The Psychological Record (2023) 73:455–466
https://doi.org/10.1007/s40732-023-00558-w

ABAI Association for Behavior Analysis International

ORIGINAL ARTICLE

The Respondent-Type Matching-to-Sample Procedure: A Comparison of One-to-Many and Linear Procedure for Establishing Equivalence Responding

Jonathan Todd[1] · **Mickey Keenan**[1] · **Stephen Gallagher**[1]

Accepted: 11 August 2023 / Published online: 5 September 2023

Abstract

Stimulus equivalence research is dominated by operant conditioning procedures that require the active responding of a participant to establish relations between arbitrary stimuli. In comparison, there has been relatively little research using respondent-type procedures, which only require the participant to view relations that appear on screen. This presentation describes two experiments using a respondent-type matching-to-sample procedure to examine the effect of the one-to-many (OTM) training procedure and the linear procedure on equivalence class formation. The OTM procedure was extremely effective in generating equivalence responding, however the linear procedure was not. These findings are discussed in the context of previous research comparing the two training procedures, as well as the effectiveness of previous respondent procedures.

Keywords Stimulus equivalence · Respondent-type matching-to-sample procedure · One-to-many · Linear procedure

Introduction

If one was to choose a random article on stimulus equivalence research that has been published over the last 50 years, chances are the matching-to-sample (MTS) procedure is what was used to establish the relations between arbitrary stimuli (e.g., Pilgrim, 2020). A sample stimulus (A1) is presented along with multiple comparison stimuli (B1, B2, B3, etc.), one of which is deemed by the experimenter to be the correct comparison (in this case B1). The participant is tasked with selecting the correct comparison stimulus in relation to the sample, with a correct response being reinforced to establish the relation between the two stimuli. Tests for emergent relations then follow to determine whether equivalence classes have been established (Pilgrim, 2020). In contrast, there has been relatively little research that has used respondent conditioning methods to establish equivalence responding (e.g., Brown et al., 2023; Leader et al., 1996). A typical respondent-type (ReT) procedure doesn´t require or reinforce selection responses, the participant is simply required to view stimulus pairings that appear on screen; for example, stimulus A1 precedes the appearance of stimulus B1. Although Leader et al. (1996) and Leader and Barnes-Holmes (2001) found the ReT design to be as effective as the MTS procedure, other research has come to the opposite conclusion. That is, although the ReT procedure can train equivalence relations successfully, it is simply not as effective as the traditional MTS procedure (Clayton & Hayes, 2004). More recent research by Amd et al. (2017) compared the ReT procedure to the MTS procedure as well as a "stimulus pairing with response (SPresp)" procedure and a "stimulus pairing with orientation (SOresp)" procedure. The SOresp training trials involved presenting a cross in one of the four corners of the screen for each trial, and the participant was tasked with first clicking the cross to initiate the presentation of the stimuli, which appeared in the same corner as the cross. They found this procedure to be the most effective procedure out of the four at establishing transitive relations, including the MTS procedure. These findings suggest that an orientation response, something typically absent in ReT procedures, may be required to facilitate the establishment of emergent relations in ReT procedures.

MTS procedures are built using different sequences of conditional discriminations, with variations in the presentation and arrangement of the stimuli in relation to each other (Saunders & Green, 1999). These training structures are known as the one-to-many (OTM), linear series and

✉ Jonathan Todd
todd-j15@ulster.ac.uk

[1] School of Psychology, Ulster University, Coleraine, NI, UK

many-to-one (MTO) structures (Plazas & Villamil, 2016). In the context of three-member equivalence classes using an OTM training structure, each equivalence class is established by training each comparison stimulus to a single sample stimulus, typically the A stimulus (e.g., train A-B, followed by A-C). Following training, participants are tested for the emergent relations of symmetry (B-A, C-A) and equivalence (B-C, C-B; Keenan et al., 2015; Delgado & Rodriguez, 2020). Linear structures train relations sequentially by training A to B, followed by training B to C, and then testing for the emergence of relations between the A and C stimuli (Green & Saunders, 1998; Reilly et al., 2005). The third training structure, the MTO structure (Saunders et al., 1993; Hove, 2003) involves training multiple samples to one comparison (e.g., training B-A, followed by training C-A). Research has shown that the potential success of generating equivalence classes is determined by the training structure employed (Saunders & Green, 1999), with the linear structure appearing to be the least effective of the three training structures in terms of establishing the baseline relations required for equivalence responding (Arntzen et al., 2010; Holth & Arntzen, 1998). It is worth noting that research has been divided on whether OTM or MTO are equally effective; Arntzen and Holth (1997) and Arntzen and Nikolaisen (2011) found OTM to be more effective, whereas Saunders et al. (1999) and Hove (2003) came to the opposite conclusion. Amd et al. (2017) also found the OTM procedure to be the most effective procedure, however, this was not to a significant margin.

A new experimental design by Delgado and Rodriguez (2020), which they termed the respondent matching-to-sample procedure (hereafter known as the respondent-type matching-to-sample procedure, or ReTMTS procedure) combines elements of both designs. Like the traditional ReT procedure, it does not involve overt reinforcement at any point in training. The training trials appear in a similar fashion to the MTS procedure, with a sample stimulus appearing followed by multiple comparison stimuli. Unlike the MTS procedure, however, the participant is simply shown the relation by virtue of a prompt (red boxes appearing around the sample and related comparison). Another difference is the presence of probe/verification trials following each training block of trials to determine whether responding is under the control of the highlighted relations. ReT experiments tend to have the participants progress to the test trials after viewing a set number of trials; the implementation of these probe trials, as well as a strict criterion required for progression through the experiment, makes it possible for the participant to be retrained at the part of the training that they have struggled with, rather than either being retrained from the beginning after failing the test phase, or not being retrained at all. Their study was successful in establishing equivalence responding, however it was conducted only using the OTM procedure, and did not examine how other training structures may affect the effectiveness of the training.

The purpose of the present study was to compare OTM and linear training structures using a variation of Delgado and Rodriguez's (2020) procedure to investigate whether the findings on the difference in effectiveness of the two training structures seen in earlier studies (e.g., Arntzen et al., 2010; Smeets et al., 1997) also applied to this experimental design, in regards to both successful establishment of equivalence relations and the amount of exposures to the baseline relations that each participant required in order to progress through the training phase. Experiment 1 used an OTM training structure similar to the original Delgado and Rodriguez (2020) study, whereas Experiment 2 used a linear structure.

Experiment 1—OTM Structure

Method

Participants

Participants were 11 nonvulnerable adults over 18 years of age who could use and access a device that was connected to the internet. They were recruited through opportunistic sampling from both inside and outside the student body of Ulster University. There was no incentive to take part in the experiment. Participants read an information sheet and completed a consent form before beginning the experiment proper. All participants for this experiment were recruited online and took part in the experiment remotely.

Setting and Apparatus

By using the online experiment building website "Gorilla" (http://www.gorilla.sc; Anwyl-Irvine et al., 2019), the study could be conducted on desktop computers, laptops, and handheld smart devices like iPads or smartphones. The dimensions of the stimuli described here are given in the context of a 32 x 20.5cm laptop screen. The stimuli used in the task consisted of bold black Cyrillic, Hebrew, and Japanese Katakana letters presented against a white background (Fig. 1). Note that the stimuli were similar to the stimuli used by Delgado and Rodriguez (2020). The participants were not asked if they were familiar with the stimuli before beginning the experiment.

Phase 1. A-B Training and Probe Trials

After participants gave their consent to take part, they were taken to Phase 1 in which A-B relations were

	A Stimuli	B Stimuli	C Stimuli
Class 1	Ю	ש	テ
Class 2	ф	ב	ク
Class 3	б	א	ネ

Fig. 1 Stimuli Used in this Study Arranged in Their Equivalence Classes. *Note:* The A and B stimuli were both in the Calibri font, while the C stimuli were in the Yu Gothic font due to the Katakana alphabet not being supported by the Calibri font. The C stimuli were in bold to give them a similar thickness to the other stimuli. The stimuli were the same symbols used by Delgado and Rodriguez (2020)

trained. On screen, they were presented with the instructions:

You will now see a figure or symbol on the top centre of the screen and three other symbols below. Only one of these three symbols is related to the figure above. Observe these relations carefully. If you focus your attention on these relations, you will be able to finish the task faster than if you get distracted.
Click continue below to begin.

The instructions presented in this experiment were identical to the instructions used by Delgado and Rodriguez (2020) and were presented on the device the participant was using. The trials were presented sequentially in order of class; 4 A1-B1 trials were presented first, followed by 4 A2-B2 trials, and then 4 A3-B3 trials, for a total of 12 A-B training trials. For each training trial, at first, the sample stimulus (A; 1.5 cm tall and 1.5 cm wide) was first presented alone at the top center of the screen (7.2 cm from the top of the screen, 11.5 cm from the bottom and 15.5 cm from each side of the screen). After 2 s, the comparison stimuli (B1, B2, and B3) appeared in a line 6.3 cm below the sample stimulus. One comparison appeared directly below the sample, with the other two comparisons 5 cm to the left and right of the center comparisons. The left and right comparisons were both 10.7 cm away from their respective side of the screen, and the center comparison was 15.5 cm away from both edges. The comparison stimuli were each 0.6 cm tall by 0.6 cm wide (smaller than the sample stimulus) and appeared 5.3 cm from the bottom of the screen. After another 2 s had passed, red boxes appeared around the sample stimulus and the related comparison stimulus. For example, if A1 was the sample, red boxes would appear around it and B1. The red box that appeared around the sample stimuli was 6.2 cm wide and 4.5 cm tall, with a line thickness of 0.2 cm. The center of the sample stimulus was 3 cm from the sides of the red box, and 2 cm from the top and bottom of the red box. The comparison stimuli red box that appeared was 3.6 cm by 2.6 cm, with a 0.1 cm thickness. The comparison stimuli were 1.7 cm from the sides, and 1.2 cm from the top and bottom of the red boxes. This screen with all the stimuli remained for 3 s before being followed by the inter-trial interval (ITI), a white screen. After 2 s, the next training trial began, following the same sequence as above (Fig. 2 illustrates this sequence below).

Following the presentation of the training trials, new instructions were given to the participant:

Now that you have observed, please select the symbol that corresponds with the one presented on top. Click to begin.

The participants were presented with a screen containing the sample and comparison stimuli in a similar arrangement to the training trials, albeit with no red boxes. A major difference was that the comparison stimuli now had grey borders as they were now "buttons" for the participant to interact with (see Fig. 3). The grey borders were 3.6cm wide by 2.6

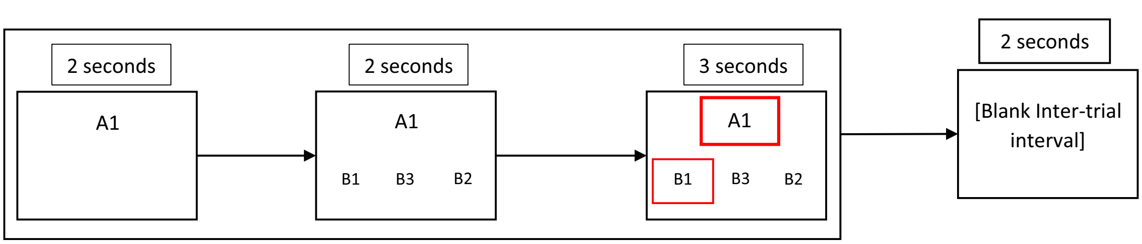

Fig. 2 Schematic Diagram of an A1-B1 Training Trial. *Note:* This schematic diagram shows a single training trial in Phase 1 which lasted a total of 7 s. The timings before the next screen was displayed are shown. After the intertrial interval, the next training trial began immediately. All training phases in both experiments used this format for establishing the baseline relations. For Phases 2 and 3 of Experiment 2, were B-C relations were trained, the B stimuli appeared as the samples

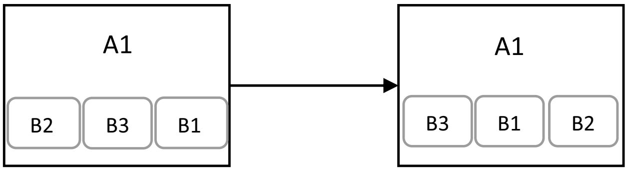

Fig. 3 Schematic Diagram of A1-B1 Probe Trials. *Note:* This schematic diagram shows two A1-B1 probe trials used in Phase 1. All probe trials were presented in a semi-random order. Once a participant selected one of the comparison stimuli, the next probe trial was presented. After the second A1-B1 probe trial, an A2-B2 probe trial was presented. This format for presenting probes was used for all A-B relations. Furthermore, during all probe trials, participants were not told if their selection was correct. This format was also used for the probe trials of the A-C relations, and the B-C relations in Experiment 2. Phase 3, the mixed training phase, contained a mix of probe trials for both A-B and A-C relations (A-B and B-C relations in Experiment 2)

cm tall, with rounded edges and a thickness of less than 0.1 cm. Clicking on any of the comparison stimuli progressed the experiment to the next trial. As in the training trials, the probe trials were presented sequentially by class. After six probes (two for each class), participants who demonstrated 100% accurate responding progressed to Phase 2, whereas those who did not meet this criterion were taken to the start of a shortened version of Phase 1 for retraining. Retraining involved viewing a block of nine training trials (three per class) and then once again completing the same six probe trials. Failure to successfully answer all six probe trials after a second retraining session resulted in the experiment concluding, followed by debriefing. Participants were presented with a screen displaying "Well Done" after answering the six probe trials regardless of performance.

Phase 2. A-C Training and Probe Trials

The procedure for Phase 2 was identical to the procedure of the previous phase, except that this phase trained relations between the A and C stimuli. A notable difference between this phase and Phase 1 was that participants did not receive instructions before beginning this test phase. This was due to an error, as participants were meant to see the same instructions that were presented for Phase 1. In Experiment 2, this error is corrected. Once again, advancement to the next phase of the experiment was contingent on participants achieving 100% mastery of the relation being trained.

Phase 3. Mixed A-B and A-C Training and Probe Trials

Again, the procedure for this phase was identical to the previous phases, with the exception of the presence of both A-B and A-C training trials in this phase, and the presence of 12 probe trials in this phase instead of just 6. Advancement to the final test phase of the experiment relied on the participant demonstrating 92% accurate responding in the probe trials in this phase. Failure to meet this criterion resulted in the participant being retrained, like in the other phases. Retraining cycles in this phase contained 12 training trials, instead of 9. Again, participants did not receive instructions at the beginning of this phase.

Phase 4. Test for Emergent Relations

Upon reaching this phase, participants were informed that they had reached the final test trial phase and were to try the trials that followed. Participants were presented with 36 test trials divided into four 9-trial (three test trials per class) blocks, each block testing for a different emergent relation. The first test block tested for B-A symmetry, the second tested for C-A symmetry, the third tested for B-C equivalence and the fourth C-B equivalence relations. To display successful equivalence class formation, 90% accurate responding had to be demonstrated in each of the four test blocks (eight of the nine trials in each block); participants completed all four test blocks regardless of below criterion performance in one or multiple blocks. Once the participants completed the final test block, they were debriefed, and the experiment was concluded.

Results and Discussion

Table 1 displays the number of training trials in each training phase for participants who successfully reached Phase 4 of Experiment 1. With the exception of P5, all participants successfully reached Phase 4 having been exposed to 36 training trials. P5 was retrained twice, unsuccessfully, in the first phase, meaning they were only exposed to 30 A-B training trials (nine extra trials per retraining block) before the experiment was concluded for them. Of the remaining 10 participants, only 2 participants required retraining at some point in the experiment. P3 and P11 were both retrained once in Phase 2, and as a result were exposed to 45 training trials in total.

Table 2 shows the participants' performances in the four tests for derived relations in Phase 4. The criteria for demonstrating successful generating of the three equivalence classes was at least 90% accurate responding in each of the test blocks (eight out of nine trials answered correctly) All the participants who reached Phase 4 met the criterion required to demonstrate successful equivalence class formation. Seven of these participants completed the phase errorless (100% correct responding in each block), with the other three participants each having made a single error at one point in the phase.

Table 1 Number of Training Trials for Each Participant in Order to Achieve Mastery of Each Baseline Relation in the Three Training Phases of Experiment 1

Participant	Phase 1: A-B Trials		Phase 2: A-C Trials		Phase 3: Mixed A-B, A-C Trials		Total Trials
	Trials	Retraining Cycles	Trials	Retraining Cycles	Trials	Retraining Cycles	
P1	12	0	12	0	12	0	36
P2	12	0	12	0	12	0	36
P3	12	0	21	1	12	0	45
P4	12	0	12	0	12	0	36
P6	12	0	12	0	12	0	36
P7	12	0	12	0	12	0	36
P8	12	0	12	0	12	0	36
P9	12	0	12	0	12	0	36
P10	12	0	12	0	12	0	36
P11	12	0	21	1	12	0	45

Under each training phase, trials refers to the number of training trials each participant was exposed to in each phase, and retraining cycles showcases the amount of retraining cycles for each participant in each phase if required (to a maximum of two retraining cycles per phase). The rightmost column displays the total number of training trials each participant was exposed to across the three training phases

Table 2 Results Obtained for Each Participant in Phase 4 of Experiment 1

Participant	B-A Symmetry	C-A Symmetry	B-C Equivalence	C-B Equivalence
P1	100%	100%	100%	100%
P2	100%	100%	100%	90%
P3	100%	100%	100%	100%
P4	100%	100%	100%	100%
P6	100%	100%	100%	100%
P7	100%	100%	100%	100%
P8	100%	100%	100%	100%
P9	90%	100%	100%	100%
P10	100%	100%	100%	90%
P11	100%	100%	100%	100%

Each test block contained 9 test trials for a total of 36 trials. The criterion for successful responding was 90% in each of the four blocks, including the symmetry test blocks. As the experiment used the OTM training structure, the first two blocks tested for symmetry, with the other two blocks testing for equivalence (combined symmetry and transitivity)

The objective of this experiment was to investigate the effectiveness of the ReTMTS procedure with a OTM training structure. With the exception of a single participant, all participants successfully displayed equivalence responding across the three classes. These findings concur with those of Delgado and Rodriguez (2020) who also used the OTM structure with the ReTMTS procedure and found that the majority of their participants were able to successfully demonstrate equivalence responding across all three classes. In fact, the current experiment was arguably more successful than the findings of Delgado and Rodriguez (2020) where only 68% of participants met the criterion for equivalence responding.

Given the extremely high success rate of this procedure, it was decided to conduct an experiment with the same procedure but with a linear training structure (and with errors rectified), to see if the training structure would have an impact on both the establishment of the initial baseline relations, and the establishment of equivalence relations.

Experiment 2—Linear Structure

This experiment investigated if using a linear training structure (i.e., train A-B and B-C) in the ReTMTS procedure had an influence on the successful establishment of relations between the stimuli and subsequent equivalence class formation.

Method

Participants

Thirteen participants were recruited from the student body of Ulster University through opportunistic sampling and use of the university's online participant pool known as SONA. Participants recruited through SONA did receive an incentive to take part in this experiment in the form of course credit, whereas those recruited through opportunistic sampling did not receive any incentive, as in Experiment 1. Participants once again received an information sheet and completed a consent form prior to beginning the experiment. Although the first four participants were recruited through SONA and took part in the experiment remotely, the remaining nine participants completed the experiment within an experimental suite on the university campus. However, these participants took part in the experiment on their own devices rather than using a device provided by the researchers. The dimensions of the experimental suite are given below.

Setting and Apparatus

As Experiment 2 was also hosted on "Gorilla" (http://www.gorilla.sc; Anwyl-Irvine et al., 2019), the setting and apparatus was the same as described in Experiment 1. The stimuli used were also the same as the stimuli used in Experiment 1. The dimensions for the experimental suite were as follows: the room was 230 cm by 230 cm, containing a desk 140 cm wide by 60 cm tall and a chair. The desk had a computer on it; however, it was not switched on and the participant was directed to sit away from it by where the QR code was set on the desk. Participants sat with their backs approximately 130 cm from the door, and approximately 40 cm from the desk. The QR code sheet displayed the QR code, along with the university's logo, name of the study and names of the researchers. It also contained basic information on the study that was also provided in the information sheet, specifically the length of time the study took to complete and that it could be completed on any device.

Procedure

Experiment 2 used a linear training structure, meaning that A-B and B-C relations were established in this experiment, and meant that the B stimuli were presented as the sample stimuli for B-C training trials. Phase 1 was identical to the same phase of the Experiment 1, whereas Phase 2 trained and established B-C relations. Phase 3 presented participants with a mix of A-B and B-C training trials, and Phase 4 tested for A-B and C-B symmetry A-C transitivity and the equivalence relation of C-A. The error with the instructions in Experiment 1 was corrected in this experiment; participants saw the same instructions before each training phase. The wording of the test phase instructions was also changed and tells the participant that now that they have learnt which symbols go together to try the following trials. Although they were still informed that this was the final phase, they were not explicitly informed that was a test phase like in Experiment 1.

Results and Discussion

Table 3 showcases the number of training trials each participant was exposed to across Experiment 2's three training phases, again omitting the results of the participants who did not meet criterion to progress to the test phase. P18 and P19 both did not reach the test phase, with P18 being unsuccessfully retrained twice in Phase 2 and P19 was trained unsuccessfully in Phase 1. Of the remaining 11 participants, 5 progressed to Phase 4 with no retraining having taken place. Three participants (P16, P22, and P24) were retrained once at one point in the three training phases, P16 and P22 were both only retrained once in Phase 2, and were exposed to 45 training trials in total, whereas P24 was retrained once in Phase 3, the only participant to require retraining in this phase. This resulted in P24 being exposed to a total of 48 training trials, as retraining cycles in Phase 3 contain 12 training trials. The remaining three participants were each retrained twice in one of the three training phases (P14 and P21 were exposed to two retraining cycles in Phase 1, whereas P15 was successfully retrained twice in Phase 2), resulting in the three participants being exposed to 54 training trials.

Table 4 displays the scores the participants obtained in each of the test blocks in Phase 4 of Experiment 2. Like in Experiment 1, at least 90% accurate responding in each test block was required to demonstrate successful generation of the equivalence classes. Three of the participants met this criteria (P12 doing so errorless), the remaining eight did not. P15, P17, and P22 were all one correct responses below criterion, having responded above criterion in all but one of the test blocks. The number of correct responses in each of the test blocks among the other unsuccessful participants varied. It should be noted that both P14 and P20 made no correct responses in the A-C transitivity test block yet made met criterion in the subsequent C-A equivalence block.

The goal of Experiment 2 was to investigate whether the ReTMTS procedure would be effective when it was employed with a linear training structure. The results above show there was a much higher level of variation in responding within the participants compared to the participants who were exposed to the OTM structure in Experiment 1. Only 3 of the 11 participants (27%) who reached the test phase met the criterion to display equivalence responding across the three classes. These

Table 3 Number of Training Trials for Each Participant in Order to Achieve Mastery of Each Baseline Relation in the Three Training Phases of Experiment 2

Participant	Phase 1: A-B Trials		Phase 2: B-C Trials		Phase 3: Mixed A-B, B-C Trials		Total Trials
	Trials	Retraining Cycles	Trials	Retraining Cycles	Trials	Retraining Cycles	
P12	12	0	12	0	12	0	36
P13	12	0	12	0	12	0	36
P14	30	2	12	0	12	0	54
P15	12	0	30	2	12	0	54
P16	12	0	21	1	12	0	45
P17	12	0	12	0	12	0	36
P20	12	0	12	0	12	0	36
P21	30	2	12	0	12	0	54
P22	12	0	12	0	24	1	48
P23	12	0	12	0	12	0	36
P24	12	0	21	1	12	0	45

Under each training phase, trials refers to the number of training trials each participant was exposed to in each phase, and retraining cycles showcases the number of retraining cycles for each participant in each phase (to a maximum of two retraining cycles per phase). The rightmost column displays the total number of training trials each participant was exposed to across the three training phases

Table 4 Results Obtained for Each Participant in Phase 4 of Experiment 2

Participant	B-A Symmetry	C-B Symmetry	A-C Transitivity	C-A Equivalence
P12	100%	100%	100%	100%
P13	100%	90%	40%	40%
P14	90%	70%	0%	100%
P15	100%	100%	90%	80%
P16	100%	100%	90%	100%
P17	100%	100%	80%	90%
P20	100%	100%	0%	90%
P21	100%	100%	90%	100%
P22	100%	90%	90%	80%
P23	90%	30%	30%	80%
P24	40%	100%	100%	80%

Each test block contained 9 test trials for a total of 36 trials. The criterion for successful responding was 90% in each of the four blocks, including the symmetry and transitivity blocks. As the experiment used the linear training structure, the first two blocks tested for symmetry, the third training block tested for transitivity, with the final block testing for equivalence

findings agree with the literature regarding the linear structure being the least effective structure when it comes to establishing equivalence responding. Of note, regarding the fact that the first four participants completed the experiment remotely, the data displayed above does not appear to show any significant difference between the results of the remote participants and the participants is completed the experiment on campus.

General Discussion

The objective of this study was to compare the effects of two respondent-type matching-to-sample (ReTMTS) experiments, one with a one-to-many (OTM) training structure and another with a Linear structure. Compared to the OTM procedure, the linear procedure was much less effective at both the initial establishment of the baseline relations and the establishment of the subsequent emergent relations. This is illustrated in Fig. 4 below.

These findings are in line with previous research on training structures conducted using operant procedures (Arntzen et al., 2010), but not in line with the findings of respondent-type (ReT) literature, which found the three structures to be equally effective when tested with preschool children (Leader et al., 2000; Smeets et al., 1997). However, it is worth mentioning that the results of the first experiment by Smeets et al. (1997), which examined the effectiveness of MTO and OTM training structures in a respondent-type

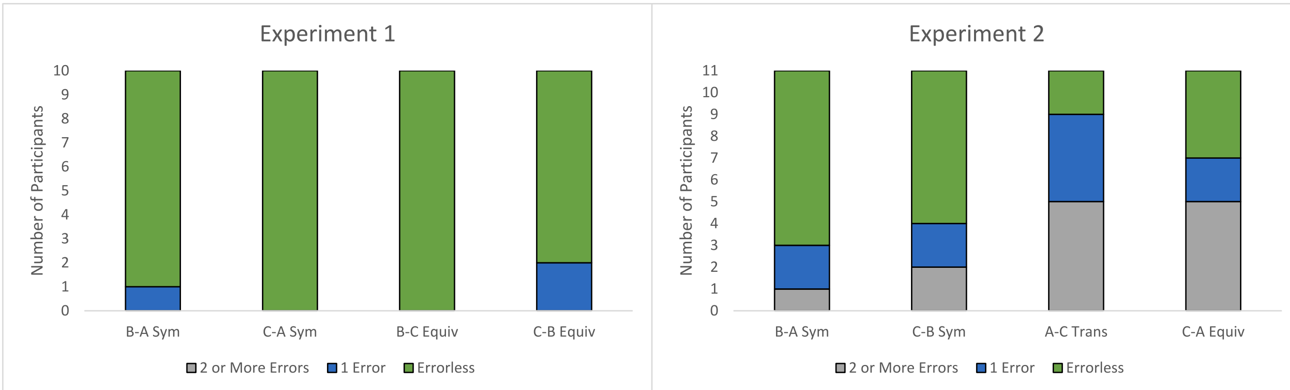

Fig. 4 Bar Charts Displaying Number of Errors Participants Made in Each Test Block in Phase 4 of Experiments 1 and 2. *Note:* The two bar charts above display the results of Phase 4 of Experiment 1 on the left and Phase 4 of Experiment 2 on the right. Errorless performance in the test block is represented by the green color, while performance where a single error was made is represented by the color blue. The gray color represents performance below criterion where participants made more than one error in the test block. Note that Experiment 2 tested one more participant than Experiment 1

procedure, alongside the results of Condition 2 (linear) in Leader et al. (1996), provided limited evidence that the OTM was more effective than the linear structure (the OTM being, in turn, less effective than the MTO procedure). These two experiments were conducted with adult participants, as was the present study. The current findings also concur with the findings of Kinloch et al. (2013), who compared the effectiveness of the MTS procedure and the stimulus pairing observation procedure (SPOP, another term for the ReT procedure) as well as the different training structures, also conducted with adult participants, found the OTM to be the most effective compared to both the MTO and linear structures with both procedures. They also found the MTS procedure to be marginally more effective than the SPOP procedure.

Regarding the relative ineffectiveness of the linear procedure, Fields et al. (1984) theorized that nodal distance was primarily responsible. They felt that the larger the distance between stimuli indirectly linked through training, the less robust the procedure is at subsequently establishing equivalence relations. What this means in the context of Experiment 2 is that the nodal distance between stimuli A and C is larger than the nodal distances of all the other relations in Experiment 1, resulting in less instances of successful equivalence responding. Saunders and Green (1999) disagreed with this analysis of nodal distance effects and hypothesized that the number of *simple discriminations* embedded in the training structures was more important. Simple discriminations refer to when responding occurs in the presence of S+ but not in the presence of S- and can be either simultaneous (such as discriminating between comparisons) or successive (discriminating between samples). They felt that as linear structures do not present all possible simultaneous discriminations during training, a higher failure rate during testing is then to be expected when using this procedure. In fact, they stated that the only structure to present all simple discriminations during training that are required for the emergence of equivalence relations was the MTO structure. The OTM procedure, not allowing for a simple discrimination to occur between the B and C stimuli, is subsequently less likely to be successful compared to the MTO. However, all participants tested in Experiment 1 displayed equivalence responding. This finding suggests that a future experiment comparing OTM and MTO structures should be conducted with the procedures outlined here.

It is notable that almost all the participants who took part in OTM procedure were able to reach the test phase in 36 training trials, the minimum amount of training required. This suggests that the OTM ReTMTS experiment was successful in establishing the baseline relations with exposures to only the minimum amount of training trials. In comparison, 7 of the 19 participants tested by Delgado and Rodriguez (2020), also using an OTM structure, required exposure to the minimum amount of training trials, the remaining participants requiring retraining. A potential reason for this difference could be the fact that the training phases in the current study presented both the training and probe trials sequentially (as in, four A1-B1 training trials followed by four A2-B2 trials and then four A3-B3 trials) rather than having them presented semi-randomly as in the Delgado and Rodriguez's (2020) study. It was only during Phase 4 that the trials were presented in a semi-random order. It is possible that the repeated pairings of the same stimuli in one block, perhaps in essence giving the participant more time to view the pairing before seeing the next pair of related stimuli, was a key factor in making the training more effective in establishing the baseline relations.

The presence of the probe trials appeared to facilitate the establishment of equivalence relations in both experiments, though to different degrees, presumably due to the differences of training structures. Another aspect that demonstrates the probe trials importance to the current procedure is the fact that equivalence responding can be reliably established with the ReTMTS procedure (most notably with the OTM structure) in as little as 36 training trials. This is quite a low number of training trials (indeed, it is only 18 trials per relation across the three training phases) compared to both ReT procedures and MTS procedures. For example, in Leader et al. (1996), 60 ReT training trials were presented, with each individual relation (A1-B1, B1-C1, A2-B2, B2-C2, A3-B3, B3-C3) presented 10 times. In comparison, each individual relation was only presented six times in each experiment, four in either Phases 1 or 2, and twice in Phase 3, not counting retraining. For a recent MTS example, Keenan et al. (2020) presented each individual baseline relation (A1-B1 etc.) 12 times for a total of 48 trials (without retraining). The inclusion of the probe trials, and the apparent requirement of their presence to establish equivalence responding, raises an interesting point, especially in regard to the ReT training component of this procedure. However, the relative contribution of operant and respondent components is not clear. From a respondent perspective, participants were exposed to more training trials than probe trials during training, and they did not receive any feedback for their responses beyond the presentation of the next trial and progression through the experiment (and a "Well done" regardless of performance). On the other hand, the probe trials appear to be necessary for the experiment to be more successful, by at the very least allowing for retraining of an individual relation to occur. However, the selection of B1 in the presence of A1 in the probe trials did not involve any feedback for a correct response. At the very least, it can be concluded that the participants should be given the option to demonstrate mastery of the baseline relations during training, in order to guarantee that the participants are entering the test phase with responding under the control of the relations highlighted by the training trials. It is also worth noting that this is not the only example of a ReT procedure introducing responding during training. Carnerero et al. (2019) also implemented probe trials following the ReT training of a relation. Amd et al. (2017), although not implementing responding in order to demonstrate mastery of the relations, did implement an orienteering response before each training trial was presented. Their findings, mentioned in the introduction above regarding the stimulus pairing with orientating procedure (SOresp), suggest that an orientation response in a ReT procedure is required to guarantee consistent success in the ReT procedure. Although an orientating response is not required in the current procedure, the participants are still required to orientate to the appropriate comparison stimuli during training.

Two of the participants were unable to respond correctly to any of the A-C transitivity trials in Phase 4 in Experiment 2. Despite this, both participants subsequentially met criterion in the C-A equivalence block. In fact, P20 met criterion in all blocks except the transitivity block. Perhaps Stimulus Control Topography Coherence Theory (SCTCT; Dube & McIlvane, 1996; Fields & Paone, 2020; McIlvane & Dube, 2003) may account for these results. Stimulus Control Topography refers to the differences between the stimuli in a class, such as their physical features, and their controlling properties, and these different control topographies may control the same measured response. SCTCT refers to the coherency between the experimenter defined stimuli that control behavior and the stimuli that actually gain control of the participant's behavior. For example, in the current study a participant might select B1 in the presence of A1 in a probe trial because of the control of the red box prompts during training, but in the next trial may select B1 in the presence of A1 because of the physical space B1 inhabits in this particular probe trial. It appears that the current study, and by extension the ReTMTS procedure in general, seems to display high levels of experimenter-defined stimulus control. The results of Experiment 1 attest to that due to the consistently accurate responding displayed by all participants in the emergent relations test phase. Indeed, even in Experiment 2 high levels of accurate responding are seen, however, there is much greater variation in the correct responses across the test blocks. That being said, what caused the failure of the emergence of transitivity in two participants who were still able to display the emergence of equivalence? Possibly it was the position of the stimuli on the screen. Apart from the training phases, the transitivity block was the first time that the C stimuli were presented as comparisons and the A stimuli as samples, when prior to that in the symmetry test blocks the A stimuli were presented as comparisons and the C stimuli as samples. Limited evidence to suggest that this physical placement of stimuli on screen was controlling responding may possibly be found in the results of P23, who, despite not reaching criterion overall, responded somewhat accurately in the first symmetry block and the final equivalence block, which both presented the A stimuli as comparisons, compared to the other test blocks. Tonneau et al. (2006) found something similar, with transfer of function effects occurring between stimuli depending on shared stimulus correlations, such as their location on screen.

Another potential reason could be that the physical shape of the stimuli, especially the C stimuli, affected control in some way, and rendered the experimenter-defined relations between the C stimuli and the other class members weak. From a cursory glance, the C stimuli do appear to be more physically different compared to the A and B stimuli. They are more complicated shapes and in the case of C1 and C3 are made up of more than one line. Perhaps the complicated shape made it much more difficult for the participants to respond accurately when the C stimuli were presented together as comparisons, as this resulted in the participants struggling to differentiate between the three compared to when they were presented alone as samples. Clayton and Hayes (2004) used similarly complex Chinese characters in two of their experiments (albeit all their stimuli where from this alphabet, rather than each member of the class belonging to a different alphabet). It was felt that these stimuli were less nameable and therefore less likely to facilitate establishment of equivalence responding compared to their more "nameable" counterparts. Perhaps this also occurred here, with the C stimuli, being less nameable than the A or B counterparts. It is also possible that this an example of delayed emergence of equivalence (Holth & Arntzen, 1998; Sidman, 1994). Delayed emergence is a phenomena reported in equivalence literature that is showcased by an increase in accurate responding as a result of continued testing (Arntzen & Mensah, 2020). In short, these participants, although scoring very low in the transitivity block (selecting C in the presence of A), were effectively trained in this block, in the absence of feedback or overt displays of the defined relations, to respond accurately in the equivalence block (selecting A in the presence of C). It should be reiterated, however, that the OTM procedure produced consistent accurate responding across all four test blocks compared to the linear experiment.

The findings reported here in regard to the use of the respondent-type matching-to-sample procedure agree with the research of Hayes (1992), Tonneau (2001, 2002) and Minster et al. (2011), who all suggested that emergent relations are influenced more by respondent-type relations, also known as stimulus pairings/correlations, than they are influenced by reinforcement contingencies established by the operant procedures more commonly used in the wider literature. These findings challenge the traditional view that stimulus equivalence is a direct result of operant contingencies, specifically reinforcement of correct responses during training (Delgado & Hayes, 2013). Tonneau (2001) suggested that functional equivalence is derived from what he called stimulus correlations and argued for a shift away from MTS procedures to stimulus correlations. He considered that the origins of symbolic behavior lay in transfer of function across stimuli that was established by stimulus pairings that were sensitive to the effects of operant feedback. In short, function transfer within MTS experiments occurs due to the stimulus pairings implicit in the MTS procedure, not through the actual matching-to-sample procedure itself. The behavioral effects of a stimulus (such as a selecting behavior) transfer between members of an equivalence class when they have been temporally, and directly or indirectly, paired (Tonneau, 2001). For example, Minster et al. (2011) were able to establish three 5-member equivalence classes despite two members of each class having an established history of extinction. The current procedures simply demonstrated that the appearance of red boxes around the related sample and comparison stimuli (which were required to highlight the relation between the two stimuli), along with the probe trials that did not provide feedback, was sufficient to establish the relations between the stimuli. The red boxes function as orienting stimuli, which is a respondent process, albeit one that has required a prior learning history to establish that boxes around the two stimuli signify a relation between them. This learning history is established both by the instructions of the experiment (which explicitly refer to stimuli "going together") and the wider environment of the participant where similar stimuli function to draw attention to certain objects (such as signs and labels). Perhaps the red boxes also function as a shared stimulus feature that is delayed in its appearance, and simply function as an extension of the stimulus-pairing procedure.

To conclude, the findings reported here demonstrate that the respondent-type matching-to-sample procedure is effective at establishing equivalence responding, with the one-to-many training structure being much more effective at doing so than the linear structure. Perhaps with more exposures to the baseline relations, the effectiveness of the linear structure would increase. Although the current study cannot speak to the procedure's effectiveness in comparison to the typical MTS procedure (as in Kinloch et al., 2013; and this should certainly be the basis of a future experiment, alongside an experiment comparing the current procedure to the ReT procedure), the fact remains that the respondent contingencies arranged by the ReTMTS procedures used here were able to successfully establish equivalence responding. However, relational frame theory (RFT) and naming theory may suggest other contingencies at play, such as contextual control or that respondent-type relations rely on an operant process such as relation framing or verbal behavior (Barnes, 1994; Horne & Lowe, 1996). Although these contingencies were unfortunately not fully addressed in the current procedures (such as the training instructions explicitly informing the participant the stimuli were to be related), the fact remains that equivalence responding was reliably established in Experiment 1 using a procedure that appears to involve mostly respondent contingencies.

Funding No funding was received for conducting this study.

Data Availability The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations This research was conducted in partial fulfilment of the first author's self-funded PhD in psychology under the supervision of Mickey Keenan and Stephen Gallagher at Ulster University. Parts of this article were presented at the 16th Annual Division of Behaviour Analysis conference for the Psychology Society of Ireland in April 2023. This article has not been previously published or submitted elsewhere.

Conflict of Interest The authors have no competing interests to declare that are relevant to the content of this article.

Ethical Approval All procedures in the current study were in accordance with the ethical standards of the institutional research committee, and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Ethical approval was approved by the School of Psychology staff and Postgraduate Research Ethics Filter Committee, Ulster University, Northern Ireland, UK. Research involved human participants from nonvulnerable populations. Informed consent was obtained for all participants.

References

Amd, M., de Almeida, J. H., de Rose, J. C., Silveira, C. C., & Pompermaier, H. M. (2017). Effects of orientation and differential reinforcement on transitive stimulus control. *Behavioural Processes, 144*, 58–65. https://doi.org/10.1016/j.beproc.2017.08.014

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2019). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47*, 309–320. https://doi.org/10.1007/BF03395227

Arntzen, E., & Mensah, J. (2020). On the effectiveness of including meaningful pictures in the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 113*(2), 305–321. https://doi.org/10.1002/jeab.579

Arntzen, E., & Nikolaisen, S. (2011). Establishing equivalence classes in children using familiar and abstract stimuli and many-to-one and one-to-many training structures. *European Journal of Behavior Analysis, 12*(1), 105–120. https://doi.org/10.1080/15021149.2011.11434358

Arntzen, E., Grondahl, T., & Eilifsen, C. (2010). The effects of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *The Psychological Record, 60*, 437–461. https://doi.org/10.1007/BF03395720

Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record, 44*(1), 91.

Brown, K., Rosales, R., Garcia, Y., & Schneggenburger, S. (2023). A review of applied research on pairing procedures to facilitate emergent language. *The Psychological Record, 73*, 221–236. https://doi.org/10.1007/s40732-023-00543-3

Carnerero, J. J., Pérez-González, L. A., & Osuna, G. (2019). Emergence of naming relations and intraverbals after auditory stimulus pairing: Effects of probing the listening skill first. *The Psychological Record, 69*(2), 239–252. https://doi.org/10.1007/s40732-019-00336-7

Clayton, M. C., & Hayes, L. J. (2004). A comparison of match-to-sample and respondent-type training of equivalence classes. *The Psychological Record, 54*, 579–602. https://doi.org/10.1007/BF03395493

Delgado, D., & Hayes, L. J. (2013). The integration of learning paradigms by way of a non-causal analysis of behavioral events. *Conductual, 1*(2), 39–54.

Delgado, D., & Rodríguez, A. (2020). Stimulus equivalence using a respondent matching-to-sample procedure with verification trials. *The Psychological Record, 72*, 1–9. https://doi.org/10.1007/s40732-020-00438-7

Dube, W. V., & McIlvane, W. J. (1996). 11 Some implications of a stimulus control topography analysis for emergent behavior and stimulus classes. *Advances in Psychology, 117*, 197–218. https://doi.org/10.1016/S0166-4115(06)80110-X

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*(2), 293–305. https://doi.org/10.1007/s40732-020-00384-4

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. https://doi.org/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). Springer.

Hayes, L. J. (1992). Equivalence as process. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 97–108). Context Press.

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

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. https://doi.org/10.1901/jeab.1996.65-185

Hove, O. (2003). Differential probability of equivalence class formation following a one-to-many versus a many-to-one training structure. *The Psychological Record, 53*(4), 617–634. https://doi.org/10.1007/BF03395456

Keenan, M., Porter, I., & Gallagher, S. (2015). Merging separately established functional equivalence classes. *The Psychological Record, 65*, 435–450. https://doi.org/10.1007/s40732-015-0118-3

Keenan, M., Schenk, J., Coyle, C., Reid, L., & Gallagher, S. (2020). The effects of social labels on the allocation of resources to equivalent stimuli: does one rotten apple spoil the whole barrel? *The Psychological Record, 71*, 17–39. https://doi.org/10.1007/s40732-020-00423-0

Kinloch, J. M., McEwan, J. S. A., & Foster, T. M. (2013). Matching-to-sample and stimulus-pairing-observation procedures in stimulus equivalence: The effects of number of trials and stimulus arrangement. *The Psychological Record, 63*(1), 157–174. https://doi.org/10.11133/j.tpr.2013.63.1.012

Leader, G., & Barnes-Holmes, D. (2001). Matching-to-sample and respondent-type training as methods for producing equivalence relations: Isolating the critical variable. *The Psychological Record, 51*, 429–444. https://doi.org/10.1007/BF03395407

Leader, G., Barnes, D., & Smeets, P. M. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record, 46*, 685–706. https://doi.org/10.1007/BF03395192

Leader, G., Barnes-Holmes, D., & Smeets, P. M. (2000). Establishing equivalence relations using a respondent-type training procedure III. *The Psychological Record, 50*, 63–78. https://doi.org/10.1007/BF03395343

McIlvane, W. J., & Dube, W. V. (2003). Stimulus control topography coherence theory: Foundations and extensions. *The Behavior Analyst, 26*(2), 195–213. https://doi.org/10.1007/BF03392076

Minster, S. T., Elliffe, D., & Muthukumaraswamy, S. D. (2011). Emergent stimulus relations depend on stimulus correlation and not on reinforcement contingencies. *Journal of the Experimental Analysis of Behavior, 95*(3), 327–342. https://doi.org/10.1901/jeab.2011.95-327

Pilgrim, C. (2020). Equivalence-based instruction. In J. O. Cooper, T. E. Heron, & W. L. Heward (Eds.), *Applied behavior analysis* (pp. 496–540). Pearson.

Plazas, E. A., & Villamil, C. W. (2016). Effects of between-classes negative relations training on equivalence class formation across training structures. The Psychological Record, 66, 489–501. https://doi.org/10.1007/s40732-016-0189-9

Reilly, T., Whelan, R., & Barnes-Holmes, D. (2005). The effect of training structure on the latency of responses to a five-term linear chain. *The Psychological Record, 55*(2), 233–249. https://doi.org/10.1007/BF03395508

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. https://doi.org/10.1901/jeab.1999.72-117

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*(4), 725–744. https://doi.org/10.1007/BF03395909

Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of the Experimental Analysis of Behavior, 71*(2), 195–214. https://doi.org/10.1901/jeab.1999.71-195

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

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

Smeets, P. M., Leader, G., & Barnes, D. (1997). Establishing stimulus classes in adults and children using a respondent-type training procedure: A follow-up study. *The Psychological Record, 47*, 285–308. https://doi.org/10.1007/BF03395226

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

Tonneau, F. (2002). Who can understand relational frame theory? A reply to Barnes-Holmes and Hayes. *European Journal of Behavior Analysis, 3*(2), 95–102. https://doi.org/10.1080/15021149.2002.11434209

Tonneau, F., Arreola, F., & Martínez, A. G. (2006). Function transformation without reinforcement. *Journal of the Experimental Analysis of Behavior, 85*(3), 393–405. https://doi.org/10.1901/jeab.2006.49-05