The Psychological Record, 1998, 48, 93-115

ERRORS AND RESPONSE LATENCIES AS A FUNCTION OF NODAL DISTANCE IN 5-MEMBER EQUIVALENCE CLASSES

R. P. BENTALL, R. M. JONES, and D. W. DICKINS
University of Liverpool, UK

Two experiments were conducted to investigate the relationship between nodal distance, response accuracy, and response latency during testing for emergent relations. In both experiments, undergraduate subjects first learned A-B, B-C, C-D, and D-E constituent relations of six 5-member equivalence classes. In Experiment 1, only selected tests of trained and of 0-, 1-, and 2-node tests of emergent relations were carried out in order to avoid testing of 0- or 1-node relations that might form constituents of the 2-node relations which were tested. In Experiment 2, all possible trained and derived relations were tested in random order. Although considerable individual variability was observed in both response times and accuracy for the 14 subjects completing Experiment 1, latencies for correct responses generally increased and response accuracy decreased as a function of nodal distance. There was no nodal distance effect for latencies of incorrect responses. In Experiment 2, these relationships between response times, response accuracy, and nodal distance were observed for 3-node relations in 5 out of 6 subjects. Analysis of error response latencies for 2 subjects who made sufficient errors revealed a nodal distance effect for 1 subject but not for the other. In both experiments, response times decreased as testing progressed, but response accuracy increased during testing only in the second experiment.

In experiments on stimulus equivalence human subjects first learn a minimal number of relations between individual stimuli in a set. Each stimulus has a trained relation with at least one other in the set, so that all the stimuli are minimally interlinked. If the subjects are then able to demonstrate, without further training, additional relations between the

All correspondence should be addressed to David Dickins, Department of Psychology, The Eleanor Rathbone Building, Liverpool University, PO Box 147, Liverpool L69 3BX, UK. E-mail: dickins@liverpool.ac.uk

We acknowledge the help of R. J. Bowles, M. Brown, D. Chadwick, D. Hughes, A. Garside, B. Rogers, and S.Turner who contributed to the collection of data for Experiment 1. We also thank Tom Critchfield and William Dube for helpful comments on an earlier draft of this paper.

stimuli, which have not been taught, these are known as 'emergent' relations. If all the stimuli become fully interrelated in this way they are said to form an equivalence class. Sidman and Tailby (1982) have proposed a number of criteria for the identification of such classes: (i) *reflexivity:* given a sample stimulus A^n subjects must be able to select an identical stimulus A^n from an array of comparison stimuli; (ii) *symmetry:* having acquired the trained relation A^n-B^n between a sample stimulus A^n and a comparison stimulus B^n, subjects must be able to relate the stimuli in the reverse order, B^n-A^n, without further training; (iii) *transitivity:* subjects should be able to select stimuli that are related to each other by means of their separate relations with a common stimulus, that is, having learned the relations A^n-B^n and B^n-A^n, subjects must be able to select the stimulus C^n from a group of comparisons when presented with the stimulus A^n. The combined relation of transitivity with symmetry (having learned the relations A^n-B^n and B^n-C^n between three stimuli, subjects must be able to select the stimulus A^n from a group of comparisons when presented with the stimulus C^n) has been described as *the* equivalence relation by Fields and Verhave (1987).

The phenomenon of stimulus equivalence has generated considerable interest because of the apparent difficulty of demonstrating equivalence classes with preverbal humans (Devany, Hayes, & Nelson, 1986) and nonhuman animals (Hayes, 1989; Saunders, 1989; but see Cerutti & Rumbaugh, 1993, and Schusterman & Kastak, 1993), and because of the formal similarity between equivalence phenomena and language (Hayes & Hayes, 1992), where the relations between a spoken word, a heard word, and their common referent have been said to form an equivalence class (Sidman & Tailby, 1982). However, the precise relationship between equivalence classes and verbal behavior remains controversial. Whereas some authors have argued that the capacity to form equivalence classes is a necessary condition for language acquisition (Sidman, 1990), others have argued that verbal behavior is required for the formation of equivalence classes (Dugdale & Lowe, 1990), or that equivalence classes are but a subset of the arbitrary relations that humans are able to synthesize as a consequence of language learning (Hayes & Hayes, 1992).

Fields and Verhave (1987) have described the mediating stimuli or common links that allow transitive and equivalence performance as 'nodes.' The number of nodes which separate two stimuli between which an emergent relation is formed can be described as the 'nodal distance.' These authors suggested that the control exerted by one stimulus over another during tests of equivalence relations will be a function of the nodal distance between them, with control decreasing as the nodal distance increases. In an experiment designed to test this hypothesis, Fields, Adams, Verhave, and Newman (1990) trained subjects in A-B, B-C, and C-D relations to form 2-node, 4-member equivalence classes. The number of correct choices (selection of comparisons compatible with the existence of equivalence relations) made in initial tests of emergent relations was found to decrease with the number of nodes implicated in the tests. However, learning was observed to take place during the testing of emergent relations, so that accuracy of responding increased on further testing. Consistent with the findings of Fields et al. (1990), Kennedy (1991) demonstrated decreasing accuracy as a function of nodal distance in tests of emergent relations in 7-member equivalence classes.

We have suggested that response latencies might provide a further indication of the control exerted by stimuli involved in transitive relations, such that the time taken to select a comparison stimulus compatible with equivalence relations should increase with increasing nodal distance (Bentall, Dickins, & Fox, 1993). We have also argued that this relationship between nodal distance and response latencies should be absent when responses in tests of emergent relations are under the control of a common naming response to the constituent members of an equivalence class. Under these latter circumstances, the common name might be regarded as a node linking the constituent members, so that all relations between stimuli have a nodal distance of 1.

Wulfert and Hayes (1988) reported that, following the establishment of 3-member equivalence classes, response latencies on transitivity trials (1-node) were longer than those for trained relations. We also investigated response latencies on tests for 3-member equivalence classes, comparing all trained and emergent relations following training with visual stimuli (Bentall et al., 1993). In one experiment, stimuli were selected which fell into common semantic categories so that the members of each potential equivalence class could readily be labeled with a common name. In this experiment there were equal response latencies on tests for trained, symmetrical, transitive, and equivalence (transitivity with symmetry) relations. However, when training was conducted with abstract stimuli or stimuli that could not be readily labeled with a common name, response latencies on tests of 1-node relations (transitivity and equivalence) were longer than response latencies on tests of 0-node relations (trained relations and symmetry). No differences in response times could be attributed to the direction of tests (symmetrical versus in the direction of training, such as B^n-A^n versus A^n-B^n) at either 0 or 1 nodes.

As in the study by Fields et al. (1990), evidence of learning during testing of emergent relations was observed, so that response latencies decreased and accuracy of responding increased as testing progressed. In a third experiment only abstract stimuli were used, but subjects in one group were pretrained to label the stimuli using 18 individual names whereas in another group they were taught a common name for all of the stimuli belonging to each potential equivalence class (6 names in all). With individual names, longer response latencies were found for 1-node relations than for 0-node relations. With the common names, response latencies for all trained and emergent relations were equal.

The response latency observations of Wulfert and Hayes (1988) and of Bentall et al. (1993) were limited by the fact that testing was carried out only with 3-member equivalence classes. For 3-member classes it is possible that the differences observed between 0-node and 1-node relations reflect a special novelty of the 1-node relations rather than a systematic functional relationship between nodal distance and response times. For this reason it is important to establish whether error rates and response latencies are an increasing function of nodal distance as predicted. Although the answer to this question is known with respect to response accuracy (Fields et al., 1990; Kennedy, 1991), no studies have yet been reported in which response latencies have been observed for equivalence classes larger than 3 members. In this paper we therefore report an investigation of the relations between response times, accuracy, and nodal distance for 5-member classes.

Certain observations indicate that tests of relations which are constituents of larger transitive relations might affect subjects' responses on tests of the larger relations. Fields et al. (1990) found that the probability of responding correctly on tests of equivalence relations increases as a function of the repetition of those tests. Lazar, Davis-Lang, and Sanchez (1984) and Sidman, Kirk, and Willson-Morris (1985), working mainly with children, found that, with most subjects, multi-nodal emergent relations were difficult to establish without first establishing, by testing, emergent relations involving 1 node. Therefore, in the first experiment reported here, only selected trained and emergent relations were tested following training. No tests were carried out of 0- or 1-node relations which formed part of the pathway of those 2-node relations which were tested (e.g., testing of the relation B^n-C^n was avoided when the relation A^n-D^n was to be tested).

Experiment 1

Method

Subjects

Nineteen undergraduate students from the Department of Psychology at Liverpool University served as subjects in the experiment. Five were eliminated from the study, 1 because of failure to reach the criterion for acquisition of trained relations during the the training phase, 3 because of computer errors causing loss of data, and 1 because he responded randomly during tests of trained relations as revealed both by error scores and his self-report. Therefore data from 14 subjects are reported here. At the time of testing all subjects were naive about the purpose of the experiment. Seven participated as part of an undergraduate practical class (see acknowledgements) and the remaining subjects were volunteers who did not receive course credits or payment for taking part.

Apparatus and Test Stimuli

Testing took place in small rooms in the Department of Psychology at the University of Liverpool. Test stimuli were presented and responses were recorded using Apple Macintosh IIci and LSII computers according to the method described by Bentall et al. (1993) and Dickins, Bentall, and Smith (1993).

Thirty stimuli from the Mobile picture font were used to establish six potential equivalence classes (see Figure 1). The stimuli were chosen and sequenced in order to minimize the possibility of preexperimental associations between some of the stimuli within potential equivalence

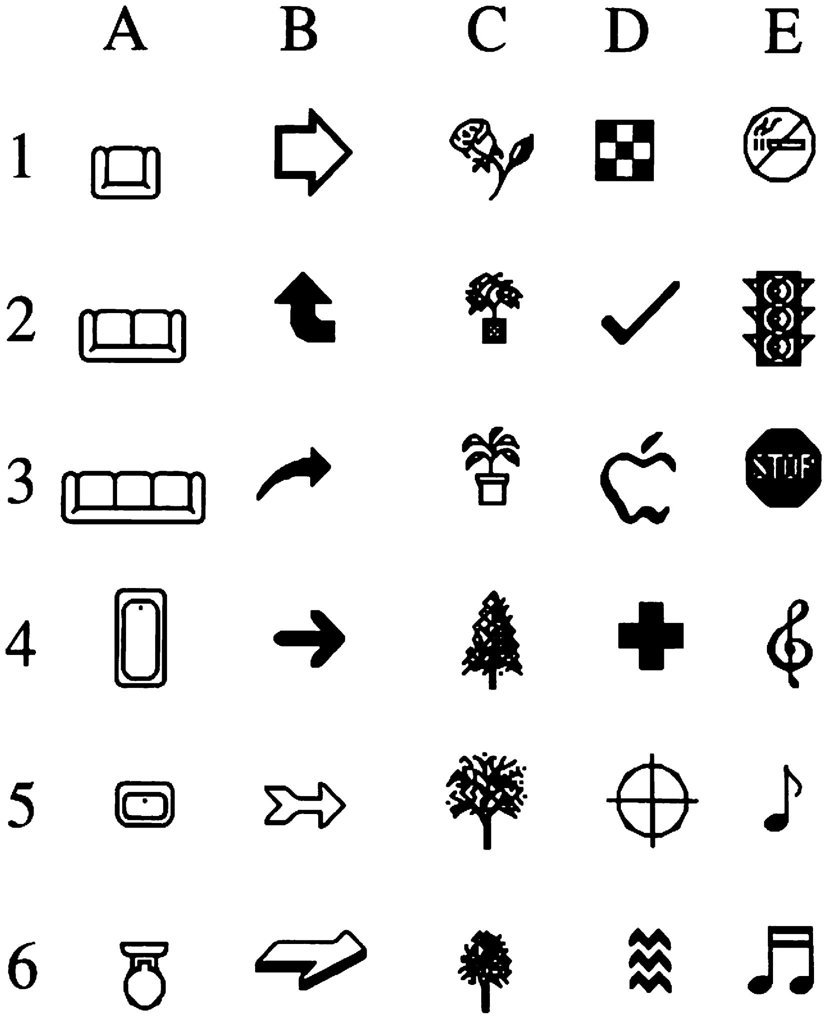

Figure 1. Stimuli used in Experiment 1 and their assignment to potential equivalence classes for one group of subjects. (For a second group of subjects the same stimuli were assigned differently to the six equivalence classes to control for the possible effects of particularly salient stimulus pairings.)

classes, and also to minimize the extent to which subjects could use verbal rules to mediate stimulus relations. Therefore items that were judged to be similar to each other were grouped in a way that was orthogonal to the potential equivalence classes. To further minimize the possibility of such confounding effects, two sets of potential equivalence classes were created with the same stimuli, allowing subjects to be randomly allocated to one or other of the two sets.

Procedure

Training: A matching-to-sample (MTS) paradigm was used to teach AB, BC, CD, and DE trained relations. In each training trial, the sample stimulus was presented on the left side of the computer monitor for 2 seconds. There was a zero delay between the removal of the sample and appearance of the six comparison stimuli. These were presented in a 2 x 3 lattice of windows on the right of the screen. Subjects chose a comparison stimulus by pressing one of an array of keys on the extended Macintosh keyboard which had been marked with colored tape. These keys (Keys 7, 8, 4, 5, 1, and 2) lay in a configuration similar to the comparison stimuli on the screen. Correct responses were followed by a text window containing the caption, "Well done - you got it right!" accompanied by an audible tone, whereas incorrect responses were followed by a different audible tone and a text box containing the caption, "You got it wrong! Bad luck, try again." Subjects then pressed the return key to start the next trial. No correction procedure was employed.

During initial AB training an 'errorless' procedure was employed to ensure maximum experimental control over the acquisition of trained relations. During the first stage of training, the presentation of the sample (say A^n) was followed after the zero delay by the presentation of the single correct comparison (say B^n) in one of the six windows, the others being empty. A criterion of 19/20 correct responses was required before subjects could move on to the next stage in which two comparison stimuli, one the correct stimulus (say B^n again) and one a foil drawn at random from the remaining five (B) stimuli, were presented following the sample. Again a criterion of 19/20 correct responses was required before subjects could move on to the next stage in which a second foil was added. This procedure was repeated by increasing the number of comparison stimuli, in a series of stages, until subjects chose between six stimuli (one correct and five foils). Following completion of AB training, the same multistage errorless method was used to establish BC, and then CD, and then DE relations. After a 10-minute break, as a final test of the adequate acquisition of trained relations, subjects were required to repeat the final 6-comparison stage of training, again in successive separate training episodes for each of the four types of trained relations, AB, BC, CD and DE, and again to the same criterion of 19/20 correct choices. During the whole of this training all correct and incorrect responses were followed by explicit feedback.

Testing: The method of presenting stimuli and collecting responses was similar to that employed during the training procedure except that no feedback was given. Test trials were selected in such a way as to exclude the possible effects of testing 0- or 1-node relations that were potential components of relations requiring greater nodality. The actual test trials administered for this purpose are represented in Figure 2. In each test block, 4 tests were presented for each of the following: trained relations, symmetrical relations, 1-node transitive relations, 1-node equivalence (transitivity with symmetry) relations, 2-node transitive relations, and 2-node equivalence relations. The rationale of this design precluded of course the testing of any 3-node relations. Thus, for each test block, subjects were required to make 24 sample-comparison choices. Four test blocks were presented, each consisting of the same 24 tests, although the order of presentation of the tests varied between the test blocks.

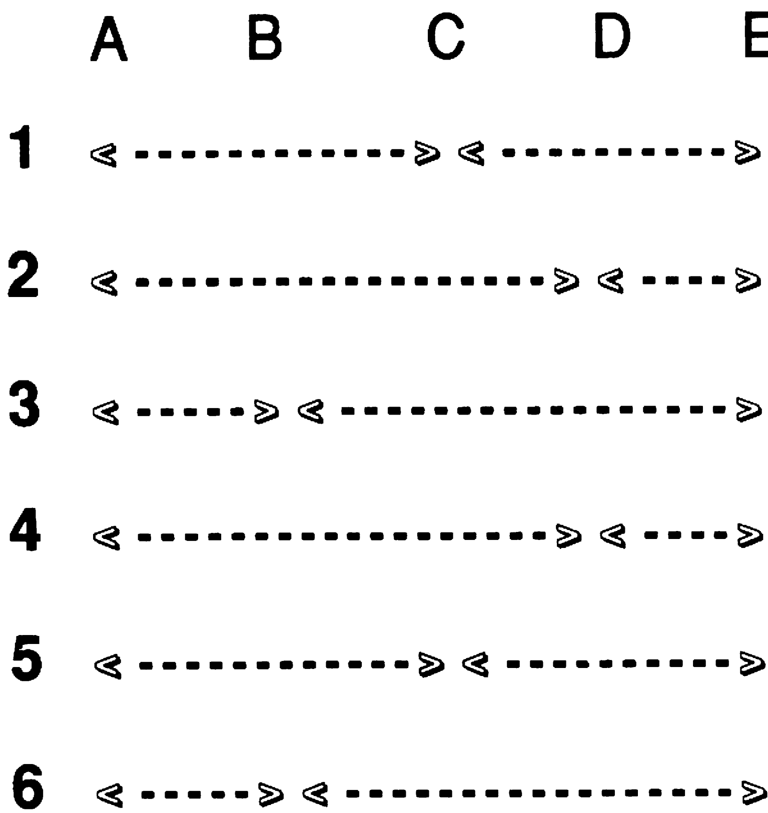

Figure 2. Specific relations tested from the six potential equivalence classes in Experiment 1.

Results

Statistical analyses comparing subjects' responses to the two alternative sets of equivalence classes derived from the 36 stimuli showed no significant differences and therefore data from the two sets have been pooled in all further analyses. The number of correct responses (out of a maximum of four in each case) for each subject on each type of test on each of the four test blocks are shown in Appendix 1, with subjects listed in decreasing order of their overall response accuracy. Individual data averaged across test blocks can be seen in the top panel of Figure 3. Considerable variability can be observed in these data, with some subjects (e.g., S1 and S2) performing at near 100% accuracy across all of the different tests whereas other subjects (particularly S9-14) produced a high proportion of errors. However, with the exception of S1, all subjects showed some evidence of a nodal distance effect. This was most evident in early test blocks and in the 6 subjects whose performance was least accurate. There is no clear indication of an effect due to the directionality of the tests (that is, whether or not the tests required symmetry).

The mean number of correct responses made by the subjects for the tests involving different nodal distances are shown in Figure 3, bottom. A three-way ANOVA (nodes x directionality x test blocks, all as within-subjects variables) on these data revealed a significant effect for nodes, $F(2, 26) = 26.63$, $p < .0001$, but nonsignificant effects for direction of tests, $F(1, 13) = 0.17$, $p = .68$, and for test blocks, $F(3, 39) = 2.18$, $p = .11$. The only significant interactions were for blocks x direction, $F(3, 39) = 4.81$, $p < .01$, and for nodality x direction, $F(2, 26) = 5.19$, $p < .02$. *T* tests used to carry out post hoc comparisons between the number of correct responses at nodal distances of 0, 1, and 2 nodes revealed that the differences between all nodal distances were significant, p at least $< .01$ for each comparison. Tests of simple effects revealed that error scores were higher on symmetrical tests than on tests that did not require symmetry on Block 1, $p < .02$, but not on subsequent trial blocks.

Median response latencies for each individual for each type of test for the four test blocks are given in Appendix 2. Some response latencies exceeded 20 seconds. Group means of individual median latencies for correct responses at different nodal distances are shown in Figure 4. With the exception of S1 and S6, there is a general tendency for response latencies to increase markedly as a function of nodal distance. However, there is no clear evidence of a consistent effect caused by the directionality of tests. These observations were confirmed by three-way ANOVAs (nodes x direction x test blocks) which were carried out separately for the 7 subjects who met a stringent criterion for the acquisition of equivalence classes (84% correct responses overall and a minimum of 83% correct responses in the final test block) and for the 7 subjects who did not meet this criterion for equivalence. The ANOVA for those who met the criteria for equivalence revealed

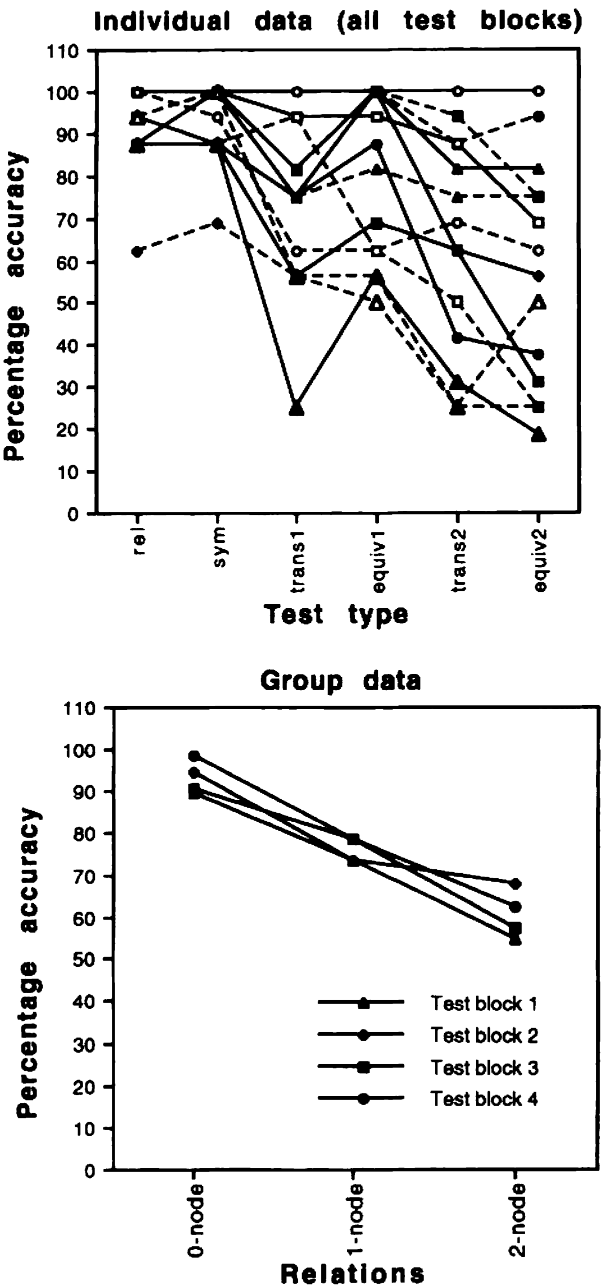

Figure 3. Individual percentage of correct responses averaged across test blocks (top panel) and mean percentage of correct responses (bottom panel) from Experiment 1.

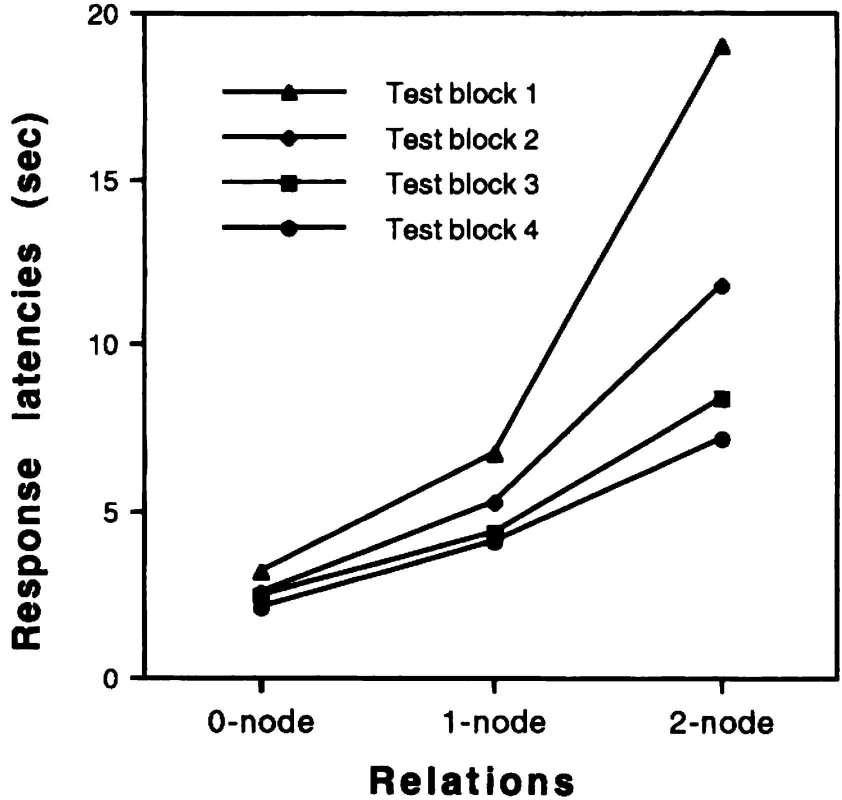

Figure 4. Mean of individual median response latencies for correct responses for 0-node, 1-node, and 2-node tests in the four test blocks of Expeniment 1.

significant effects for test blocks, $F(3, 18) = 3.30$, $p < .05$, and for nodality, $F(2, 12) = 6.68$, $p < .02$. The effect for direction, $F(1, 6) = 2.79$, $p = .15$, and all interactions failed to reach significance. Post hoc *t* tests showed that response latencies on Test Block 4 were significantly faster than on Test Blocks 1, $p < .01$, and 2, $p < .05$. Response latencies on 2-node tests were significantly longer than those on 1-node tests, $p < .05$, and 0-node tests, $p < .01$. The response latency difference between 0-node and 1-node tests did not reach significance.

Because of the low number of correct responses on some test blocks, the analysis of latencies for correct responses by those subjects who failed to meet the criteria for equivalence was carried out on data collapsed across test blocks. A highly significant effect was found for nodality, $F(2, 12) = 36.00$, $p < .0001$, but the effect for direction, $F(1, 6) = 0.65$, $p = .44$, and the interaction, $F(2,12) = .406$, $p = .68$, failed to reach significance. Post hoc *t* tests revealed significant differences between all three levels of nodality, *p* at least $< .05$.

These observations were supported by two further analyses. First, to control for the possibility that increasing latency was simply a function of increasing accuracy, a repeated measures ANCOVA was carried out on the latency data from all subjects, using the number of correct responses at each nodal distance as covariates; in this analysis subjects' data were collapsed across test blocks and direction of testing. The main effect for nodal distance remained significant, $F(2, 25) = 6.91$, $p < .005$. Second, an analysis was also conducted of incorrect responses. Sufficient data were available from only the 7 subjects who failed to meet the criteria for equivalence, and only for 1-node and 2-node relations, these subjects making very few errors on the 0-node tests. These data are given in Appendix 3. A two-way ANOVA conducted on the error latencies failed to reveal a significant effect for nodality, $F(1, 6) = 0.82$, $p = .40$, for direction, $F(1, 6) = 0.12$, $p = .74$, or for the interaction, $F(1, 6) = 1.63$, $p = .25$. Hence, there was no evidence of a systematic relationship between nodality and latencies for incorrect responses.

Discussion

In contrast with the findings of Lazar et al. (1984) and Sidman et al. (1985), who worked mainly with children, the majority of the adult subjects in the present experiment were able to make responses consistent with equivalence even on multi-node relations without prior testing of constituent 1-node relations. Consistent with the account of equivalence responding by Fields et al. (1990) and the previous data reported by Bentall et al. (1993), for most of the subjects in this experiment both error rates and response latencies increased in relation to the number of nodes implicated in the tests for emergent relations. Although error rates were low for many of the subjects, very long response latencies were nonetheless recorded for 2-node tests. The two exceptions were S1 and S6.

Further analyses of the response latency data revealed that a nodal distance effect was not evident for incorrect responses, and that the relationship between nodal distance and latency for correct responses was not merely a function of response accuracy. The nodal distance effect therefore appears to be a robust characteristic of the behavior of individuals who are learning equivalence relations, and is even evident in the correct responses of individuals who do not meet strict criteria for equivalence.

Although no systematic attempt was made to debrief subjects about their strategies during the experiment, both S1 and S6 spontaneously reported using a mnemonic strategy which consisted of assembling images of the items in each training set into a single mental image. Responding was more accurate on tests not requiring symmetry on the first test block only but, with this exception, there was little evidence that error rates and response latencies were affected consistently by the direction of tests. Overall, tests involving symmetry seemed to be completed about as accurately, and in a similar time, as the corresponding nonsymmetrical tests.

Except in the case of S1 and S6 there was a marked reduction in response latencies over successive test blocks as we had previously found (Bentall et al., 1993). However, in contrast to the findings of Fields et al. (1990), statistical analysis did not reveal evidence that responding became more accurate with repeated testing during the present experiment. Two factors intrinsic to the design of the experiment may have limited the opportunities for learning during the test phase: first, the small number of trials in each test block; and second, the exact structure of the testing procedure which was designed to eliminate the possibility that repeated testing of component trained relations or 1-node relations would influence performance on single- or bimodal equivalence relations within the same equivalence class.

A second experiment was therefore conducted in which all possible relations were tested during the testing of multi-node equivalence relations. The method of testing equivalence relations was therefore similar to that employed in our previous research (Bentall et al., 1993), in that all possible relations were tested following training on AB, BC, CD, and DE relations. This design also allowed the testing of the 3-node relations AE and EA.

Experiment 2

Method

Subjects

Six undergraduate students served as subjects. All were volunteers who did not receive course credits or payment for their participation, and all were naive to the purpose of the experiment.

Apparatus and Test Stimuli

The same apparatus and test stimuli were used as described in Experiment 1. Four subjects were assigned to one set of equivalence classes and two were assigned to the second set. However, no attempt was made to assess differences in responding to the two sets because these differences had been found not to be significant in Experiment 1, and because of the small number of subjects employed.

Procedure

Because of the large number of trials the experiment was conducted over 2 consecutive days.

Training: This was identical to that carried out in Experiment 1 except that the final group of training trials (in which subjects were tested to criterion on AB, BC, CD, and DE relations) was carried out at the beginning of the second day instead of after a short break.

Testing: The method of presenting stimuli and the responses required were identical to those in Experiment 1, but the structure of the test blocks was somewhat different, with each possible trained and

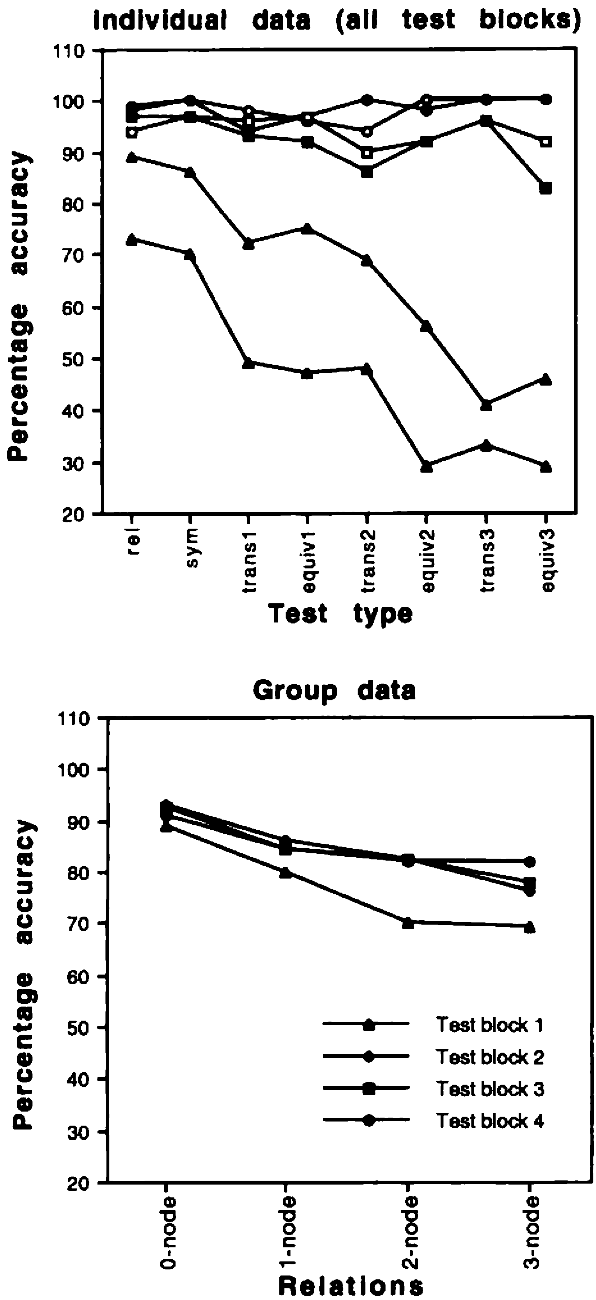

Figure 5. Individual percentage of correct responses averaged across test blocks (top panel) and mean percentage of correct responses (bottom panel) from Experiment 2.

emergent relation being tested once within each test block in a pseudorandom order. Thus, each test block consisted of 24 tests of trained relations, 24 symmetry tests, 18 1-node transitive relations, 18 1-node equivalence (symmetrical) relations, 12 2-node transitive relations, 12 2-node equivalence relations, 6 3-node transitive relations, and 6 3-node equivalence relations, making 120 trials in total. There were four test blocks so that subjects completed 480 trials by the end of the experiment. Testing took approximately 2 hours.

Results

Percentage correct responses for each subject on each type of test and in each test block are given in Appendix 4. The top panel of Figure 5 shows for each subject the mean percentage of correct responses at each type of test averaged across the four test blocks, and in each test block for the subjects as a group. The bottom panel of Figure 5 shows group mean latencies for correct responses for each of the test blocks.

Four of the subjects performed with near 100% accuracy on many of the tests. However, for the remaining two subjects, S5 and S6, there is evidence of decreasing response accuracy on tests requiring increasing transitivity, and poor accuracy was observed on all tests in the case of S6. In the case of most of the subjects, when responding was not close to 100% accurate on the first test block, some evidence can be discerned of increasing accuracy as testing progressed. However, S5 and S6 showed no such evidence of learning during the test trials. A three-way ANOVA (nodes x direction x test blocks) carried out on the accuracy data revealed a significant effect for blocks, $F(3, 15) = 3.80$, $p < .05$, and a marginally significant effect for nodes, $F(3, 15) = 3.26$, $p = .051$, with the effect for direction of tests failing to reach significance, $F(1, 5)$ 1.44, $p = 0.28$. There were no significant interactions. *T* tests revealed that accuracy of responding was significantly greater on 0-node tests than on 2-node tests, $p < .05$ and on 3-node tests, $p < .05$.

Median response latencies for correct responses made by individual subjects are given in Appendix 5. Group means for these data are shown in Figure 6. All subjects with the exception of S6 showed clear evidence of increasing response latencies for responses requiring increasing numbers of nodes in transitivity. Gross inspection of the data from S6 revealed that response latencies for 3-node relations appeared to be no longer than those for 2-node relations.

A three-way ANOVA (nodes x direction x test blocks) was carried out on these data. (No subjects had to be excluded from this analysis because all made some correct choices in all nodes x direction x test blocks conditions.) Significant main effects were observed for nodal distance, $F(3, 15) = 17.04$, $p <.0001$, and for test blocks, $F(3, 15) = 44.35$, $p < .0001$, but not for direction of tests, $F(1, 5) = 0.491$, $p = .52$. There was also a significant interaction between nodes and test blocks, $F(9, 45)$ 5.85, $p < .0001$, but all other interactions were nonsignificant.

Post hoc *t* tests revealed significant differences between all test blocks with the exception of Test Block 3 versus Test Block 2, all comparisons p at least < .05. Significant differences were also observed between response latencies for 3-node relations and both 1- and 0-node relations, and between 2-node relations and 0-node relations, $p < .01$ for each comparison. Tests of simple effects revealed significant effects for nodality within each test block, p at least < .005.

As in Experiment 1, the latency data were subjected to further analyses to control for confounding variables and in order to explore the

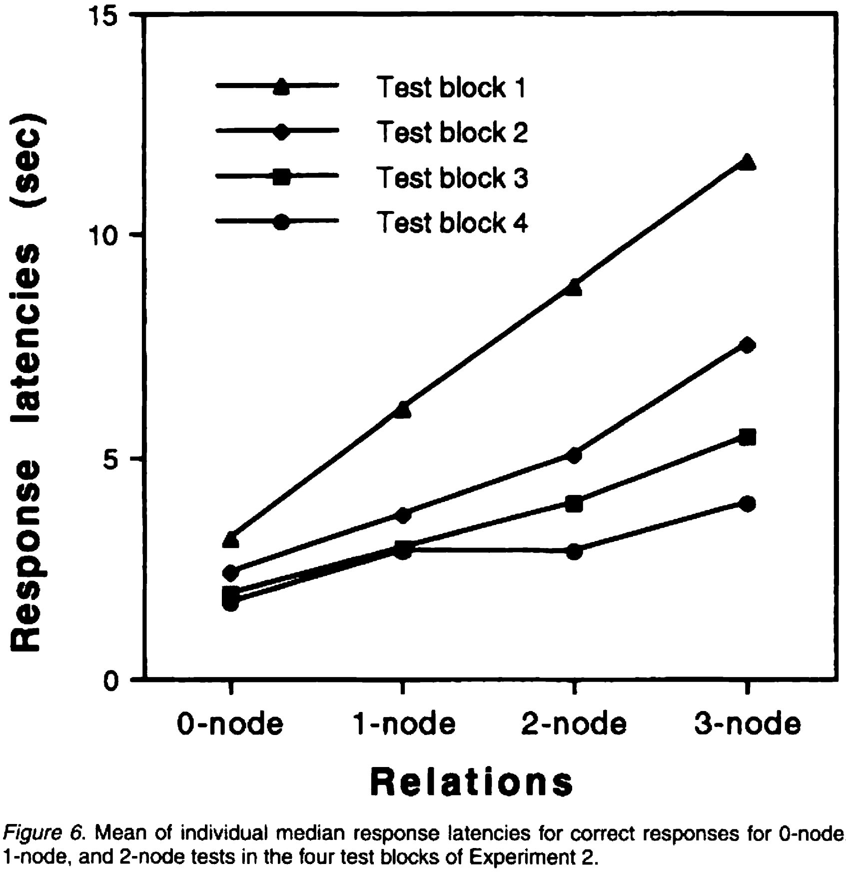

Figure 6. Mean of individual median response latencies for correct responses for 0-node, 1-node, and 2-node tests in the four test blocks of Experiment 2.

relationship between response latency and nodality for incorrect responses. A repetition of the three-way ANOVA including only those subjects who met a criterion of at least 85% correct responding at each nodal distance confirmed significant effects for nodality, $F(3, 9) = 15.55$, $p < .001$, test blocks, $F(3, 3) = 46.10$, $p < .0001$, and for the nodality x test block interaction, $F(9, 9) = 13.32$, $p < .0001$. Post hoc *t* tests on these data revealed significantly different latencies between all nodal distances, p at least < .05, except for the difference between 0-node and 1-node relations which failed to reach significance.

As in Experiment 1, a repeated measures ANCOVA was also calculated for all subjects collapsing the data across test blocks and directionality and using the percentage accuracy scores at each nodal distance as covariates. The effect for nodal distance remained significant in this analysis, $F(3, 14) = 8.45$, $p < .005$, indicating that the nodal distance effects observed were not simply a function of response accuracy.

Finally, error latency data for S5 and S6 averaged across test blocks are shown in Appendix 6. Gross inspection of these data revealed an apparent nodal distance effect for incorrect responses for S5 but not for S6. Because sufficient error data were available from only two subjects they were not subjected to a statistical analysis.

Discussion

In this study we tested the prediction that the control over responding exerted by equivalence relations is a function of nodal distance. In both Experiments 1 and 2, response latencies increased as a systematic function of nodal distance. This effect was evident even in the correct choices of most subjects who did not meet strict criteria for equivalence and was not simply a function of response accuracy. The nodal distance effect for latencies was particularly evident in early test blocks but remained even after substantial testing, during which response latencies decreased overall. Conversely, with the exception of 1 subject in Experiment 2, there was no evidence of a response latency effect for incorrect choices.

There was less evidence of a relationship between nodal distance and the probability of choices consistent with equivalence, especially in Experiment 2 in which 4 out of 6 subjects responded with a high degree of accuracy from the outset. However, in both experiments a systematic relationship between accuracy and nodal distance was observed in those subjects who made a relatively high number of errors in the early test blocks. Individual and group data indicated that some subjects' response accuracy improved across test blocks in Experiment 2 but this effect was absent in Experiment 1. Several authors (e.g., Fields et al., 1990; Saunders, Wachter, & Spradlin, 1988) have previously observed that substantial learning, as reflected by responses consistent with the formation of equivalence classes, occurred during testing without formal reinforcement of emergent relations.

The results of both of the present experiments indicate that the previous finding of longer latencies for 1-node relations in comparison with 0-node relations (Bentall et al., 1993; Wulfert & Hayes, 1988) also holds for relations involving a larger number of nodes and therefore cannot be accounted for simply by the novel conjunction of stimuli in transitivity tests. The present results are more consistent with the account of equivalence learning offered by Fields and his colleagues

(Fields et al., 1990; Fields & Nevin, 1993; Fields & Verhave, 1987) in which it is the number of nodes in a tested relation that determines both accuracy and the promptness of responding. However, it should be noted that such linear relations between nodal distance and response latencies and errors have not always been observed. In three previous experiments (Bentall et al., 1993) subjects responded with equal speed and accuracy on 0- and 1-node relations if the stimuli within potential equivalence classes belonged to clear natural groups, or if subjects had been taught common names for the stimuli within each training class. Thus, people are sometimes able to use strategies during equivalence testing which circumvent the effects of nodality.

A number of questions about the processes involved in the learning of equivalence classes remain unresolved. First, in the present study very little evidence was found of an effect for directionality (symmetrical versus nonsymmetrical tests), although symmetrical choices were less accurate on Test Block 1 of Experiment 1. We have previously suggested that human subjects acquire the ability to perform correctly on tests of symmetrical relations in the course of learning trained relations (Bentall & Dickins, 1994), and the present findings can be interpreted as consistent with this. Fields, Adams, Newman, and Verhave (1992) found that the order in which symmetrical and transitive relations are acquired varies from individual to individual, indicating that symmetrical and transitive responding may be to some degree dissociated. Further studies are required to throw light on specific processes involved in symmetrical responding.

Second, the long response latencies observed in the present study raise important questions about the behavior of the subjects during these intervals and the psychological processes involved in the acquisition of equivalence relations. In Experiment 2, for example, each node in a tested emergent relation resulted in an average increase in response latency of approximately 3 seconds during initial testing. Such latencies are of a different order of magnitude from those reported in studies of semantic memory (Anderson, 1985; Collins & Quillian, 1969) which have sometimes been compared with equivalence studies (Fields et al., 1990; Hayes & Hayes, 1992). The observed relations between latencies and nodality suggest that each node evokes an iteration of some serial mediating process, and that the iterations of this process are systematically related to the component relations of the emergent relations being tested. On this view, each internodal link requires (a) the generation of a mediating response that specifics the next node and (b) identity matching of this node with the available choices on the stimulus display. In some cases it may be that such mediating responses are verbal, for example naming, as suggested by Dugdale and Lowe (1990) and Horne and Lowe (1996).

Third, it is interesting to note that response accuracy did not improve across test blocks in Experiment 1, whereas improved accuracy was observed in Experiment 2 and in previous studies (Bentall et al. 1993;

Fields et al., 1990). Improved accuracy during testing has been characterized as 'delayed emergence' by Sidman (1994) but has never been adequately explained. Specific design characteristics of Experiment 1 which prevented delayed emergence may therefore provide an important indication of the mechanisms responsible for this phenomenon. As a unique feature of the design was that subjects were prevented from learning short-node relations which later became components of longer-node relations, it is likely that the acquisition of such component relations is responsible for improved performance during testing. Consistent with this hypothesis, Lazar et al. (1984) and Sidman et al. (1985) found that longer-node relations are often acquired after component short-node relations. However, these studies mostly involved children, and it would be useful to attempt to replicate their findings in adults and to conduct further studies of possible mechanisms involved in delayed emergence.

References

ANDERSON, J. R. (1985). *The architecture of cognition.* New York: W. H. Freeman.

BENTALL, R. P., & DICKINS, D. W. (1994). Differences between animal and human learning: Implicit and explicit processes. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages.* London: Academic Press.

BENTALL, R. P., DICKINS, D. W., & FOX, S. R. A. (1993). Naming and equivalence: Response latencies for emergent relations. *Quarterly Journal of Experimental Psychology,* 46B, 187-214.

CERUTTI, D. T., & RUMBAUGH, D. M. (1993). Stimulus relations in comparative primate perspective. *The Psychological Record,* 43, 811-821.

COLLINS, A. M., & QUILLIAN, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior,* 8, 240-247.

DEVANY, J. M., HAYES, S. C., & NELSON, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior,* 46, 243-257.

DICKINS, D. W., BENTALL, R. P., & SMITH, A. B. (1993). The role of individual stimulus names in the emergence of equivalence relations: The effects of paired-associates training between names. *The Psychological Record,* 43, 713-724.

DUGDALE, N., & LOWE, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.), *Behavior analysis in theory and practice* (pp. 115-138). Hove: Lawrence Erlbaum.

FIELDS, L., ADAMS, B. J., NEWMAN, S., & VERHAVE, T. (1992). Interactions among emergent relations during equivalence class formation. *The Quarterly Journal of Experimental Psychology,* 45B, 125-138.

FIELDS, L., ADAMS, B. J., VERHAVE, T., & NEWMAN, S. (1990). The effects of nodality on the formation of equivalence sets. *Journal of the Experimental Analysis of Behavior,* 53, 345-358.

FIELDS, L., & NEVIN, J. A. (1993). Stimulus equivalence: A special issue of *The Psychological Record,* 43(4), 541-844.

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

HAYES, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, 51, 385-392.

HAYES, S. C., & HAYES, L. J. (1992). Verbal relations and the evolution of behavior analysis. *American Psychologist*, 47, 1383-1395.

HORNE, P. J., & LOWE, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, 65, 185-241.

KENNEDY, C. L. (1991). Equivalence class formation influenced by the number of nodes separating stimuli. *Behavior Processes*, 24, 219-245.

LAZAR, R. M., DAVIS-LANG, D., & SANCHEZ, L. (1984). The formulation of visual stimulus equivalences in children. *Journal of the Experimental Analysis of Behavior*, 41, 251-266.

SAUNDERS, K. J. (1989). Naming in conditional discrimination and stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, 51, 379-384.

SAUNDERS, R. R., WACHTER, J. A., & SPRADLIN, J. E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedures. *Journal of the Experimental Analysis of Behavior*, 49, 95-115.

SCHUSTERMAN, R. J., & KASTAK, D. (1993). A california sea lion (Zalophus californianus) is capable of forming equivalence relations. *The Psychological Record*, 43, 823-839.

SIDMAN, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. LeJeune (Eds.), *Behavior analysis in theory and practice* (pp. 93-114). Hove: Lawrence Erlbaum.

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

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, 5-52.

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.

Appendix 1

The number of correct responses (out of a maximum of 4 in each case) for each subject on each type of test (columns) on each of the four test blocks (rows) in Experiment 1.

	Rel	Sym	Trans1	Equiv1	Trans2	Equiv2
S1	4	4	4	4	4	4
	4	4	4	4	4	4
	4	4	4	4	4	4
	4	4	4	4	4	4
S2	4	4	3	4	3	4
	4	4	3	4	3	4
	4	4	3	4	4	4
	4	4	3	4	4	3
S3	4	4	4	3	4	1
	4	4	4	4	3	4
	4	4	4	4	3	2
	4	4	3	4	4	4
S4	4	4	4	4	4	2
	4	4	3	4	4	4
	3	4	3	4	4	2
	4	4	3	4	3	4
S5	3	4	2	4	3	1
	3	4	2	4	4	4
	4	4	4	4	3	4
	4	4	4	4	3	4
S6	4	4	3	3	3	3
	4	4	3	3	3	3
	4	4	3	4	3	3
	4	4	3	3	3	3
S7	4	4	4	4	3	2
	4	4	3	4	2	2
	4	4	2	4	2	2
	4	4	4	4	3	3
S8	4	3	2	3	2	3
	4	4	2	2	2	2
	4	4	3	2	3	2
	4	4	3	3	4	3
S9	3	2	3	3	2	0
	4	4	2	3	3	3
	4	4	4	4	2	2
	4	4	3	4	0	1
S10	4	3	4	2	2	0
	4	4	4	3	3	3
	2	3	3	2	1	1
	4	4	4	3	2	0
S11	4	3	1	2	3	0
	4	4	3	2	3	3
	3	4	3	3	2	1
	3	3	2	4	2	1
S12	3	4	2	2	1	1
	4	4	1	2	1	2
	4	4	3	2	1	3
	4	4	3	2	1	2
S13	4	4	1	2	2	2
	3	3	1	3	0	0
	3	3	1	2	1	0
	4	4	1	2	2	1
S14	2	2	3	2	1	1
	3	2	2	3	1	2
	1	3	3	2	1	0
	4	4	1	2	1	1

Appendix 2

Median response latencies for correct responses for each type of test (columns) in the four successive test blocks (rows) for subjects in Experiment 1.

	Rel	Sym	Trans1	Equiv 1	Trans2	Equiv2
S1	2.87	2.14	2.16	2.44	2.29	1.88
	1.49	1.95	2.03	2.07	2.05	1.77
	2.44	1.87	1.42	1.90	1.69	1.60
	1.20	1.58	1.46	1.45	1.64	1.89
S2	1.85	2.57	6.45	5.37	6.42	7.62
	1.42	1.87	3.33	2.85	4.20	5.28
	1.74	2.25	2.73	2.87	4.53	4.33
	2.53	1.70	2.90	6.73	5.66	3.18
S3	2.55	2.30	14.35	4.55	44.50	19.98
	4.30	2.45	6.80	4.50	14.06	17.80
	1.93	2.54	7.20	3.35	14.48	13.20
	3.14	1.64	3.74	2.76	9.03	12.26
S4	4.17	6.59	8.04	5.86	20.77	21.04
	2.74	2.88	9.62	10.81	13.31	10.66
	1.55	1.88	3.87	5.81	8.94	5.78
	2.61	2.17	7.15	4.93	11.98	6.54
S5	4.31	4.80	6.80	11.00	115.28	14.70
	3.55	3.26	8.63	4.72	27.73	15.67
	4.80	2.34	4.39	4.23	12.56	11.70
	2.08	1.73	3.98	3.80	6.85	10.07
S6	2.20	1.44	4.80	3.13	3.60	2.01
	1.71	1.54	2.61	2.61	2.85	1.58
	2.55	1.57	3.45	3.67	2.85	2.31
	2.75	1.91	2.45	2.08	2.01	1.75
S7	2.24	3.41	7.22	3.83	12.50	8.48
	2.94	2.94	7.78	4.84	22.95	25.47
	3.00	2.66	12.29	4.39	14.23	11.06
	1.95	1.95	2.23	3.64	6.95	10.47
S8	1.78	5.43	13.15	7.58	15.35	25.05
	1.06	4.72	3.83	5.43	16.73	12.73
	2.47	4.20	3.00	7.27	7.00	16.81
	1.75	3.16	2.75	16.82	9.07	10.05
S9	2.55	2.30	14.35	4.55	44.50	19.98
	3.90	7.76	21.66	6.05	20.52	25.42
	4.88	4.60	4.52	4.89	7.45	20.53
	3.74	2.55	3.72	3.60	-	14.67
S10	5.60	6.72	6.30	18.17	34.66	-
	9.90	4.19	6.79	18.63	13.47	36.35
	6.79	6.77	6.55	23.33	10.95	42.93
	6.63	7.23	8.41	21.33	12.80	-
S11	1.53	2.83	4.02	5.54	8.62	-
	3.23	1.95	5.13	3.96	9.15	10.57
	1.30	1.67	4.97	6.83	9.18	13.33
	1.33	0.97	2.20	6.05	5.66	9.15
S12	3.11	3.31	8.00	7.28	13.11	7.85
	1.83	3.00	3.00	9.53	8.01	9.16
	2.41	2.22	3.50	4.05	9.30	8.75
	1.51	2.18	1.95	3.80	9.75	9.90
S13	1.97	7.98	3.00	5.85	14.92	16.74
	2.11	5.03	4.85	6.91	-	-
	1.73	1.55	2.15	4.85	8.45	-
	3.42	1.93	2.63	6.89	30.36	7.78
S14	6.53	3.03	10.53	7.26	52.30	21.68
	5.23	2.08	7.81	17.45	23.48	19.04
	3.00	2.75	12.35	19.01	15.13	-
	3.21	4.66	4.82	10.99	22.65	11.07

Appendix 3

Latencies for incorrect responses (1-node and 2-node relations only) for Experiment 1 collapsed across test blocks.

	Trans 1	Equiv 1	Trans 2	Equiv 2
S8	10.62	25.98	31.44	21.68
S9	18.96	16.54	22.72	18.57
S10	39.43	24.91	34.53	20.97
S11	17.71	10.43	11.26	9.56
S12	4.67	9.62	13.34	8.96
S13	5.03	16.43	8.92	12.46
S14	13.49	16.98	16.09	22.05

Appendix 4

Percentage accurate responses for correct responses for each type of test (columns) and for each test block (rows) for subjects in Experiment 2.

	Rel	Sym	Trans1	Equiv1	Trans2	Equiv2	Trans3	Equiv3
S1	95.83	100.00	100.00	100.00	91.67	100.00	100.00	100.00
	100.00	100.00	94.44	88.89	83.33	100.00	100.00	100.00
	100.00	100.00	100.00	94.44	100.00	100.00	100.00	100.00
	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
S2	100.00	100.00	94.44	94.44	100.00	100.00	100.00	100.00
	95.83	100.00	94.44	94.44	100.00	100.00	100.00	100.00
	100.00	100.00	88.89	100.00	100.00	100.00	100.00	100.00
	95.83	100.00	100.00	100.00	100.00	91.67	100.00	100.00
S3	91.67	95.83	94.44	94.44	66.67	75.00	100.00	83.33
	95.83	91.67	100.00	94.44	91.67	91.67	100.00	83.33
	91.67	100.00	94.44	100.00	100.00	100.00	83.33	100.00
	100.00	100.00	94.44	100.00	100.00	100.00	100.00	100.00
S4	95.83	95.83	83.33	83.33	66.67	75.00	83.33	50.00
	95.83	100.00	100.00	100.00	91.67	100.00	100.00	100.00
	100.00	91.67	88.89	94.44	91.67	100.00	100.00	83.33
	95.83	100.00	100.00	88.89	91.67	91.67	100.00	100.00
S5	91.67	79.17	72.22	83.33	66.67	41.67	33.33	33.33
	95.83	91.67	83.33	77.78	83.33	66.67	50.00	50.00
	87.50	91.67	72.22	72.22	66.67	58.33	33.33	50.00
	79.17	79.17	61.11	66.67	58.33	58.33	50.00	50.00
S6	66.67	58.33	38.89	22.22	33.33	25.00	33.33	16.67
	75.00	75.00	55.56	50.00	50.00	33.33	16.67	16.67
	75.00	75.00	44.44	66.67	50.00	25.00	33.33	50.00
	75.00	70.83	55.56	50.00	58.33	33.33	50.00	33.33

Appendix 5

Response latencies for correct responses for each type of test (columns) and for each test block (rows) for subjects in Experiment 2.

	Rel	Sym	Trans1	Equiv1	Trans2	Equiv2	Trans3	Equiv3
S1	1.80	2.15	2.86	4.43	5.78	6.20	5.71	11.98
	1.97	1.55	1.60	2.09	2.91	3.57	3.60	4.51
	1.22	1.47	1.62	1.78	2.02	2.13	2.82	2.75
	1.38	1.28	1.39	1.74	1.68	1.32	2.15	2.21
S2	2.79	2.55	8.35	9.12	14.91	11.46	13.95	15.61
	2.22	2.44	3.52	3.20	8.07	5.73	6.49	9.77
	1.76	2.93	2.05	3.69	5.68	7.50	6.92	7.75
	1.58	2.46	3.25	3.90	3.13	4.32	5.40	5.58
S3	2.35	3.30	4.53	5.78	6.05	4.72	9.33	9.55
	1.85	1.69	2.94	2.30	4.73	2.23	3.76	3.27
	1.99	1.82	2.33	2.35	3.02	2.60	4.28	2.55
	1.49	1.68	2.05	2.69	2.75	2.17	2.81	3.30
S4	3.41	4.25	4.75	5.23	6.69	7.93	16.75	11.47
	2.27	2.29	2.66	3.10	4.47	4.30	3.95	7.86
	2.00	2.08	2.21	2.53	3.18	3.57	3.34	4.28
	1.75	1.86	2.90	2.02	2.98	1.88	3.92	3.79
S5	2.32	3.42	5.12	6.83	17.69	6.70	16.41	14.02
	2.43	1.95	4.25	4.60	5.88	5.22	5.45	26.08
	1.65	1.51	2.77	4.11	2.97	4.93	14.82	2.58
	1.65	2.12	3.73	4.75	3.43	4.82	5.33	6.13
S6	5.41	3.85	9.38	6.91	8.47	9.68	9.84	5.42
	3.67	4.07	5.52	8.57	5.74	7.99	8.93	6.93
	2.32	2.12	4.54	5.58	6.63	3.22	6.67	7.07
	1.56	1.68	2.92	2.85	3.43	3.06	4.30	2.88

Appendix 6

Error latency data for S5 and S6 in Experiment 2, averaged across test blocks.

	Rel	Sym	Trans1	Equiv1	Trans2	Equiv2	Trans3	Equiv3
S5	6.54	5.85	7.19	10.79	12.15	10.22	12.53	18.24
S6	7.05	7.22	6.87	6.64	7.02	5.15	6.17	6.75