Measuring Relational Preferences Within an Equivalence Class

by

Erica A. Doran

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2012

APPROVAL PAGE

This page has been read and accepted for the
Graduate Faculty in Psychology in satisfaction of the
Dissertation requirement for the degree of Doctor of Philosophy.

Dr. Lanny Fields

______________________ ______________________________
Date Chair of Examining Committee

Dr. Maureen O'Connor

______________________ ______________________________
Date Executive Officer

Dr. Lanny Fields____

Dr. Robert Lanson____

Dr. Alicia Alvero____
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

MEASURING RELATIONAL PREFERENCES WITHIN AN EQUIVALENCE CLASS

By

Erica A. Doran

Adviser: Dr. Lanny Fields

Two experiments used post-class formation within-class relational assessment test performances to evaluate whether participants demonstrated preference for certain members of an equivalence class based on the type of relation that existed between class members. This research also examined certain procedural factors that influenced the percentage of participants who formed classes, referred to as yield. In Experiment 1, two 5-node 7-member equivalence classes, consisting entirely of nonsense syllables, were established using the simultaneous protocol. After class formation, the effects of the different relations between stimuli were evaluated using within-class relational assessment tests. Only one of the six participants in Experiment 1 successfully formed classes, but that one participant showed an absolute preference for transitive relations over equivalence ones, and for baseline relations over symmetrical ones. Experiment 2 was identical to Experiment 1, except that one of the nonsense syllable stimuli in each class was replaced by a pictorial stimulus. Under these conditions, class formation was enhanced, with classes being formed by 5 of 13 participants. During the relational assessment tests, each of these participants demonstrated essentially complete preferences for transitive relations over equivalence relations and for trained baseline relations over symmetrical relations. Thus, this research demonstrates that the members of equivalence classes are differentially related to each other based on relational type.

Acknowledgements

This dissertation would not have been possible without the support and assistance of many people. The author wishes to express her gratitude to her advisor, Dr. Lanny Fields, who was abundantly helpful, insightful, supportive, and patient. Deepest gratitude is also due to the members of the supervisory committee, Dr. Robert Lanson and Dr. Alicia Alvero, whose guidance, encouragement, and insight were invaluable. The author would also like to express her appreciation to her dissertation readers, Dr. Erik Arntzen and Dr. Francis Mechner, for their very helpful suggestions, comments, and criticisms. Without the invaluable input from the aforementioned, this dissertation would not have been possible.

Special thanks also to all her fellow graduate students for their ongoing support and camaraderie, but especially to Michael Marroquin, Patricia Moss-Lourenco, and Kathie Mangiapanello, each of whom at various times managed to provide just the right words of encouragement needed by the author to endure and complete this process.

The author would also like to convey thanks to her good friends who have been so supportive throughout her graduate studies and particularly tolerant of her long absences while completing her dissertation.

Finally, the author wishes to express her love and gratitude to her family, most particularly her incredibly patient and long-suffering children, Casey and Mookie, for their understanding and endless love and for all the hours good-naturedly spent entertaining themselves in the lab, through the duration of her studies.

Table Of Contents

Tables

Figures

Measuring Relational Preferences Within an Equivalence Class

Introduction

"All animals are equal, but some are more equal than others." Like the characters inhabiting the Orwellian dystopia of *Animal Farm,* individual members of an equivalence class may be equally related and yet, paradoxically, not equally related. The precise parameters under which members of an equivalence class display properties suggesting non-equivalence have not been fully explored, nor has this phenomenon been widely recognized by researchers in this field. The goal of this research is to further the understanding of the differences in relational strength among class members by disentangling the complex relations that exist within an equivalence class.

A discussion of equivalence classes typically states that an equivalence class contains a finite number of stimuli (N) that bear no overt perceptual similarity to one another (Fields and Reeve, 2001; Sidman, 1994; Fields and Verhave, 1987; Sidman and Tailby, 1982). Each becomes related to the others through training of a limited number (N-1) of baseline conditional discriminations among some of the stimuli. Following baseline training, tests that contain untrained combinations of stimuli in a set are presented to determine the emergence of untrained, or derived, stimulus-stimulus relations. When presentation of each stimulus in a test evokes selection of the other stimuli in the class without the direct training of that relation, an equivalence class has been formed (Fields & Verhave, 1987). This mutual selection of stimuli indicates that the stimuli in an equivalence class are functionally interchangeable, making each one the equivalent of the others (Sidman, 1994). Thus, it has been postulated that within an equivalence class, the members are each equally related to each other. This view has also been supported by tests that demonstrated functional equivalence, in which a function acquired by one member of the class generalizes to all other members of that class but not to members of other equivalence class (Fields, Adams, Verhave, & Newman, 1993; Rehfeldt & Hayes, 1998; Sidman & Tailby, 1982; Barnes & Keenan, 1993, Barnes, Browne, Smeets, & Roche, 1995; Saunders, Wachter, & Spradlin, 1988; Sidman, Wynne, Maguire, & Barnes, 1989).

This view has been challenged, however, by researchers who argue that, under at least some conditions, different pairs of stimuli can have different levels of relatedness. Fields and Verhave (1987) identified four atemporal variables that defined the structures of all equivalence classes: class size, nodes, training directionality, and nodal density. It is possible that any or all of these variables, either alone or in conjunction with each other, influence the level of relatedness between stimuli in an equivalence class.

To date, most of the inquiries into this possibility have focused on the effects of nodal differences on the relatedness of stimuli in equivalence classes. For example, Fields and Verhave (1987) proposed that the relatedness of stimuli is an inverse function of the nodal distance that separates the stimuli in the class. Nodal distance is defined as the number of nodes that link two stimuli within an equivalence class (Fields & Verhave, 1987). Thus, in a 5-member equivalence class established by training A→B, B→C, C→D, and D→E, resulting in an equivalence class with the structure of A→B→C→D→E, one node separates the A and C stimuli, while two nodes separate A and D and three nodes separate A and E. Because A and C are separated by fewer nodes (nodally proximal) compared with A and D or A and E, one would predict that the relational strength of A→C would be greater than A→D or A→E.

Research into the phenomenon of delayed emergence of derived relations supports this prediction. Delayed emergence is observed when, following the training of baseline relations, a participant does not immediately demonstrate the emergence of all derived relations but, with continued testing in the absence of any additional training or informative feedback, these derived relations gradually emerge. When delayed emergence of classes occurred, researchers found that the level of class-consistent responding in the initial test block was an inverse function of the number of nodes that separated stimuli in derived relations. Furthermore, derived relations reached mastery in an order that was a direct function of nodal distance (Fields, et al., 1990; Sidman, et al. 1985; Kennedy, 1991; Kennedy, et al., 1994; Bentall, et al., 1998; Spencer & Chase, 1996).

These findings with regard to delayed emergence of equivalence classes raise the question of whether these nodal distance effects continue to exist after equivalence classes have been formed. That is, are the effects of nodal distance merely transient, such that they no longer influence relatedness once classes have been fully formed, or is the effect of nodal distance permanent, even if not always expressed? This question has been examined in a variety of manners, using post-class formation tests.

One such approach to determining whether nodal distance influences performances following full class formation is the function transfer test (Fields, Adams, Verhave, and Newman, 1993). In this approach, following full formation of equivalence classes, e.g., ABCDE, participants are trained to emit a response to one member of the class. This training is followed by the presentation of all members of that class, each presented separately and without informative feedback, to assess the extent to which each of these responses transferred to other members of the class. Fields, Adams, Verhave, and Newman (1993) found that for some participants, the response generalized gradually (i.e., with repeated testing) to the other class members based on nodal distance. Specifically, responding generalized to nodally proximal stimuli earlier than to nodally distal stimuli. With repeated testing, however, responding eventually generalized to all members of the class. That is, participants in these experiments eventually reached mastery criterion for all relations, regardless of nodal distance. Thus, while these results indicate that nodal distance effects continue beyond full class formation, this effect eventually dissipated, leaving unanswered the question of whether any such differences between stimuli within a class are transient or permanent but no longer expressed. To answer this question within the context of function transfer would require the training of a new response and administration of a new function transfer test. This approach would likely lead to an endless cycle of testing, without any definitive conclusions being reached as to the permanence of the effects of nodal distance on differential relatedness of stimuli in an equivalence class. Thus, to achieve a greater understanding of nodal distance effects, and the transient or permanent nature of such effects, some other approach is needed.

Another approach taken by researchers to measure the influence of nodal distance on behavior has been to look beyond the likelihood of responding and instead look to response speed (the reciprocal of reaction time), measured from the onset of a trial to the selection of a comparison. When classes emerge on a delayed basis, response speed is an inverse function of the nodal distance that separated the stimuli in derived relations probes (Bentall, et al., 1998; Kennedy, 1991; Spencer & Chase, 1996; Wulfert & Hayes, 1988; Tomanari, Sidman, Rubio, & Dube, 2006). This result occurred in the initial test blocks during delayed emergence of equivalence classes, again raising the issue of whether these nodal distance effects are temporary or permanent. When response speed was examined after full class formation, the results have been varied and, thus, inconclusive. In some of these studies, the effect of nodal distance dissipated completely with full emergence of the classes, while in others, response speed continued to be a direct function of nodal distance even after completion of equivalence class formation (Tomanari et al, 2006; Kennedy, 1991; Spencer & Chase, 1996; Wulfert & Hayes, 1988).

A more fruitful approach to this issue has used post-class formation steady-state measures to assess the permanence of nodal distance effects. One such approach used to assess the continuing influence of nodal distance on stimuli within an equivalence class is the dual option function transfer test. Fields, Landon-Jimenez, Buffington, and Adams (1995) used this approach and determined that the frequency with which the B, C, and D stimuli evoked either the A-based response or the E-based response was an inverse function of nodal distance. That is, the stimuli that were nodally proximal to the A stimulus more frequently occasioned the A-based response, while those stimuli that were nodally proximal to the E stimulus more frequently occasioned the E-based response. Participants continued to respond in this manner in the course of repeated test administrations, providing a post-class formation steady-state measure of the effects of nodal distance on responding. Such differences in performance indicate that nodal distance continued to influence the relatedness of different stimuli within an equivalence class. Such an outcome would not be predicted based on the theory that all members of the class are substitutable. Rather, each of the B, C, and D stimuli would be expected to evoke the A-based response 50% of the time and the E-based response 50% of the time. In the above-referenced research, however, there was no measure of class integrity following completion of the dual option function transfer test. Thus, it was not known at the end of that experiment whether the division of the 5-member classes led to the degradation and elimination of the originally formed classes.

Fields and Watanabe-Rose (2008) systematically replicated and extended this research using two 4-node 6-member equivalence classes and also addressed the question raised above.

Classes were formed by training A→B, B→C, C→D, D→E, and E→F using a simultaneous protocol. Once participants had demonstrated the emergence of all derived relations, the researchers trained participants to emit different, incompatible responses to the C and D stimuli of each class. Specifically, participants were trained to respond by hitting a key a certain number of times in the presence of the C stimuli and a different number of times in the presence of the D stimuli. Following this training, the 12 stimuli from the two 6-member classes were presented repeatedly, one at a time, and without any informative feedback. Among those participants who formed classes, the response to the C stimulus commonly generalized to the A and B stimuli but rarely to the D, E, and F stimuli in the class. Similarly, the response to the D stimulus commonly generalized to the E and F stimuli but rarely to the A, B, and C stimuli in the class. That is, a bifurcation occurred within each of the trained equivalence classes, with the A, B, and C stimuli in one half and the D, E, and F stimuli in the other. This bifurcation indicates that the members of the class were differently related to each other, providing further support for the position that not all members within a class are truly equivalent or interchangeable. Post-transfer test blocks that included both baseline and derived relations, which were used to assess the intactness of the original equivalence classes, however, indicated that the original 6-member classes did indeed remain intact. This outcome indicates that, despite the fact that the stimuli in the 6-member class were differently related to each other, they continued to act as an equivalence class, and that participants' responding was strongly influenced by the testing context. That is, participants altered their responding between differential relatedness (on within-class trials) and equal relatedness (on cross-class trials).

Another approach to the assessment of the continued influence of nodal distance on stimuli within an equivalence class is to use within-class preference tests (Alligood & Chase, 2007; Fields, Adams, & Verhave, 1989, Fields, Adams, Verhave, & Newman, 1993; Moss-Lourenco & Fields, 2011). In this approach, following the formation of equivalence classes, participants are presented with trials consisting of one sample and two comparisons, all from the same stimulus set. A participant's choice of a given comparison stimulus indicates a preference by the participant for that stimulus over the other comparison stimulus in the presence of that sample. Such preference tests allow for direct comparisons of the relative strengths of any two relations that exist among stimuli in the equivalence class. For example, the B1 stimulus (that is, the B stimulus from class 1) might be presented as a sample, with the D1 and E1 stimuli as comparisons. Both the B1→D1 and B1→E1 relationships are transitive ones, but the B1→D1 relationship contains one node, while the B1→E1 relationship contains two nodes. Thus, a participant who selects the D1 stimulus rather than the E1 stimulus in this scenario is showing a preference for the nodally proximal stimulus over the nodally distal stimulus. Using this approach, researchers have found that preferences were an inverse function of nodal distance (Fields et al., 1989; Fields et al., 1993). That is, the greater the nodal distance between the sample and a comparison, the lower the participants' preference for that stimulus.

This result, however, could have been due to other factors. For example, Fields et al. (1993) used a simple-to-complex protocol to train equivalence classes that had a linear structure. That is, first the A→B relation was trained, followed by the B→C relation, then the C→D relation, forming two equivalence classes with the structure of A→B→C→D. Thus, the baseline relations were established in an order that perfectly correlated with nodal distance in the transitive relations. Thus, if A1 is presented as the stimulus, with C1 and D1 as the comparisons, the preference for C1 might reflect a preference for the earlier-trained stimulus rather than the nodally proximal stimulus. Furthermore, the number of training trials presented during class formation was also correlated with nodal distance, such that participants' selections might have reflected a preference for the stimulus to which they had received greater exposure during training, rather than nodal proximity.

One of these confounds was addressed in subsequent research by Alligood and Chase (2007). In training baseline relations, these researchers held constant the number of training trials for each baseline relation. After forming classes, participants were given within-class preference tests in which they typically selected the nodally proximal stimuli over the nodally distal stimuli. Thus, this research supported the view that the relational strengths among stimuli in an equivalence class are inversely related to nodal distance. Like Fields et al. (1993), however, these researchers also used a serialized training and testing protocol in establishing equivalence classes. Thus, participants' preferences might still have been reflecting the effects of serial order rather than nodal distance.

Moss-Lourenco and Fields (2011) addressed this issue by using the simultaneous protocol to establish equivalence classes. In the simultaneous protocol, all of the baseline relations are trained together in random order in a single block, after which all emergent relations probes are similarly presented together. By using the simultaneous protocol, these researchers ensured that nodal distance was not correlated with either the number of training trials presented for a relation or the order in which those relations were trained or tested. These researchers found that participants demonstrated a preference for nodally proximal stimuli and that this preference was indeed based on nodal distance between the stimuli.

It can be seen, then, that in an equivalence class different pairs of stimuli have different levels of relatedness, based upon differences in the number of nodes or nodal distance that separate any two stimuli in the class, and that these differences are not transient. This is not to suggest, however, that these same stimuli are not also still functioning as an equivalence class. Moss-Lourenco and Fields (2011) directly addressed this issue by administering a post-preference equivalence test. Following completion of the within-class preference tests, participants were presented with test blocks containing all the directly trained and derived relations for the two classes. These trials used the simultaneous matching-to-sample format with three comparisons: one comparison from each of the two equivalence classes and a third null comparison that was from neither class. Participants virtually always chose the comparison from the same class as the sample in these tests, demonstrating that the classes remained intact. This result indicates that the participants' responding was influenced by the testing context, responding in a class-consistent manner on cross-class trials but demonstrating a nodal preference among stimuli on within-class trials.

To summarize, prior research has convincingly demonstrated that nodal distance influences the relatedness of stimuli within an equivalence class. Of the four factors that define the structure of an equivalence class identified by Fields and Verhave (1987), however, nodal distance is the only one that has been systematically manipulated and researched in this manner. The other structural factors, which are class size, training directionality, and nodal density, might also influence the relatedness of stimuli. Of interest in the present research was training directionality; would this factor influence the relatedness of stimuli in an equivalence class? Even when the nodal distances separating two class members are held constant, the directionality of training produces different derived relations between stimuli within the class. It follows that the strength of the relation between the stimuli in, for example, a transitive relation might differ from that between two stimuli in an equivalence relation, even though the nodal distances were held constant.

According to Sidman and Tailby (1982), once a class is formed, all of the stimuli are functionally substitutable for each other. Thus, in their view, the predicted answer to the question of whether training directionality would influence the relatedness of stimuli within a class would be “no.” On the other hand, Sidman and Tailby also presented a logical analysis of the relations among the stimuli in an equivalence class that would yield another prediction, based upon the number of derived or logical operations required to link any two stimuli within a given equivalence class based upon their positions within the training structure of the class (Sidman & Tailby, 1982). For example, if a 3-member equivalence class is formed by training A→B and B→C, no logical operations are required to evoke B when presented with A; it is a directly trained relation. On the other hand, one logical operation, symmetry, is required to evoke A when presented with B. Similarly, one logical relation, transitivity, is required to evoke C when presented with A. To evoke A when presented with C, however, requires two logical operations; a participant must use both the symmetrical property to connect C→B and B→A and the transitive property to then link C with A through the nodal stimulus, B. This need for these two logical operations has led some researchers to refer to equivalence relations as combined symmetry/transitivity relations (Spencer & Chase, 1996). Thus, because equivalence relations require two logical operations whereas transitive relations require only one, it might be said that transitive relations are the “less complex” relations and the equivalence relations are the “more complex” relations. Similarly, baseline relations can be viewed as less complex than symmetrical relations.

Based on the number of logical operations required to link the two stimuli, or the complexity of the stimulus-stimulus relations, one might predict that participants would show a within-class preference for the stimuli that represent the less complex relations in each of these pairs. For example, if strength were correlated with fewer logical processes, it follows that the relation between two stimuli linked by transitivity would be stronger than that between two stimuli linked by equivalence. In a within-class relational assessment test administered after participants have completed class formation, this differential strength would be manifested by participants' selection of the comparison stimulus representing the transitive relation over that representing the equivalence relation.

This possibility has been supported to some extent by research addressing differences in response time based on relational type. Spencer and Chase (1996) reported that participants responded faster on baseline and symmetry trials than on transitivity and equivalence trials. Imam (2001) similarly found that response time was an inverse function of relational type. These experiments, however, contained certain confounds. For example, baseline and symmetry trials differ from transitivity and equivalence trials not only by relational type but also by nodal distance; baseline and symmetry are necessarily 0-node relations while transitivity and equivalence are necessarily relations involving 1 or more nodes. Similarly, in their research into nodal distance effects, Alligood and Chase (2007) also examined the differential relatedness of stimuli based upon relational type, but this research did not hold nodal distance constant, nor did it control for the serial order of the training in the simple-to-complex protocol. Additionally, no attempt was made in the prior research to hold constant the number or type of functions served by the comparison stimuli during training. An additional problem in this research was that the numbers of training and testing trials were not necessarily balanced, so that participants experienced, for example, more transitivity trials than equivalence trials, which one would expect to influence reaction time. It is thus difficult to ascertain whether any differences in responding found in prior research were due to relational differences or other factors. It does not appear that any research to date has directly addressed the issue of whether stimuli within a class have differential strength of relatedness as a function of the different stimulus-stimulus relations without any such confounding variables.

One possible reason that such research has not been forthcoming is due to pragmatic difficulties. First, relatively large classes would need to be used for such research. As alluded to above, because of the size of the classes used in previous research, the relative strengths of the different within-class relations could not be tested without some confound, such as prior trained function of the stimulus or nodal distance. For example, if a class containing A1, B1, C1, D1, and E1 stimuli were used, a within-class relational assessment test could be presented using the C1 stimulus as the sample and the A1 and E1 stimuli as comparisons. The C1→A1 relation would be a 1-node equivalence relation, while the C1→E1 relation would be a 1-node transitive relation. Selection of the E1 comparison might indicate a preference for the transitive relation. It is also possible, however, that selection of the E1 comparison reflects the fact that participants had previously, in training, seen the E1 stimulus presented as a comparison, whereas the A1 stimulus was presented only as a sample.

This confound can be avoided by using comparison stimuli that had served identical functions in training. To do so, however, would require a class with at least 7 members. For example, in a stimulus set with a linear structure and consisting of only 6 members (A1→B1→C1→D1→E1→F1), it is not possible to test for a within-class preference for transitivity or equivalence while keeping nodal distance and function of comparison stimuli equal. Thus, to effectively test participants' preferences for different within-class relations between stimuli, one must first establish larger equivalence classes with more nodal stimuli. The present research overcomes this limitation by using larger classes with more nodal separations.

The present research consists of two experiments. Experiment 1 used two 5-node, 7-member classes consisting entirely of nonsense syllables. Experiment 2 is a replication of Experiment 1 in which one member of each stimulus set was replaced with a nameable pictorial stimulus in an attempt to increase the number of participants who formed classes.

Experiment 1

In Experiment 1, two 5-node 7-member classes consisting entirely of nonsense syllables were trained. These nonsense syllables are referred to as CVCs because each consisted of a consonant-vowel-consonant string. Training was done using the simultaneous protocol and a linear structure. The 5-node, 7-member classes (A→B→C→D→E→F→G) allowed for the use of within-class probes that held constant both nodal distance and stimulus function while evaluating whether participants demonstrated a preference for certain stimuli based upon the directionality of training. If, as has been argued, members of an equivalence class are fully substitutable for each other, participants should not demonstrate any relational preference but instead would be expected to select each of the within-class comparison stimuli on approximately 50% of trials. If, on the other hand, the stimuli within an equivalence class were differentially related based upon relational types, one would expect participants to show a different response pattern. Specifically, it was expected that participants would more frequently select the comparison stimulus that represents the baseline relation (which requires no derived or logical operations) over the one that represents the symmetrical relation (which requires one derived or logical operation). Following this logic, it was also predicted that participants would more frequently select the comparison stimulus that represents the transitive relation (which requires one derived or logical operation) over the one that represents the equivalence relation (which requires two derived or logical operations).

Method

Subjects. There were 6 participants in Experiment 1, all of whom were students enrolled in a psychology course at Queens College. Participants were given partial course credit upon completion of the experiment. Each participant completed the experiment in one session, lasting approximately 2-3 hours.

Setting. The experiment was conducted entirely within a laboratory on the Queens College campus. The participant was placed in a room with a closed door to minimize distractions. The room contained a chair and a table, with a computer monitor, keyboard, and printer on the table.

Apparatus

Hardware and Software. The experiment was conducted on a personal IBM-compatible computer that displayed all stimuli on a 15" VGA monitor. Responses consisted of touching specific keys on a standard QWERTY keyboard. The experiment was controlled by custom software that programmed all stimulus presentations and recorded all keyboard responses. A separate monitor, which was not in the experimental cubicle, allowed the experimenter to monitor the participants' progress.

Stimuli. This experiment used stimuli, produced with custom-developed software, of two different types: (a) words believed to be semantically linked and readily recognizable to subjects fluent in the English language; and (b) nonsense syllables consisting of a consonant-vowel-consonant string that do not have any apparent "meaning" or little associative value, as determined by a reference/word association value test. Specifically, prior to beginning this research, a separate pool of participants, containing 30 college students, received lists of twenty-five nonsense syllables and were asked to organize the nonsense syllables into groups, based upon which they believed should go together. Participants were told a group could consist of as few as one and as many as twenty-five nonsense syllables, and they could form as many or few groups as they saw fit. These lists were then analyzed and if any nonsense syllables were paired together in the same group by 5 or more of the 30 participants, one of those nonsense syllables was removed and replaced with a new nonsense syllable. This process of testing for pre-existing associations (with a new group of participants each time) and then revising the list continued until the list contained twenty-one nonsense syllables, none of which were systematically paired with any other nonsense syllable on the list by more than 4 of the 30 participants.

The words were used in the keyboard familiarization and response training phases of the experiment only. The nonsense syllables were used in the equivalence class training and testing portions of the experiment, as well as the feasibility and relational assessment tests. These stimuli were grouped into three sets, two equivalence sets and a "null" set, each of which contained seven nonsense syllables.

The nonsense stimuli are represented by the letters An, Bn, Cn, Dn, En, Fn, and Gn. Numbers are used to indicate the different sets, with the number following each letter indicating the set membership of each stimulus. The two sets used in Experiment 1 consisted of the following stimuli: Set 1 contained *LEQ* (A1), *TYW* (B1), *MEV* (C1), *GIP* (D1), *HUK* (E1), *FOM* (F1), and *YUF* (G1), while Set 2 contained *XAH* (A2), *PYV* (B2), *JEP* (C2), *BUH* (D2), *GAZ* (E2), *BEW* (F2), and *XOL* (G2). The null comparison set contained *NIR* (A3), *TIJ* (B3), *ZUC* (C3), *WEX* (D3), *RAB* (E3), *XOL* (F3), and *COH* (G3). The null comparison was included for two reasons. First, some researchers have argued that presenting two comparisons is insufficient because participants might be forming only one equivalence class and a "reject" class, rather than forming two distinct classes (Sidman, 1987). The use of three comparisons, one from each class and a third comparison that is not from either of the two classes, averts this potential outcome. Second, prior researchers have found that inclusion of a null comparison appears to sensitize participants to the underlying structure of the equivalence classes they form (Moss-Lourenco & Fields, 2011). The stimuli used were tested prior to the experiment, using a different group of participants, to ensure that there were no systematic, pre-existing associations between any of the stimuli.

Procedure.

Trial format and contingencies. All trials in the experiment were presented using the simultaneous matching-to-sample format (Cumming & Berryman, 1965). In each trial, one sample and either two or three comparisons were presented. The samples were drawn from one of the two sets. When two comparisons were presented, the positive comparison was from the same set as the sample on that trial and the negative comparison was from the other set. When three comparisons were presented, the positive comparison was from the same set as the sample on that trial, one of the negative comparisons was from the other set, and the other negative comparison was a “null” comparison that belonged to neither set.

In each trial, the sample was presented on the upper portion of the monitor and was centered horizontally. The comparisons were on the lower portion of the monitor, evenly spaced horizontally. The location of the positive comparison was randomly assigned.

Trial block organization and feedback reduction. Each phase of the experiment contained blocks of trials. The number of trials per block varied across phases and is set forth in the detailed discussion of each phase. In all phases, the trials in a block were presented in a randomized order without replacement. A trial began when "Press ENTER" appeared on the screen. Pressing the ENTER key cleared the screen and a sample stimulus was displayed. Pressing the “B” key on the keyboard added the comparison stimuli to the display. During a trial, participants selected a comparison stimulus by pressing the corresponding key. The placement of the positive comparison was randomized across trials so that there was an equal likelihood of it appearing in any of the positions. In trials where two comparisons were presented, participants pressed the “1” key to select the comparison on the left of the display and the “2” key to select the comparison on the right. In trials where three comparisons were presented, participants pressed the “1” key to select the comparison on the left, the “2” key to select the middle comparison, and the “3” key to select the comparison on the right. A comparison selection cleared the screen and immediately displayed a feedback message centered on the screen.

A feedback message was presented after each selection, but this message could be either informative or non-informative. When informative feedback was presented, either the word "RIGHT" or “WRONG” appeared on the monitor, depending on the accuracy of the comparison selection. The message remained on the screen until the participant pressed the R key (R for RIGHT) in the presence of the word “Right” or the W key (W for WRONG) in the presence of the word “Wrong.” These responses demonstrated that the participant discriminated the feedback information. During some training and all testing trials, non-informative feedback was instead presented after a comparison selection. The non-informative feedback consisted of a dashed line that bracketed the letter E (i.e., - - E - -). This non-informative feedback remained on the screen until the participant pressed the E key, which served as an observing response for the non-informative feedback. The pressing of the appropriate key in response to the feedback message cleared the feedback message from the monitor, ending the trial and enabling the start of the next trial (Fields et al. 1995).

Experimental Phases. Experiment 1 consisted of 6 experimental phases, presented in the following order.

Phase 1: Instructions, keyboard familiarization, and response training. The experiment began with the presentation of the following on-screen instructions:

> Welcome. In this experiment you will be shown many trials.
> Each trial starts with the presentation of one cue at the top of the screen
> followed by the presentation of two cues at the bottom of the screen.
> Each cue is either a word, a picture or a 3-letter string.
> Your task is to learn which cues go together.
> The first cue will appear after a short delay.
> Then press the B key to see the second two cues.
> If the cue at the top of the screen goes together with the cue on the bottom left,
> pressing the 1 key is correct.
> If the cue at the top of the screen goes together with the cue on the bottom right,
> pressing the 2 key is correct.
> If the feedback is "right", press the R key.
> If the feedback is "wrong", press the W key.
> If the feedback is "no feedback" press the E key.
> CALL THE EXPERIMENTER IF YOU HAVE QUESTIONS.
> If not, press "Enter" to continue.

After pressing the "Enter" key, participants were trained to emit the appropriate keyboard responses to complete a trial. This phase involved the repeated presentation of a block of 16 trials. Trials contained three English words, such as KING, QUEEN, and CAMEL. The semantic relation between the sample word (e.g., KING) and one of the comparisons (e.g., QUEEN) was used to prompt the selection of the correct comparison. The word RIGHT or WRONG followed each comparison selection. Correct responding to the stimuli in a trial was facilitated by the presentation of instructional prompts (e.g., "Make your choice by pressing 1 or 2," "Press R to continue," or "Press W to continue"), which was systematically deleted across trials as long as the participant made correct responses. For detailed information regarding the prompt-fading procedure used, see Fields et al. (1990) or Fields et al. (1997). Phase 1 ended when the sample and comparison stimuli were presented without prompts, and performance was at 100% accuracy during a single block.

Phase 2: Establishment and maintenance of baseline relations using the simultaneous protocol. Training involved the establishment of the baseline relations necessary to form two 5-node 7-member equivalence classes. These baseline relations were all presented in a single block of 36 trials, using the simultaneous protocol. Participants were trained in the following relations for each class: A→B, B→C, C→D, D→E, E→F, and F→G. Each trial contained one sample and three comparisons (one from set 1, one from set 2, and a null comparison that was not from either set). The simultaneous protocol was used to avoid any confounding of order of training the baseline relations with the nodal distance that separated the stimuli in the class or stimulus preference (for example, a preference for the stimuli repeatedly encountered earlier in the training). It also equalized exposure to each baseline relation. This block was presented repeatedly until participants attained the mastery criterion of 100% accuracy in responding.

At the start of training, all trials in a block resulted in the presentation of informative feedback; that is, participants received 100% feedback for each block. The block was presented repeatedly with 100% feedback until participants reached the mastery criterion for that block. Thereafter, the percentage of trials in a block that produced informative feedback was reduced to 75%, 25%, and finally to 0%, as long as the mastery criterion was maintained in a block. During feedback reduction, a random generator determined the trials that produced informative feedback.

If the mastery criterion was not achieved within three blocks at a given feedback level during training, the participant was returned to the previous feedback level in the next block. Once all baseline relations were established and maintained with no informative feedback, the participant was presented with the derived relations test. Participants were allowed to repeat blocks and the overall feedback reduction/back-up procedure as many times as necessary to achieve the mastery criterion with no informative feedback, although participants who had not attained the 100% mastery criterion in Phase 2 after two hours were excused from the experiment.

Phase 3: Equivalence class formation probes. Participants who completed Phase 2 were next presented with the derived relations probes. In this phase, all trained and derived relations were presented with non-informative feedback. As in the previous phase, each trial consisted of the presentation of one sample and three comparisons. With two 7-member classes, this involved 84 stimulus-stimulus relations in total: twelve trained baseline relations, twelve derived symmetrical relations, ten 1-node transitive relations, ten 1-node equivalence relations, eight 2-node transitive relations, eight 2-node equivalence relations, six 3-node transitive relations, six 3-node equivalence relations, four 4-node transitive relations, four 4-node equivalence relations, two 5-node transitive relations, and two 5-node equivalence relations. Because a single block containing 84 trials would be unduly burdensome for participants, these trials were divided into three blocks of 28 trials, with the different relations distributed equally across these blocks to the extent feasible. For example, each of the three blocks contained four 1-node transitivity probes and four 1-node equivalence probes. The three-block sequence was repeated up to four times, in the same order, such that any three sequential blocks would contain all 84 stimulus-stimulus relations. To demonstrate equivalence class formation, participants were required to respond correctly to at least 96% of the trials in a block for any three consecutive blocks. Participants who failed to demonstrate equivalence class formation after completing the 12 blocks in this phase were excused from the remainder of the experiment.

Once participants had demonstrated formation of the two 5-node 7-member classes, a series of within-class assessment tests were presented to assess participants' preference for certain stimuli within a class based upon relational types. Two kinds of within-class tests were presented: feasibility (detailed in Phase 4, below) and relational (detailed in Phase 5, below). The ultimate goal of these within-class assessment probes was to determine whether participants demonstrated a preference for stimuli representing certain stimulus-stimulus relations when nodal distance was held constant.

Participants were provided with up to three blocks in each of these within-class assessment tests. On the within-class assessment tests, the mastery criterion was 93% accuracy, meaning that participants were required to select the predicted comparison in 15 of 16 trials. After a maximum of three blocks, participants were moved on to the next relational assessment test, regardless of whether they reached the mastery criterion on the prior test.

Given the size of the equivalence classes and the number of within-class relations, myriad within-class feasibility and relational assessment tests were available for presentation. Presenting participants with each such test was, however, impractically time consuming. Therefore, a small subset of these tests, selected to provide an optimal set of information, was presented to all participants as follows.

Phase 4: Feasibility Tests. Once participants passed the derived relations test, they were each presented with a set of feasibility tests designed to assess whether participants would demonstrate any preference for one relation relative to another where the two relations differed in many ways. Unlike the previous trials, these trials contained two comparisons rather than three, and the sample and both comparisons were all from the same stimulus set. These tests presented the most extreme contrasting relations possible in terms of complexity and nodal distance.

Each of the feasibility tests (as well as the within-class relational assessment tests, discussed below) was presented in probe pairs. Each combination of a specific sample and two comparisons is referred to as a probe type. Each probe type is designated with three letters, the first representing the sample, the second designating the comparison that is nodally proximal to the sample, and the third designating the comparison that is nodally distal to the sample. For example, a probe type designated as ABG would contain A as the sample with B and G as the comparisons, where B is nodally proximal (0 node) to the sample and G is nodally distal (5 nodes) to the sample. The same probe types were presented for both Class 1 and Class 2; thus, the above probe would be presented using the A1-B1-G1 stimuli and also the A2-B2-G2 stimuli.

Two different feasibility tests were administered. Each used broad nodal spreads between stimuli, which also varied by relational type. These feasibility tests were designed to assess participants' ability to discriminate between any of the relations in a class. Each test block contained 16 trials: two types of probes in each test block were each presented four times for each of the two classes. These 16 trials were presented in a randomized sequence. The feasibility probe blocks were as follows.

Feasibility Test 1: ABG and GFA probes: The first feasibility test contained probes ABG and GFA. In the ABG probe, the A stimulus was presented as a sample, with the B and G stimuli as comparisons. A→B is a directly trained, 0-node baseline relation while A→G is a derived, 5-node transitive relation. Thus, one would expect participants to demonstrate a preference for B rather then G for two reasons: (1) A→B was a directly trained rather than a derived relation (and thus requires no derived or logical operations) and (2) B is nodally proximal to A (0 nodes) whereas G is nodally distal to A (5 nodes).

In the GFA probe, the G stimulus was presented as the sample, with the F and A stimuli as comparisons. G→F is a 0-node symmetrical relation while G→A is a 5-node equivalence relation. In this probe, it was expected that participants would prefer F over A for two reasons: (1) G→F is a symmetrical relation (requiring only one logical or derived operation) whereas G→A is an equivalence relation (requiring two logical or derived operations), and (2) F is nodally proximal to G (0 nodes) compared with A (5 nodes).

In this first feasibility test, one probe used B and G as comparisons, while the other used F and A. Because the comparisons in each probe type differed, the possibility that participants' selection of a particular stimulus demonstrated unconditional preference for the stimulus itself, rather than a preference for the nodally proximal stimulus or the stimulus representing the less complex relational type, cannot be ruled out. Additionally, the two comparison stimuli in each pair differed from each other in terms of the number and type of functions they each served during training. For example, in the ABG probe, the B comparison served two functions, sample and comparison, during training, while the G comparison served only as a comparison during training. Thus, selection of B over G may reflect a preference for the comparison that served more functions (and consequently was presented in more trials) during training. Likewise, in the GFA probe, the F comparison served two functions, sample and comparison, during training, whereas the A comparison served only as a sample during training. Thus, selection of F over A may reflect a preference for the comparison that served more functions during training, or it may reflect a preference for the comparison that had previously been seen as a comparison during training. The second feasibility test was designed to control for any of these possibilities.

Feasibility Test 2: ABF (target) and GFB (companion) probes. Upon completion of the first feasibility test, either by achieving the mastery criterion or by completing three blocks without achieving mastery, participants were presented with a second pair of feasibility probes, ABF and GFB. In the ABF probe, the A stimulus was presented as the sample, with the B and F stimuli presented as comparisons. A→B is a directly trained, 0-node baseline relation, while A→F is a derived, 4-node transitive relation. There are two reasons one would anticipate participants to demonstrate a preference for B over F: (1) B is nodally proximal (0 node) to A compared with F (4 node); and (2) B represents a trained relation (and thus requires no logical or derived operations) as opposed to F, which represents a transitive relation (and thus requires one derived or logical operation).

Note that in this probe, both comparisons served two functions during training, serving as both sample and comparison, and therefore preference for one comparison cannot reflect a preference based upon the number or type of training functions. It is possible, however, that selection of a given comparison reflects an unconditional preference for that stimulus. The second probe in this pair (GFB), however, controlled for that possibility. Here, the G stimulus was presented as the sample, with the F and B stimuli again serving as comparisons; G→F is 0-node symmetry relation while G→B is 4-node equivalence relation. In this probe, there are two reasons one would expect participants to demonstrate a preference for the F stimulus over the B stimulus: (1) F is nodally proximal (0 node) compared with B (3 nodes); and (2) G→F is a symmetrical relation (requiring only one derived or logical operation) whereas G→B is an equivalence relation (requiring two derived relations, both symmetry and transitivity).

If mastery was achieved on this feasibility test, this then was a clear indication that participants were discriminating between the members of the class based on either relational types or, more likely, nodal distance. As stated previously, in this probe pair, the comparison stimuli both served two functions in training, and thus a participant cannot be selecting one comparison over another on this criterion. Moreover, because the comparisons were the same but a different comparison selection was predicted in each of the probes, mastery performance in this second feasibility test would rule out control by an unconditional preference for one of the stimuli. Thus, a participant who reached the mastery performance on these probes must necessarily have been discriminating between members of the class based on nodal distance between the comparison and the sample, the relation of the comparison to the sample, or both.

Upon completion of these two within-class feasibility tests, participants were presented with the within-class relational assessment tests, which were designed to evaluate preferences for different relational types. As stated previously, with classes of this size and structure, the number of derived relations available for testing is significant; to be exact, a 7-member 5-node class has 36 derived relations as well as 6 trained relations available for use in these preference probes. Due to time constraints, three probe pairs were selected for presentation: one pair determined whether 2-node transitive relations were preferred to 2-node equivalence relations, one determined whether 1-node transitive relations were preferred to 1-node equivalence relations, and finally, one determined whether 0-node trained baseline relations were preferred to 0-node symmetrical relations. These three tests were presented in Phase 5.

Phase 5: Within-Class Relational Assessment Tests. Following completion of the feasibility probes, the effects of different within-class stimulus relations were assessed. As with the feasibility probes, each within-class relational assessment test contained a pair of probes, with three probe pairs presented in total. These relational assessment tests were designed to assess participants' preference for certain stimuli based on the complexity of the relation with the presented sample stimulus. That is, participants were probed for a preference for transitive relations over equivalence relations where both relations shared the same nodal spread, and for trained baseline relations over symmetrical relations where both relations shared the same nodal spread. As in the earlier feasibility tests, throughout this phase, participants were presented with a sample and two comparisons, all from the same stimulus set. In these tests, however, the comparisons were of equal nodal distance from the sample but bore different relations to the sample.

As with the feasibility tests, each probe type is designated with three letters. Again, the first letter represents the sample, but in these probes the second letter now indicates the comparison with the less complex relation to the sample (that is, the one requiring fewer derived or logical operations to link the sample to the comparison), and the third letter indicates the comparison with the more complex relation to the sample. For example, a probe type designated as DGA would contain D as the sample with G and A as the comparisons, where G represents a 2-node transitive relation to the sample and A represents a 2-node equivalence relation to the sample. The same probe pair was presented in a block for both Class 1 and Class 2. Trials for each probe type for a given class were presented 4 times for a total of 16 trials per test block. Trials were presented in a randomized sequence within a test block. As with the feasibility tests, the mastery criterion was 93%. That is, if a participant selected the comparison representing the less complex relation to the sample on 15 of the 16 trials, this was deemed to be a demonstration that the different members of the class had different relational strengths and that the participant preferred the less complex relation. Upon reaching the mastery criterion, participants were moved to the next test. If a participant completed three blocks without reaching the mastery criterion, he or she was then advanced to the next test. These test blocks were presented in the following order.

DGA (target) and BAG (companion) probes: 2-node transitive vs. 2-node equivalence relations. The first of the relational assessment tests assessed preference for a 2-node transitive relation over a 2-node equivalence relation. Again, a pair of probes was presented; the target probe directly assessed relational preference while the companion probe controlled for certain confounds. First, in the target DGA probe, the D stimulus was presented as a sample, with the G and A stimuli presented as comparisons. D→G is a 2-node transitive relation, while D→A is a 2-node equivalence relation. The comparison stimuli were identical not only in their nodal distance from the sample but also in the number of functions served by that stimulus during training. Specifically, both A and G served only one function during training. Those functions, however, were different. During training, the A stimulus served as a sample only, while the G stimulus served as a comparison only. Thus, selection of the G stimulus in these trials might have indicated control by the type of relation (transitivity as opposed to equivalence), by the function served by the comparison during baseline training (comparison as opposed to sample), or by the particular stimulus itself (G as opposed to A) regardless of its relation to the sample stimulus. Selection of the G stimulus because it served as a comparison in training and continued to do so in this test (unlike the A stimulus) or due to an unconditional preference for the G stimulus would not support the hypothesis that the specific relation between given stimulus pairs influences responding. In isolation, then, this probe type could give rise to a significant confound.

The possible control exerted by these confounding factors was eliminated with the inclusion of a companion probe, BAG, with B serving as the sample and A and G serving as comparisons. B→A is a 0-node symmetrical relation, while B→G is a 4-node transitive relation. Given the previous findings that participants have a preference for the nodally proximal stimulus over the nodally distal one (Moss-Lourenco & Fields, 2011), one would expect participants to demonstrate a preference for the A stimulus on these trials. In this instance, the selection of A would rule out the possibility that the selection of G in the prior probe was controlled by a preference for G or a preference for the comparison stimulus that had served as a comparison during training. Selection of A in this probe would thus confirm that the selection of G in the DGA probe indicated a preference for transitivity over equivalence. Selection of the G comparison in this probe, however, would indicate control by a preference for the G stimulus or the stimulus that served the comparison function during training.

Thus, the probes in this set, in combination, would assess whether the different relations between stimulus pairs influences participants' responding. The two probes used identical comparisons (A and G) that served an identical number of functions during training (A as a sample, G as a comparison) and had identical relations to the sample (transitivity for G and equivalence for A). Because nodal separation between the sample and each of the comparisons was constant, nodal distance could not be influencing the test performances. If, as some have argued, all members of a stimulus set are fully interchangeable or substitutable, participants would not demonstrate any preference, but instead would select each of the comparisons approximately 50% of the time in each probe type. If, then, participants consistently select the comparison stimulus that represents the transitive relation over the one representing equivalence, this result would provide a compelling demonstration of the different relational strengths of class members based upon the stimulus-stimulus relation types.

DFB (target) and CBF (companion) probes: 1-node transitive vs. 1-node equivalence relations. The next pair of preference probes examined participants' preferences for transitivity versus equivalence relations when the stimuli were separated by one node. This smaller nodal distance allowed for the use of comparisons that each served two functions, both sample and comparison, during training. This test avoided the potential confound of participants selecting the stimulus that previously appeared as the comparison and avoiding the stimulus that served only as a sample during training. In the target probe, the D stimulus served as the sample and the F and B stimuli served as comparisons. D→F is a 1-node transitive relation while D→B is a 1-node equivalence relation. As with the 2-node relational assessment test, one would expect participants to demonstrate a preference for the transitive relation and thus select the F comparison stimulus. While this probe avoided the different single function confound of the prior test, it is possible that selection of the F stimulus might have reflected an unconditional stimulus preference. To control for this possibility, a companion probe was presented with the same comparisons, but in which the nodal distance was no longer equal. In this probe, the C stimulus was presented as the sample. In this probe, one would predict selection of the B stimulus, based on nodal proximity, rather than the F stimulus. Selection of the B stimulus here would rule out the possibility that selection of the F stimulus in the target probe was due to an unconditional preference for that stimulus.

Thus, as with the prior relational assessment test, if a participant responds by selecting the comparison representing the less complex relation on the target probes, the only factor that could be controlling responding would be a preference based upon the different relational types represented by the comparisons. According to the argument that all members of a stimulus set are fully interchangeable or substitutable, one would expect that participants would not demonstrate any preference, but instead would select each of the comparisons approximately 50% of the time in each probe type. If, then, participants did select the comparison stimulus that represents the transitive relation over the one representing equivalence, this result would provide a compelling demonstration of the different relational strengths of class members based upon the stimulus-stimulus relation types.

DCE (target) and BCE (companion) probes: 0-node baseline vs.0-node symmetrical relations. The final within-class relational assessment test presented a pair of probes that assessed preferences with respect to 0-node baseline and symmetrical relations. In the target probe, the D stimulus was used as the sample, with the E and C stimuli presented as comparisons. As in the prior probe pair, both comparison stimuli served two functions during baseline training. Again, nodal distance between the sample and each of the comparisons was equal; D→E is a 0-node trained baseline relation while D→C is an untrained 0-node symmetrical relation. In this probe, one would expect participants to demonstrate a preference for the E stimulus over the C stimulus. The companion probe in this preference pair, BCE, was used to assess whether such selection reflected a preference for the particular stimulus itself by presenting B as the sample with the same comparison stimuli (C and E), while also further assessing the effects of relationship on participants' preferences by holding directionality even while manipulating nodal distance. B→C is a 0-node trained baseline relation, while B→E is a 2-node transitive relation. Given that B→C is both the directly trained and nodally proximal relation, one would expect selection of the C stimulus over E. The BCE probe thus served as a foil that used the same comparison pair as the previous probe while controlling for an unconditional stimulus preference.

As with the previous two feasibility tests, if in fact all members of an equivalence class are fully substitutable, each comparison should be selected in approximately 50% of the trials. If, however, participants selected the comparison stimulus that represents the trained baseline relation over the one representing symmetry, this result would provide additional compelling evidence of the different relational strengths of class members based upon the stimulus-stimulus relation types.

Phase 6: Derived Relations Test. Regardless of performance on the within-class feasibility and relational assessment tests, it is possible that the previously established equivalence classes were disrupted by these tests. Moreover, to the extent that participants demonstrate a preference for one relational type over another in these tests, it is necessary to confirm that the original 7-member classes are intact before one can concluded that the members of a class have different relational strength based on relational types; if the classes no longer exist, then the prior tests cannot properly be said to have evaluated preferences between class members. It is thus crucial to determine the intactness of the original equivalence classes. This issue was evaluated through the re-administering of the original derived relations test described in Phase 3. As in Phase 3, participants who achieved a mastery criterion of 96% on any three consecutive test blocks were deemed to have intact equivalence classes.

Results and Discussion

Of the 6 participants in this experiment, three did not learn the baseline relations during the training procedure described in Phase 2 and were dismissed from further participation in the experiment. Two more participants acquired the baseline relations in Phase 2 but did not demonstrate equivalence class formation in Phase 3. The remaining participant, NW, acquired the baseline relations, formed two 7-member, 5-node equivalence classes (as demonstrated by the immediate emergence of all derived relations in Phase 3), completed the within-class feasibility and relational assessment tests in the manner predicted above, and then demonstrated that the original classes were still intact by immediately reaching the mastery criterion on the re-administered derived relations test.

Phase 2: Acquisition of Baseline Relations. The number of blocks required for each participant to complete each phase of Experiment 1 are presented in Table 1. Participant NW needed 8 blocks to acquire the 12 trained baseline relations necessary for two 7-member classes, less than either of the other two participants in this experiment who acquired the baseline relations. Feedback was then systematically reduced. A minimum of 3 blocks is necessary to complete the feedback reduction procedure in Phase 2, and Participant NW required 9 blocks, indicating some transient disruptions of the baseline relations during feedback reduction. Interestingly, NW required more blocks to complete the feedback reduction phase, and thus received more baseline training trials, than either of the other two participants, who acquired baselines but did not form equivalence classes.

Phase 3: Emergence of equivalence classes. Participant NW, the only participant in this experiment to form classes, demonstrated immediate emergence of the derived relations indicative of class formation, reaching mastery in the first three blocks, the minimum required, in Phase 3.

Phases 4: Feasibility Tests. Figure 1 depicts the outcomes of each of the within-class feasibility and relational assessment tests for participant NW, the only participant who formed the two 5-node 7-member equivalence classes. The data are presented as five pairs of bars. Moving from left to right, the pairs indicate the performances occasioned by the (1) first feasibility probe; (2) second feasibility probe; (3) 2-node transitivity vs. equivalence relational assessment test; (4) 1-node transitivity vs. equivalence relational assessment test; and (5) 0-node baseline vs. symmetry relational assessment test. Each pair of bars indicates performances (that is, selection of the comparison stimulus listed above the pair of bars) produced by the probe types listed in the labels beneath the bars on the abscissa. Data were aggregated across stimulus classes. Since a probe for a given class involved the presentation of four trials and two classes were included in a test block, performances were based on the 8 presentations of each type of probe. Preference for a given relation was demonstrated by the selection of the same comparison on at least 7 of the 8 trials.

Feasibility Probes. In each probe, as discussed previously, the two sample-comparison relations were distinguished from each other by two factors: number of derived properties needed to link the stimuli in each relation and the number of nodes that separated the stimuli in each relation. The first feasibility block contained ABG and GFA probes. In the ABG probe, Participant NW demonstrated complete preference for the AB relation, selecting the B stimulus on every ABG probe in both classes. This outcome was the predicted performance; the fact that the AB relation was preferred to the AG relation likely demonstrated sensitivity to differential strength of relations in the same class and/or the nodal separation of the stimuli in each relation.

In the GFA probe, Participant NW again demonstrated complete preference for the GF relation by selecting the F stimuli on every GFA probe for both classes. The fact that the GF relation was preferred to the GA relation likely demonstrated sensitivity to differential strength of relations in the same class and/or the nodal separation of the stimuli in each relation. It is possible, however, that the selection of the B comparison in the ABG probes and the F comparison in the GFA probes was controlled by unconditional preferences for the selected comparison stimuli. Although this possibility could not be ruled out in the first feasibility test, it was addressed in the second feasibility test.

The second feasibility test block contained two probes, ABF and GFB, in which the comparison stimuli in both probes were identical. The first, target probe (ABF) evaluated the participant's preference for the AB relation (which had also been presented in the first feasibility test block) over the AF relation. Participant NW demonstrated a consistent preference for the AB relation over the AF relation. This pattern of responding likely demonstrated sensitivity to differential strength of relations in the same class and/or the nodal separation of the stimuli in each relation. To control for the possibility that the selections were controlled by unconditional preferences for the B stimuli instead of the F stimuli in these probes, however, a second, companion probe (GFB) was presented in this block. Participant NW again demonstrated a consistent preference for the GF relation. The fact that the GF relation was preferred to the GB relation likely demonstrated sensitivity to differential strength of relations in the same class and/or the nodal separation of the stimuli in each relation and eliminated the possibility that her selection of the B stimulus in the target probe demonstrated an unconditional stimulus preference for B.

To summarize, on the two feasibility tests, NW's performance suggested that members of an equivalence class have different levels of relatedness. Her performance might have been controlled by either (1) the difference in the nodal distances between the sample and each of the stimuli; or (2) the different number of derived properties needed to link (or the type of relationship between) the sample and comparison stimuli. Thus, Participant NW's performance on the two feasibility tests indicated she was sensitive to some difference in the relations between the stimuli within a given equivalence class. It cannot be ascertained, however, whether she demonstrated a preference based upon nodal separation, the nature of the stimulus-relations, or a combination of these two factors. Regardless of the reason, however, the performances occasioned by the feasibility probes were not predicted by the assumption of stimulus interchangeability or equal relatedness among the members of an equivalence class.

Phase 5: Preference probes. Following completion of the above feasibility tests, the participant completed the three within-class relational assessment tests, in the following order and manner.

2-node transitivity vs. equivalence: DGA/BAG probes. The first within-class relational assessment test block assessed the participant's preference for the comparison stimulus representing the 2-node transitive relation over that representing the 2-node equivalence relation. On the target DGA probe, participant NW consistently selected the G comparison, which defined the transitive relation, instead of the A comparison, which defined the equivalence relation. On the companion BAG probe, she again consistently selected the nodally proximal A comparison over the nodally distal G comparison. Her performance on the companion BAG probe ruled out two possible sources of control in the target probe: (1) control by either a preference for the stimulus that had previously served as a comparison stimulus during baseline training or a rejection of the stimulus that never appeared as a comparison during baseline training; and (2) control by an unconditional stimulus preference for the G stimulus. In light of this, her performance on the target DGA probe indicated a relational preference for the less complex transitive relation, which involves only one logical or derived operation, over the more complex equivalence relation, which involves two logical or derived operations.

1-node transitivity vs. equivalence: DFB/ABF probes. The next test block assessed preference for 1-node transitivity over 1-node equivalence. This test block again consisted of two different, complimentary probes, DFB and ABF. In this test block, Participant NW consistently selected the F comparison in the target DFB probe and the B comparison in the companion ABF probe. On the target DFB probes, the participant consistently selected the comparison stimulus that represented the less complex transitive relation over the one that represented the more complex equivalence relation. Participant's performance on the companion ABF probe rules out the possibility that her selection was based on an unconditional stimulus preference. Moreover, this performance cannot be attributed to differential functions served by the two comparison stimuli because the comparison stimuli each served the same two functions during training. Therefore, the only interpretive option is that transitive relations are preferred to equivalence relations. That is, stimuli related by transitivity alone are stronger than stimuli related through equivalence, which is symmetry in combination with transitivity.

0-node baseline vs. symmetry: DEC/BCE probes. The final within-class relational assessment test assessed whether participants preferred a 0-node baseline relation over a 0-node symmetrical relation. Participant NW consistently selected the E comparison in the DEC probe and the C comparison in the BCE probe. As with the prior relational assessment test, this performance cannot be explained by an unconditional stimulus preference because a different comparison was selected on each of the probes. Her performance also could not be attributable to a preference for the comparison that previously served as a comparison during baseline training because both comparisons served as both samples and comparisons during training. Therefore, her performance can only be due to her preference for baseline relations, which require no derived operations, over symmetrical relations, which require one derived operations.

Cumulative assessment of relational preference. This experiment was designed to evaluate whether participants, having formed equivalence classes, would demonstrate a preference for certain stimuli based upon the relation between those stimuli and the sample stimuli. On each of these probes, participant NW demonstrated a clear preference for the less complex relation. That is, she consistently selected the comparison that was linked to the sample by one derived property (transitivity) over the one linked by two (equivalence, which requires both transitivity and symmetry) and the comparison that required no derived properties (baseline, a trained relation) over that requiring one derived property (symmetry). For each probe, which contained eight trials per test, the likelihood of performing in this manner by chance was 0.5 to the 8th power. Overall, then, Participant NW's performance on these within-class preference probes provides strong support for the argument that the stimuli within an equivalence class are differently related based upon the number of derived operations required to form the relations. This performance is contrary to the view advanced by various researchers, such as Sidman (1994, 2000) and McIlvane & Dube (2003), that all members of an equivalence class are interchangeable. If the stimuli in an equivalence class were interchangeable, all of the relations among the class members would be of equal strength. It necessarily follows from this premise that there should not be a preference for any relation over any other relation. In fact, however, this participant did demonstrate a preference for one relation over another, an outcome consistent with the view that the stimuli in an equivalence classes are not equally related to each other (Fields & Verhave, 1987; Moss-Lourenco & Fields, 2011; Fields & Moss, 2008; Fields & Watanabe-Rose, 2008; Fields, Landon-Jimenez, et al., 1995). The data from the one participant who completed all phases of this experiment thus provide support for the hypothesis that the individual stimuli within an equivalence class are differentially related, based on the complexity of the stimulus-stimulus relations linking the different class members.

Phase 5: Post-preference derived relations test. Following completion of the three within-class preference probes, the equivalence class formation test was re-administered to Participant NW, who passed it in the first three blocks, the minimum number of blocks required to demonstrate mastery in this phase. This performance demonstrated that the original equivalence classes were intact and had not been disrupted by exposure to the within-class tests.

The performances evoked by the cross-class emergent relations tests thus documented the substitutability of the stimuli in each equivalence class and, by implication, their equal relatedness. The performances evoked by the within-class tests documented consistent preferences for the less complex relations, showing that the individual members of an equivalence class are differentially related based on the type of relation that exists between any two stimuli in an equivalence class. Finally, the integrity of the equivalence classes was demonstrated by the performances observed during the re-administration of the emergent relations tests. These results imply that the differential relations among stimuli in a class are always present but their expression depends on the type of test administered to a participant.

To summarize, in Experiment 1, which used the simultaneous protocol for both training and testing, 6 participants attempted to form two 5-node 7-member classes. Only one participant formed classes. During the three within-class relational assessment tests, this participant demonstrated consistent preferences for the less complex relations. That is, she invariably selected the comparison stimuli that represented the transitive relation over that representing equivalence, and the baseline relation over the symmetrical one. Such an outcome is not predicted by the view that all members of an equivalence class are substitutable (Sidman, 1994, 2000). Despite the participant's discrimination between class members, however, the class remained intact, as evidenced by the fact that she immediately reached mastery criterion on the re-administration of the equivalence class formation test.

As noted above, however, only one of six participants completed Experiment 1. Overall, this participant's performance provides support for the argument that the stimuli in an equivalence class are differentially related based on the nature of the inter-stimulus relations, challenging the theory of substitutability of equivalence class members. With only one participant providing data, however, it would be over-reaching to draw any firm conclusions regarding the generality of this effect. Given the low yields obtained in Experiment 1, either an untenably large number of participants would be required or else a somewhat different experimental protocol should be used to properly assess the reliability and generality of the relational preferences observed in 7-member, 5-node classes.

Experiment 2

The low yield (1 of 6) obtained in Experiment 1 was likely due to the use of the simultaneous protocol. While it was necessary to use the simultaneous protocol to minimize confounds involving structural variables within the classes, this protocol has been noted as generally producing low yields (Fields, Landon-Jimenez, et al., 1995; Buffington, Adams, and Fields, 1997). Thus, while the results of this experiment provide support for the view that the members of an equivalence class are differentially related based upon the complexity of the stimulus-stimulus relations, this support is limited. To extend these findings and better evaluate the outcomes from the one subject in Experiment 1 who completed all phases, changes to the experimental protocol that would boost the yield obtained using the simultaneous protocol were needed.

When using the simultaneous protocol, some researchers have successfully increased yields by presenting pre-training to participants. For example, Fields, Reeve, Rosen, Varelas, Adams, Belanich, and Hobbie (1997) showed that previously establishing multi-nodal equivalence classes using the simple-to-complex protocol could enhance the formation of new equivalence classes using the simultaneous protocol. Such a pre-training procedure, however, would require participants to complete at least two separate sessions to complete the experiment. Given the time constraints and the risk of participant attrition if a second session was required, the use of this procedure was not considered optimal for this experiment.

Other researchers have increased the percentage of participants who formed equivalence classes by manipulating the stimuli used in the stimulus sets. Specifically, Fields, Arntzen, Nartey, & Eilefsen (in press) demonstrated that, when using the simultaneous protocol, yields could be substantially increased through the inclusion of a single pictorial stimulus in each stimulus set. These researchers obtained relatively low yields (20%) using the simultaneous protocol when the stimulus sets consisted of 5 nonsense syllables. When one of the nonsense syllables was removed from each set and replaced with a pictorial stimulus, however, 80% of participants formed equivalence classes. As discussed in greater detail in the General Discussion section of this paper, one possible explanation for the increased yields is that the pictorial stimuli were members of pre-existing equivalence classes, which participants then expanded to include the 5 nonsense syllables. A number of studies have shown that once an equivalence class has been formed, its expansion is probable (Saunders et al., 1988; Fields, Newman, Adams, & Verhave, 1992).

If participants were expanding existing classes rather than forming new ones, it follows that using this approach could increase yields in the present experimental paradigm. Although Fields et al. (2011) used 5-member sets, it followed that such an approach might also enhance participants' ability to form 7-member equivalence classes so as to allow relational preferences to be assessed. Therefore, a nameable pictorial stimulus was included in each set used in Experiment 2. Experiment 2 was otherwise identical to Experiment 1 in all respects.

Method

Subjects. Thirteen Queens College students served as participants in Experiment 2. Each subject completed the experiment in one session, lasting approximately 2-3 hours.

Stimuli. This experiment again used stimuli that were produced with custom-developed software and were of three distinct types: (a) words believed to be semantically linked and readily recognizable to subjects who are fluent in the English language; (b) nonsense words consisting of a consonant-vowel-consonant string that do not have any apparent "meaning" or little associative value, as determined by a reference/word association value test described previously; and (c) pictorial representations of commonly-recognizable objects. The words were again used in the keyboard familiarization and response training phases of the experiment only. The nonsense words and pictures were used in the equivalence class training and testing portions of the experiment, as well as the feasibility and relational assessment tests. These stimuli were again grouped into two equivalence sets and a "null" set, each containing six nonsense syllables and one picture.

The nonsense stimuli are represented by the letters An, Bn, Cn, En, Fn, and Gn, while the pictorial stimuli are represented as Dp. Numbers are used to indicate the different sets, with the number following each letter to indicate the set membership of each stimulus. The two sets used consisted of the following stimuli: Set 1 contained *LEQ* (A1n), *TYW* (B1n), *MEV* (C1n), *Church* (D1p), *HUK* (E1n), *FOM* (F1n), and *YUF* (G1n), while Set 2 contained *XAH* (A2n), *PYV* (B2n), *JEP* (C2n), *Crown* (D2p), *GAZ* (E2n), *BEW* (F2n), and *XOL* (G2n). The null comparison set contained *NIR* (A3n), *TIJ* (B3n), *ZUC* (C3n), *Mailbox* (D3p), *RAB* (E3n), *XOL* (F3n), and *COH* (G3n).

Results and Discussion

Of the 13 participants, 3 did not acquire the baseline relations in Phase 2 and were excused from the remainder of the experiment. Of the remaining 10 participants, 5 did not demonstrate equivalence class formation, while 5 formed equivalence classes and completed the within-class preference probes. These 5 participants then completed the re-administration of the equivalence class probes, demonstrating that the equivalence classes had been maintained throughout the experiment. Each phase is analyzed separately below, and the data for all participants in Experiment 2 who reached that phase are included in the analysis. See Table 2.

Phase 2: Acquisition of Baseline Relations. Ten of thirteen participants (77%) acquired the baseline conditional discriminations in Experiment 2. By contrast, of the six individuals who began Experiment 1, in which all the stimuli were nonsense syllables, only three (50%) acquired the baseline conditional discriminations. Overall, then, the performance of participants during this phase in Experiment 2 differed from that of participants in Experiment 1. As the only difference between the two experiments was the stimuli used, the increase in participants achieving mastery performance during the baseline acquisition phase must be attributable to the use of a single pictorial stimulus in each class in Experiment 2.

The baseline relations that were the prerequisites of the equivalence classes to be formed in the present experiment were represented symbolically as AB, BC, CD, DE, EF, and FG. With the exception of the D stimuli, which were nameable pictures, the remaining stimuli were nonsense syllables. Thus, the baseline relations could be divided into two types; those that included a nameable picture as a member of a conditional relation (i.e., CD and DE) and those that did not (AB, BC, EF, and FG.) Acquisition of the conditional discriminations was possibly influenced by the inclusion of a nameable pictorial stimulus in that participants were able to more readily acquire the D-based relations than the relations that did not include the nameable pictorial stimulus. Examining the data for all participants, however, showed that this was not necessarily the case. While some participants did indeed acquire the D-based relations before the other relations, others did not, and there did not appear to be any clear and consistent differences in performance between those participants who acquired all baseline relations and those who did not, or between those participants who went on to form classes and those who did not. Thus, while it appears that the inclusion of a nameable pictorial stimulus enhanced participants' acquisition of the baseline relations, the mechanism by which this occurred is uncertain. Further research is required to more fully understand this phenomenon.

Phase 3: Emergence of equivalence classes. Of the ten participants in Experiment 2 who acquired all baseline relations, five also demonstrated emergence of all the derived relations between stimuli, indicating formation of two 7-node equivalence classes. Of these five participants who formed equivalence classes, three showed immediate emergence, achieving mastery in the minimum of 3 blocks required to probe all possible derived relations within the 2 7-member classes. The other two participants showed delayed emergence, requiring between 5 to 9 blocks to reach mastery. By contrast, the five participants who acquired the baseline relations but failed to demonstrate equivalence class formation did not reach criterion after 12 blocks, the maximum number of blocks available in this phase.

Phases 4 and 5: Feasibility and Relational Assessment Tests. Figure 2 depicts the outcomes of each of the within-class feasibility and relational assessment tests for the 5 participants who formed the two 5-node 7-member equivalence classes. The data are presented in a manner similar to that in Figure 1. Data for each participant are presented in separate panels.

As in Experiment 1, preference for a given type of relation was demonstrated by the selection of the same comparison on at least 7 of the 8 trials.

Within-class Feasibility Tests. Participants were presented with two feasibility tests, identical to the ones presented in Experiment 1. In the first feasibility test, containing ABG and GFA probes, virtually all trials in the ABG probe resulted in the selection of the B comparisons while virtually all trials in the GFA probe resulted in the selection of the F comparisons. Similarly, in the second feasibility test, containing ABF and GFB probes, virtually all trials in the ABF probe resulted in the selection of the B comparisons while virtually all trials in the GFB probe resulted in the selection of the F comparisons. Thus, as in Experiment 1, participants' performance on the two feasibility tests provides clear indication that participants were discriminating between the members of the class based on either relational types or, more likely, nodal distance.

Within-class Relational Assessment Tests. As in Experiment 1, the differential effects of transitive and equivalence relations were assessed with two test blocks; one in which both relations contained comparisons separated from samples by 2 nodes (i.e., DGA/BAG), and one in which they were separated by 1 node (DFB/ABF). The differential effects of 0-node baseline and symmetry relations were also assessed in a separate test block.

DGA (target) and BAG (companion) probes: 2-node transitive vs. 2-node equivalence relations. The first within-class relational assessment test assessed participants' preference for 2-node transitive relations over 2-node equivalence relations. In the DGA probe, four of five participants consistently selected the G comparison, which defined the transitive relation, instead of the A comparison, which defined the equivalence relation, while in the BAG probe, all five participants typically selected the A comparison instead of the G comparison. Thus, four of the five participants demonstrated a relational preference for the less complex transitive relation, which involves only one logical or derived operation, over the more complex equivalence relation, which involves two logical or derived operations, transitivity in combination with symmetry. The performance of the fifth participant is discussed separately, below.

DFB (target) and ABF (companion) probes: 1-node transitive vs. equivalence relations. The next test pitted a 1-node transitive relation against a 1-node equivalence relation. In the DFB probe, all five participants consistently selected the F comparison, which defined the transitive relation, instead of the B comparison, which defined the equivalence relation, while in the ABF probe, all five participants consistently selected the B comparison instead of the F comparison. Thus, in the target DFB probe, all five participants demonstrated a preference for the comparison that represented transitivity over that representing equivalence, which preference could not be explained by an unconditional stimulus preference, as participants all rejected this same comparison in the companion ABF probe. These performances thus indicate a relational preference for transitivity over equivalence in all five participants.

DEC (target) and BCE (companion) probes: 0-node baseline vs. equivalence relations. This final test pitted a 0-node directly trained baseline relation against a 0-node symmetrical relation. In the DEC probe, four of the five participants consistently selected the E comparison, which defined the trained relation, instead of the C comparison, which defined the symmetrical relation, while in the BCE probe, all five participants typically selected the nodally proximal C comparison instead of the nodally distal E comparison. Thus, for these four participants, selection was controlled by a preference for the less complex baseline relation, which requires no logical or derived operations, over the more complex symmetrical relation, which requires one logical or derived operation.

As noted above, one participant did not perform as expected on the 2-node transitivity vs. equivalence test. This same participant, DF, also did not perform as expected in the 0-node baseline vs. symmetry test. In Figure 2, a second graph located to the right of the graph showing participant DF's overall performance details his performance on each probe type for each class on these two tests. Examination of this participant's performance indicates that positional preferences influenced selection in each of these test blocks. In the 2-node transitivity vs. equivalence test block, participant DF demonstrated a positional preference, invariably selecting the comparison stimulus presented on the right side of the display for both classes. On the companion BAG probe, however, participant DF did not demonstrate this positional preference, instead routinely selecting the nodally proximal A stimulus. Participant DF's performance on the 0-node baseline vs. symmetry test also reflected the influence of this positional preference, albeit to a lesser extent. On the target DEC probe, DF selected the E comparison, representing the baseline relation, on 75% of trials. This performance was the result of differential responding on the probes presented for classes 1 and 2. On the DEC probes containing stimuli from class 1, DF always selected the E stimulus, the comparison that defined the baseline relation, resulting in 100% correct on these probes. On the DEC probes containing stimuli from class 2, however, DF always selected the comparison presented on the right side of the display, resulting in the correct selection of the comparison representing the baseline relation on 50% of the trials. When combined, the two class performances yielded 75% selection of the comparison that was related to the sample as the baseline relation. Thus, the baseline was preferred to the sample in only one class, with a positional stimulus control topography controlling behavior in the other class.

Cumulative assessment of relational preferences. Five participants each completed three within-class relational assessment tests, for a total of fifteen within-class relational assessment tests. On thirteen of these fifteen tests, participants showed a clear and consistent preference for the comparison stimulus that, in relation to the sample stimulus, represented the relation requiring fewer derived or logical operations. These performances thus provide compelling support for the hypothesis that the individual stimuli within an equivalence class are differentially related, based on the complexity of the stimulus-stimulus relations linking the different class members.

Phase 6: Post-preference derived relations data. Regardless of preference performances, all five participants who initially formed equivalence classes and then completed the within-class tests passed the final derived relations test. Moreover, they each did so in the first three blocks, the minimum number of blocks required to demonstrate mastery in this phase. These results show that the intactness of the underlying equivalence classes was not influenced by exposure to the within-class relational assessment tests that showed differential strengths of stimuli within the class.

To summarize, in Experiment 2, 13 participants attempted to form two 5-node 7-member classes, each of which class contained one nameable pictorial stimulus. Three participants failed to acquire the baseline discriminations. Of the remaining ten participants, all of whom acquired the baseline discriminations, five failed to form equivalence classes. The five participants who did form equivalence classes each completed the within-class relational assessment tests, and four of the five demonstrated an absolute preference for the less complex relations, while the fifth participant demonstrated a partial preference for the less complex relation. As in Experiment 1, these results run counter to the view that all members of an equivalence class are substitutable (Sidman, 1994, 2000). Finally, despite participants' ability to discriminate between class members, the class remained intact, as evidenced by the performances on the re-administration of the equivalence class formation test.

General Discussion

The present research addressed two general topics. The primary question addressed by these experiments was whether participants would show a preference for different types of relations in an equivalence class based on the number of operations that link the stimuli in the relations (i.e., transitive vs. equivalence relations and baseline vs. symmetry relations. The answer to this question addresses the issue of the effects of structural variables on the relatedness of stimuli in an equivalence class. This research also examined certain procedural factors that influenced yield. The answer to this question addresses the generality of the findings relevant to the first matter. This section will first discuss these procedural considerations, before turning to address factors influencing relatedness among members of an equivalence class.

Procedural issues. As stated previously, in an effort to boost yields, the stimuli used as class members were manipulated. The replacement of one member of each class with a nameable pictorial stimulus, substantially increased yield. The various mechanisms and implications, both practical and theoretical, involved in this procedure are discussed below.

Effects of pictorial stimuli on equivalence class formation. In this experiment, the baseline relations were trained using the simultaneous protocol. The likelihood of equivalence class formation is usually very low when training and testing are conducted using the simultaneous protocol (Fields et al., 1995; Buffington, et al., 1997). Fields et al. (1995) reported that two 3-node 5-member classes were formed under the simultaneous protocol by less than 20% of participants. Fields, et al. (in press) obtained similarly low yields of participants forming three 3-node 5-member classes under the simultaneous protocol. Moreover, the proportion of individuals who form equivalence classes under the simultaneous protocol is generally an inverse function of class size (Fields, Varelas, et al., 2000). Based on these prior findings, one would predict that, using the simultaneous protocol, virtually no participants would form large multi-nodal equivalence classes such as those studied in the present experiment. This, however, was not the case.

In the present experiment, 5 of the 10 (50%) participants who acquired baseline discriminations, or 5 of the 13 (38%) participants who began the experiment formed classes, even though the classes in the present experiment were larger (5-node 7-member) than those in the above-mentioned experiments. These higher-than-expected yields were most likely due to the inclusion of a nameable picture as the D stimulus in each class in Experiment 2. See Table 3.

In Experiment 1, in which all the members of the equivalence classes were nonsense syllables, 50% of participants acquired the baseline discriminations; in contrast, in Experiment 2, in which one member of each set was a nameable pictorial stimulus, 77% of participants acquired the baseline relations. With regard to class formation by those who acquired the baseline relations, Experiment 1 yielded 33%, while Experiment 2 yielded 50%. When all participants who began the experiment are considered, only 17% of participants in Experiment 1 formed equivalence classes, in contrast to 38% in Experiment 2.

The fact that inclusion of the pictures probably enhanced both the acquisition of baseline discriminations and the formation of equivalence classes does not, however, explain how such an effect might have occurred. One possible explanation is that, with the inclusion of the pictures, participants are essentially expanding on existing classes, rather than forming completely new ones. A number of studies have shown that once an equivalence class has been formed, its expansion is probable (Saunders et al., 1988; Fields, Newman, Adams, & Verhave, 1992). In the present experiment, it is plausible that the nameable pictures were actually members of already existing generalized equivalence classes. Thus, the “formation” of the new 7-member equivalence classes might actually have involved the expansion of already existing classes by 6 more members. Not all participants in Experiment 2, however, successfully expanded the classes; in fact, the 50% yield obtained here was notably lower than the 80% yield obtained in the research conducted by Fields, Arntzen, et al. (2011), in which classes were expanded by four members. This result suggests that the likelihood of successful expansion is an inverse function of the amount by which the class is to be expanded.

Necessity of pictorial D stimuli for preference effects. It is important to note that while the inclusion of a nameable pictorial stimulus in each set increased the percentage of participants who formed classes and thus were able to complete the within-class relational assessment tests, the use of a pictorial stimulus was not necessary for the primary findings of this research that participants discriminate between members of an equivalence class and demonstrate a preference for the stimuli representing the less complex relations. This is apparent from the fact that the only participant to form equivalence classes in Experiment 1, which did not use any meaningful stimuli, performed in an identical manner on the within-class tests as participants in Experiment 2, which included meaningful stimuli. Thus, the inclusion of nameable stimuli did not influence participants' performance on the within-class preference probes but merely increased the likelihood that participants would form equivalence classes and thus be able to complete the within-class preference probes, providing data and insight into the differential relatedness of members of an equivalence class. The relatively high yields obtained in Experiment 2 support the view that my main finding, that stimulus pairs within an equivalence class have differential strengths based upon relational type, is representative of a broader population of subjects rather than the single, unrepresentative individual who was able to form the large multi-nodal equivalence classes containing only nonsense syllables.

Relatedness issues. The original intent and central focus of the present research was to assess whether members of an equivalence class have different relational strength due to differences in the stimulus-stimulus relations that exist as a result of the structure created during the training of baseline discriminations. Previous researchers have asserted that, once an equivalence class is fully formed, all members of the class are equally related and fully substitutable for each other. If this were correct, one would expect that on the within-class feasibility probes, participants would have chosen each comparison on approximately 50% of the trials. The present research, however, demonstrates that this is not the case. Moreover, while there may be many other sources of stimulus control, such as nodal distance, unconditional preference for a given comparison stimulus, or a preference for a comparison stimulus based on the number or type of functions that stimulus served during the training of baseline relations, each of these possibilities was controlled for within the experiments. If, then, these factors were exerting control, participants' still would have chosen each comparison on 50% of the trials. Therefore, as this was not the case, the only logical conclusion is that the differences in the types of relations between stimulus pairs was controlling selection, and that these stimuli, although members of an equivalence class are not equally related.

Within-class stimulus preferences based upon stimulus-stimulus relations. Prior research has shown that members of an equivalence class are differentially related based on the nodal distance between the sample and comparison stimuli; participants consistently demonstrated a preference for the nodally proximal stimuli to the nodally distal stimuli (Moss-Lourenco & Fields, 2011). The present research sought to extend this finding by ascertaining whether members of an equivalence class might also be differentially related based upon the number of logical or derived operations required to link the two stimuli. Specifically, it was expected that the number of such logical operations required would be inversely correlated with the relational strength of those two stimuli.

To test this hypothesis, participants were presented with a series of post-class formation, within-class relational assessment tests that held nodal distance constant but varied the types of relations that existed between the sample stimulus and each of the comparisons in the test probes. Selection of each of the two comparison stimuli approximately 50% of the time would indicate that participants did not discriminate between the stimuli based on relational type. If, however, participants consistently selected one comparison stimulus over the other in these probes, this pattern of responding would support my hypothesis. It was anticipated that participants would prefer the comparison stimulus that represented the “less complex” relation, that is, the one that entailed the fewest derived operations. So, for example, it was expected that participants would prefer the stimulus representing the transitive relation to the one representing equivalence, as equivalence includes not only transitivity but also symmetry. Similarly, it was predicted that participants would prefer the stimulus that represented the baseline relation, which had been trained directly, to the one representing symmetry, a derived relation. The results of the two experiments confirmed my predictions.

In total, nineteen individuals began either Experiment 1 or 2. Of these nineteen, six did not acquire the baseline conditional discriminations and were not tested for the emergence of the equivalence classes. An additional seven participants acquired the baseline relations but did not form equivalence classes. The remaining six participants, however, not only formed equivalence classes but went on to complete the within-class relational assessment tests, showing a decided preference for transitive relations over equivalence relations and for baseline relations over symmetrical relations, as discussed in more detail below. Finally, all six of these participants passed the re-administered derived relations test, documenting the intactness of the classes after the administration of the within-class relational assessment tests. The results of both Experiments 1 and 2 suggest that members of an equivalence class are differentially related to each other based on relational type. None of the results support the view that stimuli in an equivalence class are substitutable for each other in these test conditions.

Two of the within-class relational assessment tests evaluated participants' preferences for stimuli representing the transitive relation to that representing the equivalence relation. On 11 of these 12 probes (one probe in each of these two tests for each of the six participants who formed classes), participants demonstrated a clear preference for the transitive relations, which require only one logical operation, over the equivalence relations, which required two logical operations. For each probe, which contained eight trials per test block, the likelihood of selecting comparisons that indicated a preference for the transitive relations by chance was 0.5 to the 8th power for a given probe. This preferential selection occurred for 11 test blocks across six participants. The likelihood of obtaining these results by chance would be .5 to the 88^{th} power. Thus, this consistent selection of the stimuli that defined the less complex transitive relations provides compelling evidence that participants differentiated between the stimuli within an equivalence class based upon the nature of the derived stimulus relations within that class.

The relational strength of stimulus pairs that pitted trained baseline relations against symmetrical relations was also evaluated. On five of these six probes (one for each of six participants), participants demonstrated a definite preference for the trained baseline relation over the symmetrical relation. For each probe, which contained eight trials per test block, the likelihood of selecting comparisons that indicated a preference for the trained baseline relation by chance was 0.5 to the 8^{th} power. This occurred for 5 test blocks, so the likelihood of obtaining these results by chance would be .5 to the 40^{th} power. Thus, the consistent selection of the stimuli that defined the trained baseline relation provides further compelling evidence that participants differentiated between the stimuli within the equivalence classes based upon their relation to the sample stimuli.

For all six participants on all three within-class relational assessment tests, sixteen of eighteen of these probes resulted in participants consistently selecting the comparison stimulus that was linked to the sample stimulus by fewer logical operations. Thus, the type of relation that existed between the stimuli in an equivalence class influenced the relatedness of the stimuli. Specifically, for class members separated by the same nodal distance, participants demonstrated a near-absolute preference for the stimuli that represented the less complex relation to the sample stimuli. These results then imply that the stimuli in the less complex relation are more strongly related to each other than are the stimuli in a more complex relation and that the members of an equivalence class are not, in fact, functionally interchangeable as theorized by Sidman & Tailby (1982).

Permanence of effects of structure on stimulus relatedness. The results of the present experiments extend support for the model of differential relatedness among stimuli in an equivalence class to include not only nodal distance but now also the types of relations that exist among stimuli in the equivalence class. This result does not, however, imply that the stimuli in a class cannot also function in an interchangeable or substitutable manner. As noted by Moss-Lourenco and Fields (2011), although all of the stimuli in an equivalence class are differentially related in terms of nodal distance, all of the stimuli in the class are more closely related to each other than to any stimuli in other classes. This relatedness is guaranteed by the contingencies of differential reinforcement used to establish conditional relations among the stimuli in and between classes. The same argument is valid for relations of different types. When measured, the responses produced during within-class tests express the differential strengths of different types of relations. Those performances, however, did not disrupt the integrity of the underlying equivalence classes, as evidenced by participants' performance on the final derived relations tests.

The fact that participants performed in this manner suggests that control over responding changed with the types of test trials presented, in the absence of any explicit instructions or training in this regard. That is, between- and within-class stimulus control topographies co-existed but were evoked differentially, depending on whether the trial presented within- or cross-class stimuli. Another way to look at this phenomenon would be to examine the effects of cross- and within-class contingencies throughout the various phases of these experiments. During training of baseline relations, the trained contingencies allowed for cross-class discrimination. The training structure, however, resulted in different levels of within-class relatedness. During between-class tests in this phase, the trained contingencies overshadowed the structural effects, resulting in responding in a class-consistent manner indicative of equal relatedness among class members. Similarly, during the phases that tested for equivalence class formation and maintenance, these contingencies again overshadowed the training structure within the class, resulting in performance on between-class tests that indicated equal relatedness among class members. On the within-class tests administered during the feasibility and preference phases of these experiments, however, the training structure overshadowed the trained contingencies, resulting in performance that indicated differential relatedness among class members. Thus, depending on the nature of the trial, the same stimulus can act as a discriminative stimulus for either equal relatedness (on between-class trials) or differential relatedness (on within-class trials). This account explains why some research has appeared to support the theory of equal relatedness and complete substitutability, even though the current research demonstrates that this theory is correct only in certain testing conditions; in the present experiments, the stimuli were functioning as members of an equivalence class, as evidenced by participants' performances on the derived relations tests, while still being distinguishable and not wholly interchangeable, as evidenced by participants' performances on the within-class feasibility probes. To paraphrase George Orwell, all members of an equivalence class are equal (as evidenced by performance on cross-class tests), but some are more equal than others (as evidenced by performances on within-class tests).

To summarize, Fields and Verhave (1987) identified four variables that defined the structures of all equivalence classes: class size, nodes, training directionality, and nodal density. Two of these, nodes and training directionality, have now been shown to influence the level of relatedness between stimuli in a class. With regard to nodes, the greater the nodal distance between two stimuli, the less relational strength the stimuli have towards each other. With regard to directionality of training, the present research demonstrates that the extent to which the stimuli represent a derived relation that follows the same directionality as the trained relations have greater relational strength than those that require a reversal of directionality in order to be linked. That is, training directionality affects the number of derived operations required to link any two members of the class, so that where the two stimuli are of an equal nodal distance from the sample stimulus, the one that requires the reversal of directionality necessarily requires one additional logical operation to link the stimuli. The present research demonstrated that there is an inverse function between the logical operations needed for linkage and the strength of the stimulus-stimulus relation. These effects are present on a continuing, rather than temporary, basis.

Despite the differences in relatedness, however, the stimuli within a class are still more closely related to each other than to stimuli in other classes. Therefore, cross-class emergent relations probes produce selection of stimuli from the same class, which appears to support the theory of equal relatedness, even though the stimuli in the class are in fact differentially related to each other. The current experiment also extends the findings of prior research that found that the use of a single nameable pictorial stimulus in each class can aid in the acquisition of baseline relations and formation of equivalence classes.

The present research and that conducted by Moss-Lourenco and Fields (2011) both contained a procedural parameter that might have influenced the results obtained in each of the experiments. Prior to the administration of the within-class relational assessment tests, participants received the within-class feasibility tests. It is possible that these feasibility tests primed participants to respond in certain ways during the within-class tests. For example, in the first feasibility test, one of the probes contained the B and G stimuli as comparisons. The G stimulus was also one of the comparisons presented during the first within-class relational assessment test. It is possible that simply having seen the G stimulus as a comparison during the within-class feasibility tests enhanced the likelihood that participants would select the G stimulus during the within-class relational assessment tests. Future research, in which the feasibility tests are not presented before the relational assessment tests, would be informative on this point. If participants in such an experiment performed similarly to those in the present two experiments, this performance would indicate that the feasibility tests did not influence participants' performance in the present experiment, bolstering the findings of differential strength among members of a class based on relational differences. If, on the other hand, participants no longer demonstrate a preference for the comparison stimuli that represent the less complex relations, this performance would indicate that the feasibility tests somehow influenced responding on the relational assessment tests. The possible effect of this parameter may cast additional light on the factors that influence the expression of structural variables on the relatedness of stimuli in equivalence classes.

As a final note, while the current research undermines the universality of the argument for substitutability of members within an equivalence class, it leaves the following question unanswered. It has now been established that stimuli within an equivalence class can be differentially related on two dimensions, nodal distance and relational type. The relative strengths of these two structural parameters have not yet been explored. This task can be accomplished in future research that pits transitive relations of various nodal distances against equivalence relations with nodal distances that do not match those for the transitive relations. If a 4-node transitive relation is stronger than a 4-node equivalence relation, and a 2-node relation is stronger than a 4-node relation, will a 2-node equivalence relation be as strong as a 4-node transitive relation? That is, if a within-class relational assessment test were to use as comparisons a nodally proximal stimulus that represented an equivalence relation and a nodally distal stimulus that represented a transitive relation, which stimulus would participants prefer? Future research examining the relative preferences for nodal distance and relational type would further advance knowledge in this area and characterize how nodal distance in combination with relational type influences the relatedness of stimuli in an equivalence class. This would make more comprehensive our knowledge of factors that influence relatedness of class members alone and in combination.

Table 1. Number of blocks per phase required by participants in Experiment 1.

	Phase 1 KB trng	Phase 2 BL Acq	Phase 2 FB Red	Phase 3 ECF	Phase 4 Feasibility	Phase 5 Preference	Phase 6 ECF re-administration
Min req'd	4	2	3	3	2	3	3
Max allowed	N/A	N/A	N/A	12	6	9	12
NW	4	8	9	3	2	3	3
JP	4	12	3	12	---	---	---
AB	6	22	5	12	---	---	---
SS	4	27	---	---	---	---	---
CD	4	61	---	---	---	---	---
EF	6	23	---	---	---	---	---

Legend:

KB trng	Instructions, keyboard familiarization, and response training
BL Acq	Establishment of baseline relations using the simultaneous protocol
FB Red	Maintenance of baseline relations during feedback reduction
ECF	Equivalence class formation probes
Feasibility	Within-class feasibility tests
Preference	Within-class relational assessment tests
ECF re-administration	Re-administration of derived relations test

Table 2. Number of blocks per phase required by participants in Experiment 2.

	Phase 1 KB trng	Phase 2 BL Acq	Phase 2 FB Red	Phase 3 ECF	Phase 4 Feasibility	Phase 5 Preference	Phase 6 ECF re-administration
Min req'd	4	2	3	3	2	3	3
Max allowed	N/A	N/A	N/A	12	6	9	12
CL	4	5	3	4	2	3	3
DP	5	4	3	3	2	3	3
MG	5	5	3	3	3	4	3
DF	5	6	5	8	2	8	3
JV	6	19	5	5	3	6	3
TL	4	14	3	12	---	---	---
FB	6	44	4	12	---	---	---
JE	6	7	3	12	---	---	---
ES	7	39	10	12	---	---	---
YM	14	8	3	12	---	---	---
AB	5	27	---	---	---	---	---
LO	6	15	---	---	---	---	---
AA	7	27	---	---	---	---	---

Table 3. Differences in results for participants in Experiment 1 (all CVCs) vs. Experiment 2 (one pictorial stimulus).

	Experiment	
	1	2
N	6	13
# BL acquisition	3	10
% BL acquisition	(3/6) 50%	(10/13) 77%
# BL acq and no class formation	2	5
# BL acq and class formation	1	5
% BL acq and no class formation	(1/3) 33%	(5/10) 50%
% ecf /N	(1/6) 17%	(5/13) 38%

Figure 1. Performance on within-class feasibility and relational assessment tests in Experiment 1.

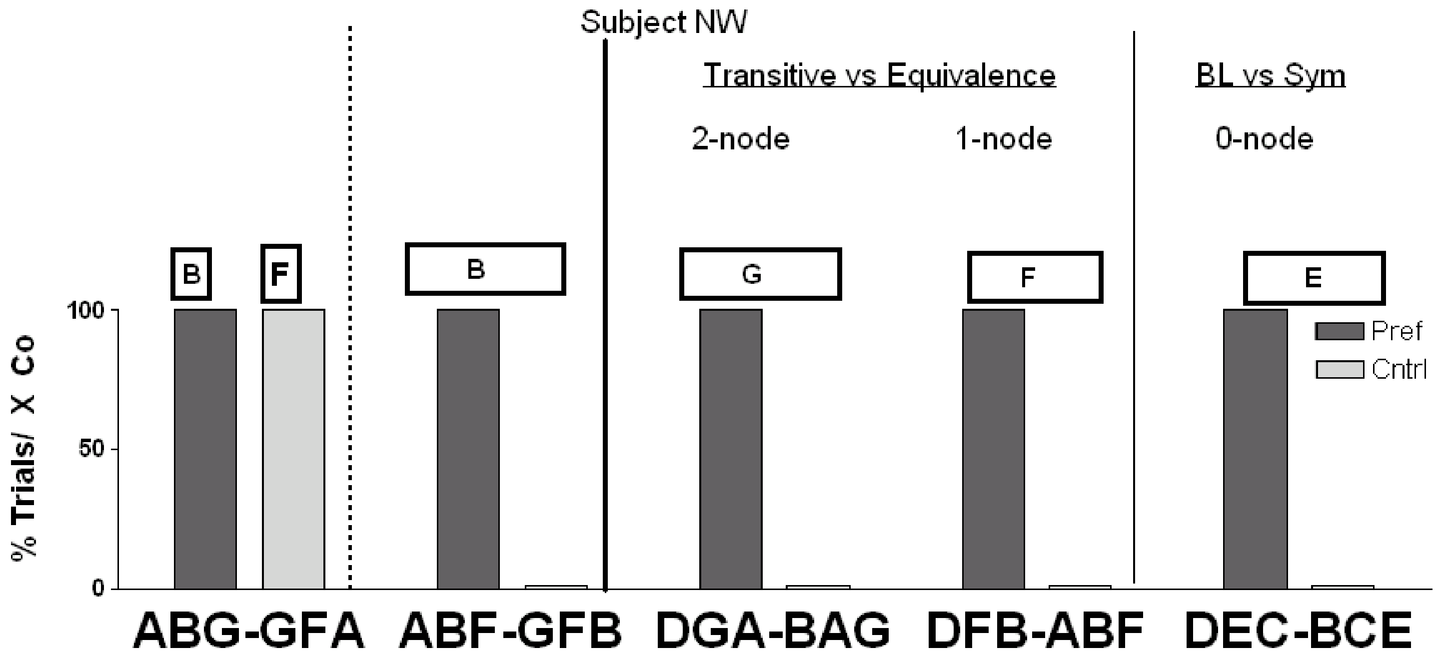
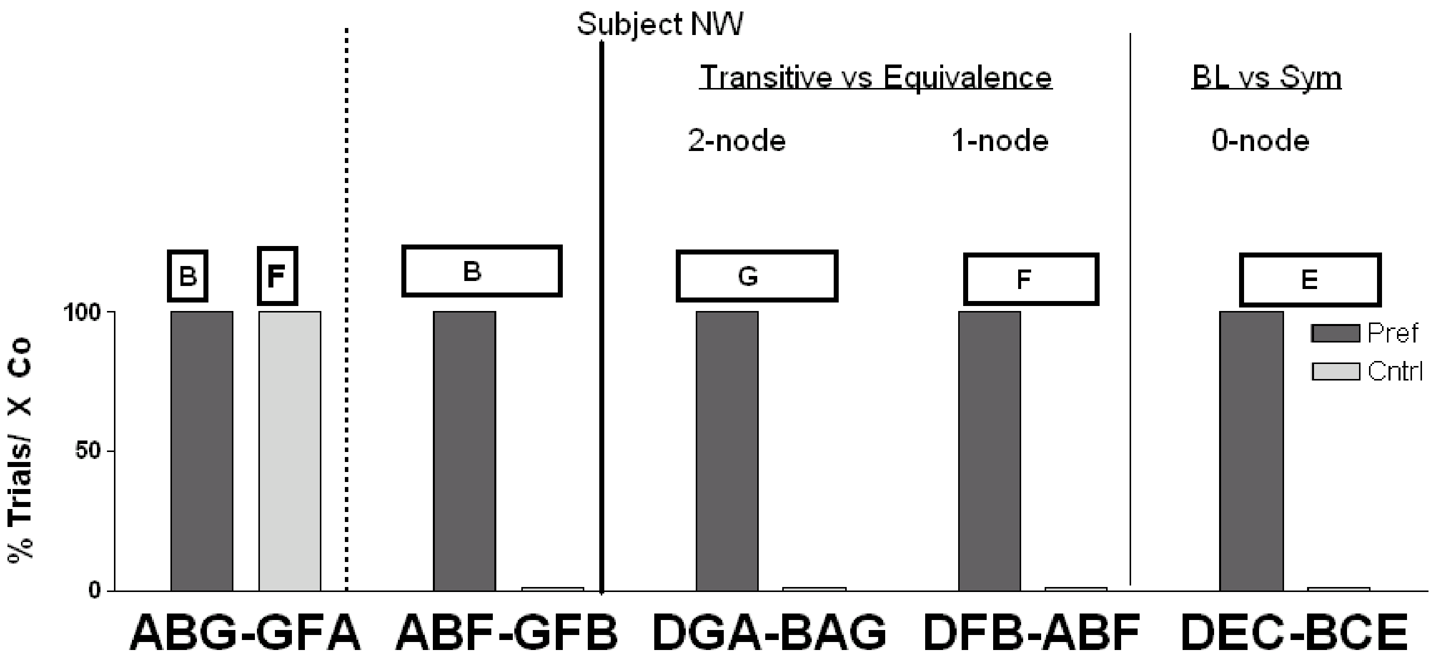


Figure 2: Performance on within-class feasibility and relational assessment tests in Experiment 2.



References

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

Barnes, D., Browne, M., Smeets, P., & Roche, B. (1995). A transfer of functions and a conditional transfer of functions through equivalence relations in three- to six-year-old children. *The Psychological Record, 45,* 405-430.

Barnes, D. & Keenan, M. (1993). A transfer of functions through derived arbitrary and non arbitrary stimulus relations. *Journal of the Experimental Analysis of Behavior, 59,* 61-82.

Bentall, R, Jones, R., & Dickins, D. (1998). Errors and response latencies as a function of nodal distance in 5-member equivalence classes. *The Psychological Record, 49,* 93-115.

Buffington, D.M., Adams, B.J., & Fields, L. (1997). Enhancing equivalence class formation by pretraining of other equivalence classes. *The Psychological Record, 47,* 69–96.

Cumming, W.W. & Berryman, R. (1965). The complex discriminative operant: Studies of matching-to-sample and related problems. In D.I. Mostofsky (Ed.), *Stimulus generalization* (pp. 284-330). Stanford, CA: Stanford University Press.

Fields, L., Adams, B.J., & Verhave, T. (May, 1989). Relative strength of symmetrical, transitive and equivalence relations after equivalence classes are fully formed. Paper presented in a symposium at the Annual Meeting of the Association for Behavior Analysis, Milwaukee, WI.

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

Fields, L., Adams, B.J., Verhave, T., & Newman, S. (1993). Are stimuli in equivalence classes equally related to each other. *The Psychological Record, 46,* 85-105.

Fields, L., Arntzen, E., Nartey, R. & Eilefsen, C. (accepted). Effects of a meaningful, a discriminative, and a meaningless stimulus on equivalence class formation. *Journal of the Experimental Analysis of Behavior.*

Fields, L., Landon-Jimenez, D.V., Buffington, D.M., & Adams, B.J. (1995). Maintained nodal distance effects after equivalence class formation. *Journal of the Experimental Analysis of Behavior, 64,* 129-146.

Fields, L., & Moss, P. (2008). The formation of partially elaborated and fully elaborated generalized equivalence classes. *Journal of the Experimental Analysis of Behavior, 90*, 135-168.

Fields, L., Newman, S., Adams, B., & Verhave, T. (1992). The expansion of equivalence classes through simple discrimination training and fading. *The Psychological Record, 42,* 3-15.

Fields, L. & Reeve, K. (2001). A methodological integration of generalized equivalence classes, natural categories, and cross modal perception. *The Psychological Record, 51*, 67-88.

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

Fields, L. 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, 443-466.

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

Fields, L. & Watanabe-Rose, M. (2008). Nodal structure and the partitioning of equivalence classes. *Journal of the Experimental Analysis of Behavior, 89,* 359-381.

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

Kennedy, C.H. (1991). Equivalence class formation influenced by the number of nodes separating stimuli. *Behavioural Processes, 24,* 219-245.

Kennedy, C.H., Itkonen, T., & Lindquist, K. (1994). Nodality effects during equivalence class formation: An extension to sight-word reading and concept development. *Journal of Applied Behavior Analysis, 27*, 673-683.

McIlvane W.J. & Dube, W.V. (2003). Stimulus control topography coherence theory: Foundations and extensions. *The Behavior Analyst, 25,* 195-213.

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, 343-368.*

Rehfeldt, R.A. & Hayes, L.J. (1998). Untrained temporal differentiation and equivalence class formation. *The Psychological Record, 48,* 481-510.

Saunders, R.R., Wachter, J., & Spradlin, J.E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discriminations procedures. *Journal of the Experimental Analysis of Behavior, 49,* 95-115.

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

Sidman, M. (1994). Equivalence relations and behavior: A research story. Boston, M.A.: Author's Cooperative.

Sidman, M. (1987). Two choices are not enough. Behavior Analysis, 22 (1), 11-18

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,* 21-42.

Sidman, M. & Tailby, W. (1982). Conditional discrimination vs. matching-to-sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37,* 5-22.

Sidman, M., Wynne, C.K., Maguire, R.W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the Experimental Analysis of Behavior, 52,* 261-274.

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

Tomanari, G., Sidman, M., Rubio, A., & Dube, W. (2006). Equivalence classes with requirements for short latencies. *Journal of the Experimental Analysis of Behavior, 85,* 349-369.

Wulfert, E. & Hayes, S.C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50,* 125-144.