

Preserved nodal number effects under equal reinforcement

Ting Wang · Charlotte Dack · Louise McHugh · Robert Whelan

Published online: 2 February 2011
© Psychonomic Society, Inc. 2011

Abstract The present set of experiments tested the hypothesis that the nodal number effects observed in previous studies of stimulus equivalence were due to the confounding factor of training structure that resulted in unequal reinforcement across trial types. In [Experiment 1](#), two 5-member equivalence classes were trained across equal and unequal reinforcement conditions, both with and without a limited hold. A significant nodal effect, as measured by response speed, was found in the equal reinforcement, no-limited-hold condition. In [Experiment 2](#), two 6-member equivalence classes were trained in equal and unequal reinforcement conditions without limited hold. In a transfer-of-function test, clear nodal effects were observed in the equal reinforcement condition. [Experiment 3](#) replicated and extended the findings of [Experiments 1](#) and [2](#) with an increased number of baseline training trials. The results of the present study suggest that the effects of nodal number are independent of differential reinforcement. Furthermore, a transfer-of-function test was most sensitive to nodal effects, response speed was the next most sensitive measure, and response accuracy was the least sensitive measure of nodal effects.

Keywords Equivalence · Reinforcement · Nodal number · Training structure · Transfer of function · Adult humans · Button press

Several laboratory studies have demonstrated that when a number of interrelated conditional discriminations are trained, *derived* (untaught) relations often emerge, even though the stimuli do not necessarily share any physical properties in common with one another (e.g., Sidman, 1971). Typically, in these studies, interrelated conditional discriminations are trained, and then the derived relations are tested (e.g., train $A \rightarrow B$ and $B \rightarrow C$, then test $C \rightarrow A$). One of the simplest examples of a derived relation is the *equivalence* relation. For example, if a participant was trained that A equals B and B equals C, one would expect the participant to select A in the presence of C. That is, if the AB, BC relations are *directly trained*, the emergent CA relation will be derived. The A, B, and C stimuli are said to participate in an *equivalence class*. A *node* has been defined as a stimulus that is linked by training to at least two other stimuli (Fields & Verhave, 1987; Fields, Verhave, & Fath 1984), and the number of nodes that link any two stimuli in a set of trained conditional relations is described as the *nodal number* (Sidman, 1994; also known as *nodal distance*). For example, a 5-member equivalence class (A, B, C, D, and E) contains six 1-node relations (e.g., B–D, with C as the node), four 2-node relations (e.g., B–E, with C and D as nodes), and two 3-node relations (e.g., A–E, with B, C, and D as nodes).

The literature on nodal distance has focused on systematically manipulating training structure and protocols in order to examine their impact on equivalence formation and the relatedness of class members (e.g., Adams, Fields, & Verhave 1993; Fields et al., 1997; Saunders & Green,

Sections of the manuscript were presented at the annual meetings of the Association for Behavior Analysis, Atlanta, May 2006, and San Diego, May 2007. Portions of the manuscript were presented in partial fulfillment of the requirements for the B.A. Honors Degree, PhD in Psychology and the M.Sc. Degree in Research Methods, Swansea University by Ting Wang and Charlotte Dack, respectively.

T. Wang (✉) · C. Dack · L. McHugh
Psychology Department, Swansea University,
Singleton Park,
Swansea SA2 8PP, United Kingdom
e-mail: t.wang@swansea.ac.uk

R. Whelan
Trinity Centre for Bioengineering, Trinity College Dublin,
Dublin, Ireland

1999). The term *training protocol* refers to the sequence of conditional discriminations presented in baseline training and testing (Fields et al., 1993a; Imam, 2006). Three training protocols that are commonly used in equivalence studies are simple-to-complex (STC), complex-to-simple (CTS), and simultaneous protocol (SP). In STC, one baseline relation (AB) is trained, followed by a test for symmetry (BA); then the new baseline relation (BC) is introduced, and symmetry (CB), transitivity (AC), and equivalence (CA) relations are tested sequentially. In contrast, the test of equivalence relations is prior to the test of symmetry and transitivity relations after baseline training in CTS. As its name suggests, all emergent relations are tested simultaneously in one mixed block after baseline conditional discriminations are simultaneously trained in SP. The term *training structure* refers to the arrangement of linking stimuli presented in baseline training (Saunders & Green, 1999). For example, a linear training structure or serial training structure involves training A–B, B–C, and C–D, whereas a one-to-many structure, or sample-as-node, involves training A–B, A–C, and A–D, and a many-to-one structure, or comparison-as-node, involves training B–A, C–A, and D–A. Accuracy of responding to unreinforced probes for derived relations has been the most common measure of the relatedness of stimuli. In addition, supplemental measures of derived relational responding may shed more light on the nature of the relations among stimuli (Dymond and Rehfeldt (2001)). Several researchers have reported that response speed (RS) was a function of nodal number (Bentall, Dickins, & Fox 1993; Bentall, Jones, & Dickins 1998; Holth & Arntzen, 2000; Spencer & Chase, 1996; Wulfert & Hayes, 1988), even when accuracy remained intact. Fields, Landon-Jimenez, Buffington, and Adams (1995; see also Fields et al., 1993b) adopted another, alternative measure of stimulus relatedness—namely, a *transfer-of-function* test. In this study, two 5-member equivalence classes were trained using a structure that ensured equal reinforcement across trial types. All (i.e., 12) participants passed baseline discriminations. However, only 2 participants formed equivalence classes. After these 2 participants demonstrated the formation of equivalence classes, new responses were trained to the end (i.e., A and E) stimuli in each group. Transfer of function was measured in terms of the relative frequency with which responses trained to A and E stimuli were evoked by all stimuli in both classes. In general, transfer of function was an inverse function of nodal number.

The status of nodal number as an independent variable has been questioned, however, and it has been suggested that apparent nodal number effects are a function of other variables (Imam, 2001, 2003, 2006; Sidman, 1994, 2000). Imam has noted that Sidman's account of equivalence does not include the notion that test outcomes should vary as a

function of training structure, providing that extraneous stimulus control is prevented (Carrigan & Sidman, 1992; Sidman, 1994). Sidman (2000) has suggested that “procedural factors . . . might account for the results of experiments that have given rise to notions of directionality and nodal distance” (p. 145).

Imam (2001, Experiment 1) trained three 5-member equivalence classes in a serial manner across two conditions: accuracy only and accuracy with a *limited hold* (LH). LH involved the participant's receiving positive feedback only to correct choices that occurred within a specific time period. In Experiment 1, a serial training structure was used. In this way, the number of reinforcers scheduled for responses to particular trial types was deliberately unequal. Response latency in Experiment 1 tended to be an inverse function of nodal number on transitivity trials, but only in the accuracy-only condition on equivalence trials. That is, a nodal number effect was generally observed, although a time–accuracy trade-off seemed to have occurred. In Experiment 2, a simultaneous structure, which ensured an equal numbers of reinforcers across trial types, was used. The participants' response times in Experiment 2 tended *not* to decline as the nodal number between stimuli increased. According to Imam (2006), “By equalizing reinforcement history, the confound noted in the first experiment was eliminated, and the nodal number effect observed in the second experiment thus was greatly diminished for one- through five-node trials” (p. 109). Imam (2003) replicated Imam's (2001) Experiment 2 with a single participant.

There were a number of factors that likely contributed to the significant effect in Experiment 1 and the lack of a significant effect in Experiment 2 (Imam, 2001; see also Imam, 2003, 2006). In Experiment 1, three conditions (one, two, and three nodes) were compared using an analysis of variance (ANOVA). In Experiment 2, however, five conditions (one, two, three, four, and five nodes) were compared, resulting in a power reduced from that in Experiment 1. Another potential confounding factor in Imam's (2001) study was that the training structure varied across experiments. In Experiment 1, AB relations were trained to criterion, followed by a mix of AB and BC trial types, and so on. In Experiment 2, all trial types were introduced simultaneously (i.e., AB, BC, DE, etc. were all presented in a mixed training block from the beginning). It is possible that it was the differential training structure that was responsible for the elimination of nodal number effects in Imam's (2001) Experiment 2.

If nodal effects are indeed a by-product of differential reinforcement, they should disappear when reinforcement is equal across baseline trial types. In the present research, the aim was to systematically manipulate training structure and, thereby, reinforcement in order to examine whether nodal

effects would indeed disappear under equal reinforcement. A group design was employed (e.g., Fields et al., 1997) in Experiments 1 and 2. In the *unequal reinforcement* condition, each baseline trial type was introduced serially in training, whereas in the *equal reinforcement* condition, all baseline trial types were introduced simultaneously. The equal reinforcement without LH condition was replicated in Experiment 3 with an increased baseline training criterion. Nodal number was measured as a function of response accuracy (RA) and RS and in a response transfer test (Experiments 2 and 3 only) across both two 5-member (Experiment 1) and two 6-member (Experiments 2 and 3) equivalence classes.

In Experiment 1, an LH was employed, in order to see whether varying reinforcement and training structure might influence accuracy and speed of responding as a function of nodal number with a time contingency in effect. In order to avoid ceiling effects in RS, Imam (2001, 2003, 2006) employed an LH; however, with this in place, a time–accuracy trade-off was observed, and thereby, modal effects may have been masked (e.g., Dickins, 2005; Imam, 2001, 2003, 2006). If nodal effects were apparent under no LH but disappeared or were weakened under an LH, this might explain previous research that failed to report a nodal number effect when an LH was in place. If nodal number effects are indeed a function of a particular training structure (and consequently, a differential reinforcement history), RA, RS (Experiments 1–3), and transfer of function (Experiments 2 and 3) should be observed only in the unequal reinforcement condition.

Experiment 1

Method

Participants Twenty-one adults participated in Experiment 1 (11 male; 10 female), ranging in age from 18 to 62 years (mean age = 27.26 years, $SD = 12.05$ years). Nineteen participants were students (11 undergraduates, 8 postgraduates) at Swansea University; 1 participant was retired, and the remaining participant was a human resources officer. Participants were recruited through e-mail and word of mouth from the second author, and all were naïve as to the purpose of the experiment. In return for their participation, participants earned £3 (approximately US \$5) for each session, which was not contingent upon performance. The experiment was approved by the Department of Psychology, Swansea University, Ethics Committee.

Apparatus, setting, and stimuli The experiment was conducted in a quiet room, containing only a desk, a chair, and a personal computer with a 550-MHz processor, a 14-in. color

monitor, and a standard computer mouse. Each participant sat at a table facing the computer monitor and keyboard. The computer controlled all trial presentations and trial and phase order and recorded all responses and RSs.

The stimuli were obtained from Massaro, Venezky, and Taylor (1979) and were letter permutations derived from the most frequent 150 six-letter English words as listed in Kučera and Francis (1967). These pseudoword stimuli met the following criteria with reference to the English language: (1) They were orthographically regular; (2) they were pronounceable; (3) they contained common vowel and consonant spellings; and (4) they had no more than three letters for a medial consonant cluster, if one occurred (see Table 1). The assignment of stimuli was randomized across participants. The stimuli were in black Times New Roman font, set against a white background.

Procedure A short questionnaire was administered to record participants' age, gender, occupation, and previous knowledge of the research topic, and each participant was also given a consent form to read and sign before beginning the experiment. All the participants were exposed individually to the experimental procedure across four sessions (defined below), irrespective of performance on the experimental task. These sessions were scheduled over a 1-week period, and each lasted between 20 and 35 min.

Each participant was randomly assigned to one of the four conditions, which differed in terms of training structure and presence or absence of an LH and are described using the following nomenclature: unequal reinforcement no LH, unequal reinforcement LH, equal reinforcement no LH, and equal reinforcement LH.

At the start of the experiment, the following instructions appeared on the computer monitor:

In a moment a word will briefly appear in the middle of the screen. It will disappear and two other words will appear. Choose 1 of the 2 words in the corner of the screen by pressing the Z key for the left word and the M key for the right word. During some stages of the experiment, the computer will NOT tell you if your choices are correct or wrong. However, based on what you have learned so far, you can get all of the

Table 1 Stimuli employed in Experiments 1, 2, and 3. Assignment of stimuli was randomized across participants

boceem	vartle	rettes	drager
lewoly	lorald	rigund	surtle
matser	betret	copher	casors
ronkeb	samolt	desund	cachen
troper	haveen	murben	Wollef
gedeer	siflet		

tasks correct. Please do your best to get everything right. Thank you and good luck.

For the two experimental conditions that included an LH (participants had to respond within 2.5 s), instructions also included the phrase “It is important that you respond as quickly as possible!”

Each trial started with a 1.7-s presentation of a sample stimulus at the center of the screen, which disappeared and was replaced by the two comparison stimuli that appeared after a 1-s interval. Participants pressed the “Z” or “M” key on the computer keyboard to select the comparison on the left or the right, respectively. When feedback was provided, choosing the correct comparison produced a 1-s display of the word “Correct.” Choosing the incorrect comparison produced a 1-s display of the word “Wrong.” Both were displayed in brown in the middle of the computer screen and were followed by a 1.5-s intertrial interval (ITI), during which the screen was blank. In the LH conditions, if participants failed to choose one of the comparison stimuli within 2.5 s, the phrase “Timed Out” appeared in brown at the top of the computer screen.

Two 5-member equivalence classes were established by training AB, BC, CD, and DE relations (i.e., a linear structure). All trial types were presented randomly within a block. All types of training block were followed by informative feedback on the participant’s choice of comparison. In the unequal reinforcement no LH and unequal reinforcement LH conditions, the AB trials were first trained to the mastery criterion of eight consecutive correct responses. Next, mixed AB and BC trials were presented until the same mastery criterion was reached, whereupon a new trial type was introduced, and so on until all trial types were presented in a mixed block. Eight consecutive correct responses were required on the final mixed block to proceed to the test phase. In the equal reinforcement no LH and equal reinforcement LH conditions, the two equivalence classes were established by training AB, BC, CD, and DE relations on a simultaneous basis, so that each relation was presented the same number of times. That is, these trial types were presented in a random manner in a mixed block from the beginning of the experiment. Eight consecutive correct responses were required to proceed to the test phase.

Once the criterion for the training session had been met, the test phase commenced without warning, and the corrective feedback (i.e., not including the “Timed Out” feedback) terminated. All baseline conditional relations, tests for symmetry, and one-, two-, and three-node trials were presented in a single randomized block. Each type of relation was presented the same number of times, with 40 trials in total (see Table 2).

A *cycle* was defined as training all relations to criterion and testing all possible derived relations. Participants were

exposed to two cycles in each session across a total of four sessions in a 1-week period to examine the delayed emergence of equivalence classes, regardless of their performance.

Data analysis The reaction time data were expressed as RS (inverted latency), since this minimized variance due to long latencies, which are more likely to be due to processes other than those of interest (e.g., due to inattention; Whelan, 2008). Due to the large within-participants variability of RSs, group means were presented. RS was analyzed from the first test block until the block on which participants reached above 90% accuracy. The rationale for excluding test blocks after participants had reached above 90% accuracy was to control for the emergence of a ceiling effect. According to Fields, Landon-Jimenez, Buffington, and Adams (1995), “the effects of nodality on test performance in all of the demonstrations were transient; they disappeared with the repeated presentation of trials in a given test” (p. 130).

A nodal number effect was deemed to have occurred on a particular test block when RA was highest on one-node trial types, with a lower RA on two-node trials and lowest RA on three-node trials. In order to test for a nodal number effect, a *Page’s trend test* (L) was conducted on each participant’s data. Page’s (1963) test is considered more powerful than Friedman’s test when the predicted order has a specific sequence (i.e., in the present experiment, RS is fastest in one-node relations, slower in two-node relations, and slowest in three-node relations). Page’s test is particularly suitable for small sample sizes. Test blocks on which RA was at 100% for two or more nodal numbers were not included, because the ceiling effect precluded an analysis.

Results

Twenty-one participants began [Experiment 1](#), 5 participants in unequal reinforcement no LH (4 passers), 4 participants in unequal reinforcement LH (4 passers), 8 participants in equal reinforcement no LH (4 passers), and 4 participants in equal reinforcement LH (3 passers). The data from the participants who left the experiment before reaching 90% correct on the equivalence test are not discussed here (interested readers can obtain these data by contacting the corresponding author). [Appendix 1](#) shows the mean numbers of cycles needed in each condition to reach the 90% criterion in the test phase (this will be revisited later in the [Discussion](#) section).

The mean numbers of reinforcers delivered across all participants in each condition in [Experiment 1](#) are presented in [Fig. 1](#). In the unequal LH and no-LH conditions, the highest number of reinforcers were delivered during AB trials, the next highest amount during BC trials and then CD trials, and the lowest number of reinforcers during DE trials. In contrast, the mean numbers of delivered rein-

Table 2 Trial types per relation type that were presented in Experiment 1

Relation Type	Trial Type					
Directly trained	A1B1	B1C1	C1D1	D1E1		
	A2B2	B2C2	C2D2	D2E2		
Symmetry	B1A1	C1B1	D1C1	E1D1		
	B2A2	C2B2	D2C2	E2D2		
One node	A1C1	B1D1	C1E1	C1A1	D1B1	E1C1
	A2C2	B2D2	C2E2	C2A2	D2B2	E2C2
Two nodes	A1D1	B1E1	D1A1	B1E1		
	A2D2	B2E2	D2A2	B2E2		
Three nodes	A1E1	E1A1				
	A2E2	E2A2				

forcers were kept approximately equal across all trial types for all participants in the equal reinforcement LH and no-LH conditions. These data indicate that the procedures employed were successful in manipulating the number of reinforcers across trial types in all the conditions. The equal reinforcement LH group had 26 time-out responses. The unequal reinforcement LH group had 55 time-out responses.

Figure 2 displays the mean RSs, with one standard error across the four groups. A series of Page's trend tests were performed to examine the nodal effect across all participants in each of the four conditions. A significant nodal effect was found in the equal reinforcement no-LH group only ($L = 54, p = .05$). There was little difference among the RSs for one-, two-, and three-node relations in the unequal reinforcement LH, equal reinforcement LH, and equal reinforcement no-LH groups.

Response accuracy Mean percentages of RA, with one standard error, across all participants in each condition are presented in Fig. 3. A series of Page's trend tests were calculated in each condition. No significant nodal effect was found in each condition; however, in the equal no-LH condition, it approached significance ($L = 53$; a significant value at the .05 level should be 54).

Discussion

In Experiment 1, the training structure—and hence, the number of reinforcers per trial type—was manipulated. In the equal reinforcement group, all trial types were introduced simultaneously, and the number of delivered reinforcers was approximately equal across all baseline trial types. In contrast, in the unequal reinforcement group, the trial types

were introduced serially, and the number of delivered reinforcers was different across trial types (see Fig. 1). The number of participants who eventually reached 90% accuracy for derived relations was higher in the unequal reinforcement group than in the equal reinforcement group. The number of cycles needed to reach the mastery criterion for class-consistent responding is displayed in Appendix 1.

The key result in Experiment 1 was that a significant nodal number effect was found in RS for the equal reinforcement no-LH group. This suggests that a serial training structure—resulting in unequal reinforcement—is *not* a prerequisite for nodal effects to emerge (cf. Imam, 2001, Experiment 2). Results obtained from the analysis of the RA data indicated that the nodal effect did not disappear under the equal reinforcement no-LH condition. In addition, the nodal effect was found in the no-LH condition, but not in the LH condition, in both measures of RS and RA. This finding provided further evidence that a time–accuracy trade-off occurred, thus possibly obscuring any nodal number effects (see Dickins, 2005) from emerging in the LH groups. In Experiment 1, the start of the test phase was not signaled. Appendix 2 shows the mean percentages of correct responses on unreinforced baseline probes. These data suggest that the baseline conditional discriminations were disrupted by the sudden termination of reinforcement, which resulted in erratic performance in initial test blocks and a gradual increase in accuracy in later blocks.

Experiment 2

The results of Experiment 1 provided evidence in support of the prediction that nodal number effects are preserved under equal reinforcement and a simultaneous training in

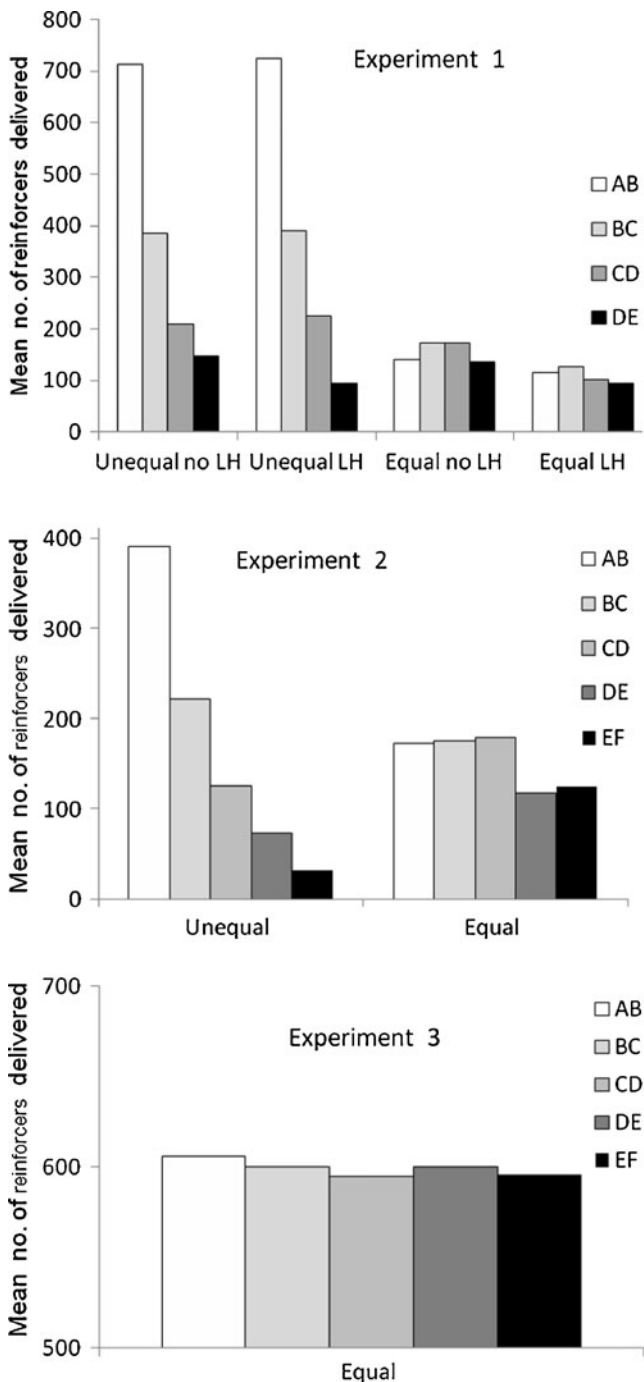


Fig. 1 Mean number of reinforcers delivered for all baseline relations across all participants in each condition in all three experiments

terms of RS. However, the results were less convincing in terms of RA. One would argue that if the nodal effect is genuine, it should not disappear after repeated testing. The absence of nodal effects under repeated tests appears to be temporary; that is, it reemerged after the introduction of a different test paradigm, known as the *transfer paradigm* (e.g., Fields, Adams, et al., 1993; Fields et al., 1995), in

which a function is trained to a particular stimulus or stimuli in a class and the degree of transfer to other members is measured. According to Fields, Adams et al. (1993), "... if the degree of transfer was a systematic function of a variable such as nodal distance ... that variable would account for the relatedness of the stimuli in the class" (p. 86). Experiments 2 and 3 included a transfer paradigm.

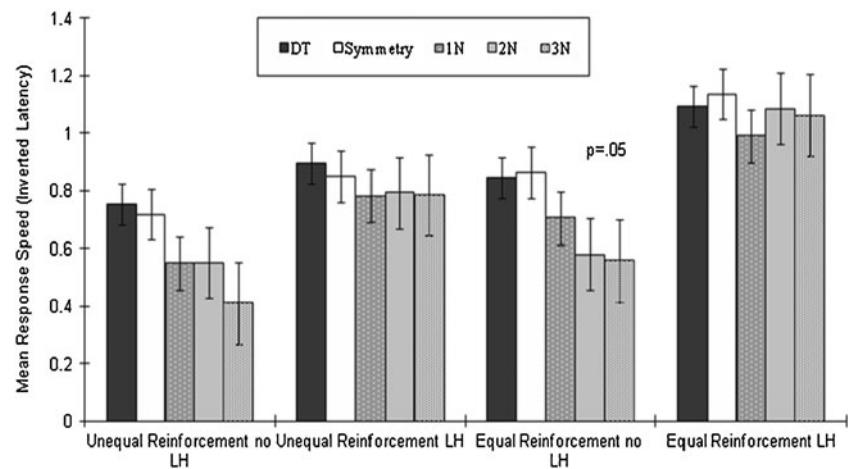
Experiment 2 employed the same manipulation of reinforcement and training structures as in Experiment 1. In addition to this, a transfer-of-function test was included after the test phase, and class size was expanded to 6 members. In order to test for the transfer of function, the introduction of an increased number of class members was necessary. Time-accuracy trade-off is a common problem (e.g., Imam, 2001, Experiment 1) that occurs under the LH condition, although the present data do not explicitly show signs of a time-accuracy trade-off; the erratic performance might be a result of insufficient training on baseline trials. Therefore, in order to control for this unwanted factor, an LH was not employed in Experiment 2, since it would likely mask the nodal number effect on RA or transfer-of-function test performance. In the transfer-of-function training phase, differential responses were trained, using corrective feedback, to the C and D stimuli in each class. Next, the A, B, E, and F stimuli were presented in the absence of corrective feedback, and the number of responses to each was observed.

If the notion that equal reinforcement will eliminate a nodal effect is correct, responses to A, B, E, and F should be distributed equally following the equal reinforcement training. In contrast, if the nodal account is correct, the A and B stimuli should evoke the response trained to the C stimulus, and the E and F stimuli should evoke the response trained to the D stimulus, despite the differential level of reinforcement. In addition, if unequal reinforcement is indeed a confounding variable, then, following unequal reinforcement training, responses trained to the C stimulus should transfer to the A and B stimuli more readily than responses to the D stimulus transfer to the E and F stimuli. That is, B–C and A–B relations should be more strongly established, whereas D–E and E–F relations should be weak.

Method

Participants Eight participants began Experiment 2 (5 male, 3 female), ranging in age from 21 to 29 years (mean age = 24 years, $SD = 2.4$ years). All the participants were students (1 undergraduate, 7 postgraduates) at Swansea University. Participants were recruited through personal contacts by the first author (2 British nationals and 6 Chinese nationals). All were naïve as to the purpose of the experiment. The experiment was approved by the Department of Psychology, Swansea University, Ethics Committee.

Fig. 2 Mean response speed (Inverted latency), with one standard error, across all relation types for all participants in Experiment 1. DT, directly trained; 1N, one node; 2N, two nodes; 3N, three nodes



Apparatus, setting, and stimuli The apparatus and setting were identical to those employed in Experiment 1. Two 6-member equivalence classes were trained to accommodate the transfer paradigm.

Procedure A short questionnaire was administered to record participants' age, gender, occupation, and previous knowledge of the research topic, and each participant was also given a consent form to read and sign before beginning the experiment. All the participants were exposed to the experimental procedure individually across a number of sessions, each lasting between 35 and 45 min and scheduled across 2 days.

Each participant was randomly assigned to one of two conditions that differed in terms of level of reinforcement (equal vs. unequal). The procedure was broadly similar to that in Experiment 1, with the following exceptions. In the unequal reinforcement condition, two 6-member equivalence classes were established by training AB, BC, CD, DE, and EF relations in a serial manner. The criterion to

proceed to the next training phase, or to the test phase, was 10 consecutive correct responses. In the equal reinforcement condition, the two equivalence classes were established by training AB, BC, CD, DE, and EF relations on a simultaneous basis, so that each relation was presented the same number of times. The criterion to proceed to the test phase was 10 consecutive correct responses.

Once the criterion for the training had been met, the test phase commenced, and the corrective feedback terminated. All baseline conditional relations, tests for symmetry, and one-, two-, three- and four-node trials were presented in a single block. The mastery criterion for testing was at least 90% class-consistent selection across the block of 60 test trials (see Table 3). The criterion for progressing to the function-training phase was originally defined as two consecutive correct test blocks. However, for participants 17, 18, 20, 22, and 23, the criterion was inadvertently set at three consecutive correct test blocks. Upon reaching the criterion across either two or three test blocks, participants were immediately exposed to the function training.

Function training and response transfer testing were conducted entirely by means of the computer and began with the presentation of the following instructions on the computer monitor (adapted from Fields et al., 1995):

In this phase, each spacebar press will produce a brick on the screen. Look at the word at the top of the screen. Your task is to learn how many bricks you should build, either 3, 5, 7 or 9 depending on what word is displayed at the top. Press the "Finish" button when you want to complete a trial. You may start a trial again, if you wish, by pressing the "Start Again" button. Sometimes you will receive feedback and sometimes you will not. Please try your best on all tasks.

The instructions cleared when a button with the caption "Press to start," which was underneath the statement, was

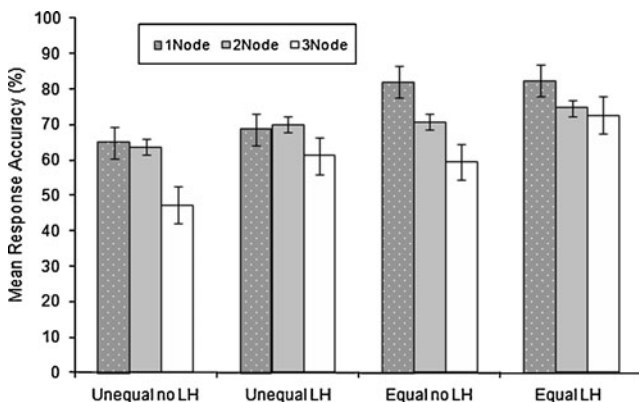


Fig. 3 Mean response accuracy, with one standard error, across all participants in each condition in Experiment 1. Unequal no LH, unequal reinforcement no limited hold; Unequal LH, unequal reinforcement limited hold; Equal no LH, equal reinforcement no limited hold; Equal LH, equal reinforcement limited hold

Table 3 Trial types per relation type that were presented in Experiments 2 and 3

Relation Type	Trial Type								
Directly trained	A1B1	B1C1	C1D1	D1E1	E1F1				
	A2B2	B2C2	C2D2	D2E2	E2F2				
Symmetry	B1A1	C1B1	D1C1	E1D1	F1E1				
	B2A2	C2B2	D2C2	E2D2	F2E2				
One node	A1C1	B1D1	C1E1	D1F1	F1D1	C1A1	D1B1	E1C1	
	A2C2	B2D2	C2E2	D2F2	F2D2	C2A2	D2B2	E2C2	
Two nodes	A1D1	B1E1	C1F1	D1A1	B1E1	F1C1			
	A2D2	B2E2	C2F2	D2A2	B2E2	F2C2			
Three nodes	A1E1	B1F1	E1A1	F1B1					
	A2E2	B2F2	E2A2	F2B2					
Four nodes	A1F1	F1A1							
	A2F2	F2A2							

pressed. In the training phase, 2 members from each equivalence class functioned as discriminative stimuli (S^D s). The stimuli were identical to those in the equivalence training and testing phases. Each S^D was presented on the top center of the screen of the monitor against a white background. Pressing the space bar produced a picture of a brick, which appeared at the center bottom of the screen. Each brick was a dark red rectangle (1 cm in width and 5 cm in length). Clicking the red “Start Again” button at the left bottom of the screen made all bricks on the screen disappear and set the response counter to zero. Clicking the green “Finish” button at the right bottom of the screen produced corrective feedback (“Correct” or “Wrong”; identical to the equivalence training phase), followed by a 1.5-s ITI (a blank screen) during the training stage or only the 1.5-s ITI during the test phase (see Fig. 4). The objective was to create three, five, seven, or nine bricks, depending on the stimulus displayed at the top of the screen. If a participant made more than 12 responses, the bricks disappeared and began forming again at the bottom of the screen.

The following responses were reinforced during the discrimination training: producing three bricks in the presence of the C1 stimulus, five bricks in the presence of the C2 stimulus, seven bricks in the presence of the D1 stimulus, and nine bricks in the presence of the D2 stimulus. Feedback was presented on all trials until the training criterion—eight consecutive accurate responses—was reached, whereupon feedback was stopped without warning. A 72-trial test block was then presented in which each of the six stimuli from each equivalence class was presented six times.

Following the first function transfer phase (i.e., training and testing), each participant was reexposed to equivalence training to criterion. The equivalence test phase was then presented, and upon passing this test, participants were reexposed to transfer-of-function training and testing. The experiment was concluded following this second function transfer test. Each participant was then thanked for participating and was debriefed.

Results

Eight participants began Experiment 2; 4 of them passed the unequal reinforcement condition, and the other 4 passed

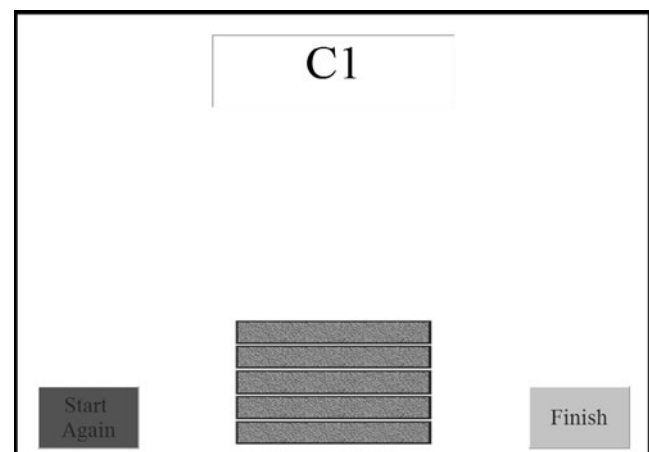


Fig. 4 A screenshot of the function training screen in Experiments 2 and 3

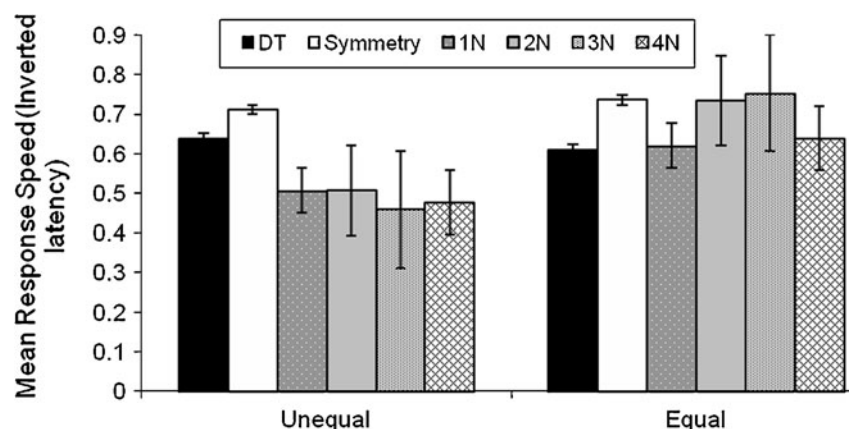
the equal reinforcement condition. Appendix 1 shows the mean numbers of cycles needed across all participants in each condition to reach the mastery criterion in the test phase. The mean number of reinforcers delivered, across all participants in Experiment 2 across each condition, are presented in Fig. 1. These data indicate that the procedures employed were successful in manipulating the number of reinforcers across trial types in both conditions, with the exception of DE and EF trials in the equal reinforcement group. Further analysis using a Wilcoxon's rank test suggested that the difference between DE, EF trials and AB, BC, CD trials in the equal reinforcement group was not significant ($p > .05$).

The data in Appendix 2 displays the mean percentages of correct responses on unreinforced baseline probes for all the participants in each test cycle. Participants from the equal reinforcement group demonstrated much more disrupted baseline conditional discriminations than did the unequal reinforcement group immediately after the first training block.

Response speed The same data treatment was employed as in Experiment 1. Figure 5 depicts mean RSs, with one standard error, across all relation types for all participants in each condition for Experiment 2. No significant nodal effect was found in each condition.

Response accuracy A nodal number effect was deemed to have occurred on a particular test block when RA was highest on one-node trial types, lower on two-node trials, lower again on three-node trials, and lowest on four-node trials. Figure 6 depicts the mean correct responses for node-related trial types, with one standard error, across all participants in each condition in Experiment 2. The Page's trend test was calculated for the mean RAs in each condition. No significant nodal effect was found; however, the trend approached significance in the equal reinforcement condition ($L = 109$; a significant value at the .05 level should be 111).

Fig. 5 Mean response speed (inverted latency), with one standard error, across all relation types for all participants in each condition in Experiment 2. DT, directly trained; 1N, one node; 2N, two nodes; 3N, three nodes; 4N, four nodes



Response transfer Response transfer was assessed by measuring the relative frequency with which the responses trained to the C and D stimuli were evoked by the other experimental stimuli. For example, the C1 stimulus was an S^D for three space bar presses; if three space bar presses were evoked during all six presentations of the A1 stimulus, the relative frequency was 100%. In some cases, the sum of responses trained to the A, B, E, or F stimulus in a class did not equal 100%. This occurred because some responses other than those trained to the C and D stimuli were evoked (e.g., across-class errors). If the response trained to a C or a D stimulus was evoked by other members of the same class, the response function was deemed to have transferred in accordance with an equivalence relation.

More important, if the response trained to C was evoked by the A and B stimuli, and not by the E and F stimuli, within the same equivalence class, those responses were deemed to be also under the control of nodal number (similarly, if the response trained to D was evoked by the E and F stimuli, and not by the A and B stimuli). A series of Fisher's exact tests with one value for each participant (i.e., yes/no) (Siegel & Castellan, 1988) were performed on C and D stimuli in both the serial and simultaneous conditions to further examine the differences between the proportion of responses for the C and D stimuli. Fisher's exact test is a powerful one-tailed test used to test the exact probability from a specific distribution (in our case, how different the proportion of responses on AB is to EF when C/D is in control). It is considered more appropriate than chi-square in small-sample experiments (Siegel & Castellan, 1988, p. 103). However, Fisher's exact test does not produce a critical value; therefore, only p values are presented.

If the proportion of responses that transferred to the A and B stimuli was not the same as the proportion of those that transferred to the E and F stimuli, the null hypothesis would be rejected. In general, presentation of C and D stimuli in each class almost always occasioned the trained responses; thus, discriminative control by the C and D stimuli was maintained in the absence of explicit reinforcement.

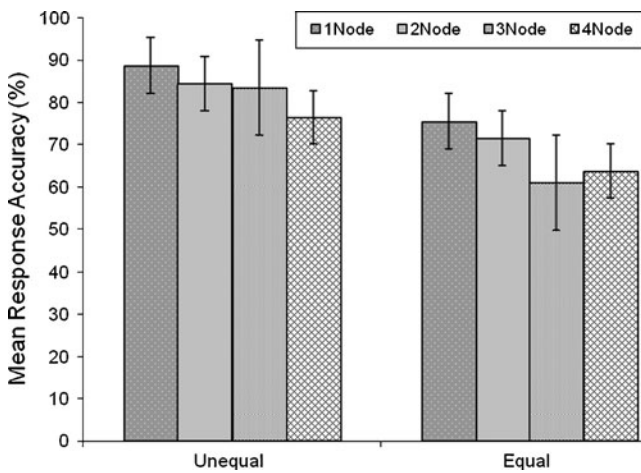


Fig. 6 Mean response accuracy for node-related trial types, with one standard error, across all participants in each condition in [Experiment 2](#)

The proportion of transferred responses was analyzed on the basis of condition. Figure 7 depicts mean relative frequency with which responses trained to C and D were evoked by stimuli in both classes across all participants in the equal and unequal reinforcement conditions in [Experiment 2](#). In the equal reinforcement condition, participants' proportion of transferred responses was significantly greater on A2 and B2 stimuli when trials were controlled by the C2 stimulus ($p = .031$, Fisher's exact) and was significantly greater on E2 and F2 stimuli when trials were controlled by D2 stimulus ($p = .031$, Fisher's exact). The proportion of transferred responses was also greater on A1 and B1 stimuli when the C1 stimulus was in control and greater on E1 and F1 stimuli when the D1 stimulus was in control; however, this trend did not reach statistical significance. In the unequal reinforcement condition, participants' proportion of responses that transferred to the A1 and B1 stimuli was significantly greater than those for the E1 and F1 stimuli when trials were controlled by the C1 stimulus ($p = .031$, Fisher's exact) and was significantly greater for E1 and F1 stimuli when trials were controlled by the D1 stimulus ($p = .016$, Fisher's exact). The proportion of transferred responses was also greater on A2 and B2 stimuli when C2 was in control and greater on E2 and F2 stimuli when the D2 stimulus was in control; however, this was not strong enough to reach statistical significance. In summary, Fisher's exact tests suggested that response transfer was a function of nodal number for the members of one of the equivalence classes (e.g., classes A1, B1, C1, D1, E1, and F1), but not the other class in both the equal and unequal reinforcement conditions.

Discussion

The results from [Experiment 2](#) suggested that level of reinforcement and training structure had no effect on the

number of cycles participants needed to reach the mastery criterion on the simple discrimination training. The response patterns in the unequal reinforcement condition did not suggest that differential reinforcement and a serial training structure resulted in differential response transfer. Transfer from the D1 stimulus to the E1 and F1 stimuli was as robust as the transfer from the C1 stimulus to the A1 and B1 stimuli. In the equal reinforcement condition, the untrained responses transferred from the C2 and D2 stimuli were under the control of nodal number, showing that unequal reinforcement was not a prerequisite for nodal effects. In summary, delivering an approximately equal number of reinforcers for each trial type and presenting all baseline trials in a simultaneous manner did not appear to have an effect on the probability of successful transfer of function for any particular relation type. Moreover, analysis of accuracy data confirmed the findings in [Experiment 1](#); that is, a nodal number effect emerged in the equal reinforcement condition only.

Experiment 3

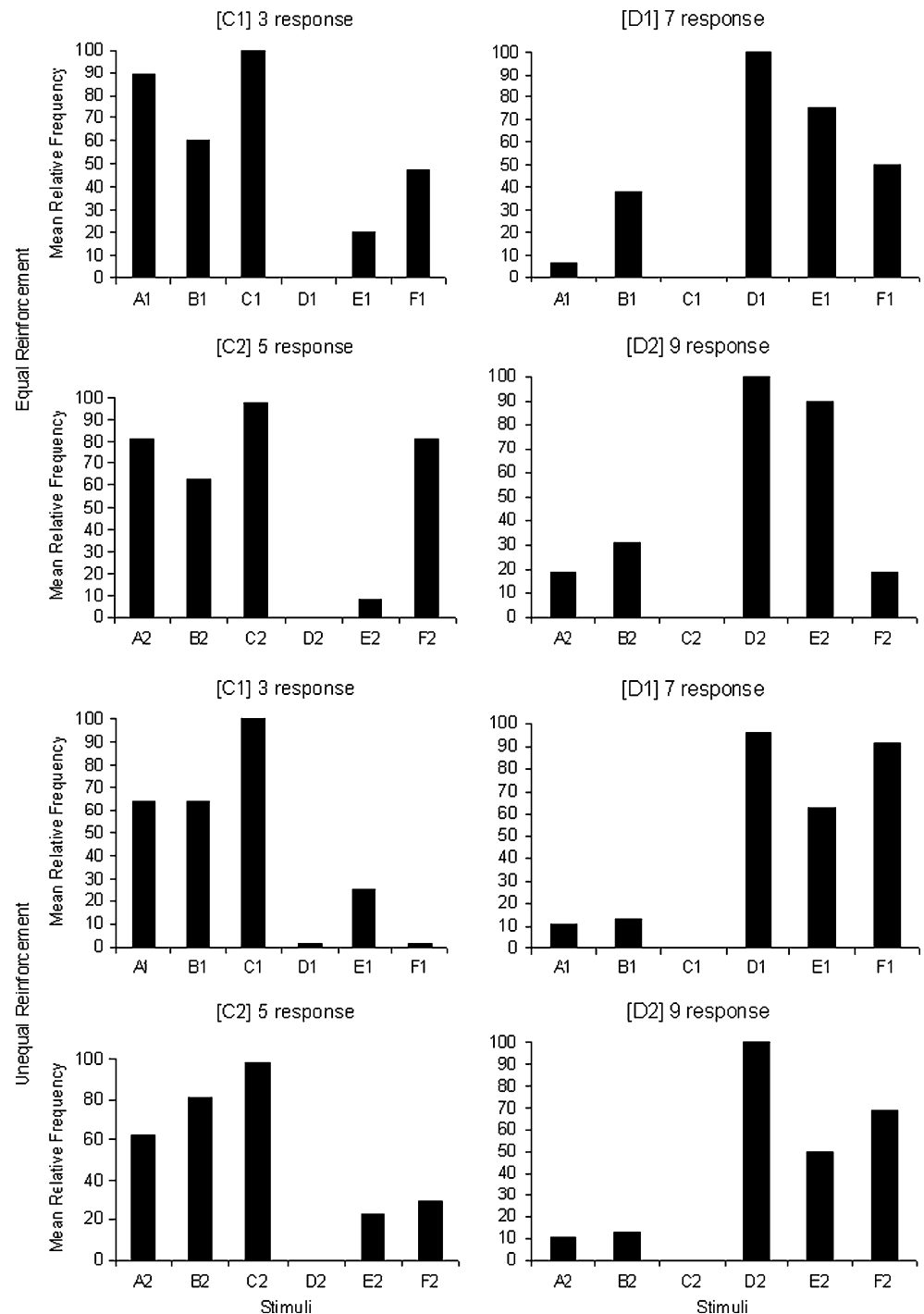
[Appendix 2 \(Experiment 2\)](#) showed a disruption of baseline relations in the first test block in the equal reinforcement condition, which might explain the lack of an emergent nodal effect in RS. [Experiment 3](#) was similar to [Experiment 2](#), with the following exceptions. The findings from [Experiment 2](#) indicated that differential reinforcement did not result in differential response transfer; therefore, unequal reinforcement and a serial training structure was not employed in [Experiment 3](#). In addition, the criterion of 10 consecutive correct responses before proceeding to the test phase was expanded to 20 consecutive correct responses in order to stabilize performance. Moreover, the number of participants was increased from 4 (e.g., [Experiments 1 and 2](#)) to 8 to determine whether a low N might have been responsible for the nonsignificant trends that emerged in [Experiment 2](#).

Method

Participants Fourteen participants began [Experiment 3](#), (8 male; 6 female), ranging in age from 20 to 27 years (mean age = 23.14 years, $SD = 2.07$ years). All the participants were students (5 undergraduates, 9 postgraduates) at Swansea University and Swansea Institute. Participants were recruited through personal contacts by the first author (all Chinese nationals). All the participants were naïve as to the purpose of the experiment. The experiment was approved by the Department of Psychology, Swansea University, Ethics Committee.

Apparatus, setting, and stimuli The apparatus, setting, and stimuli were exactly the same as those in [Experiment 2](#).

Fig. 7 Mean relative frequency with which responses trained to C and D were evoked by stimuli in both classes across all participants in equal and unequal reinforcement conditions in [Experiment 2](#)



Procedure The procedure was the same as that in the equal reinforcement condition in [Experiment 2](#), with the exception that 20 consecutive correct responses were required in the baseline training phase in order to stabilize performance.

Results

Fourteen participants began [Experiment 3](#). Eight participants formed equivalence after repeated exposure to the

training and testing cycles. The data from the participants who did not reach 90% accuracy on the equivalence test are not discussed here (interested readers can obtain these data by contacting the corresponding author). [Appendix 1](#) shows the mean number of cycles needed for all participants to reach the 90% mastery criterion in the test block.

The mean numbers of reinforcers delivered, across all participants in [Experiment 3](#), are presented in [Fig. 1](#). These data indicate that the procedures employed were generally

successful in manipulating the number of reinforcers across trial types.

Appendix 2 shows the mean percentages of correct responses on unreinforced baseline probes for all the participants in each test block. From these data, it can be seen that baseline conditional discriminations were kept intact after the first training session by increasing the pass criterion from 10 to 20 consecutive correct trials in the training block.

Response speed Mean RSs were significantly different in terms of nodal numbers ($L = 216, p < .05$). As is shown in Fig. 8, RS was inversely related to nodal number; that is, as node number increased, RS decreased.

Response accuracy Figure 9 depicts mean correct responses, with one standard error, across all participants in Experiment 3. The Page's trend test indicated that the nodal effect approached significance ($L = 210$; a significant value at .05 level should be 214).

Transfer tests Response transfer was assessed by measuring the relative frequency with which the responses trained to the C and D stimuli were evoked by the other experimental stimuli. Mean relative frequency of responses for all the participants after equivalence formation is shown in Fig. 10. A series of Fisher's exact tests was performed. The proportions of participants' responses that transferred to the A1 and B1 stimuli were significantly greater than the proportions of those that transferred to the E1 and F1 stimuli when trials were controlled by the C1 stimulus ($p = .035$, Fisher's exact) and were significantly greater than those for the E1 and F1 stimuli when trials were controlled by the D1 stimulus ($p = .013$, Fisher's exact). No significance was found for class 2 stimuli.

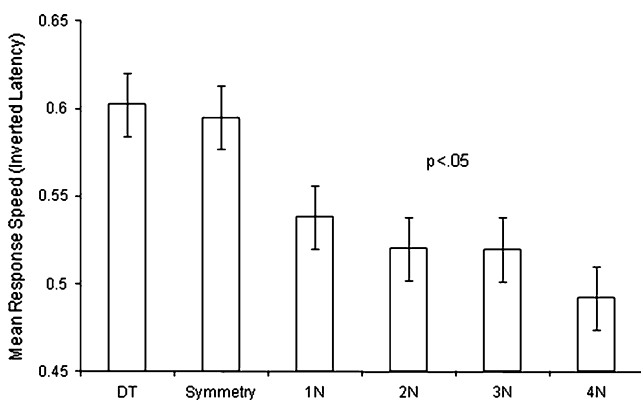


Fig. 8 Mean response speed (inverted latency), with one standard error, across all relation types for all participants in Experiment 3. DT, directly trained; 1N, one node; 2N, two nodes; 3N, three nodes; 4N, four nodes

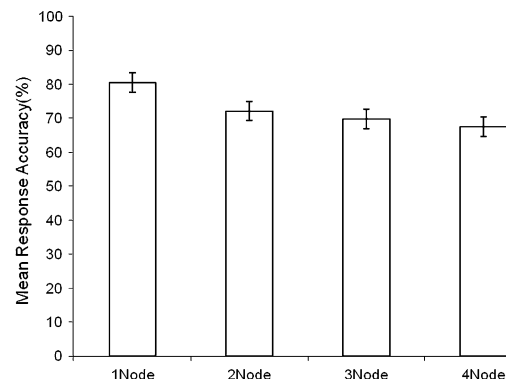


Fig. 9 Mean response accuracy in each node-related trial type, with one standard error, across all participants in Experiment 3

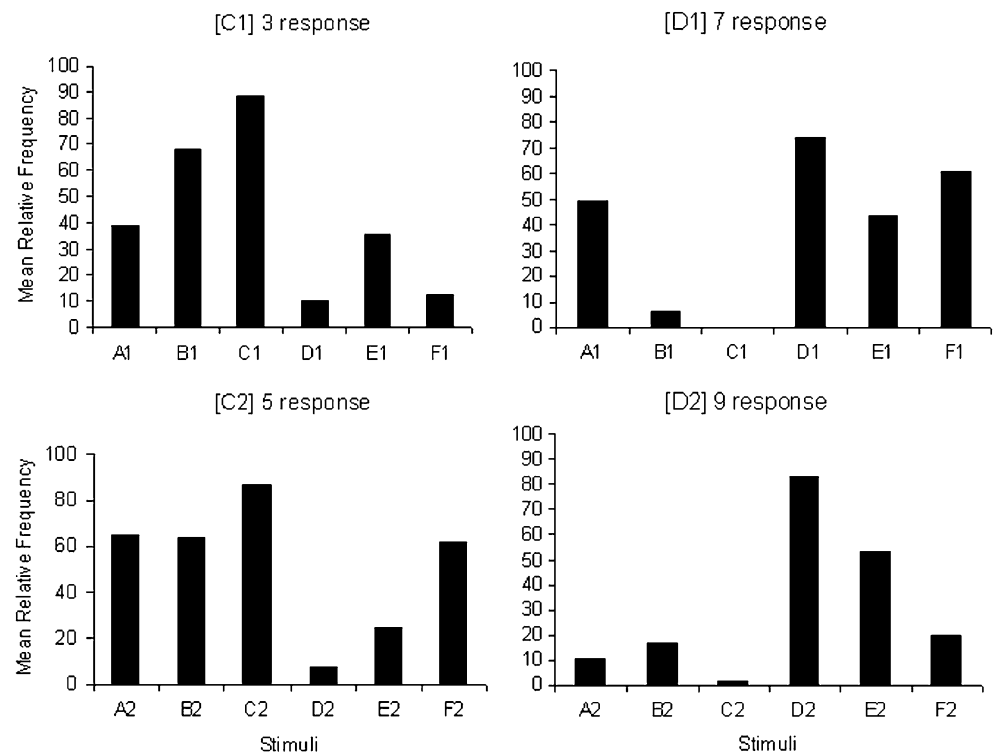
Discussion

Experiment 3 replicated the findings from the equal reinforcement groups in Experiments 1 and 2, suggesting that nodal effects remain intact under equal reinforcement and simultaneous training in measure of RSs (Experiments 1 and 3) and transfer of function (Experiments 2 and 3). The emergence of nodal effects from the analysis of RSs in Experiment 3, but not Experiment 2, suggests that performance stabilized after establishing baseline relations. However, RA seems less affected by the experimental manipulations. Even though no significant nodal effect on RA was found, the trend across all three experiments approached significance in the equal reinforcement condition, which is consistent with the RS data.

General discussion

The aim of the present study was to test the possibility that the nodal number effects reported in previous stimulus equivalence studies (e.g., Fields et al., 1997) were a function of unequal reinforcement across trial types during baseline conditional discriminations. The present study sought to provide a systematic analysis of two types of training (i.e., serial and simultaneous), which differed in terms of reinforcement delivered for particular trial types. The data in the present study suggested that nodal number was a predictor of RS even when reinforcement for responding to baseline trial types during conditional discrimination training was equalized. Moreover, the manipulation of reinforcement and training structure had little effect on the emergence of nodal effects, which is contrary to Imam's (2001, 2003, 2006) previous studies. In Experiments 2 and 3, a transfer-of-function test was employed as an additional measure to examine nodal effects. Again, nodal effects were observed in the equal reinforcement condition. Furthermore, a serial training structure that resulted

Fig. 10 Mean relative frequency with which responses trained to C and D were evoked by stimuli in both classes across all participants in Experiment 3



in unequal reinforcement did not appear to influence nodal number. That is, less reinforcement to particular trial types did not result in poorer transfer.

The findings of the present study are in accord with those in previous research by Fields and his colleagues (Fields et al., 1993a, 1993b; Fields et al., 1995; Fields & Moss, 2007; Fields & Watanabe-Rose, 2008) and conflict with research suggesting that differential reinforcement during baseline training was responsible for nodal effects (Imam, 2001, 2006; Sidman, 1994). The present research was novel in its design and analysis. Furthermore, the number of participants who demonstrated equivalence under equal reinforcement in the present study ($n = 19$) was substantially greater than in previous studies (2 participants in Fields et al., 1995, and 4 participants in Fields & Watanabe-Rose, 2008). Indeed, according to Fields and Watanabe-Rose, “additional research will be needed to determine whether a larger segment of the population would also bifurcate class membership based on nodal structure” (p. 378). The present study addressed this issue.

The present study also highlights the importance of the test format when examining the relatedness of stimuli in equivalence classes. Three measures were employed in the present study: RS, RA, and transfer of function. RS was significantly different in Experiment 1 (in the equal reinforcement condition only) and in Experiment 3. RA was not significantly different in any condition but approached significance in the equal reinforcement conditions only across the three experiments, consistent with the RS findings. Finally, it was the transfer-of-function test,

however, that demonstrated the clearest evidence of nodal effects for the equal reinforcement group. Fields and Watanabe-Rose (2008) have speculated that the format of the MTS structure itself occasions responding in accordance with class membership and discrimination *between* classes. In contrast, the format of the transfer-of-function test is such that responses occur in the presence of members of the same class and, therefore, occasions responding according to *within*-class differences, such as nodal number. On the basis of the present data, it is recommended that future research into nodal effects should include a transfer-of-function test that occasions responding according to within-class differences.

There were some differences between the present study and the previous research that varied the training structure and quantity of reinforcement (Imam, 2001, 2003, 2006). Like many other equivalence studies (e.g., Fields et al., 1993a, 1993b; Barnes-Holmes et al., 2005), the present study employed two, rather than three, comparison stimuli. The effects of nodal number have been shown to diminish with the addition of a third stimulus group (Kennedy, 1991, Experiment 2), and thus the inclusion of a third stimulus may have made interpretation difficult if nodal effects had not been observed in the present study. Although studies have demonstrated nodal effects with various types of stimuli (e.g., pictures, letters, symbols), the present study employed pronounceable letter strings, whereas Imam (2001, 2003, 2006) used graphical stimuli. Is it possible, indeed likely, that relations among stimuli were rapidly learned because pronounceable, rather

than graphical, stimuli were employed (cf. Imam, 2001, 2003, 2006). However, the use of these stimuli did not seem to facilitate learning to such a degree that the differential reinforcement delivered between unequal and equal groups was attenuated; rather, the numbers of baseline training counts.

Imam (2001, 2003, 2006) and Spencer and Chase (1996) included only those trial types involving either the most trained conditional relations or the least trained conditional relations in their analyses. According to Spencer and Chase, this exclusion procedure was designed partly to account for imbalances in “the order and differential amount of training on each baseline conditional discrimination” (p. 649). However, the aim of the present study was to measure the influence of these factors, and therefore, it seemed counterintuitive to exclude these trial types. Saunders and Green (1999, p. 132) have also noted that including only the most and least trained stimuli in the analyses introduces some additional problems in interpretation.

One possibility is that both the present results and those of Imam are reliable yet can be traced to procedural differences, such as the test format. If this is the case, caution must be exercised when interpreting findings on nodal effects in the stimulus equivalence literature. Indeed, given the reliance in the stimulus equivalence research on RA or RS, the possibility that nodal effects are differentially influenced by various test formats is an empirical question in itself worthy of future research. There are other ways in which the present study could be extended. In the present study, the separate influence of training structure and reinforcement was not assessed. Perhaps future

researchers might attempt to manipulate reinforcement while keeping constant the order in which the trial types are introduced. The inclusion of a third stimulus may also be a topic for future research.

In conclusion, the data from the present study indicate that nodal number effects are preserved and maintained even when reinforcement is kept equal. This finding was robust and was demonstrated over a relatively large number of participants. These data also emphasize the role of test format in detecting nodal differences. That is, test formats such as MTS may occasion responding in accordance with class discrimination, whereas a transfer-of-function test involving members of the same equivalence class may occasion responding in accordance with nodal number.

Appendix 1

Table 4 Mean number of cycles needed until participants in each condition reached the mastery criterion of 90% and above accuracy in class consistent responding in Experiments 1, 2, and 3

Reinforcement	Unequal		Equal	
	No LH	LH	No LH	LH
Experiment 1	4.75	3.5	3.75	3
Experiment 2	3	N/A	2.5	N/A
Experiment 3	N/A	N/A	3.25	N/A

LH, limited hold

Appendix 2

Table 5 Mean percentages of correct responses per test block for unreinforced baseline probes in each condition in Experiments 1, 2, and 3

Experiments 1	Number of Test Block								
	1	2	3	4	5	6	7	8	
Unequal no LH	75	72	84	88	94	97	94	100	
Unequal LH	78	84	97	94	97	100	100	97	
Equal no LH	84	84	81	97	94	100	97	100	
Equal LH	54	92	92	96	100	100	96	100	
Experiment 2	1	2	3	4	5	6	7	8	9
Unequal no LH	95	100	100	100	90	100	100	100	100
Equal no LH	55	100	95	100	100	100	100		
Experiment 3	1	2	3	4	5	6	7		
Equal no LH	79	89	94	95	96	100	100		

LH, limited hold

References

- Adams, B. J., Fields, L., & Verhave, T. (1993). The effects of test order on equivalence class formation. *Psychological Record*, *43*, 133–152.
- Barnes-Holmes, D., Staunton, C., Whelan, R., Barnes-Holmes, Y., Commins, S., Walsh, D., et al. (2005). Derived stimulus relations, semantic priming, and event-related potentials: testing a behavioral theory of semantic networks. *Journal of the Experimental Analysis of Behavior*, *84*, 417–433.
- Bentall, R. P., Dickins, D. W., & Fox, S. R. A. (1993). Naming and equivalence: Response latencies for emergent relations. *The Quarterly Journal of Experimental Psychology*, *46B*, 187–214.
- Bentall, R. P., Jones, R. M., & Dickins, D. W. (1998). Errors and response latencies as a function of nodal distance in 5-member equivalence classes. *Psychological Record*, *48*, 93–115.
- Carrigan, P. F., & Sidman, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior*, *58*, 183–204.
- Dickins, D. W. (2005). On aims and methods in the neuroimaging of derived relations. *Journal of the Experimental Analysis of Behavior*, *84*, 453–483.
- Dymond, S., & Rehfeldt, R. A. (2001). Supplemental measures of derived stimulus relations. *Experimental Analysis of Human Behavior Bulletin*, *19*, 6–10.
- Fields, L., Adams, B. J., & Verhave, T. (1993a). The effects of equivalence class structure on test performances. *Psychological Record*, *43*, 697–713.
- Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1993b). Are stimuli in equivalence classes equally related to each other? *Psychological Record*, *43*, 85–105.
- Fields, L., Landon-Jimenez, D. V., Buffington, D. M., & Adams, B. J. (1995). Maintained nodal-distance effects in equivalence classes. *Journal of the Experimental Analysis of Behavior*, *64*, 129–145.
- Fields, L., & Moss, P. (2007). Stimulus relatedness in equivalence classes: Interaction of nodality and contingency. *European Journal of Behavior Analysis*, *8*, 141–159.
- Fields, L., Reeve, K. F., Rosen, D., Varelas, A., Adams, B. J., Belanich, J., et al. (1997). Using the simultaneous protocol to study equivalence class formation: The facilitating effects of nodal number and size of previously established equivalence classes. *Journal of the Experimental Analysis of Behavior*, *67*, 367–389.
- Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior*, *48*, 317–332.
- Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of the Experimental Analysis of Behavior*, *42*, 143–157.
- Fields, L., & Watanabe-Rose, M. (2008). Nodal structure and the partitioning of equivalence classes. *Journal of the Experimental Analysis of Behavior*, *89*, 359–381.
- Holth, P., & Arntzen, E. (2000). Reaction times and the emergence of class consistent responding: A case for precurrent responding? *Psychological Record*, *50*, 305–337.
- Imam, A. A. (2001). Speed contingencies, number of stimulus presentations, and the nodal number effect in equivalence class formation. *Journal of the Experimental Analysis of Behavior*, *76*, 265–288.
- Imam, A. A. (2003). Assessing transfer of response speed and nodal number via conditional discriminations. *Experimental Analysis of Human Behavior Bulletin*, *21*, 1–7.
- Imam, A. A. (2006). Experimental control of nodal number via equal presentations of conditional discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of the Experimental Analysis of Behavior*, *85*, 107–124.
- Kennedy, C. L. (1991). Equivalence class formation influenced by the number of nodes separating stimuli. *Behavioural Processes*, *24*, 219–245.
- Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Massaro, D. W., Venezky, R. L., & Taylor, G. A. (1979). Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General*, *108*, 107–124.
- Page, E. B. (1963). Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Journal of the American Statistical Association*, *58*, 216–230.
- 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*, 117–137.
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, *14*, 5–13.
- Sidman, M. (1994). *Stimulus equivalence: A research story*. Boston: Authors Cooperative.
- Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, *74*, 127–146.
- Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- Spencer, T. J., & Chase, P. N. (1996). Speed analyses of stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, *65*, 643–659.
- Whelan, R. (2008). Effective analysis of reaction time data. *Psychological Record*, *58*, 475–482.
- Wulfert, E., & Hayes, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior*, *40*, 125–144.