http://researchcommons.waikato.ac.nz/

Research Commons at the University of Waikato

Copyright Statement:

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

The thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of the thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from the thesis.

Investigating Relatedness and Nodality in Stimulus Equivalence

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Applied Behaviour Analysis

at

The University of Waikato

by

Gabriel Chris Chand

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2020

Abstract

Stimulus equivalence describes how a group of stimuli can be trained to be 'equal', part of one group (equivalence class) where all stimuli are substitutable by each other. Relatedness deals with how stimuli in an equivalence class may not be equal, but rather have differing magnitudes of relations depending on how many stimuli are between two related stimuli. The stimuli in that are between two related stimuli in an equivalence class are called nodes. The issue of nodality in relatedness has been of interest to behavioural researchers. In this study, an experiment was conducted to test the strength of the relations between stimuli in trained equivalence class, which was a replication of several previous studies, with methodological improvements. The strength of stimulus-stimulus relations was tested by first using the simultaneous training method to establish two 6-member, 4-node equivalence classes, after which within-class preference testing was used. Of the 12 participants who took part in the experiment, seven participants were not able to successfully form the two 6-member equivalence classes, with the remaining five participants successfully able to form the two 6-member equivalence classes – demonstrated through achieving criterion (90%) response accuracy during the simultaneous protocol testing phase. The participants who completed the preference testing demonstrated response control by nodality, as they regularly preferred comparison stimuli that were nodally proximal to the sample; and in the case of one participant, responding was entirely consistent with the effects of nodality. That is, they always preferred the comparison stimuli that were nodally proximal to the sample stimuli. Furthermore, on the occasions when the participants did select the nodally distal comparison stimuli, they took longer to make the selection than when they selected the nodally proximal stimuli. On preference testing probes that did not differ in nodal distance (comparison stimuli were an equal nodal distance from the sample stimuli), participants showed no preference for

either of the comparison stimuli by selecting both the comparison stimuli approximately an equal number of times. These findings are consistent with earlier studies which involved using various training methods to establish equivalence classes, then testing for nodality effects on stimulus-stimulus relations within the established equivalence classes. Therefore, with the methodological improvements applied in the study, nodality was demonstrated, suggesting that stimuli in an equivalence class are related differentially.

Acknowledgements

I would like to express my deepest gratitude to all the people who have helped me throughout my year-long journey in making this project a possibility. None of this would be possible without the various contributions of all the people involved at different stages of the project throughout the past year.

My biggest thanks goes to Dr Tim Edwards, Chief Supervisor, for all the invaluable help and feedback given to refine and improve this project. Dr Edwards provided much-needed guidance and encouragement every time I needed it – Thank you so much, this would never be possible without you.

A big thanks to Andrew Malcolm, who worked tirelessly to develop the computer program that was central to this project.

Thanks to my partner who was always encouraging me and not letting me become complacent; to my family, who gave me all the support I needed, especially in moments when I was doubting myself – Thank you so much.

Table of Contents

Abstract i

Acknowledgements iii

List of figures vi

List of tables vii

Investigating relatedness and nodality in stimulus equivalence 1

Research on stimulus-stimulus relations in equivalence classes 6

Methodological effects on formation of equivalence classes and nodality 6

Nodality 11

Logic operations (relation complexity) 16

Rationale for study 19

Method 26

Results 35

Discussion 44

Nodality 44

Priming effects 45

Unequal reinforcement 47

Differential functions of nodes and singles 48

Response latencies for nodal distance probes 49

Baseline versus symmetry (Nodality or Relation complexity) 51

Response accuracy 51

Response latencies for baseline versus symmetry probes 55

Theoretical integration.....56

Limitations.....58

Future considerations.....61

Summary.....62

References 64

Appendix A 69

Appendix B.....70

Appendix C.....71

List of Figures

Figure 1: *A simple equivalence class consisting of three stimuli*.....3

Figure 2: *Layout of how the sample and comparison stimuli were displayed*..... 28

Figure 3: *Percentage of trials where the nodally proximal selection was made (shown as the letter designation in the boxes on top of the probe data bars*..... 37

Figure 4: *Percentage of trials where the baseline (0 node, forward direction) relation was selected over the symmetry (0 node, backward direction) relation*..... 40

Figure 5: *Average percentage of times the baseline or symmetry relation was chosen*.....41

List of Tables

Table 1: *The three groups of symbols (18 symbols in total) used to form the two equivalence classes* 30

Table 2: *Probe combinations presented in the preference testing phase* 34

Table 3: *Average time taken by participants when making nodally proximal versus nodally distal choices in the preference testing phase*..... 38

Table 4: *Average time taken to make baseline (0-node forward) or symmetry (0-node backward) selection* 42

Investigating relatedness and nodality in stimulus equivalence

The phenomenon where two stimuli that neither share any similarities, nor are directly related in any way, come to be viewed/responded to as if they were the same, has been of great interest to behaviour analysts. Many studies have been done to understand the processes involved when dissimilar events become equivalent, or equal such that the events become interchangeable (Green & Saunders, 1998). Stimulus equivalence is the term used to describe such a phenomenon (Kuno et al. 1974; Sidman, 1971; Sidman & Tailby, 1982). Stimulus equivalence tries to understand relations between stimuli, which share no physical properties that become related to each other. When a group of stimuli are trained to be related to each other, derived (untrained) relations start to emerge within the trained group of stimuli. For example, two stimuli become 'equal' without having any direct relation or training, on the basis of both the stimuli being trained to be related to a third stimulus. The third stimulus serves as a kind of mediator, linking the other two stimuli to each other without the two stimuli ever directly being related or trained to each other. In essence, the relationship between the two stimuli that were never directly trained is derived (from the third stimuli; if A is the same as B and B is the same as C, then A is also the same as C), and all three stimuli become associated with each other, with all of the stimuli acquiring the same function (Fields et al. 1984). When stimuli become related (either directly or derived like in the previous example) and start functioning similarly to each other (for example, each stimulus evokes the same kind of response), they are said to be in an equivalence class (Adams et al. 1993; Green & Saunders, 1998; McIlvane & Dube 2003; Sidman, 1990, 1992, 1994, 2000).

Equivalence classes are characterised by directly trained baseline relations, and untrained (derived) relations, with the type of relation depending on the complexity of the relation (see Figure 1). The types of stimulus-stimulus relations present in an equivalence

class include baseline (direct), reflexive, symmetrical, transitive, and equivalence (combination of transitive and symmetrical relations). Baseline relations refer to stimulus pairs that are directly trained (for example, selecting stimulus B being reinforced in the presence of stimulus A [A=B], and selecting stimulus C being reinforced in the presence of stimulus B (B=C). Reflexive relations are the stimulus relations where the A stimulus is the A stimulus and not the B stimulus or C stimulus and so on (A=A). Symmetry relations entail that if selecting stimulus B is reinforced in the presence of stimulus A, then without explicit training (derived), participants tend to select the A stimulus in the presence of the B stimulus (if A=B, then B=A). Transitivity relations come about when two unrelated stimuli become related through a third stimuli; If a stimulus A is trained to be equal to stimulus B, and stimulus B is trained to be equal to stimulus C (A=B, B=C), then it is derived that stimulus A is the same as stimulus C without direct training (if A=B, and B=C, then A=C); finally, equivalence is the transitive and symmetry relations combined, where if stimuli A and B are established as equal, and B and C are established as equal, the derived relation where A is equal to C (transitivity) can also be applied in reverse (transitivity + symmetry), where C becomes equal to A (if A=B, and B=C, then A=C, and also C=A) (Sidman & Tailby, 1982; Sidman et al. 1982; Sidman, 2000).

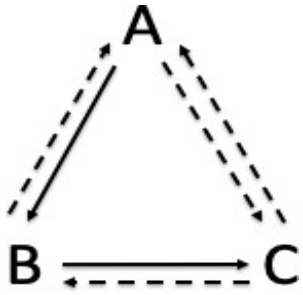


Figure 1. A simple equivalence class consisting of three stimuli. The solid arrows depict directly trained relations (in the direction of the arrow), and the dashed arrows depict untrained (or derived) relations. The relations A=B, and B=C are trained (baseline); B=A, and C=B are derived symmetry relations; A=C is a derived transitive relation; C=A is a derived equivalence (transitive + symmetry) relation.

In stimulus equivalence studies, equivalence classes, consisting of a finite number of items that do not share any physical properties, become arbitrarily related to each other through training (Doran & Fields, 2012; Fields & Reeve, 2001; Fields & Verhave, 1987; Lazar, 1977; Moss-Lourenco & Fields, 2011; Sidman, 1994; Sidman & Tailby, 1982; Wang et al. 2011). In this context, arbitrarily related stimuli are stimuli that have been related to each other by properties of the stimuli other than non-arbitrary/formal properties such as size (larger than), sound (louder than) and visual (differently coloured than) properties. Arbitrary relations can be formed based on verbal processes, or contextual cues beyond formal properties. For example, if a person is shown a New Zealand dollar coin and a 50 cent coin, and asked which one is more, they will point to the dollar coin, even though the 50 cent coin is physically larger than the dollar coin. By non-arbitrary criteria, the 50 cent coin is 'more' (bigger), but to a verbally sophisticated individual, who has been trained to treat the dollar as more valuable than 50 cents, the dollar is 'more'. That response (dollar is more) is arbitrary,

as if an individual who has no knowledge of currencies were to be asked the same question (which is more, and shown a 1 NZD coin a 50 cent coin), they would base their answer on formal properties (on their perceptions of the non-arbitrary/formal properties of the two coins; in this case the physical size) and point to the 50 cent coin (Barnes-Holmes et al. 2004; Stewart et al. 2002).

The way researchers get participants to form these equivalence classes is through training tasks involving the stimuli. Early research involved the training of paired associations, where, for example, three lists of stimuli (list A, B, and C) would be taught, then the words from list A would be paired with items from list B (by reinforcing selections of the items in list B in the presence of words from list effectively making it so that the words from list A are equal to list B), and items from list B paired with list C (B=C). Then words from list C would be shown, and participants would be asked to point out which words from list A went with the ones from list C. Participants would usually select the list A words as going with the list C words (despite the fact that they would have the option to say the words don't go together) even though they were never directly told/trained to do so. This happened due to their common relation, list B; the list B words acted as the mediator or link between list A and list C words (Jenkins, 1963; Jenkins & Palermo, 1964). This method wasn't very robust (outcomes varied a lot), and in later research, methods of training equivalence classes involved conditional discrimination training (Doran & Fields, 2012; Sidman et al. 1974; Sidman & Tailby, 1982; Sidman et al. 1985; Steele & Hayes, 1991).

Conditional discrimination involves presenting a stimulus that signals another stimulus as 'correct', and the rest of the stimuli as 'incorrect'. For example, where there are two stimulus groups (A1, B1, C1, and A2, B2, C2) the presentation of the A1 stimulus signals picking the B1 stimulus as 'correct' and B2 stimulus as 'incorrect' (in the absence of stimulus A1, both stimuli B1 and B2 are neutral). Conversely, the presentation of the A2 stimulus

signals that the B1 stimulus is now incorrect, and the B2 stimulus is now correct. The most common way the conditional discrimination method is utilised is in matching to sample procedures (MTS). Typically the correct selections are reinforced through feedback screens (showing the word 'correct'), points accumulation, pleasant sounds or pictures, or a combination of these. Incorrect selections are punished by point deductions, unpleasant sounds and pictures, or feedback screens informing that the selection was wrong.

In a typical MTS procedure, a stimulus is presented as a sample, and two or more stimuli presented as comparisons to select from (on a computer; the sample stimulus is usually in the middle of the screen, and the comparison stimuli are presented below the sample, an equal distance apart from each other); if the sample for example, is the stimulus A1 from the previous example, then the comparisons will include one correct stimulus (B1, from the same stimulus group as A1) and one or more incorrect stimuli from the other groups (in this case, B2 from the second group). MTS procedures are typically used to *a)* train baseline relations (for example, in the stimulus group A-B-C, training stimulus A to be equal to stimulus B, and training stimulus B to be equal to stimulus C, effectively getting A=B=C), and *b)* test derived relations (when the baseline relations A=B, and B=C are trained, all the symmetry, transitivity, and equivalence relations are tested). Conditional reinforcement is delivered only during training, most typically in the form of feedback screens – initially, all the training trials have feedback (letting the participant know whether their selections were correct or incorrect); as participants achieve an acceptable level of accurate responding (usually 90-100% accuracy is required), the trials are presented again, but with less feedback (from all trials having feedback to only half, to none), until participants can achieve criterion accuracy without any feedback (without reinforcement).

Following the training of the baseline relations, the stimuli are rearranged (sample and comparisons) to test for derived (untrained) relations. For example, having been trained to

pick the stimulus B1 in the presence of stimulus A1, and C1 in the presence of B1 in the baseline training, a participant who then demonstrates symmetry by selecting A1 in the presence of B1, and B1 in the presence of C1 in the testing phase, demonstrates transitivity by selecting C1 in the presence of A1, and finally, in demonstrating equivalence by selecting A1 in the presence of C1, the participant has started to form an equivalence class. The success or failure of a participant in forming the equivalence classes is usually measured through response accuracy in the absence of feedback (the percentage of trials where the correct comparison was selected based on the sample shown). If response accuracy is high enough without feedback (usually 85-90% accuracy), then there is demonstration of a formed equivalence class (Alligood, 2007; Sidman, 1990, 1992, 1994; Sidman et al. 1985).

Research on Stimulus-Stimulus Relations in Equivalence Classes

Methodological Effects on Formation of Equivalence Classes and Nodality

While matching to sample (MTS) is the most commonly used method to present the training and testing trials when forming equivalence classes, there are a number of different ways by which trials can be presented, potentially altering the way equivalence classes are acquired, and consequently affecting the relational strength of the stimuli within the equivalence class (Alligood, 2007; Arntzen, 2012; Bortoloti & de Rose, 2011; Fields, 2016). Any of the following matching to sample trial procedures can be used to establish equivalence classes – one to many (OTM), many to one (MTO), or linear, where relations are presented in a serial order. There are also different trial structures that can be used, such as the simple-to complex protocol, the complex to simple protocol, and the simultaneous protocol.

Structure of Baseline Relations

Many to One Trial Presentation

The many to one trial presentation consists of a non-linear presentation of trials, where many different samples are used to relate them to one correct comparison stimulus. For example, the selection of B1 stimulus (not B2) is reinforced in the presence of A1 stimulus. Then, selecting B1 (not B2) is again reinforced in the presence of C1, and again in the presence of D1, and so on. Essentially, many stimuli are trained to be related to one stimulus, and they all become related due to their shared association with the comparison stimulus. As an example, the baseline relations for a four-member equivalence class established using the many to one procedure will look like this: A=B, C=B, and D=B.

One to Many Trial Presentation

The one to many trial presentation is also a non-linear trial presentation method (Hove, 2003). In the one to many procedure, the selection of the A1 comparison stimulus (not A2) is reinforced in the presence of sample stimulus B1. Then the selection of the C1 comparison stimulus (not C2) is reinforced in the presence of B1 sample stimulus, and so on. In contrast to the many to one procedure, in this procedure, one stimulus is trained to be related to many other stimuli. For example, a four-member equivalence class established in this way will contain the following trained relations B=A, B=C, and B=D.

Linear Trial Presentation

The linear trial presentation consists of presenting the relations in a serial order (Hove, 2003). The linear trial presentations involves presenting the simplest baseline, forward

relations first, in order; to form a four member equivalence class in the linear trial presentation, selecting comparison stimulus B1 (not B2) is reinforced in the presence of sample stimulus A1 (A1=B1), selecting comparison stimulus C1 (not C2) is reinforced in the presence of sample stimulus B1 (B1=C1), and selecting comparison stimulus D1 (not D2) in the presence of sample stimulus C1 is reinforced, leading to the establishment of the A1=B1=C1=D1 equivalence class.

Trial Structure

Trial structure refers to the sequence in which trials are presented (one by one). There are three main ways trials are structured – The simple to complex protocol, the complex to simple protocol, and the simultaneous protocol.

Simple to Complex Protocol

This is a sequential training protocol, where the baseline relations are trained and tested in a sequence, starting from the first baseline relation (and testing the same baseline and derived relations), and adding the next baseline relation (and testing for this baseline relation, as well as the previously trained baseline relation and all possible derived relations), and so on until all the baseline relations are trained, and all baseline and derived relations are tested. For example, to form the equivalence class A=B=C in a simple to complex MTS procedure, the A=B baseline relation is trained first, and once response accuracy is achieved, testing is done on the baseline (A=B), and derived relation (B=A). Following the testing, the next baseline relation is trained (B=C, with the first baseline relation also included, A=B), and once criterion accuracy is achieved, testing is done for the all the baseline and derived relations that are possible, which includes all the relations that have been previously tested

(A=B, B=C, B=A, C=B, A=C, and C=A). This sequential method has a high chance of success (participant yield is high, as with this method, equivalence classes are easily formed), as each relation is trained one by one in a sequential order (Fields et al. 1991; Lynch & Cuvo, 1995).

Complex to Simple Protocol

A complex to simple protocol is also a sequential trial presentation, where trials are presented one by one in a sequence. In a complex to simple protocol, a pair of baseline relations are trained serially. Immediately after that, the equivalence relation is tested first. If the equivalence testing is unsuccessful (usually 4-6 trials are presented), then a step back is taken, presenting the symmetry testing trials for the pair of baseline relations that were initially trained. If responding in these trials are consistent with symmetry, then the equivalence testing trials are presented again (Fields et al., 1997). For example, to establish a three-member equivalence class (A=B=C) in a complex to simple procedure, the baseline pairs A=B and B=C are trained first. Immediately after that, the C=A equivalence test is administered. If the participant does not pass this test, then the symmetry relation B=A test is presented; with failure here resulting in the re-training of the baseline relations. If this B=A test is passed, however, the C=A equivalence trials are presented again. Failing the C=A test again leads to the testing of the C=B symmetry relation; with failure here resulting n baseline re-training once more. Passing the C=B test, leads to the presentation of the C=B and B=A test together. Passing the C=B and B=A tests now lead to the C=A test again. This time, failing the C=A test results in the presentation of the A=C transitive test; failure on this resulting in the re-training of the baseline relations again. Passing the A=C test once finally leads to the C=A trials again (Adams et al., 1993).

Simultaneous Protocol

The simultaneous protocol is not sequential, as all relations to be trained are presented randomly in one block (not one after another, like in the sequential methods), while the testing trials for all the trained and derived relations are also presented randomly in one block. Because the relations aren't presented in order nor trained one after another, it is difficult to establish equivalence classes using this method (Fields et al. 1997). However, since establishing equivalence classes is difficult with the simultaneous protocol, it makes it useful to test for the historical factors responsible for the establishment of new equivalence classes (Fields et al. 1995). "Historical factors" refer to the factors in a previously formed equivalence class (such as class size) that may influence the formation of new equivalence classes. Since the sequential method has such a high success rate of equivalence class formation, it becomes difficult to test for the historical factors that may influence formation of new equivalence classes (success rate is so high, it doesn't matter whether the participant has had previous training to form equivalence classes).

Another advantage to the simultaneous protocol, is the balanced presentation of all the trials during training and testing. Imam (2001, 2006) pointed out how, in the sequential training procedure, the number of trials are cumulative. That is, the first baseline relation is trained, then trained again when the second baseline relation is trained, and so on until the last baseline relation is trained; and by this time, the first baseline relation will have been presented many more times than the other relations. This means that the relations are reinforced unequally, influencing responding during the derived relations and preference testing procedure (Saunders & Green, 1999). In his experiments, Imam (2001, 2006) found that participants were responding most accurately and quickly to baseline relations, then symmetry, followed by transitivity relations and finally to equivalence relations. He explained that these relational complexity (logic operations) effects were due to the

differential presentation and consequently unequal reinforcement for these relations (baseline reinforced the most, followed by symmetry, and so on). However, Wang and his colleagues (Wang et al. 2011; Wang et al. 2012) resolved the issue of unequal reinforcement by using the simultaneous protocol to establish equivalence classes instead of sequential protocol, therefore minimizing differential reinforcement across the relations.

Nodality

Stimulus equivalence, by definition, implies that in order for a group of stimuli to be in an equivalence class, they must be functionally substitutable, or interchangeable with each other (Moss-Lourenco & Fields 2011; Sidman, 1994). This means that each stimulus in the equivalence class evokes the same behavioural response, or otherwise function the same as the other stimuli in the equivalence class; if the stimuli are functionally similar, it also implies that they are equally related to each other (Fields and Verhave, 1987). In this context, equally related refers to the magnitude or strength of the relations between the stimuli in the equivalence class.

An alternate view, *relatedness*, implyies that stimuli in an equivalence class are related differentially to each other, or rather, unequally related (Bortoloti & de Rose, 2011; Fields, 2016; Fields & Moss-Lourenco; 2007; Moss-Lourenco & Fields, 2011).

One of the most prominent factors, and strongest indicator of the existence of a disproportionate magnitude of relatedness between stimuli in an equivalence class, is nodality; specifically nodal distance. Nodality refers to stimuli in an equivalence class that are related by training to at least two other stimuli, usually the common relation that is linking two untrained stimuli to each other. For example, if an equivalence class is formed by training the baseline relations A=B, B=C, and C=D (so the equivalence class A=B=C=D is formed), then it can be observed that the stimuli B and C are both related to two separate

stimuli in the equivalence class; stimulus B is related to A and C; stimulus C is related to B and D. Stimuli B and C are nodes, as they are in between two other stimuli (B is in between A and C; C is in between B and D). Nodal distance refers to how many stimuli (hereafter referred to as nodes) are in between two given stimuli in an equivalence class. Using the previous example, only one node (B) is between A and C, meaning that stimulus C is one node (a nodal distance of one) away from A; whereas there are two nodes (B and C) between A and D. Stimuli in an equivalence class that aren't nodes (aren't related to at least two other stimuli) are referred to as singles (since they are only related to one other stimulus through training). Usually in an equivalence class, the singles are the first and last stimulus, while all the stimuli in between are nodes.

Studies done on nodal distance usually begin by using the matching to sample (MTS) protocol to train baseline relations. Once the baseline relations are established, the formation of the equivalence classes is inferred through testing of both the baseline and all the derived relations (as mentioned during the explanation of MTS procedures).

For example, when trying to establish two three-member equivalence classes (A1-B1-C1, and A2-B2-C2), the MTS procedure is used to train the baseline relations A1=B1, B1=C1, A2=B2, and B2=C2. Initially, feedback is given after every trial, informing whether the selection made is correct or incorrect. Once criterion response accuracy is achieved, testing commences, presenting baseline and derived relations (again looking at response accuracy). Formation of equivalence classes is confirmed when participants achieve criterion response accuracy during testing. From there, nodality is usually tested through within-class preference assessments. Within-class preference assessments consist of probe trials – probes are three letter designations that inform what stimuli are in the trial; the first letter is the sample stimulus, the second letter is the comparison stimulus selection that is nodally proximal to the sample stimulus, and the third letter is the comparison stimulus that is nodally

distal from the sampe stimulus. For example, a BCD probe means that the trial consists of the sample B, and the nodally proximal comparison stimulus, C, and the nodally distal comparison stimulus, D, being the last letter. Preference probe trials always contain stimuli from the same class (sample and the two comparison stimuli all from the same class).

Fields et al. (1990) used the matching to sample program to establish two four-member equivalence classes (A=B=C=D). The stimuli used were three letter nonsense words. During the testing phase of the procedure, they found that participants were responding more accurately on tests that had one node relations (e.g. A=C), over tests that had two node relations (e.g. A=D). Spencer and Chase (1996) established three five-member equivalence classes, and measured response accuracy along with response speed during the testing of the equivalence classes. Once again, the findings showed that participants were more accurate, and responded quicker on tests which had relations containing fewer nodes than on tests which had relations containing more nodes. Moss-Lourenco and Fields (2011) established two six-member equivalence classes in one of their experiments, and used the within-class preference assessment to test for the effects of nodality. Participants who were able to form the equivalence classes showed preference for the stimuli that were nodally closer to the sample stimulus.

Other than within-class preference assessments, transfer of function tests have also been utilized to test for the effects of nodality. In this method, after the MTS procedure is used to establish equivalence classes, two stimuli are chosen (usually the stimuli that are furthest away from each other in the class) from the equivalence class, and a different behavioural response is trained in the presence of each of the two stimuli. Then the other stimuli in the equivalence class are presented to see which response each of those stimuli evoke. Fields et al. (1995) got participants to form five-member equivalence classes (A=B=C=D=E) with three letter nonsense words as the stimuli. They then trained participants

to push the spacebar key in the presence of A1 stimulus and to push the enter key in the presence of the E1 stimulus. They then measured responses when the B1, C1, and D1 stimuli were presented. The results showed that the responses being emitted were related to nodality; The B1 stimulus was more likely to evoke the behaviour trained to A1 (spacebar key), while the D1 stimulus was more likely to evoke the behaviour trained to E1 (enter key)

The stimulus equivalence model would imply that the comparison stimuli which are from the same equivalence class will be selected roughly an equal number of times, regardless of nodal distance, since all the stimuli in that class are equally related. However in nodal distance studies, the comparison stimulus that is nodally closer to the sample stimulus has been consistently preferred over the stimulus that is nodally further away (Bentall et al. 1999; Bortoloti & de Rose, 2011; Fields et al. 1990; Fields et al. 1995; Fields et al. 1997; Fields, 2016; Moss-Lourenco & Fields, 2011; Spencer & Chase, 1996).

Fields and his colleagues (1997) tested for the influence of nodality on the formation of new equivalence classes. They wanted to see whether the number of nodes in an established equivalence class, or number of stimuli (item number) in an equivalence class that aren't nodes influenced the formation of new equivalence classes. Item number refers to the number of items in an equivalence class, with the hypothesis being that as the number of items increased (without increasing number of nodes), the strength of the relations would be altered; in other words, more items in the pre-trained equivalence class would cause participants to take longer to form new equivalence classes under the simultaneous protocol. The experiment consisted of using the simple-to-complex protocol to pre-train equivalence classes that differed in terms of the number of nodes in the class, or number of items in the class (without altering number of nodes). The simple-to-complex protocol consists of training baseline and testing derived relations one by one in a sequence (train A=B, then test A=B, and B=A; then train A=B, B=C, and test A=B, B=C, B=A, C=B, A=C, and C=A, and so on).

Participants were separated into six groups; three groups trained to form one node classes that had either three, five or seven members. Usually, in the linear method of training, the establishment of equivalence classes mean that as the stimuli in the equivalence class increase, so do the nodes in the equivalence class (a 4-member class has 2 nodes; 5 member class has 3 nodes and so on). They held the number of nodes constant by using a non-linear many-to-one trial presentation. In the many-to-one trial procedure, selection of one comparison stimulus is reinforced in the presence of all the other stimuli in the intended equivalence class. For example, for the one node five-member class, they trained the relations A=B, and C=B (three items, one node); they then trained further relations D=B, and E=B, essentially having only stimulus B in the group related to two or more stimuli (being a node) and yet the other members (A, C, D, E) only being trained to be directly related to B (other members in the class were singles). In essence, all the stimuli in the equivalence class were related through their common trained relation to stimulus B, with B being the only node in the class (related to more than one stimulus). A further two groups were trained to form three node classes with five or seven members (using the same technique described in the previous example, but with two nodes instead of one), with one group acting as a control, receiving no pre-training.

They wanted to test how quickly these participants would form new three node five-member equivalence classes in a simultaneous protocol, if they had pre-training to form equivalence classes that had either more nodes or the same number of nodes but more singles beforehand. The results showed that the speed at which the baseline relations were formed in the simultaneous protocol were an inverse function of the number of nodes in the pre-trained equivalence classes, not the number of items. In other words, the three groups that pre-trained for one-node classes were slower to acquire baseline relations in the simultaneous protocol than the two groups that had pre-training for three-node classes; furthermore, there were no

differences in acquisition speed when the item number was increased but nodes kept constant. The results further showcased the important influence of nodality, as the higher the number of nodes in a pre-trained equivalence class, the easier it becomes to form new equivalence classes. Pre-training equivalence classes helps establish new equivalence classes more quickly, and the number of nodes in the pre-trained equivalence class determines how quickly the new equivalence classes are formed.

These studies have provided insight into nodal distance; it appears that there is an inverse relationship between relational strength and nodal distance – as nodal distance increases, the magnitude of the relational strength between the stimuli decreases. For example, in a four-member equivalence class (A=B=C=D), stimulus A would have a stronger relation to stimulus B than to C, and have a stronger relation to C than to D (Fields et al. 1990; Fields at al. 1995).

Logic Operations (Relation Complexity)

Another factor that appears to be supporting relatedness in stimulus equivalence classes is logic operations (Bortoloti & de Rose, 2011; Fields, 2016). Logic operations refers to the complexity of the relations, and how much mental effort is required to arrive to deduce the relation between stimuli in an equivalence class (Cowan, 2001; Shen et al. 2018). For example, a simple symmetry relation (A=B, then B=A) is going to be stronger than a more complex transitivity one (A=B, B=C, so A=C), because the transitivity relation requires more operations to derive the relation compared to the simpler symmetry relation. Both the symmetry and transitivity relations are going to be stronger than the even more complex equivalence (symmetry + transitivity) relation (A=B, B=C, so A=C, and therefore C=A), as the equivalence relation requires even more operations to arrive to the relation when compared to the transitivity and symmetry relations.

Studies have shown that when nodal distance is kept the same, then logic operations govern the magnitude of relations between the stimuli in an equivalence class; the more complex the relation, the weaker the relation (Doran & Fields 2012; Fields, 2016). Doran and Fields (2012) demonstrated the influence of logic operations by establishing two, 5-node 7-member equivalence classes (A=B=C=D=E=F=G). They then presented only D as a sample, pitting different relations against each other (within class), using the MTS preference method to test selected preference probes that had comparison stimuli differing in logic operations/complexity. Probes refer to three letter stimulus designations (e.g. CEA), where the first letter represents the sample stimulus, the second letter representing the simpler (fewer logic operations) comparison stimulus from the same equivalence class, and the third letter representing the more complex comparison stimulus from the same equivalence class. For example, the DFB probe tested a one-node equivalence relation (DB) against a one-node transitive relation (DF). They found that preference was skewed towards the simpler (fewer logic operations) relations (in the example of the DFB probe, preference was shown for the comparison with the simpler relation to the sample, the F comparison).

Furthermore, they also found that when the logic operations were the same (e.g. transitivity vs. transitivity), then nodal distance would affect preference such that the higher nodal distance relation would be weaker (non-preferred). For example, the DFG probe (both transitive relations, but DF is one node, while DG is two nodes) would show preference for DF over DG. This also extends to baseline versus symmetry relations – If nodal distance is the primary factor in relatedness, then in a preference MTS test where, for example, a CBD probe is presented (C is sample, B and D comparisons), participants will be expected to choose equally between the two comparisons on average, as although the B and D stimuli have different relations to C (symmetry and baseline, respectively), they are nodally identical (0 node from C to B, 0 node from C to D). If logic operations is the primary influence in

relatedness however, then the same preference test should show an affinity towards selecting the baseline relation over the symmetry relation.

Alligood (2007) conducted a study on the stimulus-stimulus relations in equivalence classes, by training participants to form two 6-member, four-node equivalence classes. The within class preference tests administered to the participants yielded responding in line with nodality, meaning the participants tended to choose the nodally proximal option, and when a baseline versus symmetry probe was presented, participants tended to choose approximately equally between the baseline and symmetry relations on average. This result indicated that nodality may have a larger influence in stimulus-stimulus relations in equivalence classes than logic operations – if logic operations was the bigger influence, participants would tend to prefer trained baseline relations over untrained (derived) symmetry relations.

While there does seem to be evidence that number of logic operations play a role in the relatedness of stimuli in an equivalence class, the reality is that more research is required to clarify the influence of this factor, and it is not entirely clear how much of an influence the factor of logic operations has on the differential relatedness of stimuli in an equivalence class.

Together, these findings challenge the definition of stimulus equivalence; equivalence implies that all stimuli in a given equivalence class are interchangeablel and substitutable by one another (Sidman, 1994). Sidman stated that without the property of interchangeability or substitutability, the stimuli cannot be said to be in an equivalence class. Studies on nodality, however, have shown that not only is ‘relatedness’ supported (through witihin-class preference tests, semantic differential ratings, and post-class discrimination tests), the equivalence classes still hold the overall property of the same acquired function. This was demonstrated by Moss-Lourenco and Fields (2011), where, after the formation of equivalence classes and within-class preference testing (which showed nodality), emergent tests were done (MTS checking all the emergent relations) which showed that the equivalence class

remained intact. This suggests that the degree to which stimuli in an equivalence class acquire the same function may be under the control of nodality relative to each other. That is, stimuli in an equivalence class are equal in certain contexts (across-class), but are related differentially to each other in other contexts (within-class).

Rationale for Study

There are several methodological issues with previous research on nodality; mostly during the testing and/or preference phase when the MTS protocol is presented (Imam, 2006; Moss-Lourenco and Fields, 2011). One of the methodological criticisms made by Sidman (1994) and Imam (2001; 2006) is presenting a stimulus from the equivalence class as a sample, with only two comparison stimuli to choose from. The two comparisons are from the same class (within-class preference tests) or one from the same class and one from a different trained class (emergent test, usually straight after training protocol). As Sidman pointed out, the problem with only presenting two choices during training and testing is that participants can start choosing based on rejection or avoidance. It would be impossible to tell whether participants are selecting comparison stimuli based on positive selection (recognizing that the comparison stimulus is related to the sample stimulus), or selecting comparison stimuli based on rejection (selecting the positive comparison stimulus that is related to the sample stimulus, because they recognized that the other comparison stimulus is not part of the group, consequently rejecting that comparison stimulus, leaving only one choice left).

These issues have been dealt with by simply adding a third null option (distractor stimulus). The distractor is supposed to prevent random choice control. However, it can be argued that a third distractor option does not affect the test at all – participants will see a stimulus from neither the A1 or A2 class and dismiss it instantly, which would leave two choices and the same issues. For example, if a participant was asked to choose the ‘sweet’

item, and the choices were chocolate, a sandwich, and a hat, the hat wouldn't even be considered, meaning the participant would respond as if there were only two choices (third option may as well not exist). This is what happens if a third distractor option is given during testing.

A proposed solution that might work is, during testing for equivalence class formation, add a third option from the second trained class (one correct option from the same class, two incorrect ones from the second trained class). This way, the third option serves as a true distractor, as it is part of an established equivalence class, but just not the correct option. Another solution could be to have a third group of stimuli (similar to the two equivalence classes in type of stimuli and number of stimuli) present for training and testing, but never be the correct option. Rather than having an obvious 'wrong' choice that was never present during baseline training and has come up during testing for the formation of equivalence classes, there will be a third stimuli class that was present all the way from baseline training up to and including the testing for equivalence class formation.

This problem will be addressed in the current study by introducing a third comparison from a different class that appears in training as an option, containing the same type of arbitrary stimuli as the arbitrary stimuli in the other two classes intended to be established. For example, the two class of stimuli intended to be trained to form equivalence classes (letter denotes stimuli, number denotes the class it belongs to) could be A1, B1, and C1 for class 1, and A2, B2, and C2 for class 2. However, a third class that has similar arbitrary stimuli that are contained in the other two classes could also be added (A3, B3, and C3). The stimuli from the third class would only serve as a third comparison stimulus option in the training and testing phases - during training testing for the first class, presenting A1, then B1 (correct), and the same letter (B) but from the other two classes (B2 and B3). This ensures that there isn't any selecting based on rejection or selecting away from a particular stimulus,

and should therefore allow for true selection based on learning. Keeping the distractor choices the same letter designation from two other equivalence classes also ensures that the participants actually make a careful choice rather than encountering an obviously different stimulus and rejecting it instantly (therefore giving the impression that they have acquired the class, yet in reality they are just recognizing the stimuli that they have never seen before and rejecting it).

Another procedural issue is that of the different functions of the stimuli, specifically the nodes and the singles. A few studies (Alligood, 2007; Doran & Fields, 2012; Moss-Lourenco & Fields, 2011) have used singles (either the first or last stimulus in the linearly trained equivalence class) in their within-class preference testing phase. For example, Moss-Lourenco & Fields (2011) trained participants to establish two 5-member, three-node equivalence classes. Following that, they presented the participants with probe trials in the within-class preference testing phase that included probes such as ABE and EDA. The problem with this is the different functions stimulus A (single, only serves as a sample in training), and stimulus E (single, only serves as a comparison in training) have, compared to the rest of the stimuli in the class that are nodes. The nodes are presented as both sample and comparison in the trials, however the singles are presented only as a sample, or only as a comparison; this could lead to selection based on stimuli that have two functions (nodes) over one function (singles). Simply increasing the class size, and removing the singles during preference testing, only testing the nodes, could solve this problem. For example, in the equivalence class A=B=C=D=E=F, A has only one function (as sample), and so does F (is always a comparison). Removing the A and F stimuli during preference testing will allow for participants to choose between the remaining stimuli, which are all nodes and therefore have appeared as both the sample and comparison stimuli during training trials. This is the strategy that will be utilized in this study – increasing the class size will allow for the removal of the

very first and last stimuli in the equivalence class, and allow for testing of nodes that are all functionally equal (they all are nodes, and all have served as sample and comparison during training). In other studies that have utilized preference testing, they have usually selected a small number of probes to test (Moss-Lourenco & Fields, 2011; Fields, 2012) In this study, a further step will be taken in the preference phase, where, of the nodes being tested (4 nodes in total; B, C, D, E), all possible combinations of probes will be presented for both the established classes. This way, every combination of relation versus relation probes possible in a 4-node class will be covered during preference testing, for both the established classes.

A further issue is that of priming effects during the within-class preference testing phase. Recent studies that have utilized the post-class formation within-class preference testing procedure, have presented only a selected few probe trials, and presented them in a set order (Alligood, 2007; Doran & Fields, 2012; Moss-Lourenco & Fields, 2011). The probes have usually been presented in pairs that test the same characteristic. In recent studies, the first probe pairs that have been presented in the within-class preference testing procedure have been that of extreme differences – The probe pairs have a sample stimulus, a comparison stimulus that is the closest nodal distance possible within the equivalence class, and the second comparison stimulus that is the farthest possible nodal distance from the sample stimulus. For example, in the Moss-Lourenco and Fields (2011) study, the participants that did the within-class preference testing were first presented with the ABE and EDA preference probes from the class A=B=C=D=E. This meant that the comparison stimuli had the most difference in terms of nodal distance between each other (one comparison was closest to the sample, one comparison was farthest from the sample). Priming effects may have been possible, where participants selected the closest comparison stimuli, and consequently may have learnt to keep looking for the closer comparison stimuli in all the

preference testing probes. In this study, all possible preference probes were presented in a random order, minimizing the possible effects of priming.

An issue in general, regardless of what kind of dependent measures (within-class preference, transfer of function, rating scale, speed) are being employed, is that of unequal reinforcement (Imam 2001; 2006). It has been established that the best training method in terms of forming equivalence classes where differential reinforcement is minimized, is the simultaneous protocol (Alligood, 2007; Wang et al. 2011). Wang and his colleagues (Wang et al. 2011; Wang et al. 2012) used the simultaneous protocol to establish equivalence classes instead of sequential protocol, therefore minimizing differential reinforcement across the relations. The participants still demonstrated the effects of both nodality and relational complexity (as the nodes and logic operations increased, response accuracy decreased, and response speed got slower). Since all the relations in a simultaneous protocol are presented all at once in one block of trials, the issue of unequal reinforcement is minimized. However, this protocol is extremely difficult, and yields a low proportion of participants who achieve criterion responding. Fields and his colleagues (1997) tried to solve this issue by introducing pre-training – training an equivalence class using the sequential protocol in order to prime the participants and give them experience in forming equivalence classes. After the pre-training, the simultaneous protocol is used to form new equivalence classes. This method has been somewhat successful and seems to be the best available method. The only issue with adding pre-training is that it adds to the length of time the experiment runs. As more items are added to the equivalence class, the number of trials during the testing phase of the simultaneous protocol increases, thus increasing the time of an already long experiment. It seems more research is required to find a compromise between adding a simple-to-complex pre-training phase, number of items in the equivalence class (more items lead to more trials, and more time), and the logistical challenges of extended experiments. In the present experiment the

number of items was minimized with the aim of gathering an adequate number of participants by using the simultaneous protocol, with the equivalence classes to be formed having the least number of items possible in order to properly test for nodality (6-member, 4-node classes). There were intervals where, after a set number of trials, the instructions on the computer screen informed the participants to take a small break before continuing.

Another change that might improve the formation of equivalence classes is using symbols instead of nonsense syllables. In one of their experiments, Doran and Fields (2012) replaced one stimulus within the equivalence class with a pictorial stimulus, and increased participant yield from 1 to 5. Other studies (Arntzen & Mensah, 2020; Mensah & Arntzen, 2017) have explored the effects of meaningful pictorial stimuli amongst arbitrary stimuli in equivalence classes. Arntzen and Mensah (2020) manipulated the number of pictorial stimuli (either 1, 2, or three pictorial stimuli) and the location of pictorial stimuli (C1, B2, D3 for the three picture group, C1, B2, and C1, B3 for the two picture groups) to see how these affected the establishment of three 5-member equivalence classes. They also had a control group (no pictures, just abstract shapes), and a one picture group, that had the C stimulus location changed to a meaningful picture, and the rest of the class consisting of arbitrary shapes. The no picture and one picture groups had 6.7% and 86% of participants successfully forming the equivalence classes respectively. The three picture and two picture groups respectively, 66.7% and 50% of participants were successfully able to form the equivalence classes. They concluded that the number of pictures and the location of picture stimuli has a profound effect on the establishment of equivalence classes. One of the possible reasons for why a single pictorial stimulus can enhance the formation of equivalence classes, may be that the pictorial stimulus is the ‘unique identifier’ of that class, or the stimulus that stands out in the class. This allows the pictorial stimulus to act as an anchor of sorts. So, participants only have

to link stimuli to the one pictorial stimulus in the class, rather than trying to link each stimulus to the other stimuli in the class.

While these studies have shown that picture stimuli can enhance the formation of equivalence classes, there might be variables that can affect participant responding, such as the pictures evoking different memories and emotions for participants, which transfers to the arbitrary stimuli in the class and affecting response patterns, and consequently the formation of equivalence classes.

Since presenting pictures can potentially cause confounds in experiments like this (participants can have history and experience with the contents of the picture, therefore affecting results), the decision was made to use symbol stimuli instead of nonsense syllables or pictures. There is no relevant research around the effectiveness of using symbol stimuli instead of nonsense syllables. The symbols themselves were chosen by doing a survey on a list of potential symbols, in order to minimize potential 'familiar' symbols.

Since the experiment was likely to be very long and tedious for the participants, the font color was changed from black to red in an attempt to capture and maintain participant attention for an extended period of time (changing it up from the monotonous black font in a white background all students are used to). Some studies (Belfiore et al. 1996; Camgoz et al. 2004; Gaddy, 1996) suggest that certain colors (yellow, red etc.) can capture and maintain attention, even in students with learning disabilities. This was the study done by Belfiore and his colleagues (1996), where they tested the effects of color on the reading comprehension of students with attention-deficit/hyperactivity disorder (ADHD). They found that, while color did not enhance sight-word learning, it did have an immediate effect on attention span in longer reading comprehension tasks).

This study aims to address methodological issues, as well as make adjustments that may be helpful, in order to confirm the validity of the theory of relatedness. I predicted that,

with the proposed methodological changes in place, nodality will still emerge as a driving factor in the magnitude of relatedness within a trained and established equivalence class. In other words, in the MTS preference testing, participants (who will have acquired the two trained equivalence classes) will show preference for nodally proximal comparisons given a sample and two comparisons from the same class, separated by nodal distance. Furthermore, when participants do choose the nodally distal comparison, their response time will be longer than when they choose the nodally proximal comparison (it will take longer to select nodally distant comparisons than nodally proximal ones).

Method

Participants – A total of twelve participants took part in the experiment. Of the twelve participants, two were male, and ten were female. All the participants were enrolled in psychology papers at the University of Waikato. The participants were recruited through posters that were put up around the university notice boards (see appendix A). The individuals would contact the experimenter and set up times to come in and take part in the experiment.

At the beginning of the experiment, participants were given an information sheet (see appendix B) that gave them general information about stimulus equivalence, nodality, and what they would be doing. They were also given a consent form (see appendix C) which also included options for whether they wanted to be notified of the results of the experiment, and that they could leave at any time during the experiment with no consequences. Participants were given vouchers to a local department store, that totalled to $20 New Zealand dollars. The vouchers for each participant however, were split into two $10 vouchers – One $10 voucher was given for attending the experiment, and another $10 voucher was given when

either the program was completed (due to completing all the phases or even failing the testing phase, at which point the 'experiment ended' thank you screen would pop up), or when the timer on the program (set to two hours) reached 0 (again, the 'experiment ended' thank you screen would show).

Ethics approval was worked on and gained prior to recruiting the participants, from Human Research Ethics Committee of the Faculty of Arts and Social Sciences (HREC2019#13).

Apparatus and setting – The experiment was conducted over a 2-hour session in a standard university computer lab; a room of approximately 8 x 8 meters. The lab consisted of approximately 24 computers arranged in clusters of 4. The application was run on a Dell intel core i5 7th generation computer with a 20-inch display monitor and a screen resolution of 1920 x 1080 pixels. The operating system was Windows 10.

The computer ran an application programmed to deliver the matching to sample protocol. The application displayed the sample and comparison stimuli on the monitor. The sample stimulus appeared in the middle of the screen towards the top. The comparison stimuli appeared below the sample stimulus, an equal distance from each other and the sample stimulus; the exact locations of the comparison stimuli depended on whether there were two or three comparisons. On trials with two comparison stimuli, the stimuli appeared at the bottom on either side of the sample stimuli (one to the left, and one to the right of the sample); on trials with three comparison stimuli, two would be on either side of the sample stimulus, and a third comparison would be directly underneath the sample stimulus (see Figure 2). There were no delays between the appearance of the sample stimuli and the comparison stimuli; all the stimuli appeared at the same time in a trial.

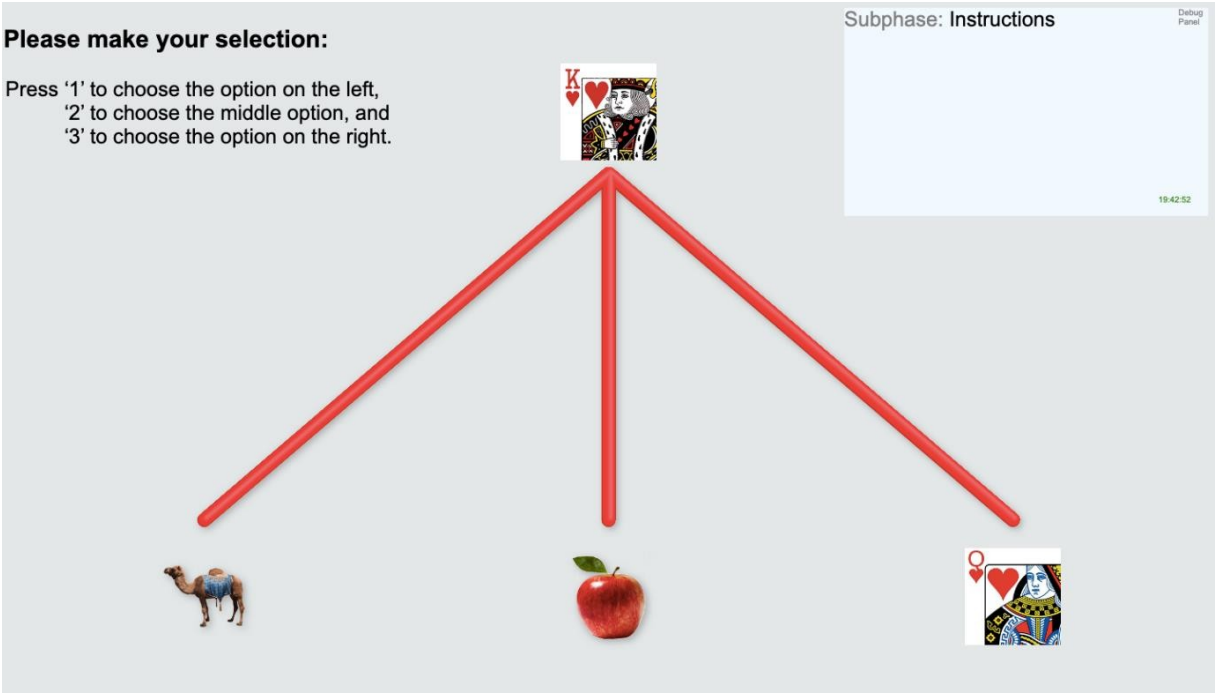


Figure 2. Layout of how the sample and comparison stimuli were displayed (middle branch removed for trials with two comparison stimuli).

Responses consisted of specific key pressing (keys 1, 2, and 3 to select left, centre, and middle options respectively, and keys W when the screen showed feedback 'WRONG', R when the feedback screen showed 'RIGHT', E when no feedback was given, or spacebar to progress to the next phase) on a standard keyboard. Keys 1, 2, and 3 corresponded to a comparison stimulus choice, and the spacebar key to start the trial, or to progress to subsequent trials. A separate program on Microsoft publisher ran concurrently with the matching to sample program, which recorded each key press for the participant, recording details including response latency, choice made, correct/incorrect, trial number and type, and phases the participants were in.

Stimuli - The stimuli used in the familiarization phase were common pictures – the sample picture was always that of a king of hearts playing card. The comparison stimuli were pictures of a camel, an apple, and a queen of hearts playing card.

Eighteen rare symbols were used for the rest of the phases (simultaneous protocol and preference testing, Table 1 and Table 2 respectively). The symbols were chosen based on surveys done on 10 students from The University of Waikato. They were presented with a table of 30 symbols, and they were instructed to circle any symbol that they felt was familiar. Through this process, familiar symbols were eliminated. Symbols were eliminated as soon as any student circled it (criteria for a symbol being eliminated was to be circled at least once by a survey taker). After this, 18 symbols were chosen from the remaining non-eliminated symbols, to form three groups of six symbols each. Of these three groups of symbols, two groups served as the two equivalence classes to be trained, while the third group of symbols served as the distractor stimuli during training and testing. All stimuli were presented in red on a white background.

Table 1.

The three groups of symbols (18 symbols in total) used to form the two equivalence classes. Symbols in classes 1 and 2 were the stimuli for the equivalence class formation training, and symbols in class 3 served as the distractor stimuli

Class 1	A1	B1	C1	D1	E1	F1
Symbol						
Class 2	A2	B2	C2	D2	E2	F2
Symbol						
Class 3	A3	B3	C3	D3	E3	F3
Symbol						

Trial format – all trials in the familiarization phase and experiment were presented in a matching to sample (MTS) format. Each trial consisted of a sample, and three comparisons, of which one was the positive comparison, with the other(s) being stimuli from different classes. The location of the positive comparison was random (left, right, or middle), with the only stipulation being that it appeared an equal number of times on the three locations.

Each phase of the experiment had blocks of trials. Blocks were a set of trials that contained the number of trials required to complete a level of the phase – e.g. in baseline training a block consisted of 45 total trials, and every time participants achieved criterion accuracy on a block, the next block would have less feedback, until there was a block of no feedback trials. The trials in a block were presented in a randomized order. Blocks of smaller number of trials were used to manage participant fatigue as well. For example, the testing phase during the simultaneous protocol had a total of 180 trials. These were split into 6 blocks of 30 trials each; after a block ended, a screen would appear informing the participant to take a rest (pressing the spacebar key would then deliver the next block of trials).

Pressing the spacebar key began the trial and a sample stimulus was displayed in the middle-top of the screen, along with the three comparison stimuli, after which specific keys could be pressed to choose a location (1 key for right side, 2 key for middle, 3 for left). Once the key was pressed and selection made, feedback was presented, which was either 'RIGHT' for the correct response, 'WRONG' for incorrect responses, and –E— for no feedback on trials in which no feedback was provided. To show that the participants had seen the feedback, R key had to be pressed to continue when the feedback was RIGHT, W key when the feedback was WRONG, and E key when there was no feedback. No feedback was

presented on some trials in some of the training blocks, and all trials in all of the testing blocks.

At the start of training trials, there was 100% feedback on all the trials in a block; this would go on until there was 100% correct responding for the block on this level of feedback. Once that was established, the next block would have the same number of trials, but feedback on only 50% of the trials. Again, once criterion (80% accuracy) was achieved, the next block would feature no feedback. Achieving 80% correct on the no feedback level would enable participants to progress to the testing level. If criterion was not met on a block with reduced feedback for three consecutive blocks, the next block would go back to the previous level of feedback (increased feedback). The training was terminated if more than 2 hours was spent on the training phase. Blocks ended with the message, 'press enter to start next block'.

Phase 1 – Instructions and familiarization – Participants saw these instructions on the screen at the beginning:

> 'Thank you for volunteering to participate in this experiment. PLEASE DO NOT TOUCH ANY OF THE KEYS ON THE KEYBOARD YET! In this experiment you will be presented with many trials. Each trial contains three or four CUES. These will be familiar and unfamiliar picture images. YOUR TASK IS TO DISCOVER HOW TO RESPOND CORRECTLY TO THE CUES. Initially, there will also be INSTRUCTIONS that tell you how to respond to the cues, and LABELS that will help you to identify the cues on the screen. The labels and the instructions that tell you which KEYS to press will slowly disappear. Your task will be to RESPOND CORRECTLY to the CUES and the INSTRUCTIONS by pressing certain keys on the computer's keyboard. The experiment is conducted in phases. When each phase ends, the screen will sometimes tell you how you did. If you want

to take a break at any time, please call the experimenter. PRESS THE SPACEBAR TO CONTINUE'.

After pressing spacebar, participants went through prompted familiarization trials – 10 to 15 trials consisting of four English words (one sample, three comparison), two of which were semantically related (e.g. KING, QUEEN, APPLE and CAMEL). Through prompts such as 'make the correct selection', and 'press R (or W, or E) to continue' the participants were familiarized with how the trials would work. These prompts were faded out over trials as long as responses were correct. Phase one ended when trials were presented without any prompts with 100% accuracy.

Phase 2 – simultaneous protocol: *baseline training*– In this phase, the simultaneous protocol was used to form two equivalence classes of four-nodes, 6 items. In this protocol, the baseline relations for all the classes were presented in one block, with 100% feedback. Each baseline relation from each class was presented three times, with the correct comparison stimulus appearing once at each of three comparison stimuli locations. For example, there were two classes of A=B, B=C, C=D, D=E, and E=F (third class being distractor stimuli, so not trained), then one training block contained a total of 30 trials (5 base relations presented 3 times for 2 classes). Training blocks were repeated until there was 100% accuracy in one full block. After that, feedback was systematically reduced (100%, 50%, 0%) per block of trials, provided that criterion accuracy was achieved every time, until 80% accuracy could be maintained in a block in the absence of feedback.

Simultaneous protocol: *derived relations testing* - Once progress was successfully made through training by achieving criterion accuracy (80%), a test was run to evaluate the trained and untrained relations in all the classes. Each relation was presented 3 times (in order to have the positive comparison appear equally at each location), from each of the classes. Each

test block contained three trials for the two classes, with the items from the third class appearing as distractor options. The maximum of 180 trials were divided into 6 blocks of 30 trials per block. The 6 blocks of 30 was one set. Within each block, the trials were presented in a random order, and the set of 6 blocks of 30 trials (180 trials total) were repeated up to 3 times, or until an accuracy of 90% was achieved in one set of 6 blocks. Response time, selection (keyboard pressing), and accuracy was recorded during testing.

Phase 3 – Within-class preference tests – Once phase 3 was completed, the within-class preference tests began. These were also matching-to-sample procedures, except that the sample, and all comparisons were from the same class. Of the comparisons from the same group, one was nodally proximal to the sample, while the other(s) were increasingly distal. The preference tests never included the stimuli in the two trained classes that were not nodes (A1, F1, A2, F2). The remaining four nodes for both classes (B1, C1, D1, E1, and B2, C2, D2, F2) were presented in probe trials (e.g. BCD probe, where B is the sample stimulus, with C and D stimuli being comparisons from the same class). All the probe combinations that were possible with the 4 nodes were presented. This worked out to be 12 different probe combinations for each of the trained equivalence classes (see Table 2 for probe combinations). There were no probes across classes (any stimulus from class 1 was never mixed with stimuli from the second trained class, and vice-versa).

Table 2.

Probe ccombinations presented in the preference testing phase. Each three-letter probe represents trials from both the trained equivalence classes (for example, the probe BCD represents both the B1C1D1 (from class 1), and B2C2D2 (from class 2) trials

Sample	Possible Probes		
B	BCD	BCE	BDE
C	CDE	CDB	CBE
D	DEC	DEB	DCB
E	EDC	EDB	ECB

Probes are identified by three letter combinations that specify which stimulus from the class is the sample (first letter), which one is the predicted or nodally proximal comparison (second letter), and which one is the non-predicted or nodally distal comparison (third letter). For example, the probe BCD indicated that the stimulus designated as B was presented as a sample, while the stimulus designated as C was the predicted/nodally proximal comparison, and the stimulus designated as D was the non-predicted/nodally distal stimulus. Since there were two options, the probes were presented twice (so that the two comparisons appeared an equal number of times in the two locations); and since there were two trained classes, that was a further two trials to accommodate the second trained equivalence class. This meant that each probe had a total of 4 trials (two trials per class, two classes), and a grand total of 48 trials presented in random order (12 probes with 4 trials per probe).

Results

Of the 12 Participants that took part in the experiment, two participants did not make it past the training phase. Both participants were able to get past the 100% feedback level of training, however they were still at the 50% feedback level of training when the experiment timed out at two hours.

Of the remaining ten participants, five participants achieved the criterion accuracy of 80% correct responding for the baseline training phase at each feedback level (100%, 50%, and no feedback). These five participants had a variable number of attempts during the 100% feedback level of baseline training, but only needed one attempt at the 50% and no feedback levels of the baseline training phase to get to the testing phase. In the testing portion of the simultaneous protocol, after three attempts at the testing level, none of the five participants had achieved the 90% accuracy required to make it to the final phase, which was the within-class preference testing phase.

The other five remaining participants all made it through the training phase, achieving the required 80% criterion accuracy for each level of feedback. They also had a variable number of attempts at the 100% feedback level, and once they made it past the 100% feedback level, they only needed one attempt at the 50% and no feedback levels of baseline training. All five participants completed the final phase (preference testing). Four of the five participants that made it to the preference testing phase only needed one attempt at the testing phase to achieve the 90% criterion accuracy, while one of the five who made it through to the end needed two attempts at the testing phase to achieve the 90% criterion accuracy. One of the participants (Participant 1) who made it through and completed the preference testing had to be excluded from data analysis, as the participant had acquired knowledge about the experiment (information that could alter their approach and attitude towards the experiment,

and consequently confound the results) before the experiment began, thus compromising the integrity of the data collected.

Figure 3 shows the percentage of trials in the preference testing phase where the participants selected the comparison choice that was predicted (or nodally proximal to the sample stimulus). The percentage of predicted (nodally proximal) comparison stimulus selections were calculated for all the preference probes, except for two probes – the CDB, and DEC probes respectively. Figure 3 only shows results from probes that involved the comparison stimuli choices differing in nodal distance (in all the probes, the only thing different between the comparison stimuli was how many nodes away each comparison was from the sample stimulus). The aforementioned probes (the CDB, and DEC probes) however, did not differ from each other nodally when compared to the sample stimulus. In terms of nodality, these two probes were predicted to have equal responding for each of the comparisons, as the comparison stimuli in both the probes are nodally equal. For example, in the CDB probe, with C as the sample, D is 0 nodal distance from C (D comes right after C) and B is also 0 nodal distance from C (B comes right before C). The only difference between the comparison stimuli in these two probes are relation complexity (or logic operations). In the aforementioned example, B and D are 0 nodal distance from C, but D is in a baseline relation with C (C=D), while B is a symmetry relation (B=C, so C=B). A separate analysis was carried out on the results from these two probes (see Figure 4).

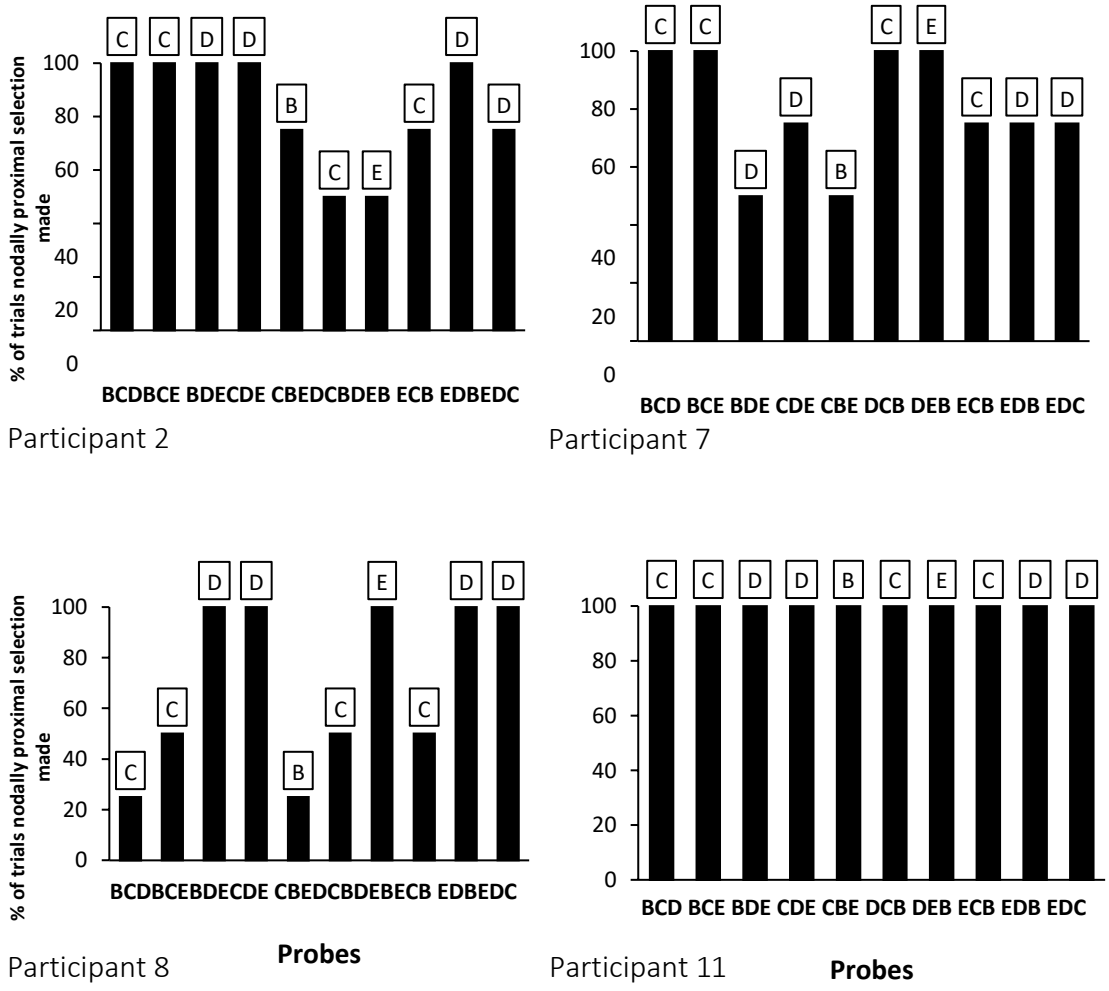


Figure 3. Percentage of trials where the nodally proximal selection was made (shown as the letter designation in the boxes on top of the probe data bars. Probes are identified as follows: First letter = sample stimulus, second letter = proximal comparison stimulus, and third letter = distal comparison stimulus.

As Figure 3 shows, in the majority of trials, participants tended to show either a full preference (selecting the nodally proximal comparison stimulus 100% of time; Participant 11) or close to full (75%, 80%, and 70% for Participants 2, 7 and 8) preference for the nodally proximal comparison stimuli. One participant (Participant 11) showed complete preference for the nodally proximal comparison stimuli, selecting the nodally proximal comparison stimulus on all the probes in the preference tests. Participant 11 also took the longest time to respond to the stimuli compared to the other three participants (average

response time of 12.84s per probe trial in the preference testing phase for Participant 11, compared to average response times of 4.49s, 2.47s, and 3.19s for Participants 2, 7, and 8 respectively). Participant 11 spent a longer time than the other participants on every trial.

The other 3 participants showed slight variances in preference. That is, they did not always choose the nodally proximal stimulus choice, selecting the nodally distal stimulus choice some of the time as well. Participant 2 selected the nodally proximal comparison stimulus an average of 82.82% of the time during preference testing. Similar averages were observed for participant 7 (nodally proximal comparison stimulus chosen 78.13% of the time), and participant 8 (nodally proximal comparison stimulus chosen 70.31% of the time). Overall, the four participants who completed the within-class preference testing, selected the nodally proximal comparison stimuli on 82.82% of the total probe trials, and selected the nodally distal comparison stimuli on 17.18% of total probe trials.

Amongst all the probes where the nodally distal comparison stimulus was selected at least once, the average time taken for the participants to select the comparison stimulus on those probe trials (the appropriate key press on the keyboard) is shown (Table 3).

Table 3.

Average Time Taken by Participants When Making Nodally Proximal versus Nodally Distal Choices in the Preference Testing Phase

Participant	Average response time (s) for nodally proximal comparison	Standard Deviation	Average response time (s) for nodally distal comparison	Standard Deviation
Participant 2	7.19	10.37	6.64	5.84
Participant 7	2.32	0.66	3.23	2.37
Participant 8	2.94	0.79	4.24	1.95

Participants 7 and 8 took longer, on average, to select the nodally distal comparison stimulus in a given probe trial. Participant 2 is the only participant that shows longer average response times for the nodally proximal comparisons than the nodally distal comparison stimuli. Further investigation shows that on the very first trial of the preference testing, most participants took much longer than average to respond, and amongst those, Participant 2 took the longest to respond on trial 1 of preference testing (65 seconds to respond to trial one). If trial 1 response time is excluded for every participant, their average response times remain similar (less time to select proximal comparison than distal comparison), except for Participant 2, whose average response time for selecting the proximal comparison drops down to 1.55 seconds (with the average response time for selecting the distal comparison remaining the same at 6.64 seconds). Participant 11 never made the nodally distal comparison stimulus selection in any of the probe trials that contained nodally proximal versus nodally distal comparisons, always selecting the nodally proximal comparison stimuli. They did however, take a lot longer to make the proximal choice than any other participant who did the

preference test (average time taken to select the nodally proximal stimuli per trial was 14.97s, SD 14.61)

In the preference testing phase, the probes that were presented had a sample, and two comparison stimuli that differed in nodal distance. However, two probes (Probes CDB, and DEC) had comparison stimuli that were nodally equal (in both the probes CDB and DEC, both comparison stimuli are 0 nodal distance away from the sample). The only difference between the two comparisons in both the probes, is direction (or complexity of relation). Directly trained baseline relations (0-node comparison, but in the forward direction from the sample) are simpler than untrained derived symmetry relations (0-node comparison, but in the backwards direction from the sample). Speaking strictly in terms of nodality, comparison stimuli that are nodally equal should evoke responding to both comparison stimuli at an approximately equal rate. In other words, participants should be selecting both comparison stimuli fairly equally. Figure 4 shows the average percentage of trials (for the 4 participants that did the preference testing) where the directly trained baseline (0-node, forward) relation was preferred.

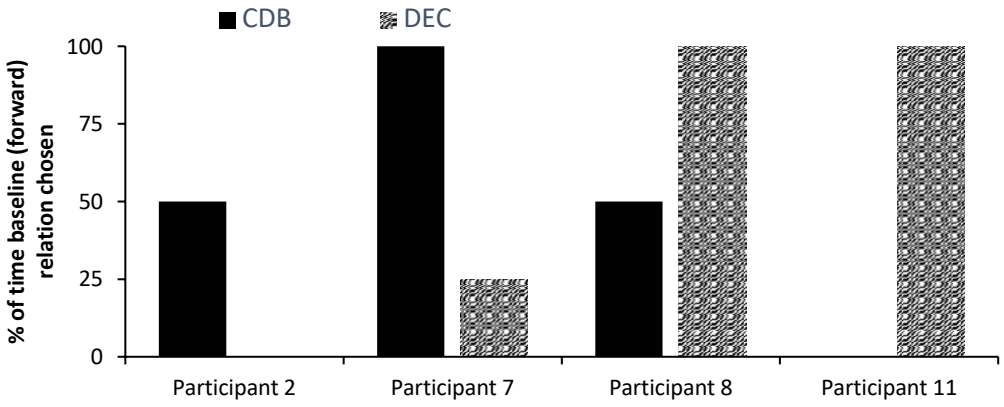


Figure 4. Percentage of trials where the baseline (0 node, forward direction) relation was selected over the symmetry (0 node, backward direction) relation for each probe that had baseline versus symmetry comparison stimuli

Participant 2 seemed to show a higher overall preference for the untrained, derived symmetry relation; selecting the trained baseline comparison stimuli on only 25% of the total trials involving the CDB and DEC probes. Furthermore, Participant 2 selected trained baseline and derived symmetry relations equally for the CDB probe trials, and showed complete preference for the derived symmetry relation on the DEC probe trials. Participants 7 and 8 show more overall preference for the trained baseline relation. Participant 7 showed complete preference for the trained baseline relation in the CDB probe trials, and only partial (25%) preference for the trained baseline relation in the DEC trials, overall selecting the trained baseline relation on a total of 62.5% of the trials involving the CDB and DEC probes. Participant 8 showed no preference for either the trained baseline or derived symmetry relation on the CDB probe trials, selecting both comparison stimuli an equal number of times (50%); and showed complete preference for the trained baseline relation in the DEC probe trials (100%), overall showing more preference for the trained baseline relation by selecting the baseline comparison stimuli on 75% of the total trials involving the CDB and DEC

probes. Participant 11 showed a complete preference for the derived symmetry relation on the CDB probe trials, but also showed complete preference for the trained baseline relation on the DEC probe trials, overall showing no preference for either the trained baseline or the derived symmetry relations, selecting them both on an average of 50% of the trials that involved the CDB and DEC probes.

Combining the responses to the baseline versus symmetry probes (CDB, and DEC) for the four participants shows that overall, the percentage of times they picked the baseline relation over the symmetry relation was very similar (Figure 5).

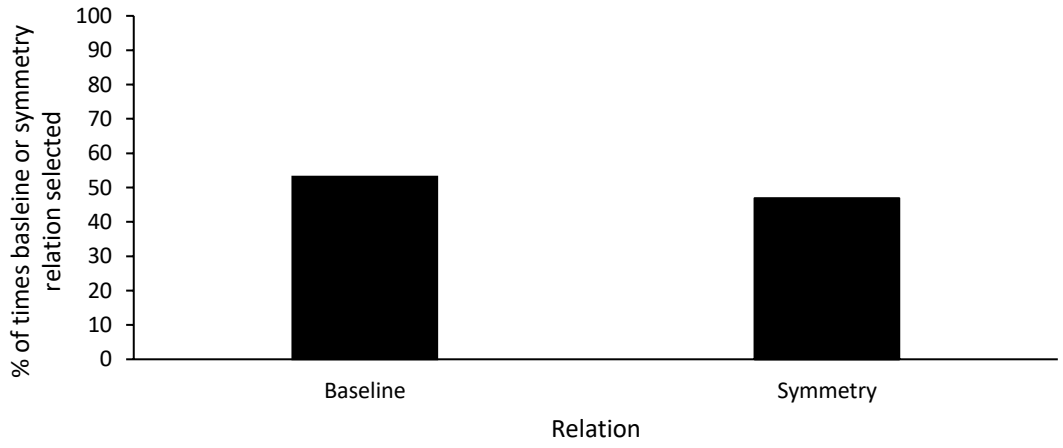


Figure 5. Average percentage of times the baseline or symmetry relation was chosen, on probe trials that had baseline versus symmetry comparison stimuli, in the preference testing phase

Amongst the 4 participants that completed the preference testing, on probes where both the comparisons were an equal nodal distance apart (0-node distance in forward and backward directions), they selected the baseline relation on 53.13% of trials, and the symmetry relation on 46.87% of the trials. Table 4 shows each participant's average response times when selecting either the baseline or symmetry relations.

Table 4.

Average Time Taken to Make Baseline (0-node forward) or Symmetry (0-node backward) Selection on Probes (CBD, DEC) that had Baseline versus Symmetry Choices

Participant	Average time taken (s) to select baseline relation	Average time taken (s) to select symmetry relation
Participant 2	1.70	2.54
Participant 7	1.31	3.46
Participant 8	2.41	1.75
Participant 11	11.94	4.41
Mean	**4.34**	**3.04**

Participant 7 took longer to select the symmetry relation, while Participant 11 took longer to select the baseline relation. Participants 2 and 8 had almost identical average response times for both the baseline and symmetry relations (0.84 seconds and 0.66 seconds difference between average response times for baseline and symmetry choices, for Participants 2 and 8 respectively), however Participant 2 took slightly longer to select the symmetry relation, while Participant 8 took slightly longer to select the baseline relation. Participant 11 had much higher average response times than the rest of the participants, as Participant 11 took much longer to respond in general for every trial in every phase than every other participant. Overall, the participants took an average of 4.34 seconds to make the baseline comparison stimuli selections, and an average of 3.04 seconds when making the symmetry comparison stimuli selections. The average times for Participant 11, however, had a big influence on the average times for the baseline selections (their average response time

of 11.94 seconds was much higher than all the other average response times, bringing the combined average response time for the baseline selections up as a result). This in turn caused a much higher difference between the two total average response times (a difference of 1.3 seconds). Accounting for the response times of Participant 11 by removing them from the calculations, the differences between the two average response times decreased (1.81s average response time for baseline selections; 2.58s average response time for symmetry selections; difference of 0.72s).

Discussion

Nodality

The results of this experiment appear to support the findings made in other studies on nodality, specifically the contention that all the stimuli in an equivalence class are equal, in terms of being substitutable by each other, but in certain contexts, they are also differentially related, under the control of nodal effects, where the strength of the relation between two stimuli in an equivalence class get weaker as the number of nodes between them increase (Alligood, 2007; Moss-Lourenco & Fields, 2011).

Generally, there has been a view that, when an equivalence class has been fully established, all the stimuli that participate in that equivalence class are equal in that, all the stimuli in the equivalence class have the same function and are substitutable by each and every member in the established equivalence class (McIlvane & Dube 2003; Sidman, 1990, 1992, 1994, 2000). If indeed all stimuli within an established equivalence class are substitutable by one another, then on tests such as the within-class preference testing (like the one presented in this study), it would be expected that responding would be approximately equal towards the comparison stimuli. In other words, participants would be expected to

choose either of the two presented comparison stimuli on approximately 50% of all the preference testing trials (if both the comparison stimuli from the same equivalence class are 'equal', then choosing either one would have the same effect). Participants in studies on nodality (including this study), however, have responded in a manner consistent with the effects of nodality among stimulus-stimulus relations in an established equivalence class (Alligood, 2007; Fields, 2016; Moss-Lourenco & Fields, 2011).

Priming Effects

In this study, participants tended to show preference for stimuli that were nodally proximal to the sample stimulus, compared to the nodally distal comparison choice. Specifically, the participants who completed the within-class preference testing, picked comparison stimuli choices that were nodally closer to the sample stimuli on an average of 82.82% of the total trials, picking the nodally distal comparison stimuli on 17.18% of total trials. This is consistent with the findings in the study by Moss-Lourenco and Fields (2011). They trained participants to establish two 3-node, five member equivalence classes using the simultaneous protocol training method. They then used the post class formation within-class preference testing procedure to check if there was stimulus control by nodality. Participants responded in a manner that demonstrated control by nodal distance, that is, they consistently showed preference for comparison stimuli that were nodally proximal to the sample stimuli compared to the nodally distal comparison stimuli. They, however, structured the preference testing in order of nodal spread. They structured the within-class preference testing such that, probes that had the largest nodal spreads were presented first. In other words, the probes which contained two comparison stimuli that were the highest nodal distance towards and away from the sample stimulus were presented first. For example, in the 3-node five member equivalence class (A=B=C=D=E), the ABE, and EDA probe trials were presented first, as the

comparison stimuli in those two probes were the most different from each other in terms of nodal distance from the sample (in the ABE probe, B is 0 nodal distance from A, and E is 3 nodal distance from A; while in the probe EDA, D is 0 nodal distance from E, and A is 3 nodal distance from E).

The issue with presenting probes with extreme nodal differences first is that there might have been priming effects, where the nodally proximal selections were made first as they were the most similar to the sample stimulus, then on subsequent probes, participants may have continued making selections based on whichever comparison stimulus was 'most similar' (nodally proximal) to the sample stimuli; participants may have been 'trained' to make the nodally proximal choices. This issue was alluded to by Moss-Lourenco and Fields (2011) themselves, acknowledging that similarity to a trained relation is indexed by nodality, and as such, initially presenting extreme nodal distance difference probes in the within-class preference testing may have 'strengthened' and trained (primed) the participants to respond to every other preference probe in such a way (participants may have assumed that the within-class preference testing phase was a test where the correct answers were the comparison stimuli selections 'most similar' (nodally proximal) to the sample stimuli, and started responding accordingly).

Furthermore, presenting a pre-determined and a limited number of preference probes meant that there was no random presentation of probes – further exacerbating the priming issue. Presenting all the possible preference probe trials in a random order may minimise the priming or training of a certain response style – something that was done in this study. Even with these changes, the results of this experiment still demonstrated within-class stimulus control by nodal distance in established equivalence classes. Figure 3 shows how each of the four participants' responded to the probes in the within-class preference testing phase. Participant 11 demonstrated total response control by nodal distance, by preferring the

nodally proximal comparison stimuli 100% of the time. Participants 2, 7, and 8 showed some variance in responding, but still preferred the nodally proximal comparison stimulus 82.82%, 78.13%, and 70.31% of the time respectively. Despite taking measures to minimise the effects of priming in the within-class preference test, the participants in this study still demonstrated response control under the effects of nodality, by consistently (and in Participant 11's case, completely) selecting comparison stimuli that were nodally closer to the sample stimuli in most trials.

Unequal Reinforcement

One of the explanations offered by Imam (2001; 2006) on the relatedness of stimulus-stimulus relations in equivalence classes was that nodality does not exist, but rather the response pattern observed is merely a function of unequal reinforcement and a disproportionate volume of trials presented during training. The simultaneous protocol (used in this study) minimizes the differential reinforcement of trained baseline relations by, as the name suggests, simultaneously presenting all the baseline training trials in a single block of trials. Due to not training each baseline relation one after the other in a linear fashion, it becomes much harder for participants to establish equivalence classes. Indeed, in this study, only 5 of 12 participants (41.67%) made it through and completed the within-class preference testing phase.

In this study, the unequal reinforcement and trial presentations were minimized by utilizing the simultaneous training protocol. In this method, all the baseline relations to be trained were presented randomly, and in one sequence (not one after another in a serial order). Amongst the 5 participants who made it through to the within-class preference testing in this study, the results still show support for the theory that stimulus-stimulus relations in an

established equivalence class may be under the control of the nodal distance topography. That is, the participants still tended to prefer nodally proximal comparison stimuli.

Differential Functions of Nodes and Singles

In recent studies on nodality and logic operations, the sizes of the classes that have been established have been quite varied. For example, Alligood (2007) used the simple-to-complex training method to establish two 6-member, 4-node equivalence classes, Doran and Fields (2012) used the simultaneous protocol to establish two 7-member, 5-node equivalence classes, while Moss-Lourenco and Fields (2011) used the simultaneous protocol to establish two 4-member, 2-node equivalence classes. They all also utilized the within-class preference testing to assess the effects of nodality or logic operations. The one common issue amongst these, was that the within-class preference testing probes contained stimuli that had differing functions during training. In a linearly-trained equivalence class, the first and last stimulus in the class are not nodes (which is why the number of nodes in a linearly-trained equivalence class is always two less than the number of items in the class). They are in fact singles, meaning they serve only as either a sample (the first stimulus) or a comparison (the last stimulus), and are only related to one other stimulus (nodes are related to two or more stimuli, meaning that they serve as both sample and comparisons during training). This is problematic during preference testing, where participants may reject the stimulus on the grounds that it has never appeared as a comparison (or sample) before.

In this study, only the nodes in the two established 6-member, 4-node equivalence classes were presented during the within-class preference testing. The singles (stimuli A and F from both the established equivalence classes) were removed from preference testing to ensure that selections were based on stimuli that had the same role (nodes) in the equivalence

classes. With this alteration, the results of this study still show support for the effects of nodality on stimulus-stimulus relations in established equivalence classes.

Response Latencies for Nodal Distance Probes

Upon further investigation, it was revealed that even though participants occasionally selected the nodally distal choice, they took longer to do so (see Table 3). This is in support of previous studies that have explored the effects of nodality through response speeds (Spencer & Chase, 1996; Wang et al., 2011). In the study by Wang and colleagues (2011), they had participants establish two 5-member, 3-node equivalence classes, and recorded response speed during derived relations testing. Their results showed that participants responded in line with the effects of nodality. That is, the participants responded fastest to comparison stimuli that were nodally proximal, and the response speeds increased as the nodal distance of the comparison stimuli increased.

In this study, Participants 7 and 8 took 2.32 and 2.94 seconds respectively to pick the nodally proximal comparison stimuli on average per trial, which was faster than the time it took them to make the nodally distal choices (3.23 and 4.24 seconds average response time per trial when making the nodally distal comparison stimuli choices for Participants 7 and 8 respectively). This suggests that the longer response times are indicative of more intervening behaviour (the more nodes in between the stimuli, the longer it took to respond). Participant 2 did have an average response time that was longer for the proximal comparison stimuli (7.19s average response time per trial) than the distal comparison stimuli (6.64s average response time per trial). However, a closer look at their data revealed that, when the program shifted from the equivalence testing phase to the within-class preference testing phase, they had a long pause while the first preference trial was already on the screen. They later reported that when the within-class preference testing began, they took a while to adjust to the new phase.

They eventually selected the proximal comparison stimulus on that trial, recording a very long response time (65 seconds), which was incongruent with the response times for the rest of the within-class preference test trials for the participant. Excluding that high response time from the analysis revealed response times consistent with the rest of the participants, that is, Participant 2 also tended to take longer to respond when making nodally distal comparison stimuli selections (average response time of 6.64s per trial) than when making nodally proximal comparison stimuli selections (average response time of 1.55s per trial).

The response times of Participant 11 could not be analysed, as they never made a nodally distal comparison stimulus selection, so there was no time to compare the average response time for the nodally proximal comparison stimuli (14.97s) selections. This meant that Participant 11 selected the comparison stimulus that was nodally proximal to the sample stimulus 100% of the time (see Figure 2). In contrast to the rest of the three participants though (Participants 2, 7 and 8), Participant 11 also had the highest average response time per probe trial in the within-class preference testing phase (12.84s per probe trial for Participant 11, compared to 4.49s, 2.47s, and 3.19s for Participants 2, 7, and 8 respectively). The longer response times for Participant 11 suggests that they were thinking a lot more about their responses in an attempt to get the 'correct' answer, suggesting that there was a 'good participant' effect. This was confirmed by the participant themselves; after the experiment, they revealed during debriefing that they were motivated to get everything correct, and this mentality extended into the within-class preference testing phase (there is no mention about getting selections right or wrong either in the information sheet or during explanations by the experimenter), where they thought there was always a 'correct' choice to make. In the within-class preference testing phase, because there were no wrong selections, they reported that they felt the comparison stimulus most 'closely associated' with the sample stimulus was the correct choice in every trial. The fact that both choices were from the same class (both

correct), made them pause and think a lot more about what selections they were making, and ultimately led to them choosing the most closely associated comparison stimuli to the sample stimuli (nodally proximal stimuli), but also increasing their response time.

Baseline versus Symmetry (Nodality or Relation Complexity)

Response Accuracy

During the within-class preference testing phase, the probes presented had a sample stimulus and two comparison stimuli that were a) from the same established equivalence class as the sample stimulus, and b) differed from each other only in terms of their nodal distance from the sample stimulus – one comparison stimulus was always nodally proximal to the sample stimulus, and the other comparison stimulus was nodally distal to the sample stimulus. Two of the preference probes, however, did not have such a dynamic (probes CDB, and DEC). Instead, these two probes had a sample stimulus, and two comparison stimuli that were an equal nodal distance from the comparison stimulus. The only difference here, was that the two comparison stimuli had different relation complexity to the sample stimulus. For example, the probe CDB had a sample stimulus of C, with both the B and D comparison stimuli being 0 nodal distance from C. However, D was a directly trained baseline relation (C=D), while B was a derived, non-reinforced symmetry relation (B=C trained, then C=B symmetrically derived).

If the participants' responses were a direct function of nodal distance, then upon the presentation of the trials of these two probes, they would be expected to pick either of the comparison stimuli on approximately 50% of the trials. In other words, participants would be expected to pick both comparison stimuli about the same number of times (as there is no nodal distinction between the comparison stimuli in the CDB and DEC probes). If however,

as Doran and Fields (2012) postulated, relation complexity may also be influencing responding, and therefore affecting the relatedness of stimulus-stimulus relations in established equivalence classes, then responding will be biased towards the simpler, directly trained relation.

In their study, Doran and Fields (2012) had participants form two 5-node, seven member equivalence classes, using nonsense syllables and one picture stimulus in each class. They administered within-class preference testing, with a select number of probes that contained baseline versus symmetrical comparison stimuli, and transitive versus equivalence comparison stimuli. They found that participants showed preference for trained baseline over derived symmetry relations, and transitive over equivalence relations. Preference for the simpler relation, especially directly trained baseline relations over derived symmetry relations, make more sense when considering what Imam (2001; 2006) had to say about differential reinforcement – by design, baseline relations are the only stimulus-stimulus relations in an equivalence class that are directly trained, and therefore are the only relations that get conditionally reinforced (phase 1 of the MTS procedure usually involves presenting all the baseline relations, and reinforcing all the correct selections through points, feedback screen, pleasant sounds and so on). The derived symmetry relations are never reinforced, and are only presented during testing, meaning they are presented far less than baseline relations. According to Imam (2001; 2006), differential presentation and reinforcement would lead to responding as a function of reinforcement – the baseline relations should be preferred more often as they are presented more, and reinforced more than the derived symmetry relations.

The results of this study, though, seems to contradict at least one part of the Doran and Fields (2012) study. Overall, the participants in this study appeared to be quite neutral to the baseline versus symmetry probes (CDB, and DEC); selecting the trained baseline relation on 53.13% of trials, and the derived symmetry relation on 46.87% of the baseline versus

symmetry probe trials (Figure 5). While this result contradicts the results obtained by Doran and Fields (2012), it aligns with the results found by Alligood (2007). In that study, Alligood 2007) used the delayed MTS, simple to complex (STC) procedure to train participants to form two 4-node, six member equivalence classes. The participants who completed the within-class preference testing (5 participants) showed responses as a function of nodality; even in the baseline versus symmetry probes, they showed as much or more preference for the derived symmetry relations (0-node) than they did to the directly trained baseline relations (0-node). These results seem to suggest that nodal distance plays a core role in the stimulus-stimulus relations of established equivalence classes; even more so than relation complexity. More studies may be needed, however, to properly assess how nodality and relation complexity interact and exert influence in trained equivalence classes. For example, in the Doran and Fields (2012) study, they presented only a select number of probes (removing the opportunity for random preference probe presentation); they also started the within-class preference testing with extreme difference probe trials, the same as in the Moss-Lourenco and Fields (2011) study on relatedness in equivalence classes. As mentioned previously, these extreme difference probes may be priming participants to respond in a certain way (choosing nodally proximal, less complex relations). Even in the Alligood (2007) study, the MTS procedure used to establish the equivalence classes was the simple to complex type. This meant that there was unequal reinforcement and disproportionate volume of trials presented – in a simple to complex procedure, the earliest trained relations get presented first, and by the end they are presented more frequently than the relations that were trained last. In the current study, all preference probes were presented, and in a random order, while the training procedure used was the simultaneous protocol, where all the trials are presented randomly and at the same time in one continuous block, minimising unequal reinforcement. With these alterations, the results show that participants tended to select

trained baseline and derived symmetry comparison stimuli on approximately 50% of the baseline versus symmetry probe trials.

Looking closely at individual participant responses to the two baseline versus symmetry probes shows a scattered response pattern for most of the participants (Figure 4). Participant 2 showed a higher overall preference for the untrained derived symmetry relation (75%). Participants 7 and 8 showed more preference for the trained baseline relations (62.5% and 75% respectively). Participant 11 was unique in that, they showed complete preference (100%) for the trained baseline relation on the DEC probe trials, but then showed a complete preference for the opposite relation (derived symmetry relation) for the CDB probe trials, bringing overall preference to 50% for baseline and symmetry relations. As mentioned before, however, when combining all the responses for the four participants to all the baseline versus symmetry probe trials, the results show an approximately equal percentage of responding towards both the trained baseline and derived symmetry relation. One reason for the scattered response patterns for individual participants may be the number of trials. There were only two probes (CDB, DEC) that were testing baseline versus symmetry relations. There were four total trials in a probe (2 trials for the first equivalence class, 2 trials for the second equivalence class), meaning that each participant was only responding to 8 total trials involving the baseline versus symmetry relations. Had there been a higher volume of trials, one possibility is that the responding pattern would fall more in line with the overall percentages that were obtained in Figure 4. One of the reasons the results in Figure 4 looks more in line with responding based on the function of nodality, is because, when the trials of all four participants are combined, there is a larger volume of trials to analyse; meaning that even though responding may have been biased towards one relation early on, as more and more trials were added, the response pattern stabilised and trended towards a pattern that

suggests nodality has a bigger influence on stimulus-stimulus relations in established equivalence classes than relation complexity.

Response Latencies for Baseline versus Symmetry Probes

The average time each participant took to respond when choosing either the trained baseline comparison stimuli or the derived symmetry comparison stimuli further supports the notion that participants' responses in this study were a function of nodal distance. Participants 2 and 7 took longer to make the derived symmetry comparison choice (2.54 seconds and 3.46 seconds on average when making the symmetry comparison stimulus selection for Participants 7 and 8 respectively) than when making the trained baseline comparison choice (1.7 seconds and 1.31 seconds for Participants 7 and 8 respectively). However, the opposite was true for Participants 8 and 11, as they took longer when making the trained baseline comparison stimuli choice than when they were making the derived symmetry comparison stimuli choice (2.41s and 11.94s for the baseline selections, and 1.75s and 4.41s for the symmetry selections for participants 8 and 11 respectively). The lack of a pattern (half of the participants took longer to make the baseline selections, while the other half of the participants took longer to make the symmetry selections) suggests that, overall, the response of the participants was indeed under the influence of nodal distance – when the nodal distance is equal, the responding is expected to be approximately equal in terms of accuracy, and in terms of average time taken to respond to either of the baseline or symmetry comparison stimuli.

Theoretical Integration

While the results of the current study extend the support for the theory of stimuli in an equivalence class being differentially related, it does not mean that stimuli in an equivalence class do not function similarly and are therefore not substitutable. For example, Moss-Lourenco and Fields (2011) within-class preference tests to assess the effects of nodality, and also administered the derived relations test (across-class) after the preference tests to check the intactness of the established equivalence classes. As mentioned earlier, the results of the preference tests revealed responding as a function of nodal distance (preference for nodally proximal comparison stimuli), and the derived relations tests showed that the equivalence classes were still intact (participants could distinguish the stimuli in one equivalence class from the other). This meant that while the members of the equivalence class were intact (all members belonged to that group, and not any other group), they were related differentially within the group itself (nodal distance effects). Indeed, Sidman (2000) pointed out the contradiction, suggesting that, by definition, an equivalence class cannot be equal and unequal at the same time. He suggested that nodality may merely be a result of the reinforcement contingencies, and structural variables when establishing equivalence classes.

Sidman (2000) explained that equivalence classes could still be formed even though two baseline relations may not share any common sample or comparison stimuli. A participant was trained to select B1 in the presence of A1, and B2 in the presence of A2 (A1=B1; A2=B2). The participant was then trained (same reinforcer) to select C1 in the presence of D1, and C2 in the presence of D2 (D1=C1; D2=C2). This effectively meant that the two relations (A=B; D=C) had no conditional or discriminative stimuli in common. The participant was still able to demonstrate the establishment of equivalence classes, for example matching A1 to C1 and A2 to C2. There was nothing common between the two trained

relations that would enable them to be in an equivalence class, except for the reinforcer. Essentially, the reinforcer was the mediator, part of the equivalence class that was linking the two unrelated relations.

Sidman (1994) explained equivalence classes using the bag analogy; where an equivalence class is said to be like a bag containing ordered pairs of events that the contingency specifies (if A1, then B1, B1, then C1 and so on); the bag can be shaken and elements mixed without regard for spatial or temporal relations. In other words, when an equivalence class is established, the stimuli that are in the class have a common feature – their class membership. This means that this group of stimuli can be differentiated from every other stimulus based on their class membership, and other stimuli not having such membership (across-class distinction). One consequence of such an analogy, is what is happening when taking a closer look inside the bag/class (within-class stimulus-stimulus relations). Sidman (2000) noted that structural variables within the equivalence class may be differentiating the members of an equivalence class from each other, by directionality (relation complexity) and nodality. It appears then, that Sidman's explanation is that nodality is the result of structural variables; similar to what Imam (2001; 2006) has suggested, when he stated that variables such as unequal reinforcement (due to the set-up of the training), unequal volume of trials presented, as well as preference testing containing stimuli of differing functions (nodes versus singles).

Several researchers have argued that nodality may not be a product of training design, but rather a defining characteristic of equivalence class membership (Alligood; 2007; Doran & Fields, 2012; Fields, 2016; Moss-Lourenco & Fields, 2011). Sidman (1994; 2000) did not dismiss nodality, but rather suggested that including nodality as a characteristic of an equivalence class contradicts the very definition of stimulus equivalence. Advocated of relatedness, however, argue that nodality should be seen as an important characteristic that

defines an equivalence class, because nodality not only appears to be present whenever equivalence classes are established, but also influences how participants respond to stimuli within the formed equivalence classes, as well as when forming new equivalence classes (Fields et al., 1997; Fields 2016; Moss-Lourenco & Fields, 2011).

In fact, as mentioned previously, the advocates of relatedness have argued that stimuli in an equivalence class are indeed equal, but while interchangeability may be the primary defining characteristic of an equivalence class, it is not the only defining characteristic – In certain contexts, there is a clear distinction between the stimuli that are participating in an equivalence class, and all other stimuli that are not in that equivalence class. This is usually shown through the derived relations testing, where participants distinguish between comparison stimuli that either belong to the class (in the same class as the sample stimulus), or don't (comparison stimuli that either came from the second class or the distractor). However, in other contexts, stimuli in the same equivalence class show differing levels of relational strength – and this difference in strength always follows the pattern of nodality. This becomes evident in the within-class preference testing, where the stimuli presented are all from the same equivalence class. Participants tend to show preference for the stimuli that are closest in nodal distance to the sample. In other words, there is an inverse nodality relationship within an equivalence class.

Limitations

There were several limitations associated with this study. One limitation was that of equivalence class size. In this study, two equivalence classes were trained, containing 6 items each (4 nodes in a class), with a third group of 6 items serving as distractor stimuli. A larger class size would have allowed for a more thorough testing for nodality, allowing for more probes and a larger volume of within-class preference testing trials. Participant yield is often

low when it comes to formation of equivalence classes, especially when using the simultaneous training method. In fact, it appears that equivalence class formation under the simultaneous protocol is generally an inverse function of class size (Fields et al. 2000). However, the simultaneous training method is important, as it allows for equal and random trial presentations during baseline training and derived relations testing, minimising differential reinforcement of relations and disproportionate presentations of trials. Adding more items to a group of items that is to be established as an equivalence class would only make it that much more difficult for participants to successfully form equivalence classes under the simultaneous training method (there are more unique relations in a larger class than in a smaller one). In the present study, one of the methodological refinements was to exclude the singles (the first and last stimulus in the equivalence class), as they only served one function (either only a sample, or only a comparison), and may influence responding in derived and within-class preference tests. A consequence of this, is that a class with six members (as in this study) only ends up with four stimuli to present in the within-class preference testing phase. In the current study, this meant that, while the available number of nodes were adequate to test for nodality, it was not possible to present probes where for example, a two node transitivity relation was tested against a three or four node transitivity relation.

Another limitation in this study was regarding participants. Experiments were run one participant at a time, however on two of the experiments, due to participant availability, experiments were run simultaneously with two participants in the same room (but on separate computers). This may have changed the way these participants were responding to the experiment. In particular, Participant 8 (who completed the experiment early as they finished the within-class preference testing) may have affected how Participant 9 was responding to the experiment. Participant 9 appeared to have been doing well in terms of training and

establishing the equivalence classes; however when they saw Participant 8 leave, they started responding very quickly and randomly. They eventually failed to get past the derived testing portion of the simultaneous protocol, and the experiment ended. Had Participant 8 not been there (and left early), Participant 9 likely would have made it through to the within-class preference testing phase. On the other occasion where there were two participants in the same room, neither of the participants made it past the training phase, with both the participants (Participants 3 and 4) taking the full two hours of the experiment before the experiment timed out. There was also the first participant, who became aware of information that completely changed their response style and compromised their data – during briefing and introductions, they found out that there were two groups of stimuli that were meant to go together, and so during training and testing, participant 1 knew to look for stimuli that went together in separate groups. It is unclear however, whether they would have made it to the within-class preference testing phase had they not known what they knew prior to starting the experiment (they spent quite a long time in the baseline training phase, needing 12 attempts before making it to the derived relations testing part of the simultaneous protocol). Having to exclude the data for Participant 1 decreased yield on a training method already well known for having a low participant yield.

Finally, due to time constraints, there was no derived relations re-testing at the end of the within-class preference testing to check for the intactness of the equivalence classes that were established. The derived relations test consisted of a total of 180 trials, and was the longest section of the experiment. Doing the derived relations testing again at the end of the experiment would not be possible to do in one session, and time, logistics, and participant availability meant that having longer sessions or spreading sessions out over different days would not be feasible. It is arguable though, that progressing past the initial derived relations testing showed that equivalence classes were indeed intact – an accuracy of 90% was

required to progress, meaning that there was little room for error, and that even if a participant made it past baseline training by chance, they would be highly unlikely to progress past the derived relations testing phase.

Future considerations

Increasing the size of the equivalence classes may lead to a diverse range of preference testing. For example, there have not been any studies on relatedness in equivalence classes that have preference probes containing comparison stimuli that pit nodal distance against relation complexity. For example, if a nine member, 7-node class were to be established (A=B=C=D=E=F=G=H=I), a possible preference probe could be an ECH probe. In the ECH probe, the comparison stimulus C, would be a one node equivalence relation to E, while the comparison stimulus H would be a two node transitive relation to E. Effectively, this would be a nodality versus relation complexity probe; if preference for C is shown despite being the more complex of the two relations (equivalence relation, compared to the simpler transitive relation between H and E), then it could be inferred that nodality exerts a greater influence in the stimulus-stimulus relations within equivalence classes. If however, preference is shown for H, despite it being a longer nodal distance from E (2 nodal distance away compared to C, which is one nodal distance away), then it can be inferred that relation complexity plays a more important role in the stimulus-stimulus relations in equivalence classes. The issue with increasing class size is that it adversely affects participant yield. However, there have been a number of methods that seem to have increased participant yield over the years – including strategies such as adding a unique 'anchor' stimulus in each equivalence class, like a pictorial stimulus amongst nonsense syllable stimuli (Arntzen & Mensah, 2020; Doran & Fields, 2012; Mensah & Arntzen, 2017), and including a simple-to-complex pre-training procedure to form equivalence classes, before establishing the target equivalence classes using the

simultaneous protocol (Moss-Lourenco & Fields, 2011). These strategies combined may allow for the formations of larger equivalence classes, making it possible to present a diverse range of preference testing probes.

Another interesting factor to consider may be the further distinguishing of separate equivalence classes by having separate reinforcers for each class during baseline training. For example, during baseline training, the first class of stimuli could be reinforced with points, while the second class of stimuli could be reinforced with a pleasant sound. This would further separate the two classes, and therefore make it easier to form the equivalence classes, even in the simultaneous protocol.

Summary

In summary, the current study attempted to replicate and extend previous work on the stimulus-stimulus relations within established equivalence classes. The results suggest that stimuli in an equivalence class are substitutable by each other in certain contexts (derived relations testing), however in other contexts, stimuli in an equivalence class are differentially related to each other as a function of nodal distance (within-class preference testing). Furthermore, and in line with the theory of nodality, 0-node symmetry relations were preferred by participants just as much as 0-node baseline relations. These findings support the existence of differential relatedness of stimuli in established equivalence classes, in the right contexts.

In the current study, some methodological refinements were made from previous studies on relatedness in stimulus equivalence (minimise unequal reinforcements, priming, and removal of stimuli with differing functions). The simultaneous protocol was used to establish two 6-member (4 node) equivalence classes, after which within-class preference tests were administered. Participants that made it through to the within-class preference

testing responded in a pattern suggesting response control as a function of nodality, where, when given a choice between a nodally proximal and nodally distal comparison stimuli, participants most often preferred the nodally proximal stimuli choices. Their response times also indicated nodality, selecting the nodally proximal comparison stimuli quicker than when selecting nodally distal comparison stimuli. Their response times were also inconsistent when selecting between stimuli that did not differ in nodal distance (baseline versus symmetry). The current study also raised more questions, and possible future research avenues, that may hopefully add to the understanding and application of equivalence relations and nodality.

References

Adams, B. J., Fields, L., & Verhave, T. (1993). Formation of generalized equivalence classes. *The Psychological Record*, *43*(4), 553-566.

Alligood, C. (2007). Choice among stimuli in equivalence classes. Unpublished doctoral dissertation, Elbery College of Arts and Sciences at West Virginia University.

Arntzen, E. (2006). Delayed matching to sample: Probability of responding in accord with equivalence as a function of different delays. *The Psychological Record*, *56*(1), 135-167.

Arntzen, E. (2012). Training and testing parameters in formation of stimulus equivalence: Methodological issues. *European Journal of Behavior Analysis*, *13*(1), 123-135.

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*.

Barnes-Holmes, Y., Barnes-Holmes, D., & McHugh, L. (2004). Teaching derived relational responding to young children. *Journal of Early and Intensive Behaviour Intervention*, *1*(1), 3.

Belfiore, P. J., Grskovic, J. A., Murphy, A. M., & Zentall, S. S. (1996). The effects of antecedent color on reading for students with learning disabilities and co-occurring attention-deficit/hyperactivity disorder. *Journal of Leaning disabilities*, *29*(4), 432-438

Bentall, R. P., Jones, R. M., & Dickins, D. W. (1999). Errors and response latencies as a function of nodal distance in 5-member equivalence classes. *The Psychological Record*, *49*(1), 93-115.

Bortoloti, R., & De Rose, J. C. (2009). Assessment of the relatedness of equivalent stimuli through a semantic differential. *The Psychological Record*, *59*(4), 563-590.

Bortoloti, R., & de Rose, J. C. (2011). An "Orwellian" Account of stimulus equivalence. Are some stimuli "more equivalent" than others?. *European Journal of Behavior Analysis*, *12*(1), 121-134.

Camgöz, N., Yener, C., & Güvenç, D. (2004). Effects of hue, saturation, and brightness: Part 2: Attention. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur*, *29*(1), 20-28.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*. http://doi.org/10.1017/S0140525X01003922

Doran, E., & Fields, L. (2012). All stimuli are equal, but some are more equal than others: Measuring relational preferences within an equivalence class. *Journal of the Experimental Analysis of Behavior*, *98*(3), 243-256.

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.

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Eperimental Analysis of Behavior*, *48*(2), 317-332.

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

Fields, L., Reeve, K. F., Adams, B. J., & Verhave, T. (1991). Stimulus generalization and equivalence classes: A model for natural categories. *Journal of the Experimental Analysis of Behavior*, *55*(3), 305-312.

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*(2), 129-145.

Fields, L., Reeve, K. F., Rosen, D., Varelas, A., Adams, B. J., Belanich, J., & Hobbie, S. A. (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*(3), 367-389.

Fields, L., Varelas, A., Reeve, K. F., Belanich, J., Wadhwa, P., Derosse, P., & Rosen, D. (2000). Effects of prior conditional discrimination training, symmetry, transitivity, and equivalence testing on the emergence of new equivalence classes. *The Psychological Record*, *50*(3), 443-466.

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

Fields, L., & Moss, P. (2007). Stimulus relatedness in equivalence classes: Interaction of nodality and contingency. *European Journal of Behavior Analysis*, *8*(2), 141-159.

Fields, L. (2016). Stimulus relatedness in equivalence classes, perceptual categories, and semantic memory networks. *European Journal of Behavior Analysis*, *17*(1), 2-18.

Gaddy, M. L. (1996). Studying from highlighted material: the role of color in memory for text. Unpublished masters thesis, San Jose State University, San Jose, CA.

Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In *Handbook of research methods in human operant behavior* (pp. 229-262). Springer, Boston, MA.

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.

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

Imam, A. A. (2006). Experimental control of nodality via equal presentations of conditional discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of the Experimental Analysis of Behaviour, 85*(1), 107-124.

Jenkins, J. J. (1963). Mediated associations: Paradigms and situations. In *Conference on Verbal Learning and Verbal Behavior, 2nd, Jun, 1961, Ardsley-on-Hudson, NY, US.* McGraw-Hill Book Company.

Jenkins, J. J., & Palermo, D. S. (1964). Mediation processes and the acquisition of linguistic structure. *Monographs of the Society for Research in Child Development*, 141-169.

Katz, J. S., Wright, A. A., & Bodily, K. D. (2007). Issues in the comparative cognition of abstract-concept learning. *Comparative Cognition & Behavior Reviews, 2*, 79.

Kuno, Hiromichi, Tsutomu Kitadate, and Takashige Iwamoto. (1994). "Formation of transitivity in conditional matching to sample by pigeons." *Journal of the Experimental Analysis of Behavior* 62(3), 399-408.

Lazar, R. (1977). Extending Sequence-Class Membership with Matching to Sample. *Journal of the Experimental Analysis of Behavior*, *27*(2), 381-392.

Lynch, D. C., & Cuvo, A. J. (1995). Stimulus equivalence instruction of fraction-decimal relations. *Journal of Applied Behavior Analysis*, *28*(2), 115-126.

Magnotti J.F., Goodman A.M., Katz J.S. (2012). Matching to Sample Experimental Paradigm. In: Seel N.M. (eds) Encyclopedia of the Sciences of Learning. Springer, Boston, MA

McIlvane, William J., and William V. Dube. "Stimulus control topography coherence theory: Foundations and extensions." *The Behavior Analyst* 26.2 (2003): 195-213.

Mensah, J., & Arntzen, E. (2017). Effects of meaningful stimuli contained in different numbers of classes on equivalence class formation. *The Psychological Record*, *67*(3), 325-336.

Moss-Lourenco, P., & Fields, L. (2011). Nodal structure and stimulus relatedness in equivalence classes: Post-class formation preference tests. *Journal of the Experimental Analysis of Behavior*, *95*(3), 343-368.

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.

Saunders, R. R., Chaney, L., & Marquis, J. G. (2005). Equivalence class establishment with two-, three-, and four-choice matching to sample by senior citizens. *The Psychological Record*, *55*(4), 539-559.

Shen, Z., Popov, V., Delahay, A. B., & Reder, L. M. (2018). Item strength affects working memory capacity. *Memory & Cognition, 46*(2), 204-215.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, *14*(1), 5-13.

Sidman, M., Cresson Jr, O., & Willson-Morris, M. (1974). ACQUISITION OF MATCHING TO SAMPLE VIA MEDIATED TRANSFER 1. *Journal of the Experimental Analysis of Behavior*, *22*(2), 261-273.

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

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

Sidman, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies*, 93-114.

Sidman, M. (1992). Equivalence relations: Some basic considerations. *Understanding Verbal Relations,* 15-27.

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

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

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

Steele, D., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior*, *56*(3), 519-555.

Stewart, I., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2002). Stimulus equivalence and nonarbitrary relations. *The Psychological Record*, *52*(1), 77-88.

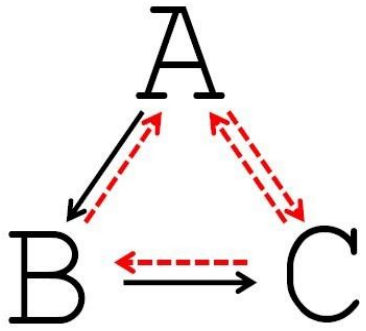
Stromer, R. (1986). Control by exclusion in arbitrary matching-to-sample. *Analysis and Intervention in Developmental Disabilities*, *6*(1-2), 59-72

Wang, T., Dack, C., McHugh, L., & Whelan, R. (2011). Preserved nodal number effects under equal reinforcement. *Learning & Behavior*, *39*(3), 224-238.

Wang, T., McHugh, L. A., & Whelan, R. (2012). A test of the discrimination account in equivalence class formation. *Learning and Motivation, 43*(1-2), 8-13.

Appendix A

PARTICIPANTS WANTED



Interested in participating in a psychology experiment? Want to learn about how we study relational learning processes?

WHAT: Studying the factors that govern how we relate pieces of information that are not directly related.

HOW: There will be up to 2 sessions, around 2 hours each, on average; you will sit at a computer and press keys following some instructions.

BENEFIT: You can learn about this area of research and gain vouchers of up to $20.

CONTACT: Gabriel (gcc9@students.waikato.ac.nz)

This research project has been approved by the Human Research Ethics Committee of the Faculty of Arts and Social Sciences. Any questions about the ethical conduct of this research may be sent to the Secretary of the Committee, email fass-ethics@waikato.ac.nz, postal address, Faculty of Arts and Social Sciences, Te Kura Kete Aronui, University of Waikato, Te Whare Wananga o Waikato, Private Bag 3105, Hamilton 3240."

Appendix B

SCHOOL OF PSYCHOLOGY

HUMAN ETHICS RESEARCH

INFORMATION SHEET

The researchers are Gabriel Chris Chand (lead researcher) and Dr Tim Edwards (supervisor). This study is part of a master's thesis project.

Gabriel can be contacted through his email (gcc9@students.waikato.ac.nz); Tim can also be contacted through his email (tim.edwards@waikato.ac.nz).

Stimulus equivalence is a theory explaining how two or more related stimuli (e.g. A=B=C) elicit the same response, meaning that each stimuli in the example is the exact same as the others in that group (even A=C, which don't actually interact in that group or been specifically trained).

Over the years, studies into stimulus equivalence have revealed the possibility of *relatedness*; the notion that stimuli in an equivalence class may not be equal – some stimuli might be *more equal* than others. This seems to have something to do with nodality, or how far apart is one stimulus from the other in the same group.

Previous studies have demonstrated relatedness and nodality, but there were some methodological issues with these studies. This project aims to rectify those issues to see if there really is a valid argument for relatedness. Are some stimuli in an equivalence class really *more equal* than others?

As a participant, you will be attending up to two sessions that will last between 2-4 hours each. You will be interacting with a software program on a computer, following the instructions on the screen and pressing specific keys to continue with the experiment. You can take a break at any time during the experiment if you feel fatigued. You will receive a $10 warehouse voucher for participating (which stays even if you later decide to withdraw and do the text-based exercise instead), and another $10 warehouse voucher for staying until the session is completed.

Your data will be anonymized and no identifying information will be used in any reports on this study. Data will be kept in secure, password-protected servers and USB drives. If you choose to be informed of the research results, you will be notified and provided with the results (via email) at the conclusion of the project.

Remember that you may withdraw from participation at any time during the project, up until three weeks after your participation in the experiment (after that the data will be analysed). Your withdrawal will not affect the bonus course marks (or voucher) you will have accumulated for the amount of time participated.

"This research project has been approved by the Human Research Ethics Committee of the Faculty of Arts and Social Sciences. Any questions about the ethical conduct of this research may be sent to the Secretary of the Committee, email fass-ethics@waikato.ac.nz, postal address, Faculty of Arts and Social Sciences, Te Kura Kete Aronui, University of Waikato, Te Whare Wananga o Waikato, Private Bag 3105, Hamilton 3240."

Appendix C

CONSENT FORM

Research Project: Investigating relatedness and nodality in stimulus equivalence

Please complete the following checklist. Tick (□) the appropriate box for each point.	**YES**	**NO**
1. I have read the Participant Information Sheet (or it has been read to me) and I understand it.		
2. I have been given sufficient time to consider whether or not to participate in this study		
3. I am satisfied with the answers I have been given regarding the study and I have a copy of this consent form and information sheet		
4. I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time		
5. I have the right to decline to participate in any part of the research activity		
6. I know who to contact if I have any questions about the study in general.		
7. I understand that the information supplied by me could be used in future academic publications.		
8. I understand that my participation in this study is confidential and that no material which could identify me personally, will be used in any reports on this study.		
9. I wish to receive a copy of the findings.		
If yes, please provide your email address:		

Declaration by participant:

I agree to participate in this research project and I understand that I may withdraw at any time. If I have any concerns about this project, I may contact the convenor of the Psychology Research and Ethics Committee (Professor Nicola Starkey, phone 07 837 9230, email: nicola.starkey@waikato.ac.nz)

Participant's name (Please print): _____

Signature: _____ Date: _____

Declaration by member of research team:

I have given a verbal explanation of the research project to the participant, and have answered the participant's questions about it. I believe that the participant understands the study and has given informed consent to participate.

Researcher's name (Please print): _____

Signature: _____ Date: _____