

Fields, L. (1996). The evidence for naming as a cause or facilitator of equivalence class formation. *Journal of the Experimental Analysis of Behavior,65*(1), 279-282. doi:210.1901/jeab.1996.1965-1279. doi:10.1901/jeab.1996.65-279

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53*(3), 345-358. doi:10.1901/jeab.1990.53-345

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1993). Are stimuli in equivalence classes equally related to each other. *Psychological Record, 43*, 85-106. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

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

Fields, L., Landon-Jimenez, V., Buffington, D. M., & Adams, B. J. (1995). Maintained nodal-distance effects in equivalence classes. *Journal of the Experimental Analysis of Behavior, 64*(2), 129-145. doi:10.1901/jeab.1995.64-129

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

Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229-262). New York, NY: Plenum Press.

Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience, 3*(8), 837-843. doi:10.1038/77754

Hayes, S. C. (1996). Developing a theory of derived stimulus relations. *Journal of the Experimental Analysis of Behavior, 65*(1), 309-311. doi:10.1901/jeab.1996.65-309

Hayes, S.C., Barnes-Holmes, D. & Roche, B. (2001). Relational frame theory: A prècis. In S.C. Hayes, D. Barnes-Holmes & B. Roche (Eds.) *Relational frame theory. A post-skinnerian account of human language and cognition.* (pp. 141-154). New York, NY: Kluwer Academic / Plenum Publishers.

Hayes, S. C., Fox, E., Gifford, E. V. & Wilson, K. G. (2001). Derived relational responding as learned behavior. In S.C. Hayes, D. Barnes-Holmes & B. Roche (Eds.) *Relational* *frame theory. A post-skinnerian account of human language and cognition.* (pp. 21-50). New York, NY: Kluwer Academic / Plenum Publishers.

Hayes, S. C., Gifford, E. V. & Wilson, K. G. (1996). Stimulus classes and stimulus relations: Arbitrary applicable relational responding as an operant. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in human and animals* (pp. 279-300). Amsterdam, Netherlands: Elsevier.

Holth, P., & Arntzen, E. (1998). Stimulus familiarity and the delayed emergence of stimulus equivalence or consistent nonequivalence. *The Psychological Record, 48*(1), 81-110. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Holth, P., & Arntzen, E. (2000). Reaction times and the emergence of class cosistent responding: A case for precurrent responding? *The Psychological Record, 50*(3), 603-628. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

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

Hull, C. L. (1939). The problem of stimulus equivalence in behavior theory. *Psychological Review, 46*(1), 9-30. Retrieved from http://psycnet.apa.org/journals/rev/46/1/

Imam, A. A. (2001). Speed contingencies, number of stimulus presentations, and the nodality effect in equivalence class formation. *Journal of the Experimental Analysis of Behavior, 76*(3), 265-288. doi:10.1901/jeab.2001.76-265

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

Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision, 3*(7), 499-512. doi:10.1167/3.7.4

Lane, S. D., Clow, J. K., Innis, A., & Critchfield, T. S. (1998). Generalization of cross-modal stimulus equivalence classes: operant processes as components in human category formation. *Journal of the Experimental Analysis of Behavior, 70*(3), 267-279. doi:10.1901/jeab.1998.70-267

Mandell, C., & Sheen, V. (1994). Equivalence class formation as a function of the pronounceability of the sample stimulus. *Behavioural Processes, 32*(1), 29-46. doi:10.1016/0376-6357(94)90025-6

McKeever, W. F. (1986). Tachistoscopic methods in neuropsychology. In H. J. Hannay (Ed.), *Experimental techniques in human neuropsychology* (pp. 167-211). New York, NY: Oxford University Press.

Myerson, J., Hale, S., Zheng, Y., Jenkins, L., & Widman, K. F. (2003). The difference engine: A model of diversity in speeded cognition. *Psychonomic Bulletin and Review, 10*(2), 262-288. Retrieved from http://pbr.psychonomic-journals.org/

Pilgrim, C. (1996). Can the naming hypothesis be falsified? *Journal of the Experimental Analysis of Behavior, 65*(1), 284-286. doi:10.1901/jeab.1996.65-284

Popper, K. (1959). *The logic of scientific discovery*. New York, NY: Routledge

Posner, M. I. (2005). Timing the brain: Mental chronometry as a tool in neuroscience. *PLoS Biol, 3*(2), 204-206. doi:210.1371/journal.pbio.0030051

Randell, T. & Remington, B. (2006). Equivalence realtions, contextual control, and naming. *Journal of the Experimental Analysis of Behavior, 86*(3), 337-354. doi:10.1901/jeab.2006.82-05

Rose, J. C. (1996). Controlling factors in conditional discriminations and tests for equivalence. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 253-278). Amsterdam, Netherlands: Elsevier.

Rousselet, G. A., Fabre-Thorpe, M. & Thorpe, S. J. (2002). *Nature Neuroscience, 5,* 629-630. doi:10.1038/nn866

Saunders, K. J., Saunders, R. R., Williams, D. C., & Spradlin, J. E. (1993). An interaction of instructions and training design on stimulus class formation: Extending the analysis of equivalence. *The Psychological Record, 43*, 725-744. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

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

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior, 50*(2), 145-162. doi:10.1901/jeab.1988.50-145

Saunders, R. R., Wachter, J. A., & Spradlin, J. E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedure. *Journal of the Experimental Analysis of Behavior, 49*(1), 95-115. doi:10.1901/jeab.1988.49-95

Schatzman, K. B., & Schiller, N. O. (2004). The word frequency effect in picture naming: Contrasting two hypotheses using homonym pictures. *Brain and Language, 90,* 160-169. doi:10.1016/S0093-934X(03)00429-2

Sidman, M. (1971). Reading and auditory-visual equivalence. *Journal of Speech and Hearing Research, 14,* 5-13. Retrieved from http://jslhr.asha.org/cgi/content/abstract/14/1/5

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213-245). Hillsdale, NJ: Erlbaum.

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

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

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

Sidman, M., Cresson, O., & Willson-Morris, M. (1974). Acquisition and matching to sample via mediated transfer. *Journal of the Experimental Analysis of Behavior, 22*(2), 261-275. doi:10.1901/jeab.1974.22-261

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

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

Skinner, B. F. (1989). *Recent Issues in the Analysis of Behavior*. Columbus, OH: Merrill Pub. Co.

Spencer, T. J., & Chase, P. N. (1996). Speed analyses of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65*(3), 643–659. doi:10.1901/jeab.1996.65-643

Spradlin, J. E., & Saunders, R. R. (1986). The development of stimulus classes using match-to-sample procedures: Sample classification versus comparison classification. *Analysis* *and Intervention in Developmental Disabilities, 6*, 41-58. doi:10.1016/0270-4684(86)90005-4

Sternberg, S. (2004). Memory-scanning. Mental processes revealed by reaction-time experiments. In D.A. Balota & E. J. Marsch (Eds.) *Cognitive psychology* (pp.48-74). New York, NY: Psychology Press.

Stromer, R. (1996). On the experimental analysis of naming and the formation of stimulus classes. *Journal of the Experimental Analysis of Behavior, 65*(1), 250-252. doi:10.1901/jeab.1996.65-250

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520-522. doi:10.1038/381520a0

Tomanari, G. Y., Sidman, M., Rubio, A. R., & Dube, W. V. (2006). Equivalence classes with requirements for short response latencies. *Journal of the Experimental Analysis of Behavior, 85*(3), 349-369. doi:10.1901/jeab.2006.107-04

Ulrich, R., Mattes, S., & Miller, J. (1999). Donder's assumption of pure insertion: an evaluation on the basis of response dynamics. *Acta Psychologica, 102*, 43-75. doi:10.1016/S0001-6918(99)00019-0

Welford, A. T. (1980). Choice reaction time: Basic concepts. In A. T. Welford (Ed.), *Reaction times* (pp. 73-128). New York, NY: Academic Press.

Wilkinson, K. M., & McIlvane, W. J. (2001). Methods for studying symbolic behavior and category formation: Contributions of stimulus equivalence research. *Developmental Review, 21*(3), 355-374. doi:310.1006/drev.2000.0526

Williams, B., Myerson, J., & Hale, S. (2008). Individual differences, intelligence, and behavior analysis. *Journal of the Experimental Analysis of Behavior, 90*(2), 219-231. doi:10.1901/jeab.2008.90-219

Wulfert, E., & Hayes, S. C. (1988). The transfer of conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50*(2), 125-144. doi:10.1901/jeab.1988.50-125

Responding in Accordance to Equivalence as a Function of Titrating Limited Hold

Silje Haugland

Akershus College

Abstract

Reaction time (RT) seems to vary as a function of relational type (e.g. Spencer & Chase, 1996). Responding given strict restrictions on time to respond, have resulted in 3 of 5 participants responding in accordance with equivalence (Tomanari, Sidman, Rubio & Dube, 2006). The current study was a systematic replication of Tomanari et al. The purpose was to establish 3 classes with 3 members, using arbitrary stimuli and a one-to-many (OTM) training structure. Touch screen and software based recordings and calculations were used. After training of the relations, a limited hold was calculated for each of the 5 participants based on their actual reaction time (RT). Then training with a titrating limited hold (LH) in 100 ms steps was conducted until all participants responded within 1200 ms and at last feedback was faded. Simultaneous testing for emergent relations had a LH of 2500 ms. Responding in accordance with equivalence was found in 1 of the 5 Participants. The endurance of the study was shorter than Tomanari et al. The pattern of increasing RT to trained, symmetry and equivalence test trials respectively, was not found.

Responding in Accordance to Equivalence as a Function of
Titrating Limited Hold

Responding in accordance with stimulus equivalence (SE) after establishment of conditional discriminations in a matching-to-sample (MTS) format, have been reported and is easily generated in humans (Catania, 1992). There are an amount of studies, showing that responding in accordance to equivalence do emerge in humans (e.g. Bentall, Dickins, & Fox, 1993). On the other hand, several studies have shown that stimulus classes do not form in all humans (e.g. Holth & Arntzen, 2000). This has lead to the search for variables crucial to the probability of equivalent formation (e.g. Arntzen, 2004). One example of this is breakdown of trained relations during testing (Saunders & Green, 1999).

The definition of SE requires the properties of reflexivity, symmetry and equivalence (Sidman & Tailby, 1982), but the percentage correct and numbers of correct test trials has not been explicit. The fact that several studies have reported learning during testing (Barnes & Keenan, 1993) has lead to a focus on immediacy, meaning that correct responding on first test trials are better indications of SE than responding correctly later in testing (Dymond & Rehfeldt, 2000). This debate over how to define SE is due to SE not existing as a thing, and thereby it can not be observed directly, but has to be inferred from the new units that emerge (Sidman, 1994). Definitional theorizing and empirical search for influencing factors have included time as a variable. Even though SE is formed to a limited degree or imperfect in some organisms or specific species, learning is always connected to time.

There are three issues connected to time as variable: 1) Time as dependent variable. 2) Time as independent variable. 3) How to interpret any temporal patterns in responding connected to other variables (Spencer & Chase, 1996).

Time as dependent variable is usually measured in reaction time (RT) meaning the time from a stimulus appears until the participant responds to it. This could either be to a sample or a comparison stimulus in a MTS task. Including measures of RT, can result in a more fine grained analysis of SE (Dymond & Rehfeldt, 2000; Holth & Arntzen, 2000; Spencer & Chase, 1996).

The tendency found is longer latency for derived relations compared to trained relations (Wulfert & Hayes, 1988). Differentiation between the different derived relations has been shown. Spencer and Chase (1996) found that RT during testing increased with relational type. RT to baseline was shortest and symmetry trials had higher RT. Transitivity and combined test trials had the highest RT, but there were no difference between the two latter. Several other studies have also found the pattern that RT increases with the relational types trained, symmetry and transitivity respectively (Arntzen, 2004; Eilifsen & Arntzen, in press; Arntzen, Grondahl, & Eilifsen, in press; Arntzen & Lian, in press; Bentall et al., 1993; Holth & Arntzen, 2000; Spencer & Chase, 1996).

The tendency that RT is shorter when trained relations are tested compared to emergent relations is quite consistent, even though the studies vary in training structure. One recent study compared many-to-one (MTO), one-to-many (OTM) and linear series (LS) training structure and found the same pattern as mentioned above (Arntzen et al., in press). There are other methodological differences as well. Wulfert and Hayes (1988) and Bentall et al. (1993) did not require overt response to sample, but in later studies (e.g. Spencer & Chase, 1996) this response has been included. This could explain why the first two studies did not find significant differences in RT to baseline trials and symmetry trials. In Spencer and Chase (1996) responding was both contingency shaped and rule based, but the same pattern of RT to relational type occurred across both conditions. There were also a difference in class sizes established, as in Spencer and Chase, (1996), where they included 7 members in each of 3 classes, while Bentall et al. (1993) had 6 classes with 3 stimuli each. Wulfert and Hayes (1988) had only 2 comparison stimuli displayed Several studies have also shown RT to

decrease during testing for all relational types (Arntzen, 2004; Eilifsen & Arntzen, in press; Arntzen et al., in press; Arntzen & Lian, in press).

How should differences in reaction time to different relations be interpreted? Fields, Adams, Verhave and Newman (1990) refers to the *nodal distance effect* meaning that in LS training the RT reflects that when stimuli trained in a series, the RT increases with increasing number of nodes. Nodes are stimuli linked by training to at least two other stimuli.

Some studies have shown RT to be an inverse function of nodal number (Bentall et al., 1993; Imam, 2001; Spencer & Chase, 1996). The findings of Imam (2001, 2006) was that there were no differences in RT when the number of trials on each relation had been equalized. The hypothesis is that RT decreases with number of presentations of a stimulus-stimulus presentation, meaning trials in training.

This nodal distance effect is assumed to reflect that stimuli are unequally related, meaning that there is more neurological effort the longer nodal distance which shows in increased reaction time (Fields, Hobbie-Reeve, Adams, & Reeve, 1999; Fields, Landon-Jimenez, Buffington, & Adams, 1995). The same logic may apply to OTM training structure, where the effort is bigger when the relation is emerged as opposed to directly trained. Symmetry will be tested with two stimuli appearing together in different positions compared to in training. Tests for equivalence will involve stimuli that have never appeared together (Green & Saunders, 1998).

While Sidman (1994, 2000) defines SE through reflexivity, symmetry and transitivity which are mutually dependent to define an equivalent class which result is that stimuli become substitutable (Green & Saunders, 1998), the notion of nodal distance effect (Fields et al., 1995) is not in line with this equality (Imam, 2006). Sidman (1994) explains the formation of SE as a result of the reinforcing contingency, like a basic and innate behavioral function probably due to phylogenies, no different from discrimination and generalization, and with no further need of explanation. From his point of view, the responding in accordance to equivalence is not dependent on learning history. In contrast Horne and Lowe (1996) includes a behavioral history of naming as a prerequisite for mediation which is necessary for formation of emergent equivalent responding.

Few studies have manipulated time as an independent variable. This is done but setting a time limit, which the participants have to respond within or a limited hold (LH), which means that the time to respond is restricted, and responding too late, will lead to no reinforcement. The thematic background have been that mediation, naming (Horne & Lowe, 1996), requires time. In Holth and Arntzen (2000) 10 participants were trained in MTS using OTM with a 2.0 s LH to comparison during last phase of training and in testing. RT to sample increased when restrictions to comparison were introduced. Five out of ten participants completed training and none of these responded in accordance to equivalence. In a second equivalence test with no restrictions on RT, equivalent responding occurred in 3 participants.

The challenge of interpretations within the topic naming, is the lack of experimental control with variables connected to this explanation. The variable could be a covert response and there are challenges regarding how to isolate and control it. One line of reasoning has been to include variables from cognitive psychology (Dickins, 2005). One cognitive field is *chronometry*, in which conclusions about human information processing are reached through measures of subjects` RT (Meyer, Osman, Irwin & Yantis, 1988, p. 3). The limitations are that these findings are just analogue, since there are no studies measuring directly the time course of behaviors involved in the MTS format used within research on SE. It is within this thematic context, that Tomanari et al. (2006) conducted the only study this author know of, manipulating time as an independent variable both to sample and comparison. The purpose according to (Tomanari et al., 2006), was to test for emergent conditional discriminations included rapid-responding contingencies. The question was if new performances indicative of equivalence would emerge even when the participants had minimal opportunity to engage in subvocal, mediating behavior. Both LH to comparison and sample was titrated, meaning that LH was reduced according to the participants' performance of accuracy and responding within the LH. Time was titrated up and down, based on improvement or worsening of correctness and velocity of responding. There were 3 training phases and a test phase. First relations were established using an OTM training structure in which trials were presented on serialized basis without LH. Then a second phase followed with LH to both comparison and sample, but now titrating was introduced until responding reached an asymptotic level. Asymptotic meant that no further reduction in velocity was observed, so that RT stabilized at a certain level. Then followed a third phase with fading of consequences and at last a phase of testing for emergent relations.

Time restrictions seemed to result in a large number of trials before criteria of mastery was reached in this study. Calculations of the titration were also done manually, and this could impact the high number of trials. Criteria in training were reached after 22-34 sessions with a total number of trials varying from 11088 to 20928. Based on the number of sessions and outline of the method, one can assume that the study was conducted across several weeks, even though this is not explicit in the article. Three of five participants showed responding in accordance with equivalence. Baseline accuracy scores were maintained in all test blocks. There was no clear tendency that subjects who responded according to equivalent relations did have the shortest RT. LH values in the phase with titration reached about 1300 ms for comparisons and 500 ms for sample.

The purpose of this study was to investigate emergent relations given restricted time to responding by keeping inte- trial-intervals at a minimum and give short time to respond to comparison during testing. The current study was a systematic replication of Tomanari et al. (2006) with the purpose of reducing number of trials needed by the participants and thereby the period where the participants were free to engage in behaviors associated to the training (e.g. rehearsal). The major changed methodological features were: (1) The LH was applied only to comparison and not to sample, (2) work sessions were kept shorter to reduce fatigue during training and testing, (3) class sizes were reduced from 4 to 3 members, (4) the study involved a slightly different training protocol (Imam, 2006) including a mixed test of equivalence instead for separate blocks of equivalence mixed with trained test trials and then symmetry mixed with trained test trials.

One of the questions was if untrained relations would emerge given rapid responding contingencies to comparison, with a learning history of shorter endurance than previous studies` experimental conditions had resulted in (Tomanari et al., 2006). This could shed light on both time as a variable in the learning history of trained relations and on the emergent relations in themselves. A second question is if the pattern of different RT to trained, symmetry and equivalence shows the same pattern as previous studies have reported (e.g. Spencer & Chase, 1996). A third question was how RT to sample was influenced by the LH to comparison.

Method

Participants

Five female participants were recruited through a university system. Age range was 24 – 63 years. Participant #1, #3, #4 and #5 were Bachelor students. Participant #2 was employed at the University. None of the participants had any learning history within the area of SE. The participants were given written information and a consent form, emphasizing that they could withdraw from the study at any point.

Setting

All sessions were conducted in a quiet 2 x 4 m room at the University. The participants were seated on a chair by a 90 x 150 cm table. The windows were covered by curtains to reduce disturbances from outside. The computer in use was placed about 30 cm into the centre of the table.

Apparatus

In the experiment a HP Compaq nc 6320 portable computer with a touch screen (Magic Touch by Keytec) of 23 x 30 cm was used. Computer software by Psych Fusion Ltd made in collaboration with Eric Arntzen, presented all experimental events and recorded responses. Participants received 100NKR a day, which involved 65 minutes of work including breaks.

Stimuli

Experimental stimuli were 9 arbitrary symbols (see Figure 1) approximately 3 x 3 cm in size, drawn in black lines and displayed on a white background. One stimulus was displayed as the sample in the center of the monitor and three stimuli were displayed as comparisons in the corners leaving one corner blank. The stimuli were presented approximately 13 cm from the center of the screen. For convenience experimental stimuli will be designated in this article as three sets (A, B and C) with three stimuli in each (A1, A2, A3, B1, B2, B3, C1, C2, C3).

Procedure

Experimental sessions were conducted once a day (with no more than 2 days between each session). Participant #1 had two sessions in one day. The text accompanying the form was:

> Now that the experiment starts, there will appear a symbol in the middle of the screen. Touch it . Then three new symbols will appear in the corners. You are supposed to choose one of these. If you choose a correct one, “good”, “well done” etc. will appear and if you choose the wrong one, the word “wrong” will appear on the screen. In the right corner correct responses will be counted. This is the way you find out what is right and what is wrong. You are supposed to touch the screen to choose. After a while you will not get feedback about what is correct and incorrect and the count will also disappear. Do your best to get it all correct. Thanks for participating and good luck.

Sessions consisted of blocks of 0-second delayed MTS trials. Each trial started with a sample stimulus presentation. When the participant touched the sample it disappeared and the three comparison stimuli were immediately and simultaneously displayed. When the participants touched one of the comparison stimuli, all three immediately disappeared. Differential reinforcement for correct and incorrect choices was programmed in some of the experimental conditions (se details below).

Following correct responses the computer flashed “good”, “correct”, “well done” etc. in the middle of the screen and points were counted in the lower, right corner of the monitor in training conditions. One point was added for each correct response. If the participant responded incorrect, “wrong” was flashed in the middle of the screen. A 0.4 s inter-trial-interval followed all responding to comparison. During all inter trial intervals the monitor screen went white and any touches on the screen had no effect.

The participant worked for 30 minutes, had a 5 minute break and then worked another 30 minutes. Within these working periods a block of 36 trials was completed based on the software flashing “take a break” on the monitor screen. Then the participant timed a break of approximately one minute and started working again by touching the screen.

Comparison stimuli appeared equally often in each location. The general procedure was elaborated in three experimental phases. The accuracy level was evaluated by the software after each block. The accuracy criterion to advance from one phase to the next was 90% correct across 3 blocks of trials.

Phase 1

The same message that the participants read attached to the consent form was presented. The training followed an OTM training structure. Using the MTS procedure and differential reinforcement the participants were taught to perform AB matching, that is to say choosing B1, B2 and B3 comparison stimuli upon sample stimuli A1, A2 and A3 respectively. When the accuracy criterion was met for AB, AC training was conducted (selecting C1, C2 and C3 conditionally upon A1, A2 and A3). When the accuracy level criterion was met for AC, equal numbers of AB and AC trials were intermixed within the same block until accuracy criterion was met again. During Phase 1, there was no time restriction for responding, although response latencies always were recorded. Response to sample RT was measured from the moment the sample appeared to the moment the participant touched the sample. Response to comparison RT was measured form the moment the three comparisons appeared until the moment one of the comparisons was touched.

Phase 2

During Phase 2 titrating LH contingency was added to the MTS procedure. Comparison stimulus display was restricted, thus establishing a limited time available for the participant to respond. If the participant responded within the available time, the next trial continued as described above. If however the participant did not respond to the comparison within the time available, the trial ended with "timed out" flashed on the screen. The next trial then continued as described in Phase 1. No correction procedure was added.

The different relations were introduced concurrent, meaning responding to AB and AC were introduced interspersed in contrast to Phase 1, where AB was introduced first and then AC. The LH started in Phase 2 on a level calculated by the mean of the RT to comparison over the five last trials of Phase 1. Then the RT titrated, meaning was increased or decreased based on the performance of the participant. Given 80 % correct responding across six trials, the available RT was reduced by 100 ms. Given responding below 80% correct over six trials resulted in increase of available time by 100 ms. Titration continued until the participants RT was 1200 ms. When responding was too late, meaning outside the LH calculated by the software program, feedback was given to the participants by flashing the word "too late" on the screen. The software program counted these responses as incorrect.

Phase 3

In Phase 3 responding with the restriction of 1200 ms continued, but the differential consequences were fades from 75%, 50%, and 25% to extinction blocks. The criterion for ending this Phase was 80 % correct across the last three blocks of 36 responses each.

Phase 4

In Phase 4 the participants were given three blocks of 18 responses each of mixed testing, meaning 54 probes in total. Within a block baseline, symmetry and equivalence probes were randomly presented. The tests were administered under extinction conditions. There was a constant LH of 2500 ms.

Results

Number of trials and RT during training

There were big differences in the amount of trials for each participant before they reached mastery criteria. Participant #5 reached this within 126 trial and Participant #3 needed 2214 trials in Phase 1 (see Table 1). There were also individual differences in Phase 2 (titration) and 3 (fading of consequences) with Participant #4, who responded to 522 and 1386 trials respectively compared to Participant #5 who has 162 trials in Phase 2 and 378 trials in Phase 3.

The starting point in seconds for each participant based on mean RT to comparison in the last five trials in Phase 1, also varied across participants. The LH started at 2.80 s for Participant #4 and 1.40 s for Participant #3. Due to the mastery criteria in the software programming, this meant that the participants would differ in number of trials Participant #4 would have 84 trials minimum, while Participant #3 would have only 12 trials if no incorrect or timed out responses occurred during this training.

The Participant (#4) with highest starting point, who would need most trials to get to the set value of LH also used most trials (522). Without mistakes the titration from 2.80 sec to 1.20 sec would include 96 trials. The tendency that low amount of trials followed low starting points of LH held for the other participants as well. All participants (except #3), needed more trials in the phases with titration than in Phase 1 with no LH. Feedback fading required 1386 for Participant #4 and were considerably lower for the other participants.

There was also a decrease in RT to both comparison and sample during Phase 1, with no LH to comparison. RT to comparison decreased from a mean of 4290 ms in first training block to 2500 ms in the last training block for participant #5. Also RT to sample decreased from a mean of 3910 ms 790 ms. This pattern held for all the other participants as well.

All participants decreased their RT to both sample during training, even though there were not restrictions on time to respond. Also RT to comparison decreased in Phase 1 without any LH, except from participant #2, who had slightly higher RT in the last block compared to first block in Phase 1.

Test for baseline, symmetry and equivalent relations

One out of five responds in accordance with stimulus equivalence, given the definition of 90% correct responding on symmetry and equivalence trials. Participant #2 had 88,89% correct on symmetry and 94,44% on equivalence testing. Participant #3 and #4 responded according to symmetry but not equivalence. Participant #1 responded in accordance with equivalence but not symmetry. All baseline relations were intact during testing for all participants.

Reaction time

Figure 3 shows RT to baseline, symmetry and equivalence during mixed testing. The graph displays all three relations separately, even though the software program administered the trials randomly. RT did not seem to decline during testing for either of the relations.

Figure 4 shows mean RT during mixed testing for Participant #1- #5. Both correct and incorrect responding is included. RT to trained relations compared to emergent relations, does not seem to follow any clear pattern. Participant #1 and #3 had highest RT on symmetry. Participant #2 and #4 shows a pattern of increased RT from baseline to symmetry and to equivalence. The opposite pattern was seen in Participant #5. The differences in mean RT to the different relational types were small. The LH was 2500 ms and the participants mean RT had a range of 910 – 1300 ms. The largest difference was for Participant #2 with a difference of 120 ms between baseline and equivalence probes based on mean RT. For the other participants the differences were smaller. RT to comparison on incorrect trials, were slightly longer for Participant #1, #2 and # 3, but shorter for Participant #4 and #5.

Distribution of incorrect, timed out and the interspersed baseline trials are displayed in Figure 5. Horizontally the 54 trials are presented with one square equals one trial, the upper row shows symmetry and the lower row shows equivalence probes. The Figure shows that there are two participants whose incorrect responding occurs during the first half of the mixed testing (Participant #1 and #3). For the other participants there are very few mistakes, and mistakes occur in the second half of the probes. There are very few instances of timed out responses (Participant #4 and #2).

Figure 6 shows RT to comparison and sample during testing for all participants. RT to sample is shorter than RT to comparison for all participants except from Participant #5. Participant #5 had a mean RT of 1.69 s compared to 0.64 s for Participant #3. There was also more variability in RT to sample for this participant. There is no clear tendency that the participants have longer RT to sample after different relational types.

Figure 8 and 9 shows incorrect and timed out responses during testing for the emergent relations symmetry and equivalence. Mistakes and timed out responses does not seem to cluster around certain stimuli for any participant or if participants responding are compared.

Discussion

The purpose of the current study was to limit participants` time to respond to comparison and look at emergent responding and RT given this limitation. One of five participants responded in accordance to stimulus equivalence given restricted time to respond to comparison. The RT to comparison to different relational types, did not seem to follow patterns from previous studies (Arntzen, 2004; Eilifsen & Arntzen, in press; Arntzen et al., in press; Bentall et al., 1993; Holth & Arntzen, 2000; Spencer & Chase, 1996). Methodological changes compared to Tomanari et al. (2006) were included to reduce the training period. The changed elements were mainly shorter work sessions, a set lower value of LH, different protocol and reduced class sizes. These changes probably contributed to less trials in training than in the replicated study of Tomanari et al.

Some of the findings in the current study are similar to Tomanri et al. (2006). The participants vary regarding how many trials needed for reaching mastery criteria (see Table 1) The way the starting point for titration was calculated, could be an explanation of the difference in trials due to the mastery criteria accumulated trials, but the differences in trials far extended the amount of trials included in the mastery criteria. The differences in the overall amount of trials in this Phase could have been due to the different starting point of the participants, which is supported by the participants using most trials also had the longest RT and the fact that none of these responded in accordance to stimulus equivalence. It was therefore not the participants who had the longest experience with the material, who did respond in accordance with equivalence.

All participants (except Participant #3), used more trials in titration and fading added, compared to establishing the relations in Phase I. Participant #3 could have a learning style that made speed easier than the training of the relations. It must be commented that this participant did claim during breaks of Phase 1 that “I do not understand anything, so I just press some buttons, but I will not give up”.

Participant #4 had considerably more trials in Phase 2 and 3 with titration and was also the one who did achieve the lowest amount of correct scores during testing. One explanation is that the high amount of trials in training is caused by the same variables as responding in accordance to equivalence.

The high trials number in Tomanari et al. (2006) is probably due to methodological differences, with LH to sample and titration to a asymptotic level in contrast to no LH to sample and a set lower value of LH in the current study.

One of five participants responded in accordance with stimulus equivalence, compared to 3 out of 5 in Tomanari et al. The difference could be due to the far more extensive training phase and the methodological changes in the current replication study, like for instance the use of mixed testing instead of blocks of equivalence and trained, then symmetry and trained test trials.

Some other suggested causes for lack of positive test results should be considered. Some studies have reported breakdown of trained relations as an explanation for lack of responding (Devany, Hayes, & Nelson, 1986; McIlvane & Dube, 1996; Saunders & Green, 1999). All directly trained trials in this study were intact during testing, so this is not a likely explanation. Note that no reinforcers followed the trained test trials as well as symmetry and equivalence test trials.

Stimulus equivalence is based on interpretation of results on tests of untrained relations (Sidman, 1994). The possibility is that using other criteria, more participants are included in the positive outcome group. The study uses a quite strict stability criteria of 90%, which measures immediate transfer by only including a first block of test trials (Dymond & Rehfeldt, 2000). It is discussable whether Participant #2 also responded in accordance to stimulus equivalence as well with as much as 88,89% correct on symmetry testing. There is a possibility that the percentage correct used to define stimulus equivalence and stimulus equivalence is not the same (Arntzen et al., in press).

Another possibility found in some studies, which threatens the validity of interpretation, is learning of symmetry and equivalence relations during testing (Fields et al., 1990). There is no tendency that incorrect responding decreased during testing, indicating that the equivalent responding of Participant #5 and the high percentage of Participant #2 were immediate, which indicates that the responding was a product of the trained relation (Dymond & Rehfeldt, 2000).

In spite of the immediacy that validates the formation of equivalent classes occurring in the study, there is still a small proportion which did form classes. Since most studies on the topic of SE have included few participants and different experimental conditions, it is not possible to find a predicted success rate of participants similar to the five included in this study. The low number though is in line with other studies where time restrictions are set up for responding in accordance to equivalence (Holth & Arntzen, 2000; Tomanari et al., 2006). Holth and Arntzen (2000) found that none of 10 participants formed equivalent classes and Tomanari et al. (2006) found that 3 out of 5 participants responded in this way. The latter result is discussed by the authors themselves, since one of these three participants scored 85% on symmetry testing in the first test block, but was included as successful due to over 90 % on the first 35 test trials in this block. This is in line with the criteria of immediacy (Dymond & Rehfeldt, 2000) which rules out delayed emergence (Bush, Sidman, & de Rose, 1989; Sidman, 1994), but nonetheless indicate that there was a low number of participants responding in accordance to equivalence in this study as well.

One explanation for the outcome of these studies could be that the participants were given too short time to engage in mediating or precurrent behavior (Holth & Arntzen, 2000). Naming could be such an intervening and necessary variable. Naming as an explanations claims that a mediating response which takes time is required to form equivalent classes (Horne & Lowe, 1996). Restrictions on time to respond could then be a result of too little time to perform this precurrent behavior (Tomanari et al., 2006). Is there evidence that no naming occurred during the present experiment? This question has to be looked at from different angels. First, there is the question of how long naming would take, which needs to draw on findings from other disciplines within psychology. One example that Tomanari et al. (2006) refers to is (Schatzman & Schiller, 2004), who reported that latencies in labeling of pictures of high frequency reached an asymptote of 580 ms after repeated exposure. The problem is that this includes one stimulus. In the MTS format in this study, there were one sample stimulus, which the participants responded to, and next three stimuli which the participant needed to choose from and then respond to. Findings like those of (Schatzman & Schiller (2004) have restricted relevance, because some stimuli may be dealt with simultaneously in parallel processing (Posner, 2005) and this makes an additive approach to the chain of responses in MTS questionable. It is therefore hard to say what time restriction gives no opportunity to engage in naming. In the current study there is no evidence that time did not allow for naming. In testing the actual RT to comparison based on mean was the range 0.91 – 1.28 s and actual RT to sample based on mean was in the range 0.91 – 1.28 s for the participants. Even if we assume that this RT represented too little time to name the stimuli, the participants were free to engage in any behavior in Phase 1 with no LH, in the 5-minute breaks and between sessions. Tomanari et al. (2006) did have LH to sample as well, which made the actual RT shorter than in the present study, but here a training phase, breaks and between sessions periods gave opportunity to rehearse. These findings support the claim that naming as an explanation of stimulus equivalence is difficult to falsify (Pilgrim, 1996).

The tendency that RT is shorter for trained compared to emergent relations (Wulfert & Hayes, 1988) or that RT increases with relational type in the order trained, symmetry and equivalence test trials (Arntzen, 2004; Eilifsen & Arntzen, in press; Arntzen et al., in press; Arntzen & Lian, in press; Bentall et al., 1993; Holth & Arntzen, 2000; Spencer & Chase, 1996) does not hold for this study. Participant #1 and #3 had highest RT on symmetry. Participant #2 and #4 shows a pattern of increased RT from baseline to symmetry and to equivalence. The opposite pattern was seen for Participant #5. Tomanari et al. (2006) did also get mixed results, with the pattern in line with e.g. Spencer and Chase (1996) in 3 out of 5 participants.

The differences in all the studies are small (Arntzen, 2004; Eilifsen & Arntzen, in press; Arntzen et al., in press; Arntzen & Lian, in press; Bentall et al., 1993; Spencer & Chase, 1996) , in the sense that such short time segments are fragile to confounding variables like sneezing of the participant (Tomanari et al., 2006). The possibility that previous studies patterns were not replicated due to such variables can not be ruled out.

Is there a possibility that actual differences in RT, was blurred, because some of the stimuli were “easier” in the sense nameable than others. Studies have shown that RT to nameable stimuli is shorter than to abstract stimuli as node and are more readily learnt (Arntzen & Lian, in press). By analyzing the mistakes to the different relations, it does not seem that mistakes cluster on some stimuli during testing in this study. This is an indication that it is unlikely that nameability covered the pattern of RT to different relational types. The fragility of RT as a variable could also increase by the training structure OTM, which seems to result in the smallest differences in RT based on different relational test trials (Arntzen et al., in press).

One possibility is that patterns were washed out as a result of time restrictions. It might be that more predictable performance under conditions of retention and application is seen when time is limited as well as accuracy (Johnson & Layng, 1992). The training using titrating LH, could then equalize the learning histories (Spencer & Chase, 1996).

Interchangeability or substitutability is a defining characteristic of stimulus equivalence. (Fields, Adams, Verhave, & Newman, 1993) suggested that the more similar the speed of accurate responding on baseline, symmetry and equivalence, the more substitutable the stimuli. It could be that the training, with extensive training with limited hold, contributed to more relatedness between the stimuli compared to studies which not included LH in training.

If differences in RT are an indication of relatedness, the development of RT in this study indicates that stimuli were equally related. Figure 3 indicates that there was no tendency to reduction in RT during mixed testing. This is contrary to studies which have found decrease in RT during testing, (Arntzen, 2004; Arntzen et al., in press; Arntzen & Lian, in press; Bentall et al., 1993; Holth & Arntzen, 2000). These studies have used far more extended testing spaced in blocks. Fields et al. (1990) did not find a decrease in RT during testing. One debatable issue in the current study, is that LH was longer for testing than for training. This means that participants getting in contact with this new contingency (meaning reinforcement in spite of more than 1200 RT like in training), would get higher RT because this shaped their responding to get slower. The relatedness could also be contradicted by the fact that there was more variability in RT to symmetry and equivalence test trials compared to trained relations during testing.

The procedural variable of including LH to comparison and not to sample, gave an opportunity to look at RT from a more molecular analytic perspective (Arntzen & Lian, in press). One might expect that even though there was no reduction in RT to comparison there would be reduced time to respond to sample during testing, but this was not the case. Another suspicion was that if the feature of longer RT to trained versus untrained relations, reflect some kind of difference in how difficult the tasks were and the participants would spend longer time looking at sample after a more difficult trial. This speculation was not supported in the data in this study.

Based on some informal testing of the experimental setup, one person did discover the contingency of no LH to sample, and had a pattern of up to 10 seconds of rest during sample presentation and short RT to comparison. No participants followed this extreme pattern. It is interesting though, that Participant #5 had the longest RT to sample and the shortest mean RT comparison, meaning that the Participant were under control of the contingencies of different time demands to the stimuli.

The inclusion of RT as a crucial variable in equivalent classes (Spencer & Chase, 1996), is in opposition to the notion of full substitutability (Green & Saunders, 1998). An interesting result found in Eilifsen and Arntzen (in press) that the participants responding in accordance to equivalence, the previously mentioned pattern according to relational type, was not present in contrast to the RT of the participants who did not form equivalent classes. The lack of clear patterns in the current study and the pattern in previously mentioned studies might in this perspective not be different outcomes, but represent RT in responding when equivalent classes are under control of contingencies other than those who establish substitutable stimuli within a class.

Neither the current study or Tomanari et al. (2006) did arrange for contingencies that made equivalent responding impossible, all the time there were participants responding in accordance to equivalence. There is a need for further search for the differences in RT to different relational types among participants responding and not responding in accordance to equivalence. Variations in LH hold to comparison and sample is one direction to follow. Future studies should try inductively to find the threshold where class formation does not occur. One way to do this is to put further restrictions on time, by doing all the training phases under restricted time to respond conditions. This requires experimental setups that minimize the opportunity for participants to engage in rehearsal without the confounding variable of fatigue.

In summary the current study did show that time restrictions can be implemented with shorter endurance of training than previous studies (Tomanari et al., 2006). Results confirm that emergent relations do not necessarily occur given limited time to respond in training and testing (Holth & Arntzen, 2000). The well documented pattern of shortest RT in baseline test trials, higher RT on symmetry test trials and longest RT on equivalence test trials was not confirmed. This calls for continued manipulations of time variables in research on SE to see if the finding of the current study was due to titrating LH or other variables.

References

Arntzen, E. (2004). Probability of Equivalence Formation : Familiar Stimuli and Training Sequence. *The Psychological Record, 54*(2), 275-291. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html

Arntzen, E., Grondahl, T., & Eilifsen, C. (in press). The effect of different training structures in the establishment of conditional discriminations and the subsequent performance on the test for stimulus equivalence. *The Psychological Record*.

Arntzen, E., & Lian, T. (in press). Trained and derived relations with pictures as nodes. *The Psychological Record*.

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

Bentall, R. P., Dickins, D. W., & Fox, S. R. A. (1993). Naming and equivalence: Response latencies for emergent relations. *The Quarterly Journal of Experimental Psychology, 46*(2), 187-214. Retrieved from http://www.eps.ac.uk/index.php/journals

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

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:10.1901/jeab.1986.46-243

Dickins, D. W. (2005). On aims and methods in the neuroimaging of derived relations. *Journal of the Experimental Analysis of Behavior, 84*(3), 453-483. doi:10.1901/jeab.2005.92-04

Dymond, S., & Rehfeldt, R. (2000). Understanding complex behaviour: The transformation of stimulus functions. *The Behavior Analyst, 23*(2), 239-254. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/journals/557/

Eilifsen, C., & Arntzen, E. (in press). On the role of trial types in tests for stimulus equivalence. *European Journal of Behavior Analysis*.

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53*(3), 345-358. doi:10.1901/jeab.1990.53-345

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1993). Are stimuli in equivalence classes equally related to each other. *Psychological Record, 43*, 85-106. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html.

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

Fields, L., Landon-Jimenez, V., Buffington, D. M., & Adams, B. J. (1995). Maintained nodal-distance effects in equivalence classes. *Journal of the Experimental Analysis of Behavior, 64*(2), 129-145. doi:10.1901/jeab.1995.64-129

Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: a methodological analysis. *Journal of the Experimental Analysis of Behavior, 42*(1), 143-157. doi:10.1901/jeab.1984.42-143

Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229-262). New York, NY: Plenum Press.

Holth, P., & Arntzen, E. (2000). Reaction times and the emergence of class consistent responding: A case for precurrent responding? *The Psychological Record, 50*(3), 603-628. Retrieved from http://thepsychologicalrecord.siuc.edu/index.html.

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

Imam, A. A. (2001). Speed contingencies, number of stimulus presentations, and the nodality effect in equivalence class formation. *Journal of the Experimental Analysis of Behavior, 76*(3), 265-288. doi:10.1901/jeab.2001.76-265

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

Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist 47*(11), 1475-1490. doi:10.1037/0003-066X.47.11.1475

Lowenkron, B. (2004). Meaning: A verbal behavior account. *The Analysis of Verbal Behavior, 20*, 77-99. Retrieved from https://apps.abainternational.org/store/ViewJournals.asp?strJournalType=avb

McIlvane, W. J., & Dube, W. V. (1996). Naming as a facilitator of discrimination. *Journal of the Experimental Analysis of Behavior, 65*(1), 267-272. doi:10.1901/jeab.1996.65-267

Meyer, D.E, Osman, A. M., Irwin, D. E., & Yantis, S. (1988). Modern Mental Chronometry. *Biological Psychology,* 26, 3-67. doi:10.1016/0301-0511(88)90013-0

Pilgrim, C. (1996). Can the naming hypothesis be falsified? *Journal of the Experimental Analysis of Behavior, 65*(1), 284-286. doi:10.1901/jeab.1996.65-284

Posner, M. I. (2005). Timing the brain: Mental chronometry as a tool in neuroscience. *PLoS Biol, 3*(2), 204-206. doi:210.1371/journal.pbio.0030051

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

Schatzman, K. B., & Schiller, N. O. (2004). The word frequency effect in picture naming: Contrasting two hypotheses using homonym pictures. *Brain and Language, 90,* 160-169. doi:10.1016/S0093-934X(03)00429-2

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

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*(1), 127-146. doi: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:10.1901/jeab.1982.37-5

Spencer, T. J., & Chase, P. N. (1996). Speed analyses of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65*(3), 643–659. doi:10.1901/jeab.1996.65-643

Tomanari, G. Y., Sidman, M., Rubio, A. R., & Dube, W. V. (2006). Equivalence classes with requirements for short response latencies. *Journal of the Experimental Analysis of Behavior, 85*(3), 349-369. doi:10.1901/jeab.2006.107-04

Wulfert, E., & Hayes, S. C. (1988). The transfer of conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50*(2), 125-144. doi:10.1901/jeab.1988.50-125

Table 1

Amount of trials the participants needed during training with no LH (Phase 1), titrating LH (Phase 2) and fading of consequences (Phase 3). The LH before titration started is displayed in Phase 2. In Phase 3 LH had been titrated to the set value of 1200 ms for all participants.

	Participants									
Phase	#1		#2		#3		#4		#5	
	LH	Trials	LH	Trials	LH	Trials	LH	Trials	LH	Trials
1		333		198		2214		180		126
2	2.00	378	1.60	180	1.40	234	2.80	522	1.90	162
3	1.20	342	1.20	433	1.20	486	1.20	1386	1.20	378

Note: LH is displayed in seconds and the trial amount is including the mastery criteria.

Table 2

Amount of correct responses to 54 test trials in mixed testing, where baseline, symmetry and equivalence trials were interspersed.

	Participants				
	# 1	# 2	# 3	# 4	# 5
Baseline	18/18	18/18	17/18	18/18	18/18
Symmetry	15/18	16/18	17/18	16/17*	18/18
Equivalence	18/18	16/17*	15/18	12/17*	17/18

Note: The numbers to the left in the columns show correct responding and the numbers to the left show total probes.

*The probes timed out during testing are subtracted from the total number of probes.

Figure captions:

Figure 1: The figure shows the experimental stimuli. The experimental contingencies are in accordance with three classes of stimuli (1,2, and 3) with three members in each class (A, B and C).

Figure 2: The left panel show RT to sample and the right panel shows RT to comparison in the first and the last block of training for Participant #1 - #5.

Figure 3: Shows RT in seconds for mixed testing of baseline (filled circles), symmetry (filled square) and equivalence (filled diamond) relations for Participant #1-#5. These were randomly presented. Timed out responses are not presented.

Figure 4: Shows mean RT to baseline probes (black), symmetry (grey) and equivalence (white) for Participant #1-#5 during mixed testing.

Figure 5: Displays of distribution of incorrect responding (black), baseline relations (grey), timed out responses (dark grey) and baseline trials (light grey).

Figure 6 : The figure shows mean reaction time to sample (filled triangles) and comparison (filled squares) during mixed testing of baseline, symmetry and equivalence trials for Participant #1 - #5.

Figure 7: The figure shows mean RT to sample after a test trial of baseline, symmetry and equivalence.

Figure 8: Response matrices show the number of responses for each type of test trial symmetry during the first block of tests. Rows correspond to sample stimuli and columns to comparison stimuli. Highlighted diagonal cells show choices consistent with experimental equivalence classes. The column designated NR (no response) shows number of trials on which the participant failed to respond to the sample or comparisons within the LH duration.

Figure 9: Response matrices show the number of responses for each type of test trial equivalence during the first block of tests. Rows correspond to sample stimuli and columns to comparison stimuli. Highlighted diagonal cells show choices consistent with experimental equivalence classes. The column designated NR (no response) shows number of trials on which the participant failed to respond to the sample or comparisons within the LH duration.

Figure 1. Experimental stimuli.

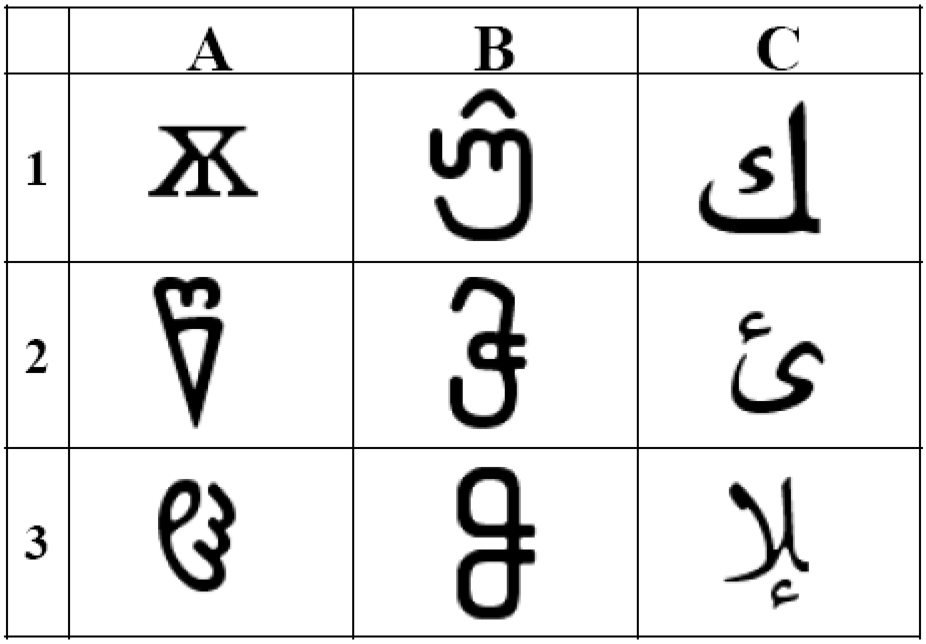
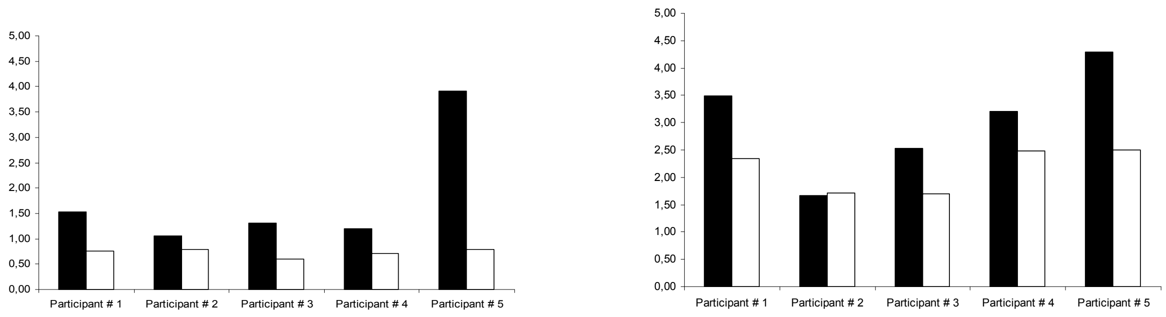


Figure 2. RT to sample and comparison bases on first and last block of training without LH.



Note: The number of blocks differs across participants.

Figure 3. RT to different relational types during mixed testing.

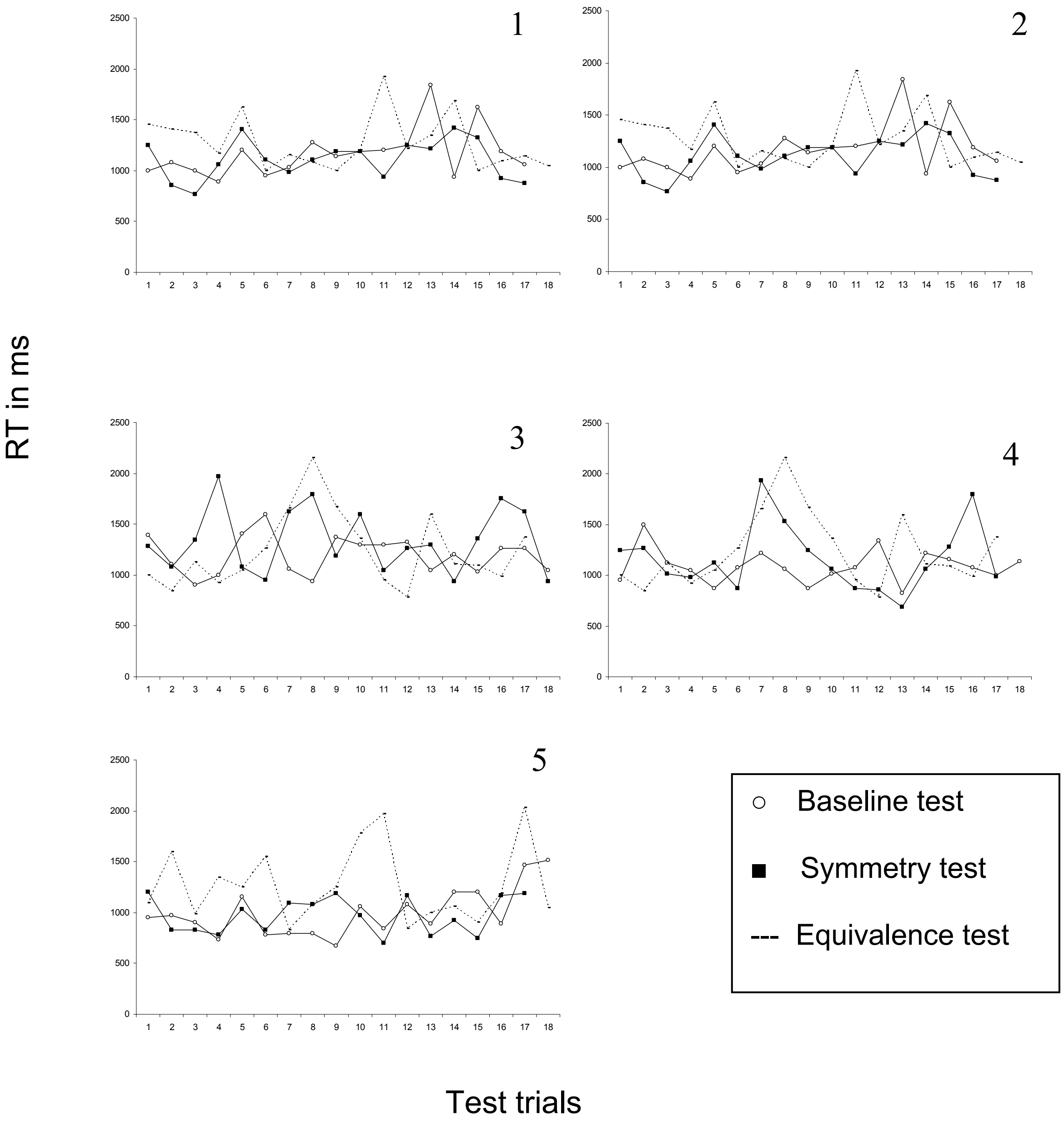


Figure 4. Mean reaction time to comparison during testing.

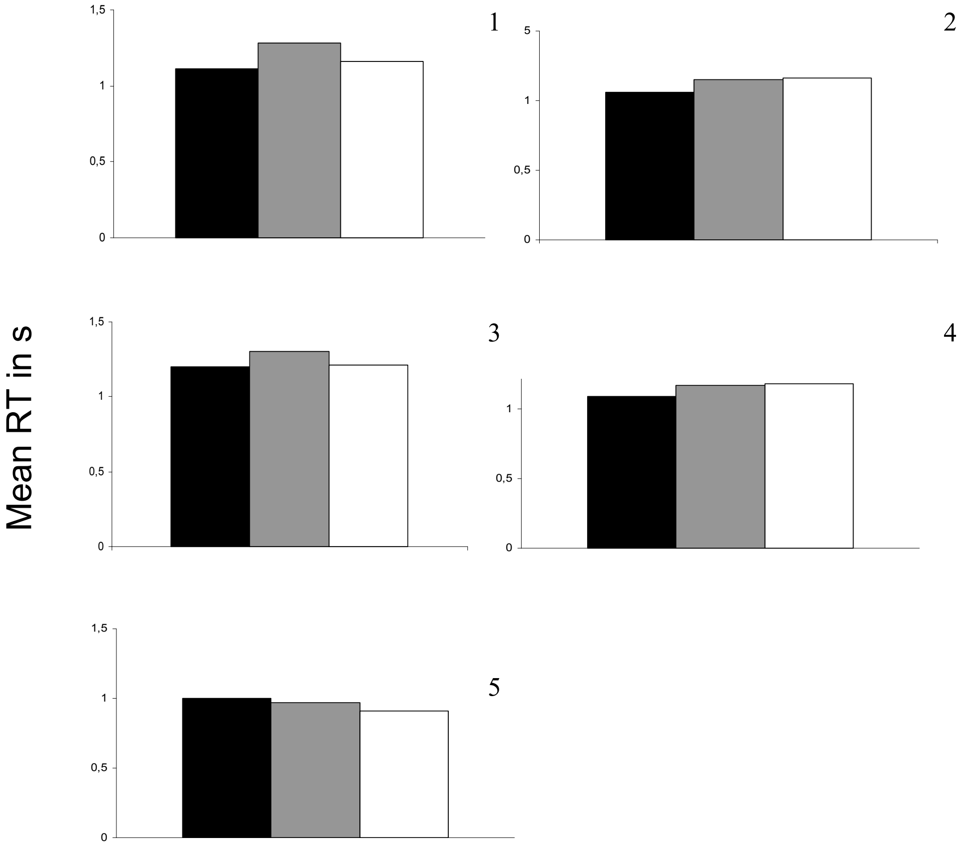


Figure 5. Distribution of timed out and incorrect responses during testing.

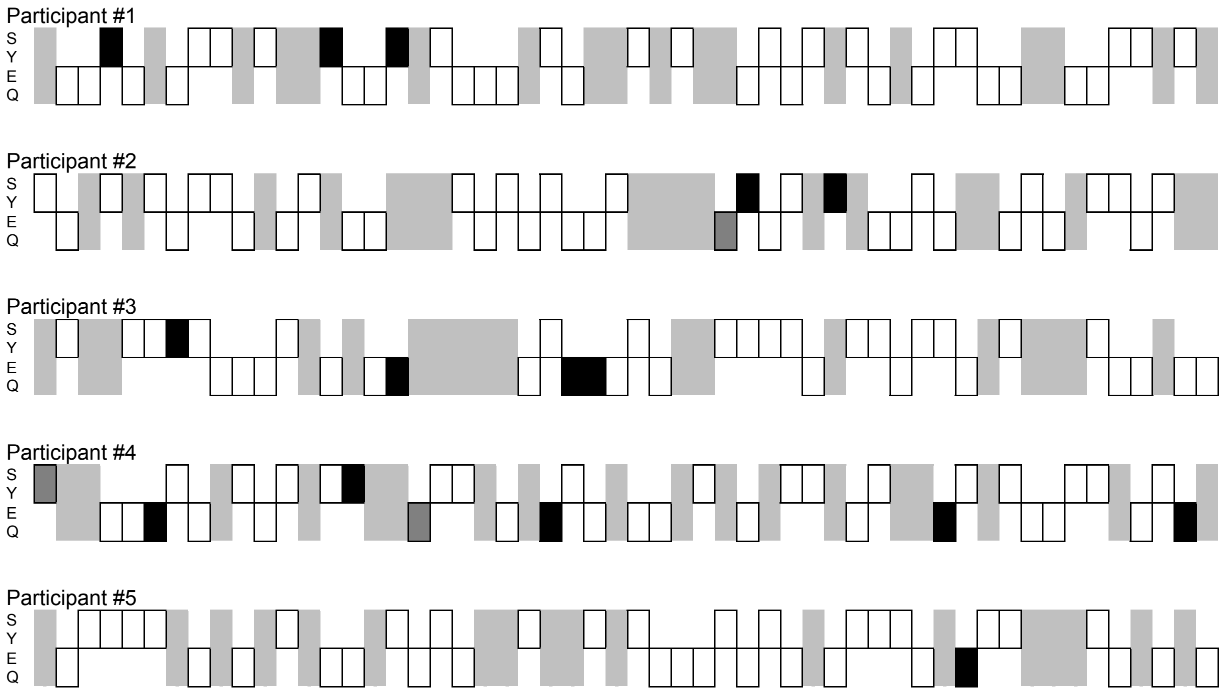


Figure 6. RT to comparison and sample during mixed testing.

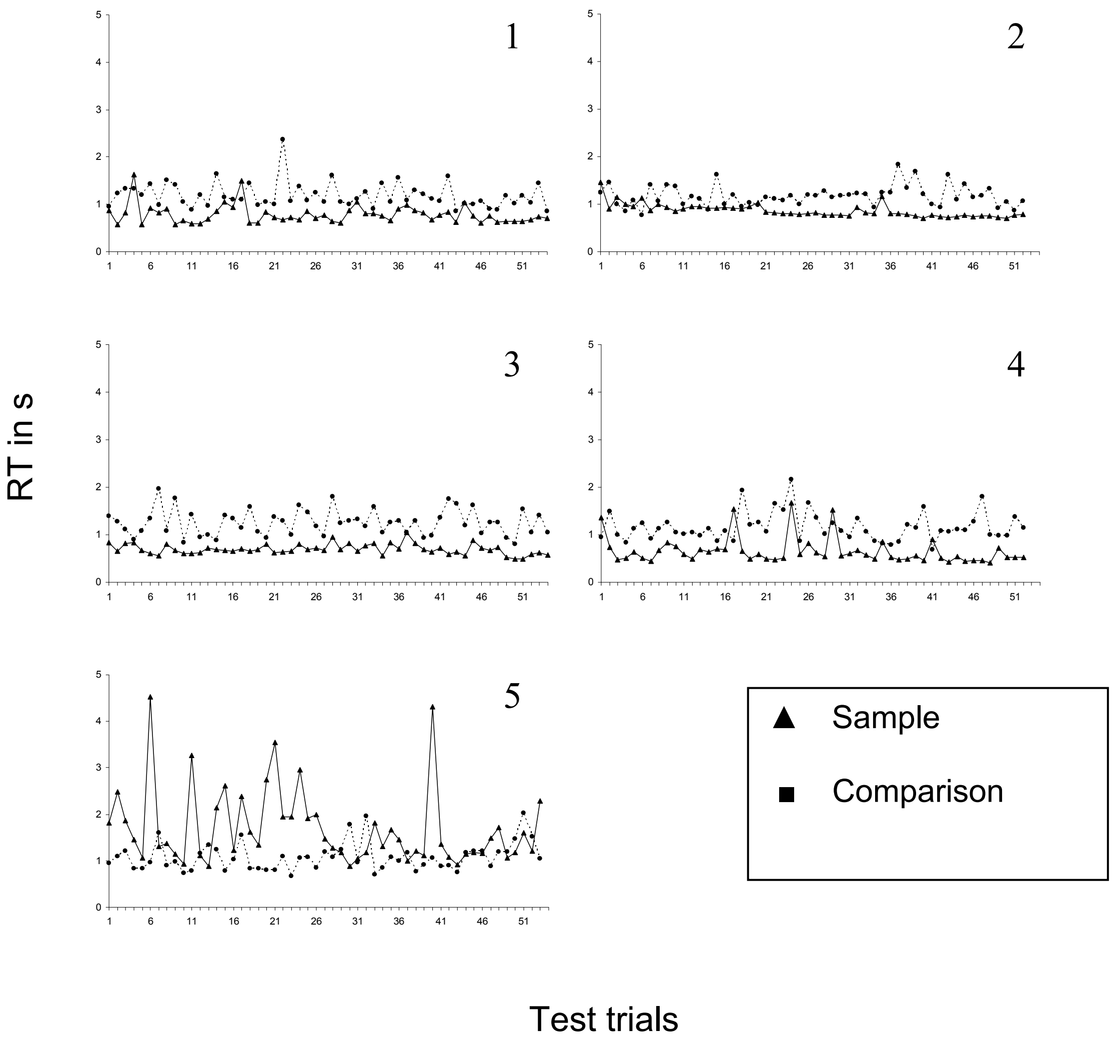


Figure 7. RT to sample the trial after different relational types during testing.

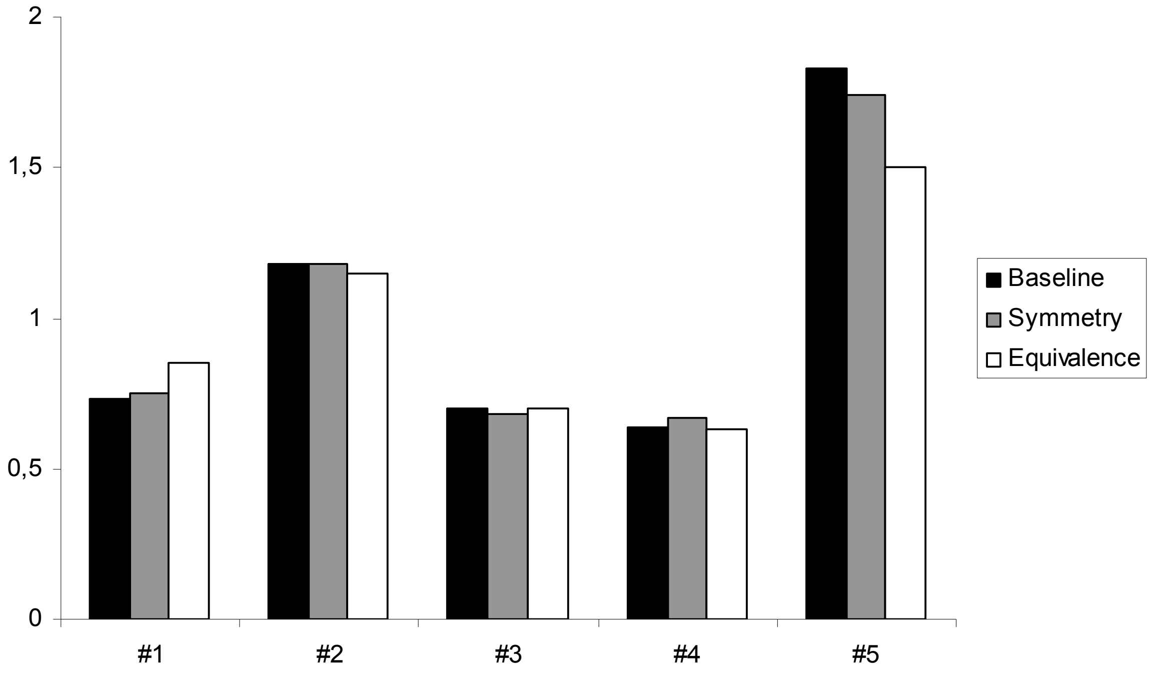


Figure 8. Incorrect, correct and timed out responses for each symmetry trial type during testing.

Participant #1

Symmetry	A1	A2	A3	NR
C1	3	0	0	0
C2	0	3	0	0
C3	1	0	2	0

	A1	A2	A3	NR
B1	2	0	1	0
B2	0	3	0	0
B3	1	0	2	0

Participant #2

Symmetry	A1	A2	A3	NR
C1	3	0	0	0
C2	0	3	0	0
C3	0	1	2	0

	A1	A2	A3	NR
B1	3	0	0	0
B2	1	2	0	0
B3	0	0	3	0

Participant #3

Symmetry	A1	A2	A3	NR
C1	3	0	0	0
C2	0	3	0	0
C3	0	0	3	0

	A1	A2	A3	NR
B1	3	0	0	0
B2	1	2	0	0
B3	0	0	3	0

Participant #4

Symmetry	A1	A2	A3	NR
C1	3	0	0	0
C2	0	3	0	0
C3	0	0	2	1

Symmetry

	A1	A2	A3	NR
B1	3	0	0	0
B2	0	3	0	0
B3	1	0	2	0

Participant #5

Symmetry	A1	A2	A3	NR
C1	3	0	0	0
C2	0	3	0	0
C3	0	0	3	0

	A1	A2	A3	NR
B1	3	0	0	0
B2	0	3	0	0
B3	0	0	3	0

Figure 9. Incorrect, correct and timed out responses for each equivalence trial type during testing.

Participant #1

Equivalence	C1	C2	C3	NR
B1	3	0	0	0
B2	0	3	0	0
B3	0	0	3	0

	B1	B2	B3	NR
C1	3	0	0	0
C2	0	3	0	0
C3	0	0	3	0

Participant #2

Equivalence	C1	C2	C3	NR
B1	3	0	0	0
B2	0	3	0	0
B3	0	0	3	0

	B1	B2	B3	NR
C1	3	0	0	0
C2	0	3	0	0
C3	0	0	2	1

Participant #3

Equivalence	C1	C2	C3	NR
B1	3	0	0	0
B2	0	3	0	0
B3	0	0	3	0

	B1	B2	B3	NR
C1	3	0	0	0
C2	0	2	1	0
C3	0	0	3	0

Participant #4

Equivalence	C1	C2	C3	NR
B1	1	0	1	1
B2	0	3	0	0
B3	2	0	1	0

	B1	B2	B3	NR
C1	3	0	0	0
C2	0	3	0	0
C3	0	2	1	0

Participant #5

Equivalence	C1	C2	C3	NR
B1	3	0	0	0
B2	0	3	0	0
B3	0	0	3	0

	B1	B2	B3	NR
C1	3	0	0	0
C2	0	3	0	0
C3	0	1	2	0