Volume 21, number 2, 2021 https://www.ijpsy.com Volumen 21, número 2, 2021

Research Articles // Artículos de investigación

Theoretical and Review Articles // Artículos teóricos y de revisión

Notes and Editorial Information // Avisos e información editorial

ISSN 1577-7057

2021, 21, 2
International Journal of Psychology & Psychological Therapy

Volume 21, number 2 June 1, 2021
Volumen 21, número 2 1 Junio, 2021
ISSN: 1577-7057

IJP&PT

International Journal of Psychology & Psychological Therapy

https://www.ijpsy.com

IJP&PT

INTERNATIONAL JOURNAL OF PSYCHOLOGY & PSYHOLOGICAL THERAPY

International Journal of Psychology & Psychological Therapy is a four-monthly interdisciplinary publication open to publish original articles, reviews of one or more area(s), theoretical reviews, or methodological issues, and series of interest to some of the Psychology areas. The journal is published for the *Asociación de Análisis del Comportamiento* (AAC) and MICPSY, and indexed and/or abstracted in:

- **Academic Search Complete** (*EBSCO Publishing Inc.*)
- **Cabell's Directory** (*Cabell Scholarly Analytics*)
- **CLARIVATE-WEB of SCIENCE** (*Emerging Sources Citation Index*)
- **ClinPSYC** (*American Psychological Association*)
- **DIALNET** (*Fundación Dialnet, Universidad de La Rioja*)
- **DICE-CSIC** (*Difusión y Calidad de las Revistas Españolas*)
- **Directory of Open Accest Journals** (DOAJ)
- **EBSCO Information Service**
- **GOOGLE Scholar Metrics**
- **IBECS** (*Índice Bibliográfico Español en Ciencias de la Salud*)
- **IN-RECS** (*Index of Impact of the Social Sciences Spanish Journals*)
- **ISOC** (CINDOC, CSIC)

International Journal of Psychology & Psychological Therapy es una publicación interdisciplinar cuatrimestral, publicada por la Asociación de Análisis del Comportamiento (AAC), abierta a colaboraciones de carácter empírico y teórico, revisiones, artículos metodológicos y series temáticas de interés en cualquiera de los campos de la Psicología. Es publicada por la *Asociación de Análisis del Comportamiento* (AAC) y MICPSY y está incluida en las bases y plataformas bibliográficas:

- **Journal Scholar Metrics**
- **LATINDEX** (*Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal*)
- **MIAR** (*Matriz de Información para el Análisis de Revistas*)
- **ProQuest Prisma Database**
- **Psychological Abstracts**(*American Psychological Association*)
- **PsycINFO** (*American Psychological Association*)
- **REBIUN** (*Red de Bibliotecas Universitarias Españolas*)
- **RESH** (*Revistas Españolas deCiencias Sociales y Humanidades*)
- **SCIMAGO** (*SCImago Journal & Country Rank -SCOPUS*)
- **SCOPUS** (*Scopus Database Elsevier*)

International Journal of Psychology & Psychological Therapy, 2021, 21, 2, 221-237

Formation of Equivalence Classes Including Emotional Functions

Jon Magnus Eilertsen*, Erik Arntzen
Oslo Metropolitan University, Norway

Abstract

Eleven participants in two experimental groups, DMTS-3s and DMTS-6s, trained conditional discriminations (AB, AC, CD, DE, and EF) with the potential emergence of three 6-member equivalence classes. The A stimuli (A1, A2, and A3) consisted of faces showing angry, neutral, and happy facial expressions, respectively. All participants responded in accordance with the experimentally defined criterion of 95% correct on two consecutive transitivity BF and equivalence FB test blocks. Next, participants rated the abstract D1 and D3 stimuli equivalent to the angry and happy faces (A1 and A3) on a Semantic Differential Rating Scale. A control group rated the facial stimuli and the abstract D stimuli on a similar rating scale. Results show that stimuli are more related when trained with DMTS-3s than DMTS-6s. Abstract D3 stimuli rated by the DMTS-3s group deviated less from control group ratings of respective faces than ratings of all other abstract D stimuli by the two experimental groups.
Key words: delayed matching-to-sample, stimulus equivalence, transfer of function, conditional discriminations.

How to cite this paper: Eilertsen JM & Arntzen E (2021). Formation of Equivalence Classes Including Emotional Functions. *International Journal of Psychology & Psychological Therapy, 21, 2,* 221-237.

Novelty and Significance

What is already known about the topic?

- Abstract stimuli in equivalence classes are more related to facial stimuli showing emotional functions when trained with DMTS-2s than with SMTS.
- Happy faces seem to cause over-rating of abstract stimuli compared to ratings of the faces by a control group.

What this paper adds?

- DMTS-3s causes abstract stimuli to be rated more similar to abstract faces equivalent to them than DMTS-6s.
- Happy faces are over-rated in the DMTS-3s group but not in the DMTS-6s group.
- Delays higher than DMTS-2s might have implications on the transfer of meaning or emotional functions in stimulus equivalence classes.

Training conditional discriminations in a matching-to-sample (MTS) format might result in the emergence of new, untaught stimulus-stimulus relations. For example, if stimulus A1 is presented, the selection of stimulus B1 and not B2 or B3 is reinforced. In the next trial, B1 could be presented as a sample stimulus, and the selection of C1 and not C2 or C3 is reinforced. When meeting the mastery criterion after n number of training trials, tests for emergent conditional discriminations are presented. If the emergent conditional discriminations share the properties of reflexivity (AA), symmetry (BA), and transitivity (AC), the relations amongst the stimuli can be defined as a stimulus equivalence class (Sidman & Tailby, 1982). There are a variety of training and testing parameters that can be manipulated to influence the formation of equivalence classes; see Arntzen (2012) for an overview. One of these parameters is arrangements of the delay between the offset of the sample and onset of the comparison (Delayed Matching-to-Sample, DMTS). The most common way to present conditional discriminations is with

* Correspondence: Jon Magnus Eilertsen, Oslo Metropolitan University, Department of Behavioral Science, PO Box 4 St. Olavs Plass, 0130 Oslo, Norway. E-mail: Jonmagnuseilertsen@gmail.com or erik.arntzen@equivalence.net. *Availability of Data and Materials*: Data can be obtained upon request.

the use of Simultaneous Matching-to-Sample (SMTS). In SMTS, the sample stimulus remains present during the presentation of the comparison stimuli. In DMTS, the sample stimulus is removed n-seconds before the presentation of the comparison stimuli. For example, an observing response (e.g., mouse click) removes the sample stimulus, and a 3s-delay elapses before the onset of the comparison stimuli (DMTS-3s).

Studies have shown how DMTS procedures influence several aspects of how conditional discriminations are trained and subsequently the formation of equivalence classes. For example, in Arntzen (2006) experiments 1, 2, and 3, increasing delays successively increased the probability of responding in accordance with stimulus equivalence. The design employed within-participant manipulations and found that participants who started with SMTS and increased delays with 0s, 2s, and 4s used a mean of 110 training-trials more than the minimum criterion in the SMTS condition. In addition, responding in accordance with stimulus equivalence increased with increasing delays, the number of trials increased to 180 in the DMTS-0s condition, then dropped to 40 trials in the DMTS-2s condition, and finally increasing again to 80 in the DMTS-4s condition. When participants started with DMTS-4s, the number of training trials above minimum criterion systematically dropped for each successive condition, including the SIM condition. Vaidya and Smith (2006) replicated the results of Arntzen (2006) with a group design where participants were randomly assigned to one out of three groups (DMTS-0, DMTS-2s, and DMTS-8s). Participants trained conditional discriminations and were tested for emergent symmetrical relations. Participants in the DMTS-8s group showed more symmetry consistent responses than participants in the DMTS-2s and DMTS-0s group.

DMTS procedures have also been used to investigate to what degree properties of a meaningful stimulus can become equivalent to the remaining members of the equivalence class. When new contingencies are applied to one member of an equivalence class or a new stimulus member with some specific properties is added to an existing class, the contingencies might subsequently transfer to the other members of the class. Class union by the inclusion of the response as one of the event pairs is by definition a defining feature of equivalence classes (cf. Sidman, 2000), but the observed process has also been termed Transfer of Function (ToF) in stimulus equivalence research (e.g., Dougher & Markham, 1994). For example, Bortoloti and De Rose (2009) found a greater degree of relatedness between properties of meaningful stimuli and other abstract class members in equivalence classes when conditional discriminations were trained with DMTS-2s than with SMTS. In Experiment 1, participants in two groups (DMTS and SMTS) trained conditional discriminations where the A stimuli consisted of faces showing different emotional expressions. A1 showed an angry face, A2 neutral and A3 showed a happy facial expression. Participants trained conditional discriminations in a mix of One-To-Many (OTM) and Linear Series (LS) training structure, training AB/AC, CD. Next, the participants were tested for emergent symmetrical BD and equivalence DB relations. Participants in the DMTS-2s group and the SMTS group were then asked to rate the D stimuli on a Semantic Differential rating scale (Osgood *et alia*, 1957) to compare the ratings with the ratings done by the control group. The Semantic Differential Scale had the D stimuli at the top with thirteen 7-point bipolar Likert Scales below. On each side of the scale, an adjective was placed (good/bad, happy/sad, heavy/light, and so on). The scales ranged from -3 to 3, with 0 as a middle point or neutral. Ratings by the experimental groups were compared to ratings done by a control group. The control group rated both the faces and the abstract stimuli. Results showed that ratings by the DMTS-2s group were more similar to the ratings of the corresponding faces by the control group than for the SMTS group. In Experiment 2, participants formed three 7-member classes. Two groups (SMTS and DMTS-2s) rated the abstract D stimuli, and two groups rated the abstract F stimuli. Findings form Experiment 1 were replicated for the D stimuli, but ratings of the F stimuli significantly deviated from control group ratings.

Findings were replicated a similar setup (two groups, SMTS and DMTS-2s) but measured with the Implicit Relational Assessment Procedure (IRAP) (Bortoloti & de Rose, 2012). Both these studies used either SMTS or DMTS-2s delay.

To summarize, DMTS procedures have been observed not only to influence the formation of equivalence classes but also how stimuli are related compared to training with SMTS. In the present experiment, we would like to investigate if different delays in DMTS will differentially influence how stimuli with emotional functions, such as faces showing angry or happy expressions (A1 and A3), affect the rating of abstract D stimuli (D1 and D3). Except from the study by Bortoloti and De Rose (2009) later replications have been done with three 4-member or three 5-member classes (Bortoloti *et alia*, 2019; Bortoloti & de Rose, 2012; Bortoloti *et alia*, 2014; Bortoloti *et alia*, 2013; Silveira *et alia*, 2016). Both Bortoloti and De Rose (2009, 2012) compared DMTS 2s with SMTS. It would be interesting to test if training with delays longer than DMTS 2s will differentially influence ratings of abstract stimuli equivalent to happy or angry faces in three 6-member classes.

Method

Participants

Sixty-one university students volunteered to participate in the study. Participants were assigned to one control group (*n*= 25) and two experimental groups (*n*= 36). Ten participants either choose to not complete the conditional discrimination training, or experienced parameter errors. Their data are not included in the data analysis. Fifteen participants, 11 in the DMTS-3s Group and four in the DMTS-6s Group, did not respond in accordance with the criterion for symmetry and equivalence, their data are included when calculation percentage of participants who responded in accordance with the experimenter defined criterion in each group. Of the remaining 11 participants, four females and one male with a mean age of 25 years (*SD*= 4) responded in accordance with the experimenter defined criterion in the DMTS-3s Group. Finally, six participants, four females and two males with the mean age of 21 years (*SD*= 1), responded in accordance with the experimenter defined criterion in the DMTS-6s Group. The participants assigned to the experimental conditions were paid 100 Norwegian kroner for their participation (approximately US $11) regardless of their test results or if they finished the conditional discrimination training or testing. The Participants in the control group were not compensated, none of the participants had any prior knowledge about stimulus equivalence. Participants in the experimental conditions were handed consent forms, which they read and signed before the experiment started. The consent forms contained general information about the experimental setting and who was conducting the experiment. In addition, the consent forms contained information about participant anonymity and their right to withdraw from the experiment at any given time. At the end of the experiment, all participants were fully debriefed. All procedures performed were in accordance with the ethical standards and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Apparatus and Setting

Experimental sessions lasted from 48 minutes to 138 minutes. The experiment was conducted in quiet lab locations and cubicles, and participants were seated in front of a blank wall. The cubicles measured approximately 200 cm x 135 cm and were furnished with a table and a chair. The MTS software was run on an HP ProBook 470 GP laptop computer with a Windows 10 64-bit operative system. The computer had a 17.4 inch screen. The program administered the conditional-discrimination training and testing.

Stimuli

Figure 1 shows the stimuli used for the conditional-discrimination training and testing. The A1, A2, and A3 stimuli consisted of faces showing angry, neutral, and happy facial expressions, respectively. The remaining stimuli consisted of abstract symbols and shapes. The size of the face stimuli was 4x4 cm, and the abstract shapes varied from 0.5 cm to 2.5 cm in height and from 0.7 cm to 3.1 cm in width. The faces expressing emotions were retrieved from The Karolinska Directed Emotional Faces (KDEF) with approval for use in non-commercial research (Lundqvist *et alia*, 1998). The Semantic Differential Scale had instructions on the first page (see Figure 2) and the two D stimuli (D1 and D3) on pages two and three. The order of the two sheets with the D stimuli was randomized for each participant. Figure 3 shows an example of the Semantic Differential Scale used to rate the D stimuli for the experimental groups and the faces

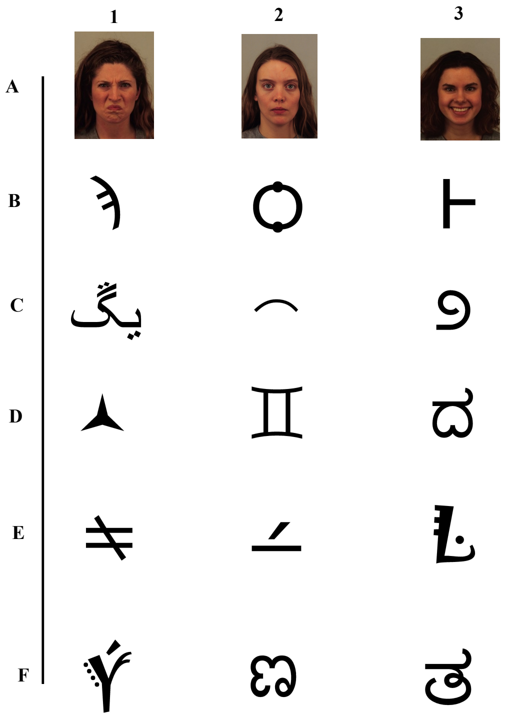
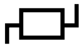


Figure 1. Stimuli used for the conditional discrimination training and testing for the DMTS-3s and DMTS-6s groups. The faces shown as A stimuli.

INSTRUCTIONS

You will find a picture on the top of each of the following sheets. Your task is to mark with an X the location of the picture in scales limited to opposite adjectives. Each scale represents a continuum from one adjective to its opposite. Thus, you will find, for ex-ample, the pair beautiful/ugly and will have to judge, based on this pair of adjectives, a figure like:



If you consider the figure above *extremely beautiful*, you should mark the space closest to beautiful, as follows:

BEAUTIFUL **UGLY**

If you consider the figure *extremely ugly*, you should mark the space closest to ugly, as follows:

BEAUTIFUL **UGLY**

If you consider the figure *quite beautiful*, you should mark the second space close to beautiful, as follows:

BEAUTIFUL **UGLY**

If you consider the figure *quite ugly*, you should mark the second space close to ugly, as follows:

BEAUTIFUL **UGLY**

If you consider the figure *slightly beautiful*, you should mark the third space close to beautiful, as follows:

BEAUTIFUL **UGLY**

If you consider the figure *slightly ugly*, you should mark the third space close to ugly, as follows:

BEAUTIFUL **UGLY**

If you consider the figure *not related* to any adjective of the pair, you should mark the central space, as follows:

BEAUTIFUL **UGLY**

If you have any doubt about these instructions, call the experimenter.

Thank you for your collaboration!

Figure 2. The front page of the Semantic Differential Rating scale with instructions handed to both experimental groups and the control group.

and the D stimuli for the control group. The scale consists of a picture of one of the three D stimuli for the experimental group. Below each stimulus, there were thirteen 7-point Likert Scales with adjectives on each side of each scale. The adjectives used in the Semantic Differential Scale have been evaluated with a factor analysis showing the adjectives into two factors (Factor 1 and Factor 2). Factor 1 adjectives are more related to the rating task and were labeled as "Evaluation," whereas Factor 2 words are more unrelated to the rating task and were labeled as "Potency" (Almeida *et alia*, 2014). The adjectives are presented in the Semantic Differential as opposites, and the Factor 1 words are "happy/sad, tense/relaxed, rough/smooth, ugly/beautiful, heavy/light, negative/ positive, hard/soft, bad/good, pleasant/unpleasant." The Factor 2 words are "fast/slow, active/passive, rich/poor, dominant/submissive." For illustrative purposes, the adjectives in Figure 3 have all positives on one side and all negative on the other side. In the forms presented to the participants, the locations of the negative and positive adjectives were randomized. The middle box always gave 0 points, whereas each box from left to right increased or decreased. Ticks towards the negative adjectives decreased with one point (-1, -2, -3), and ticks towards the negative adjectives increased with one point (1, 2, 3).

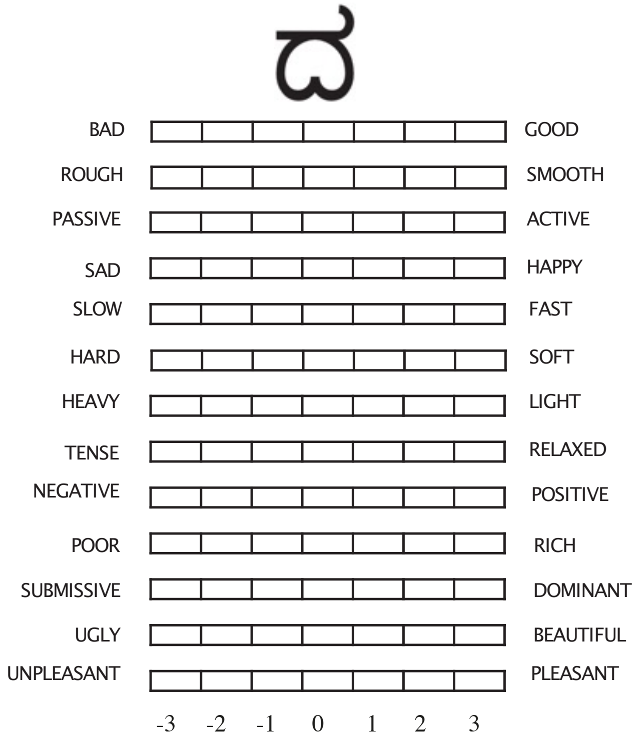


Figure 3. Example of the Semantic Differential Scale (showing the D3 stimulus) as handed to both experimental groups and the control group. Note that in the figure, the positive adjectives are placed on the right hand side and negative adjectives on the right hand side. For participants, the locations of the adjectives were randomized, and the Likert point scale at the bottom was not visible.

Procedure

The experimental phases can be divided into three phases for the experimental groups (1) conditional-discrimination training, (2) two test blocks for emergent transitive and equivalence relations (BFx2, FBx2), (3) rating the D stimuli on a Semantic Differential Scale. The control groups were not exposed to conditional discrimination training and testing. They were just given the semantic differential forms and asked to rate the stimuli according to the instructions. They rated both the abstract shapes and the faces expressing different emotions. When the participants in the experimental groups had signed the consent forms, they were seated in front of the laptop computer. The participants were asked not to click on anything before they had read the instructions. The computer screen displayed the following instruction:

> Thank you for participating in this experiment. This is an experiment within learning psychology and requires no prior computer-knowledge. In short, you should click some stimuli that appear on the screen. The goal is to get as many correct as possible. When you move the mouse cursor on the stimulus in the middle and click it, more stimuli will appear on the screen. Mouse clicks on the correct ones in the corners will be followed by the text "Correct" or similar on the screen. Clicking on one of the wrong ones will be followed by the text "Wrong." That is how you find out what is right and wrong. After a while, you will not be notified if it is correct or wrong what you click, no text on the screen. However, it will always be necessary to click on the middle one before clicking the ones in the corners.

The experimenter remained with the participants when they read the instructions. If they had any questions, the relevant part of the instruction was repeated to them. No additional information was provided. The participants were also informed that the experiment was done when the text "Congratulations, you have now completed the experiment" was displayed on the screen, and they could then get the experimenter. Participants clicked a grey button located at the bottom of the screen saying "start," and the program initiated the conditional discriminations.

A sample stimulus appeared at the center of the screen. When the participants clicked the sample stimulus, it was removed, and three comparison stimuli were presented in three of the four corners. The location of the blank corner was randomized through the experiment. Depending on the condition, there was a 3s or 6s delay between the removal of the sample stimulus and the presentation of the comparison stimuli. Clicking on one of the three comparison stimuli immediately removed all the three stimuli, and the programmed consequences were presented at the center of the screen. Clicking on the comparison stimulus defined as correct resulted in the presentation of one of the written words "Awesome," "Very Good," Excellent," and "Well Done." Clicking on one of the two stimuli defined as wrong produced the written word "Wrong" at the center of the screen. The presentation of the programmed consequences lasted for 500 ms, followed by an intertrial interval (ITI) of 500 ms before the presentation of a new sample stimulus in the middle of the screen. The conditional discriminations were presented serialized in a mix of an OTM and LS training structure. The trained relations presented in an OTM training structure were **A1**/B1-B2-B3, **A2**/B1-B2-B3, **A3**/B1-B2-B3, **A1**/C1-C2-C3, **A2**/C1-C2-C3, **A3**/C1-C2-C3. The remaining conditional discriminations were presented in an LS training structure (C→D→E→F). See Table 1 for an overview of the trained and tested relations. The trained relations were presented serialized in 15 trial blocks. In the serialized presentation of training trials, the AB trials were trained until criterion before the AC, CD, DE, and EF training trials were presented. Each training block was repeated until the mastery criterion of 100% correct was attained (15/15). Next, when the mastery criterion was attained for the last E→F block, all the trained conditional discriminations were mixed and presented in blocks with 75 trials per block, each relation type was presented five times per block in a randomized order. In the mixed block, programmed consequences were gradually reduced based on performance. If the mastery criterion of 96% (72/75) was attained, the programmed consequences were reduced to 75%, then 25%, and finally 0%. After the mastery criterion was attained in the last block of 0% programmed consequences, the test for emergent relations was

Table 1. Sequence of training and testing.

Phases	Trial types	% Program Consequences	Number of trials
Acquisition of baseline relations (All trial types presented randomly)			
1. Serialized trials	A1B1, A2B2, A3B3	100	15
2. Serialized trials	A1C1, A2C2, A3C3	100	15
3. Serialized trials	C1D1, C2D2, C3D3	100	15
4. Serialized trials	D1E1, D2E2, D3E3	100	15
5. Serialized trials	E1F1, E2F2, E3F3	100	15
6. Mixed trials	A1C1, A2C2, A3C3, A1C1, A2C2, A3C3, C1D1, C2D2, C3D3, D1E1, D2E2, D3E3 E1F1, E2F2, E3F3	100	75
7. Mixed trials	A1C1, A2C2, A3C3, A1C1, A2C2, A3C3, C1D1, C2D2, C3D3, D1E1, D2E2, D3E3 E1F1, E2F2, E3F3	75	75
8. Mixed trials	A1C1, A2C2, A3C3, A1C1, A2C2, A3C3, C1D1, C2D2, C3D3, D1E1, D2E2, D3E3 E1F1, E2F2, E3F3	50	75
9. Mixed trials	A1C1, A2C2, A3C3, A1C1, A2C2, A3C3, C1D1, C2D2, C3D3, D1E1, D2E2, D3E3 E1F1, E2F2, E3F3	0	75
Test blocl 1 for emergent 3-node transitivity relations	B1F1, B2F2, B3F3	0	15
Test block 2 for emergent 3-node transitivity relations	B1F1, B2F2, B3F3	0	15
Test block 1 for emergent 3-node equivalence trials	F1B1, F2B2, F3B3	0	15
Test block 2 for emergent 3-node equivalence trials	F1B1, F2B2, F3B3	0	15

Note: In both training and testing phases, each relation is presented five times in a randomized order per block.

initiated. The first test block tested for the emergence of symmetry relations B→F. Tests for both symmetry and equivalence were presented in two consecutive blocks, and each block consisted of 15 trials where the three relation types were presented five times each in a randomized order. The tested relations were **B1**/<u>F1</u>-F2-F3, **B2**/F1-<u>F2</u>-F3, **B3**/F1-F2-<u>F3</u>. Next, a test for emergent equivalence relations F→B) was initiated in two consecutive blocks with 15 trials per block. The tested relations were **F1**/<u>B1</u>-B2-F3, **F2**/B1-<u>B2</u>-B3, **F3**/B1-B2-<u>F3</u>. In contrast to the training-phases (except for the last phase with 0% programmed consequences), there were no programmed consequences in the testing phase, and there were no break in-between testing blocks. Thus, the participants did not notice the transition from test block one to test block two or from transitivity to equivalence testing.

When the last test block was finished, the text "Congratulations, you can now get the experimenter" appeared in the middle of the screen. Participants were handed the Semantic Differential Scale (see Figure 3) with the D1, D2, and D3 stimuli printed on each one out of three sheets. The written instructions on the first page were read aloud to the participants. The participants were told that they could ask questions if they did not understand the task. None of the participants indicated that they did not understand it. The experimenter left the participants when they filled out the forms, as not to bias the ratings in any way. When they had filled out the forms, they were thanked for their participation, debriefed, and paid for their participation.

Data Analysis

Semantic Differential scales are Likert-type ordinal data. An ANOVA (one-way) Kruskal Wallis test with a Dunnett's multiple comparison's follow up was run to compare the control group ratings of the angry and happy faces to the ratings of the D1 and D3 stimuli by the DMTS-3s and DMTS-6s groups. Significant differences in ratings would indicate that the stimulus valence of the facial expressions did not transfer to the abstract D stimuli equivalent to them. Likewise, non-significant differences would indicate that stimulus valence from facial expressions had transferred to the abstract D stimuli equivalent to them. The Kruskal Wallis test was also employed to evaluate deviations in ratings of the facial expressions by the control group and the ratings of the D stimuli by the experimental groups. We wanted to evaluate if the experimental groups had either over or under-rated the valence of the D stimuli compared to the control group's ratings of the faces. The calculations were done by comparing the median ratings of adjectives by the control group and the experimental groups. If the control group had rated the happy face with a median of 2, and the experimental group rated the D3 stimulus with a median of 1, the deviation would be calculated as -1. Likewise, if the control group had rated the happy face with a median of 2, and the experimental group rated the D3 stimulus with a median of 3, the deviation would be counted as 1. Finally, a Man Whitney test was employed to pairwise evaluate the differences in ratings of the D stimuli (D1 and D3) between the two delay groups (DMTS-3s and DMTS-6s). All statistical tests were run with Factor 1 words.

RESULTS

In total, 11 participants (five in the DMTS-3s group and 6 in the DMTS-6s group) responded in accordance with the experimentally defined criterion of 95% correct for symmetrical and equivalence relations (BF, FB). See Table 2 for an overview of the results, including trials to criterion and responses in the two consecutive tests. The average number of trials to the criterion for the DMTS-3s group was 495 (*SD*= 118.6) and for the DMTS-6s group, the average number of trials to criterion was 627.5 (*SD*=

Table 2. Results for Training and Testing.

DMTS-3sec					
		Test 1 BF	Test 2 BF	Test 1 FB	Test 2 FB
P#	Trials	Transitivity	Transitivity	Equivalence	Equivalence
13415	450	15	15	15	15
13418	705	15	15	15	15
13423	465	15	15	15	15
13424	435	15	15	15	15
132431	420	15	14	15	15
DMTS 6 sec					
		Test 1 BF	Test 2 BF	Test 1 FB	Test 2 FB
P#	Trials	Transitivity	Transitivity	Equivalence	Equivalence
13435	420	15	15	15	15
13450	660	15	15	15	15
13451	750	15	15	15	14
13482	810	15	15	15	15
13384	615	14	15	15	15
13385	510	15	15	15	15

Notes: P#= Participant number; Test1, and Test 2 columns shows the number of correct test trials out of the total number of possible correct for each test; BF, and FB columns indicates the tested B1F1, B2F2, B3F3, and F1B1, F2B2, and F3B3 transitivity and equivalence relations; Each of the test blocks consisted of 15 trials with each trial type presented five times in a randomized order.

145.9). For the participants who did not form classes, the average number of trials to criterion was 752.3 (*SD*= 215.3) for the DMTS-3s group and 525 (*SD*= 99.4) for the DMTS-6s group. A two-tailed independent *t*-test showed a significant difference in trials to criterion between the participants who formed classes (*M*= 495, *SD*= 118.6) and the participants who did not form classes (*M*= 752.3, *SD*= 215.3) $t(14)= 2.47$, $p= .02$ for the DMTS-3s group.

Figures 4 and 5 shows the Semantic Differential Ratings of the D1 and D3 stimuli for the DMTS-3s and DMTS-6s groups respectively. The ratings of the D stimuli by the experimental groups are indicated with a solid black line. The control group rated the angry and happy faces as well as the abstract D stimuli. The control-group ratings of the angry and happy faces are indicated with a stippled line marked with X's, and the ratings of the D stimuli are marked with a dotted line with boxes. The Kruskal-Wallis test showed no significant difference in ratings of the happy faces (*Mdn*= 2) by the control group compared to the D3 stimuli by the DMTS-3s group ($p >.05$). Likewise, there was no significant difference between the ratings of the angry faces by the control group (*Mdn*= -2) and the D1 stimuli by the DMTS-3s group ($p >.05$). However, for the DMTS-6s group, there were significant differences in ratings between the D3 and D1 stimuli compared to the ratings of the happy ($p <.05$) and angry faces ($p <.01$) by the control group. A Mann-Whitney test was used to compare the ratings of the D1 stimuli for the DMTS-3s group to the ratings of the D1 stimuli for the DMTS-6s group. The differences were not significant $U= 1202$, $p= .9284$. However, the ratings of the D3 stimuli (happy class) for the DMTS-3s group (*Mdn*= 3) and the ratings of the D3 stimuli (happy class) for the DMTS-6s group (*Mdn*= 1) shows a significant difference, $U= 717$, $p= .0002$. The difference between the ratings of the D3 stimuli for the DMTS-3s and DMTS-6s groups shows a decrease in positive ratings from three to six seconds delay.

The DMTS 3 group rated some of the abstract D3 stimuli more positive than how the control-group rated the happy faces. By visually inspecting the median ratings of the happy faces by the control group, the median ratings of the abstract D3 stimuli for participants in the DMTS-3s group, two of the Factor 1 words are given 1 point more compared to the control group (Smooth, and Soft). In contrast, the median ratings of the D3 stimuli by the DMTS-6s group are generally lower than the ratings of the happy faces by the control group. Four of the adjectives are rated with a median of one point lower than for the control group (Sad/Happy, Rough/Smooth, Negative/positive, and Bad/Good). Two adjectives are rated with a median of two points lower (Ugly/Beautiful, and Unpleasant/pleasant), while one adjective is rated with a median of 1.5 points lower compared to the control group (Tense/Relaxed). Only Heavy/Light was over-rated by the experimental group by a median of 0.5 points, while Hard/Soft was rated similarly. For the ratings of the D1 stimuli compared to the ratings of the angry faces by the control group, the median ratings are generally lower by both DMTS groups compared to the ratings by the control group. Figure 6 represents the results of a Kruskal-Wallis test with a Dunnett's multiple comparisons follow up. The figure shows the mean of median deviations from the ratings of the D stimuli by the experimental groups compared to the respective happy and angry faces rated by the control group. The test found a significant difference in the over/under-rating of the D3 stimuli by the DMTS-3s and DMTS-6s groups compared to the ratings of the happy faces of the control group ($p <.05$). No significant differences were found for the other deviation ratings between groups.

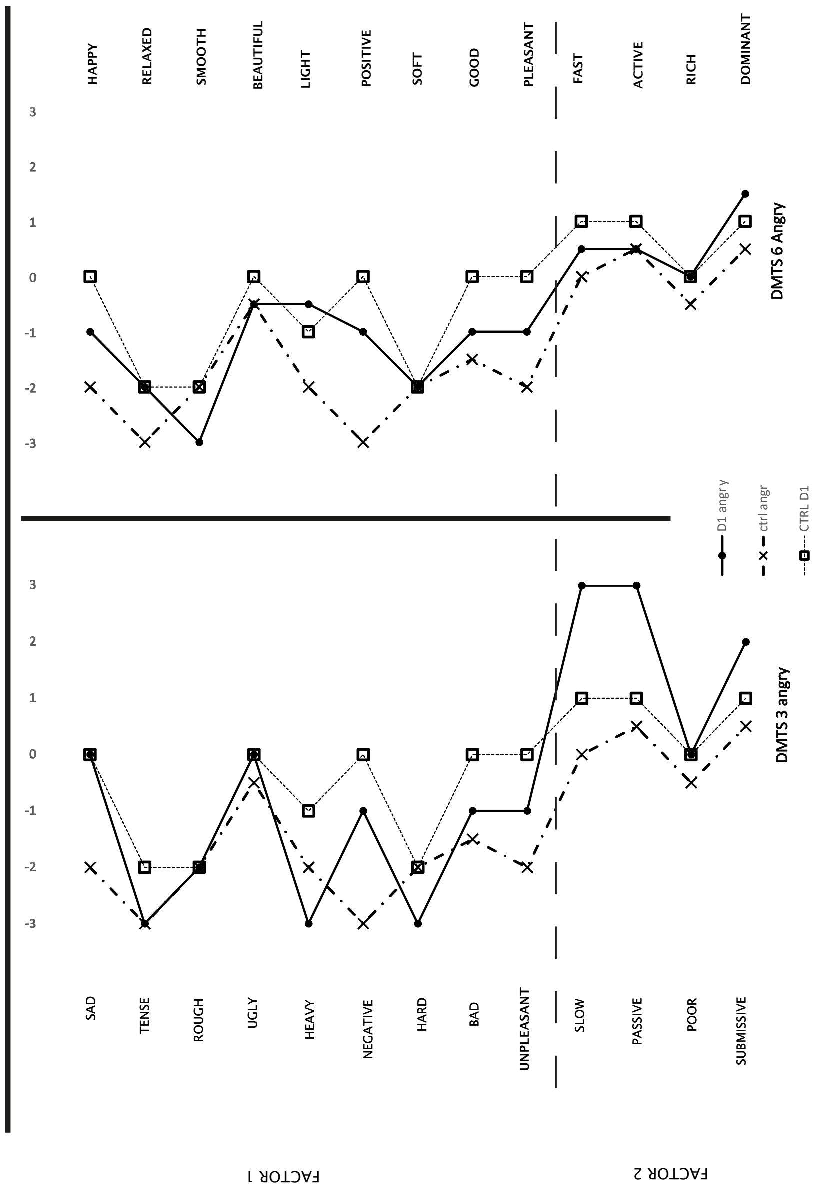


Figure 4. The solid black lines shows the Median scores on the Semantic Differential Scales of the D1 stimuli by the DMTS-3s, and DMTS-6s groups. The dotted line with squares and the stippled lines with marked with X represents the median scores of the D1 and angry face stimuli by the control group.

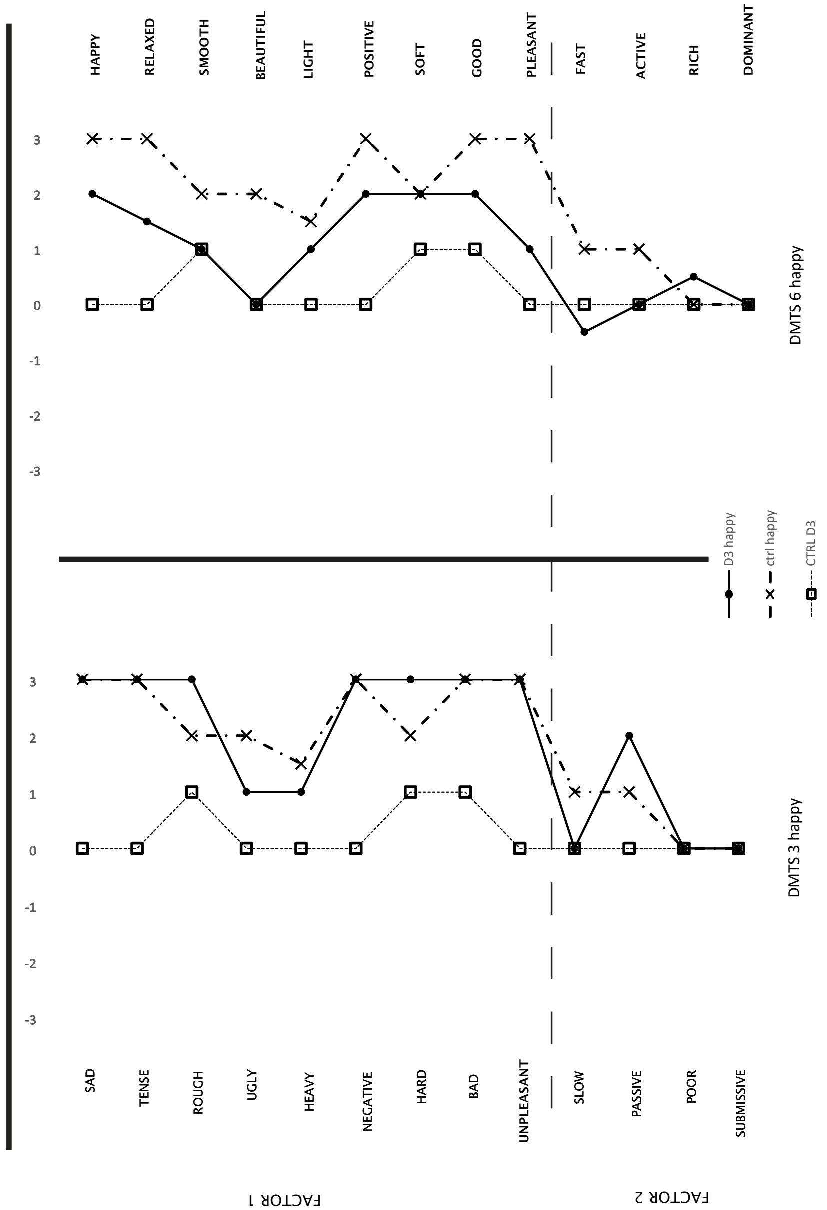


Figure 5. The solid black lines shows the Median scores on the Semantic Differential Scales of the D3 stimuli by the DMTS-3s, and DMTS-6s groups. The dotted line with squares and the stippled lines with marked with X represents the median scores of the D3 and happy face stimuli by the control group.

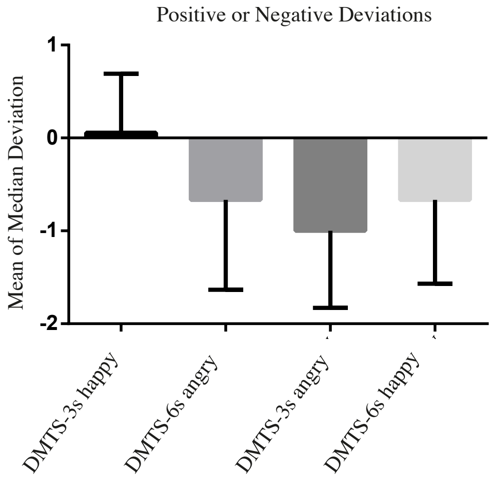


Figure 6. Mean deviations D-stimuli ratings by the experimental groups and the ratings of the respective faces by the control group. The scale also shows if the D-stimuli were under or over rated.

Discussion

Two experimental groups formed three 6-member equivalence classes where conditional discriminations were trained with either DMTS-3s or DMTS-6s. The A1, A2, and A3 stimuli consisted of angry, neutral, and happy facial expressions, respectively. Both experimental groups rated the D1 and D3 stimuli equivalent to the angry and happy faces on a Semantical Differential Rating Scale. The facial stimuli and the abstract D stimuli (D1 and D3 stimuli) were rated by a control group on a similar Semantic Differential scale. Control group ratings of the facial stimuli were compared to the ratings of the D1 and D3 stimuli by the experimental groups.

The main results showed that the DMTS-3s group formed classes were angry (A1) and happy (A3) facial expressions were more related to abstract stimuli (D1 and D3) than the DMTS-6s group. The DMTS-3s group over-rated the abstract D3 stimuli equivalent to the happy faces compared to how the control group rated the happy faces. For the other happy and angry classes in both experimental groups, the stimuli were under-rated.

In total, 11 out of 26 participants responded in accordance with the experimentally defined criterion of 95 % correct in two consecutive test blocks on transitivity (BF) and equivalence (FB) relations. For the DMTS-3s group, 31,5 % (five out of 16) participants formed classes, while 60 % of the participants (six out of 10) formed classes in the DMTS-6s group. Others have found that increasing delays facilitates the formation of equivalence classes. For example, Arntzen (2006) found that DMTS-4s resulted in a higher probability of equivalence class formation compared to DMTS 2s, DMTS 0 s, and SMTS (see also Vaidya & Smith, 2006). In comparison, Bortoloti et alia (2013) recruited 34 participants where 13 of them (38 %) did not show the emergence of the BE and EB relations. A possible reason for the high number of participants failing to respond in accordance with the experimenter defined BF and FB transitivity and equivalence relations could be that they were not tested for any of the other conditional discriminations at lower stages. As observed and pointed out by Sidman *et alia* (1985) in Experiment 3, the emergence of lower stage conditional discriminations facilitated the emergence of n-stage conditional discriminations.

In the DMTS-3s Group, both the abstract D3 stimuli (happy class) and the abstract D1 stimuli (angry class) were scored similarly to how the control group scored the corresponding faces. Thus, this was not the case for the DMTS-6s group, where neither the abstract D1 stimuli (happy class) or the abstract D3 stimuli (angry class) were scored similarly to how the control group scored the corresponding faces. The ratings of the D1 stimuli (angry class) did not change significantly from DMTS-3s to DMTS-6s, while the positive ratings for the D3 stimuli decreased from DMTS-3s to DMTS-6s. This is also in contrast to what others have observed as a happy superiority effect, where stimuli showing happy faces in equivalence classes formed using DMTS are shown to cause a higher degree of relatedness in contrast to classes containing angry faces (Bortoloti & De Rose, 2009, 2011) In addition, Silveira *et alia* (2016) found that training with DMTS-3s equivalence classes containing happy faces was more stable over time than classes with angry faces. We found no difference in how the D1 stimuli (angry class) were rated between the DMTS-3s group and the DMTS-6s group, but the ratings of the D3 stimuli decreased as a function of increasing the delay. On the other hand, we found that the deviations between the ratings of the D3 stimuli equivalent to the happy faces and the ratings of the happy faces by the control group were lower for the DMTS-3s group than the ratings of the D1 stimuli compared to the angry faces for the control group. However, the deviations increased more for the happy classes than for the angry classes for the DMTS-6s group.

We did not test if the number of nodes from the A stimuli (facial expressions) affected the ratings of the abstract stimuli. For example, Bortoloti and De Rose (2009) found an inversed effect of numbers of nodes on the similarity in ratings of the abstract stimuli for the experimental groups and the ratings of the facial stimuli for the control group.

The training and testing arrangement in the present paper differs slightly from the study by Bortoloti and De Rose (2009). Participants in the current experiment were not given any form of pre-training, and transition from baseline acquisition to the testing of emergent relations was done with a stepwise thinning of programmed consequences with blocks including all the serially trained conditional discriminations. During testing, participants in the present experiment were provided with two consecutive test blocks for emergent symmetry and equivalence relations. In contrast, participants in the Bortoloti and De Rose (2009) study were subjected to a pre-training that prompted the selection of the correct comparison stimulus for the first 12 trials. Also, the removal of programmed consequences during baseline acquisition was signaled with a written text instead of blocks with a stepwise decrease of programmed consequences. The test blocks were presented once for symmetry and equivalence testing interspersed with a mixed training block without programmed consequences. With the data at hand it is not possible to know how any of these differences might have affected the difference in number of participants who formed equivalence classes, or the ratings of the D stimuli, but these are parameters that could be tested in further research. For example, the effect of preliminary training on the acquisition of baseline relations and subsequent formation of equivalence remains to be tested.

One implication of the current study could be that there are differences in how delays above DMTS-2s influences how emotional functions are transferred to abstract stimuli in equivalence classes. In addition, happy faces were found to be over-rated compared to the control group for the DMTS-3s group, which replicates findings by others and termed as a happiness superiority effect (Bortoloti *et alia*, 2019). It has also been observed that ratings of abstract stimuli equivalent to happy faces deviate less than angry faces rated by control groups (e.g., Bortoloti & De Rose, 2009; Bortoloti & De Rose, 2011). This was also the case for the DMTS-3s group in the present study, but this effect seemed to diminish for the DMTS-6s group.

One reason for the high number of participants who did not complete the conditional discriminations might be that the relations are not mixed and trained again after each introduction of new baseline relations. Usually, the simultaneous training and testing protocol present all relations concurrently in mixed training and test blocks (e.g., Imam, 2006). The trials could also be presented on a sequential (or serialized) basis in which each relation is trained separately before they are mixed. For example, if training A→B→C→D→E, the AB relations are trained first, followed by training of the BC relations and a mixing AB and BC relations. Next, training of CD relations until mastery criterion, followed by AB, BC, and CD relations, in a mixed block. Finally, DE relations are trained to mastery and followed by a mix of AB, BC, CD, and DE relations before testing of emergent relations.(e.g., Fields & Paone, 2020; Mensah & Arntzen, 2017). Arntzen *et alia* (2014) investigated the differences between a sequential (also referred to as serialized) and concurrent presentation of conditional discriminations on the formation and maintenance of equivalence classes over time. They found the sequential arrangements of baseline trials led to both fewer trials to the acquisition and that equivalence classes were more stable over time. In the present experiment, training of the separate relations to criterion was followed by the full mix block of all relations with a stepwise thinning of programmed consequences (75 %, 50 %, and 0 %). Thus, not employing mixing of new relations with previously trained baseline relations and not testing for all of the derived relations could have caused some of the participants not to meet the criterion for the transitivity and equivalence test trials. Future experiments should investigate the implications of including a sequential introduction of baseline relations.

The few numbers of participants in each of the experimental groups could be a limitation to the present study. Thus, future studies should try to include more participants in each of the conditions. In addition, future studies should investigate the effects of other delays. For example, Lian and Arntzen (2013) investigated the effects of DMTS-0.1 s, DMTS-3s, and DMTS-12 s on the formation of equivalence classes and found that DMTS-12 s increased the likelihood of equivalence class formation over DMTS-0.1 s and DMTS-3s.

The present study found that DMTS-6s increased the likelihood of equivalence class formation over DMTS-3s, but it would be interesting to see how increasing delays over 6 s influences the transfer of emotional functions. Finally, all of the participants were trained and tested with the use of the same facial stimuli, and ratings of abstract stimuli equivalent to the faces were compared to ratings of the facial stimuli by a control group. In studies where facial stimuli showing different emotions are employed, it might be a stimulus control topography discrepancy between what is intended by experimenters and what is shown in the responding of individual participants. Tailoring the stimuli for each individual might be one way to reduce the discrepancy. For example, each participant could choose or rate their own emotional stimuli to be used in the experimental setup. Future studies should investigate the implications of tailoring individual stimuli. One method that has been shown to facilitate conditional discrimination training is errorless teaching (see Green, 2001 for an overview of errorless teaching and training conditional discriminations). In errorless teaching, opportunities to respond in accordance with experimenter defined relations is maximized while at the same time, opportunities to respond in accordance with competing stimulus control topographies (STC) is minimized. Recently, Fields and Paone (2020) found that acquisition of baseline relations was lowest with an errorless learning method compared to concurrent and serially, but the training modalities did not differ in the percentage of participants who responded in accordance with stimulus equivalence.

In the present experiment the transfer of function was tested for the D stimuli. However, transfer of function could also have been tested for the E and F stimuli, and should be included in future research.

The purpose of the present study was to investigate if facial stimuli with different emotions would be differentially related to abstract stimuli equivalent to them, as a function of different delays in DMTS (DMTS-3s and DMTS-6s). Participants in two experimental groups, five in a DMTS-3s group, and six in a DMTS-6s group, trained conditional discriminations AB, AC, CD, DE, and EF where A1 were angry faces, and A3 were happy faces. All participants responded in accordance with a mastery criterion of 95% correct on two consecutive test blocks for symmetry (BF) and equivalence (FB) relations. In the DMTS-3s group, five out of 16 participants formed the classes, whereas in the DMTS-6s group, six out of 10 participants formed the classes. Participants in the DMTS-3s group showed a greater stimulus relatedness between angry and happy faces and abstract stimuli equivalent to them than participants in the DMTS-6s group. The abstract D3 stimuli (equivalent to the happy faces) were rated more positive for the DMTS-3s group.

REFERENCES

Almeida JHd, Bortoloti R, Ferreira PRdS, Schelini PW, & Rose JCCd (2014). Análise da Validade e Precisão de Instrumento de Diferencial Semântico. *Psicologia: Reflexao e Critica, 27*, 272-281. Doi: 10.1590/1678-7153.201427207

Arntzen E (2006). Delayed matching to sample and stimulus equivalence: Probability of responding in accord with equivalence as a function of different delays. *The Psychological Record, 56*, 135-167. Doi: 10.1007/BF03395541

Arntzen E (2012). Training and testing parameters in formation of stimulus equivalence: Methodological issues. *European Journal of Behavior Analysis, 13*, 123-136. Doi: /10.1080/15021149.2012.11434412

Arntzen E, Halstadtro LB, Bjerke E, Wittner KJ, & Kristiansen A (2014). On the sequential and concurrent presentation of trials establishing prerequisites for emergent relations. *The Behavior Analyst Today, 14*, 1-8. Doi: 10.1037/h0101280

Bortoloti R, de Almeida RV, de Almeida JH, & de Rose J C (2019). Emotional Faces in Symbolic Relations: A Happiness Superiority Effect Involving the Equivalence Paradigm. *Frontiers in Psychology, 10*, 1-12. Doi: 10.3389/fpsyg.2019.00954

Bortoloti R & De Rose JC (2009). Assessment of the relatedness of equivalent stimuli through a semantic differential. *The Psychological Record, 59*, 563-590. https://doi.org/10.1007/bf03395682

Bortoloti R & De Rose JC (2011). An "Orwellian" account of stimulus equivalence. Are some stimuli "more equivalent" than others? *European Journal of Behavior Analysis, 12*, 121-134. Doi: 10.1080/15021149.2011.11434359

Bortoloti R & de Rose JC (2012). Equivalent stimuli are more strongly related after training with delayed matching than after simultaneous matching: A study using the Implicit relational assessment procedure (IRAP). The Psychological Record, 62, 41-54. Doi: 10.1007/BF03395785

Bortoloti, R., Pimentel, N., & De Rose, J., C. (2014). Electrophysiological investigation of the functional overlap between semantic and equivalence relations. Psychology and Neuroscience, 7(2), 183–191. Doi: 10.3922/j.psns.2014.015

Bortoloti, R., Rodrigues, N. C., Cortez, M. D., Pimentel, N., & De Rose, J., C. (2013). Overtraining increases the strength of equivalence relations. Psychology and Neuroscience, 6(3), 357–364. Doi: 10.3922/j.psns.2013.3.13

Dougher M & Markham MR (1994). Stimulus equivalence, functional equivalence and the transfer of function. In SC Hayes, LJ Hayes, M Sato, & K Ono (Eds*.), Behavior Analysis of Language and Cognition* (pp. 71-90). Reno, NV: Context Press.

Fields L & Paone D (2020). Training Modality and Equivalence Class Formation under the Simultaneous Protocol: A Test of Stimulus Control Topography Coherence Theory. *The Psychological Record, 70*, 293-305. Doi: 10.1007/s40732-020-00384-4

Lian T & Arntzen E (2013). Delayed Matching-to-Sample and Linear Series Training Structures. *The Psychological Record, 63*, 545-561. Doi: 10.11133/j.tpr.2013.63.3.010

Lundqvis E, Flykt A, & Öhman A (1998). *The Karolinska Directed Emotional Faces* -KDEF, CD ROM, Department of Clinical Neuroscience, Psychology section, Karolinska Institutet.

Mensah J & Arntzen E (2017). Effects of Meaningful Stimuli Contained in Different Numbers of Classes on Equivalence Class Formation. *The Psychological Record, 67*, 325-336. Doi: 10.1007/s40732-016-0215-y

Osgood CE, Suci GJ, & Tannenbaum PH (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

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

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. Journal of the Experimental Analysis of Behavior, 43(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 Behavior, 37*, 5-22. Doi: 10.1901/jeab.1982.37-5

Silveira MV, Aggio NM, Cortez MD, Bortoloti R, Rico VV, & de Rose JC (2016). Maintenance of equivalence classes and transfer of functions: The role of the nature of stimuli. *The Psychological Record, 66*, 65-74. Doi: 10.1007/s40732-015-0152-1

Vaidya M & Smith KN (2006). Delayed matching-to-sample training facilitates derived relational responding. *Experimental Analysis of Human Behaviour Bulletin, 24*, 9-16.

Received, January 25, 2021
Final Acceptance, April 30, 2021