Derived Stimulus–Response and Stimulus–Stimulus Relations in Children and Adults: Assessing Training Order Effects

Paul M. Smeets

Leiden University, Leiden, The Netherlands

and

Dermot Barnes-Holmes and Bryan Roche

National University of Ireland, Maynooth, Ireland

Preschool children and adults received training on three sets of successive discriminations: (1) A1-R1,A2-R2, (2) B1-R1,B2-R2, and (3) A1-R3,A2-R4. Then they received tests assessing derived stimulus–response relations (B1-R3, B2-R4) and stimulus–stimulus relations (e.g., A1-B1, A2-B2). Four training protocols were used. The protocols differed with regard to the order in which the sets were trained: Many-To-One (1-2-3), One-To-Many-1 (1-3-2), One-To-Many-2 (3-1-2), and One-To-One (3-2-1 or 2-3-1). The adults displayed class-consistent B-R and A-B performances over all conditions. The children displayed class consistent B-R performances more often in Many-To-One and One-To-Many than in One-To-One. Their A-B performances were highly consistent with the trained A-R and tested B-R performances. Present findings are consistent with the stimulus equivalence account rather than with the mediated response account. © 2001 by Academic Press

Key Words: children; adults; mediated generalization; stimulus equivalence; generalized conditional responding.

Previous studies on stimulus class formation have shown that animals and humans may derive novel stimulus–response relations from previously trained sets of stimulus–response relations (for an extensive review, see Zentall & Smeets, 1996). For example, in a study by Wasserman and DeVolder (1993), 20 4- and 5-year-old children were trained to place multiple exemplars of four different stimulus categories (flowers, chairs, people, and cars) at two different locations of a quadrant: flowers and chairs at the top right corner, and cars and people at the bottom left corner (initial training). Then they were trained to place flowers at the top left corner and cars at the bottom right corner of another

Address correspondence and reprint requests to Paul M. Smeets, Behavior Analysis Unit, Department of Psychology, Leiden University, P.O. Box 9555, 2300 RA, Leiden, Holland. E-mail: SMEETS@FSW.LEIDENUNIV.NL.

130

quadrant (reassignment training). During subsequent probes, 13 children (65%) placed the nonreassigned stimuli (chairs and people) at the same locations as the reassigned stimuli (chairs: top left; people: bottom right). This finding was similar to that previously obtained with pigeons (Wasserman, DeVolder, & Coppage, 1992).

Smeets, Barnes, and Roche (1997) demonstrated that the trained and/or tested stimulus–response relations may also produce class-consistent stimulus–stimulus relations. After being trained to (a) place two class-1 stimuli at one location and two class-2 stimuli at another location of a quadrant (A1-R1, B1-R1, A2-R2, and B2-R2) and (b) to place one member of each class at different locations of another quadrant (A1-R3,A2-R4), 20 4- and 5-year-old children received a test in which transfer from A to B was assessed. Of the 18 children who continued to respond accurately on the A1-R3 and A2-R4 tasks, 15 (83%) evidenced class-consistent B-R relations (B1-R3,B2-R4). When given match-to-sample probes, 11 (73%) of these children also related same class A and B stimuli conditionally with one another (e.g., A1-B1,A2-B2). These findings raised the question of how these derived relations came about.

The emergent B-R relations could be seen as a demonstration of equivalence class formation (Sidman, 1992, 1994; Sidman & Tailby, 1982). According to Sidman (1994), the initially trained A1-R1, A2-R2, B1-R1, and B2-R2 relations should lead to two three-term equivalence classes (A1-B1-R1 and A2-B2-R2) and hence to class consistent A-B matching performances (e.g., A1-B1,A2-B2). The explanation is that if the A-R relation has the property of symmetry, the A-R training produces R-A. Coupled with the transitive relation from B to R to A, B-A should emerge. Likewise, if the relation B-R has the property of symmetry (R-B), this results in the transitive relation A to R to B, so that A-B should emerge. Training A1-R3 and A2-R4 therefore may produce class expansion (R3-A1-R1-B1,R4-A2-R2-B2), as evidenced by the emergence of B1-R3 and B2-R4 relations. This account would be plausible for children but not for pigeons (Wasserman et al., 1992). Sidman's equivalence has been demonstrated with children (e.g., Barnes, Browne, Smeets, & Roche, 1995; Pilgrim, Chambers, & Galizio, 1995; Saunders, Drake, & Spradlin, 1999; Schenk, 1994; Sidman & Tailby, 1982; Smeets, Leader, & Barnes, 1997) but not in animals (Dube, McIlvane, Callahan, & Stoddard, 1993; Hayes, 1989; Lipkins, Kop, & Matthijs, 1988; Rodewald, 1974; Sidman, Rauzin, Lazar, Cunningham, Tailby, & Carrigan, 1982; but see Schusterman & Kastak, 1993).

The emergent B-R relations could also be seen as a case of response-mediated transfer. As formulated by Urcuioli (Urcuioli, 1996; Urcuioli, Zentall, & deMarse, 1995), this account implies that when new reponses are trained to a subset of stimuli, a mediating link is formed by covertly emitting the originally trained response. For example, after training A1-R1 and B1-R1, A1 and B1 are capable of producing R1 covertly ("R1"). Then, when A1-R3 is trained, the organism may covertly "see" or "feel" the originally trained response ("R1") before emitting R3: A1-"R1"-R3. When B1 is presented, the organism will emit

"R1" which, given the reinforced "R1"-R3 link established during previous trials (training A1-R3), will produce or facilitate R3 without further training (e.g., B1-"R1"-R3). Thus, although the targeted B-R relations are not explicitly trained, and hence may be considered "emergent," they constitute a unidirectional chain of trained components (Saunders, Williams, & Spradlin, 1996). This account could easily explain the B1-R3 and B2-B4 performances in children *and* pigeons (Smeets et al., 1997; Wasserman & DeVolder, 1993), but not the children's matching performances (Smeets et al., 1997).

The validity of these accounts could be tested by varying the order in which the prerequisite (baseline) stimulus–response relations are trained. The designs of the aforementioned studies permit three basic training protocols: Many-To-One, One-To-Many, and One-To-One. The *Many-To-One* protocol starts with the training of two stimulus classes: A1-R1,B1-R1 and A2-R2,B2-R2. This can be achieved by training first A1-R1,A2-R2 and then B1-R1,B2-R2 or vice versa. Then the participants are trained to emit a different response in the presence of one member of each class (A1-R3,A2-R4) and transfer from A to B is assessed (B1-R3,B2-R4). The studies by Wasserman and DeVolder (1993) and Smeets et al. (1997) employed this protocol.

The *One-To-Many* protocol starts with the training of two response classes: A1-R1,A1-R3 and A2-R2,A2-R4. This can be accomplished by training A1-R1 and A2-R2 first and A1-R3 and A2-R4 second (One-To-Many-1) or A1-R3 and A2-R4 first and A1-R1 and A2-R2 second (One-To-Many-2). Then the participants are trained to emit R1 and R2 in the presence of the B stimuli (B1-R1,B2-R2), and transfer from A to B is assessed (B1-R3,B2-R4).

In the *One-To-One* protocol, four unrelated stimulus–response relations are trained: B1-R1,B2-R2 then A1-R3,A2-R4 (One-To-One-1) or A1-R3,A2-R4 then B1-R1,B2-R2 (One-To-One-2). After training also A1-R1 and A2-R2, transfer from A to B is assessed (B1-R3,B2-R4). Note that, except for the order of the training tasks, all conditions are identical (Fig. 1).

The response-mediated transfer account predicts that, for reasons mentioned above (see also Fig. 1), Many-To-One and One-To-Many-1 will produce the untrained B-R relations (B1-R3,B2-R4) effectively, while One-To-Many-2 and One-To-One will not. Consider, for example, both One-To-Many protocols. In One-To-Many-1, A1-R1 is trained first. During the subsequent A1-R3 training, A1 will lead to the emission of (R1) first and R3 second (A1-[R1]-R3) so that not only A1-R3 but also (R1)-R3 is reinforced. Thus, when B1-R1 is trained, B1-R3 should occur as a result of the two trained components, B1-R1 and (R1)-R3. In One-To-Many-2, training of A1-R3 before A1-R1 produces A1-(R3)-R1. The subsequent training of B1-R1 does not help the organism to emit R3 because the mediating (R1)-R3 link is missing.

Sidman's position does not necessitate differential effects (Arntzen & Holth, 1997; Barnes, 1994; Saunders, Wachter, & Spradlin, 1988; Sidman, 1994; Spradlin & Saunders, 1986; Wetherby, Karlan, & Spradlin, 1983). Yet, the Many-To-One and One-To-Many protocols would be expected to be more effective than the
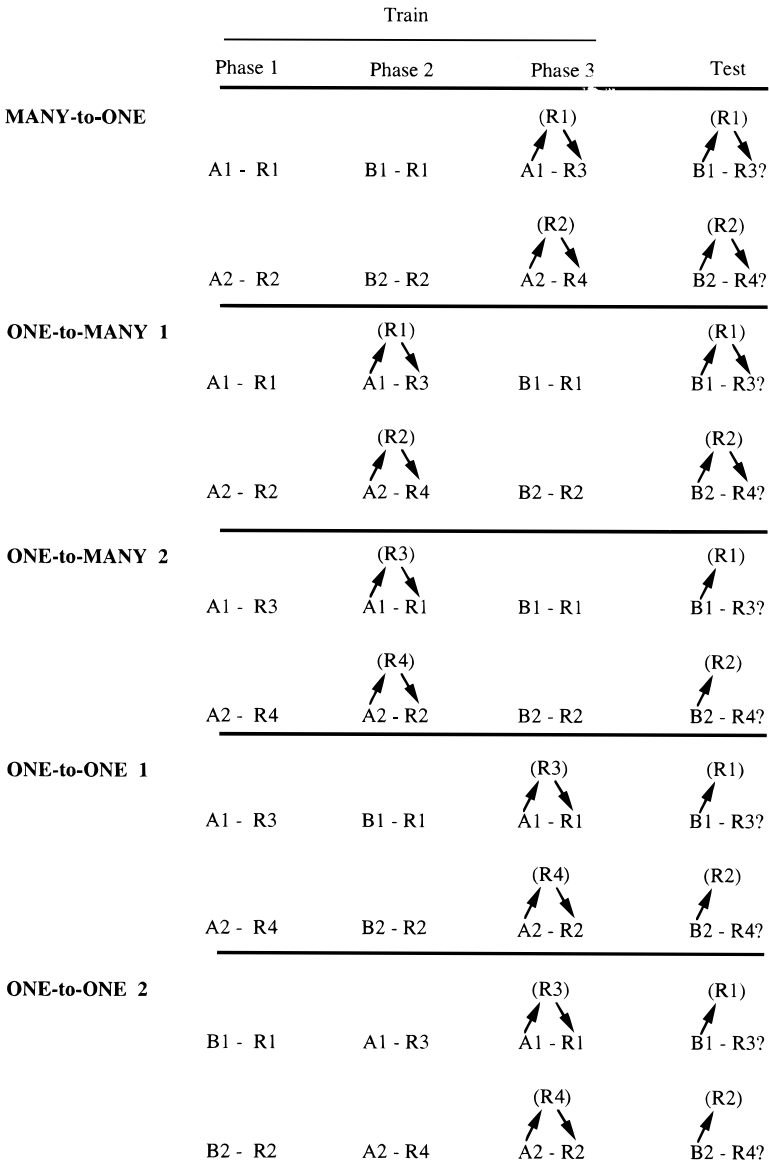
	Train			Test
	Phase 1	Phase 2	Phase 3	
MANY-to-ONE	A1 - R1	B1 - R1	(R1) ↗↘ A1 - R3	(R1) ↗↘ B1 - R3?
	A2 - R2	B2 - R2	(R2) ↗↘ A2 - R4	(R2) ↗↘ B2 - R4?
ONE-to-MANY 1	A1 - R1	(R1) ↗↘ A1 - R3	B1 - R1	(R1) ↗↘ B1 - R3?
	A2 - R2	(R2) ↗↘ A2 - R4	B2 - R2	(R2) ↗↘ B2 - R4?
ONE-to-MANY 2	A1 - R3	(R3) ↗↘ A1 - R1	B1 - R1	(R1) ↗ B1 - R3?
	A2 - R4	(R4) ↗↘ A2 - R2	B2 - R2	(R2) ↗ B2 - R4?
ONE-to-ONE 1	A1 - R3	B1 - R1	(R3) ↗↘ A1 - R1	(R1) ↗ B1 - R3?
	A2 - R4	B2 - R2	(R4) ↗↘ A2 - R2	(R2) ↗ B2 - R4?
ONE-to-ONE 2	B1 - R1	A1 - R3	(R3) ↗↘ A1 - R1	(R1) ↗ B1 - R3?
	B2 - R2	A2 - R4	(R4) ↗↘ A2 - R2	(R2) ↗ B2 - R4?

FIG. 1. Many-To-One, One-To-Many, and One-To-One sequences.

One-To-One protocol because they permit the four-term equivalence classes (e.g., A1-R1-R3-B1) to be formed in a stepwise fashion, whereas the One-To-One protocol does not. Consider One-To-Many-2 and One-To-One. In One-To-Many-2, the initially trained baseline tasks (A1-R3,A2-R4 and A1-R1/A2-R2)

may produce two three-term equivalence classes (A1-R1-R3,A2-R2-R4) before class expansion is induced by training B1-R1 and B2-R2. In One-To-One, the initially trained baseline tasks (B1-R1,B2-R2 and A1-R3,A2-R4) prevent three-member class formations. Thus, the four-term equivalence classes have to evolve all at once (i.e., after completing the final set of baseline tasks, A1-R1 and A2-R2), which may be difficult for young children (Fields, Adams, & Verhave, 1993; Fields, Adams, Verhave, & Newman, 1990; Fields, Landon-Jimenez, Buffington, & Adams, 1995).

The present study examined the effects of different training protocols on the emergence of class-consistent stimulus–response and stimulus–stimulus relations. The study consisted of five experiments, four with children (Experiments 1–4) and one with adults (Experiment 5). The design of Experiment 1 was similar to that of the previous study (Smeets et al., 1997): pretraining of a symbolic X-Y matching task; training (a) A1-R1,A2-R2, (b) B1-R1,B2-R2, and (c) A1-R3,A2-R4; testing B1-R3 and B2-R4; and, finally, testing A-B and B-A. Four protocols were used: Many-To-One, One-To-Many-1, One-To-Many-2, and One-To-One (two versions). The protocols differed only with regard to the order in which the training tasks were introduced (see Fig. 1). Experiments 2 to 5 examined whether the obtained results were related to unscheduled stimulus control and age. These experiments are introduced after Experiment 1.

EXPERIMENT 1

This experiment examined whether children's emergent stimulus–response and stimulus–stimulus relations are affected by different training protocols.

Method

Children

Thirty-two 5-year-old Dutch children participated. The children were recruited through school contacts and participated, with their parents' approval, on a voluntary basis. None had participated in similar research before. The children were quasirandomly assigned to four conditions (i.e., protocols), eight per condition. Their age and sex are listed, together with the results, in Table 3.

Sessions and Setting

The sessions were conducted in a quiet room of the school building, once a day, 5 days a week, and lasted up to 15 min. The children required 7 to 17 sessions (mean: 10.3) spread over 9 to 30 days (mean: 15.7). An adult female served as experimenter. The experimenter and child were seated at the same table facing one another. The experimenter had received extensive training on the correct execution of the training and test procedures with special emphasis on the prevention of any cues (e.g., facial expression and eye darting) that could influence the participants' responses. During training trials, the experimenter looked at the child's face when giving instructions and delivering programmed consequences. During the remainder of these trials (i.e., when presenting stimulus materials and while
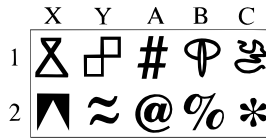
FIG. 2. Experimental stimuli.

the child responded), the experimenter gazed at a fixed location on the table. Precautions were taken to prevent the children from observing the experimenter's recordings on the data sheets.

Three other adults, one at a time, served as observers and monitored the children's responses independently of the experimenter. The observer was in the same room and was located behind and to the side of the child. The observer could clearly monitor the child's responses, but not see the experimenter's data sheet.

Materials

The stimuli consisted of eight black forms (3.0×3.0 cm). The stimuli are shown in Fig. 2 together with their assigned alphanumerical codes: X1, X2, Y1, Y2, A1, A2, B1, and B2 (the C stimuli shown in Fig. 2 were not used in this experiment). The children never saw these codes.

The stimulus configurations differed across tasks. When used in match-to-sample tasks, the stimuli were presented on white cards (15.0×21.0 cm). Each card showed three stimuli, two horizontally aligned choice stimuli 8 cm apart (e.g., Y1 and Y2) and a sample stimulus (e.g., X1) centered 3.0 cm below. When used in successive simple discrimination tasks, each stimulus was presented on a small separate card (5.0×5.0 cm). Some of the successive simple discrimination tasks (see below) involved the use of a white card (21.0×29.5 cm) showing three quadrants, one at the center, one at the top left corner, and one at the bottom right corner (three-quadrants card). All materials were laminated in clear acrylic. Additional materials were a tray with beads and a standing glass tube showing a mark. Filling the tube to the mark required 50 beads.

Tasks, Responses, and Contingencies

Three types of tasks were used: match-to-sample tasks and two types of successive discrimination tasks, one involving clapping or waving (clap/wave task), the other one placing stimuli on squares (stimulus placement task).

A trial on a matching task consisted of the experimenter presenting a stimulus card showing a sample and two choice stimuli while saying "Point" and waiting for the child to respond. A trial on a clap/wave trial involved the experimenter silently presenting a stimulus card (e.g., A1) and waiting for the child to clap or wave. A trial of a stimulus placement task consisted of the experimenter first presenting the three-quadrants card and then putting a stimulus card (e.g., A1)
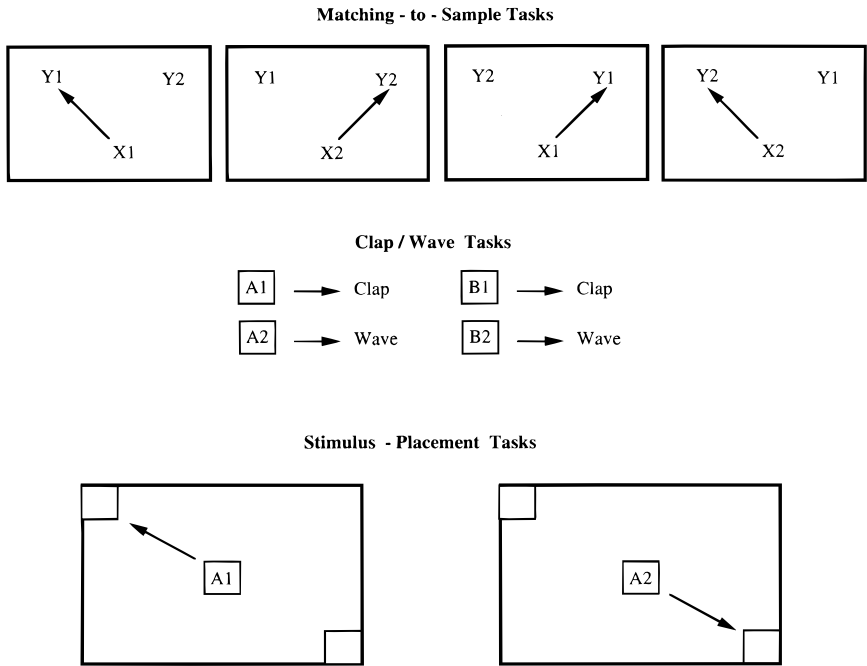
Matching - to - Sample Tasks

Clap / Wave Tasks

Stimulus - Placement Tasks

FIG. 3. Examples of the match-to-sample, clap/wave, and stimulus-placement tasks.

silently on the center quadrant and waiting for the child to place that stimulus on one of the two peripheral quadrants. Examples of each type of task are shown in Fig. 3.

Responses were scored as correct or incorrect. During *training*, correct responses were followed by verbal praise (e.g., "Good" or "Correct") and the delivery of a token (bead) into the glass tube. If the accumulated beads in the tube reached the mark (50 beads), the experimenter interrupted the training, allowed the child to exchange the tokens for a color card (e.g., animal, cartoon character, or race car), and then resumed the training. Incorrect responses were followed by "Wrong. No bead." During *testing*, each response was followed by the presentation of another trial (no programmed consequences). However, immediately after each test, the children received a fixed number of beads irrespective of their performance: 10 beads after completing 12 trials (Steps 2 and 10) or 20 beads after completing 24 trials (Steps 8, 9, and 11).

Many-To-One Condition

The training and test sequence consisted of 11 steps (see Table 1). Each step consisted of two blocks of 12 training trials (Steps 1, 3, 4, 5, 6, and 7) or two blocks of 12 test trials (Steps 2, 8, 9, 10, and 11).

TABLE 1
Many-To-One Training and Test Sequence in Experiment 1

Steps	Train or test	Task	Steps	Train or test	Tasks
1	Train	X1-Y1,X2-Y2	6	Train	A1-R3,A2-R4
2	Test	X1-Y1,X2-Y2	7	Mixed training	
		Y1-X1,Y2-X2	8	Mixed testing	
3	Train	A1-R1,A2-R2	9	Test	B1-R3,B2-R4[a]
4	Train	B1-R1,B2-R2			A1-R3,A2-R4
5	Mixed training		10	Rev test	X1-Y1,X2-Y2
			11	Test	A1-B1,A2-B2[b]
					B1-A1,B2-A2[b]

[a]Derived stimulus–response relations.
[b]Derived stimulus–stimulus relations.

Step 1: Pretraining X-Y. The children were trained on two symbolic match-to-sample tasks with X1 and X2 as samples and Y1 and Y2 as comparisons. Three substeps were used (Smeets & Striefel, 1994). Step 1a consisted of two demonstration trials followed by 12 no-help trials. During each demonstration trial, the experimenter presented a match-to-sample card, pointed to the sample (e.g., X1) while saying "Look here," pointed to the correct comparison (Y1) while saying, "Look, I point to this," and invited the child to do the same ("Now you point"). During the no-help trials, the experimenter simply presented the stimulus card while saying, "Point." During this substep, the positions of the comparisons were fixed, Y1 on the left and Y2 on the right. The samples, X1 and X2, varied quasirandomly across trials. Responding to Y1 was reinforced when given X1. Responding to Y2 was reinforced when given X2 (X1-Y1,X2-Y2). Correct responding on at least 11/12 trials (92%) of one block was required. Step 1b was the same except that no demonstration trials were used and the positions of the comparisons were reversed (Y2 on the left and Y1 on the right). In Step 1c, the locations of the Y stimuli varied quasirandomly across trials.

Step 2: Testing X-Y and Y-X. Two substeps were used. Immediately before the introduction of Step 2a, the experimenter informed the child that she would no longer state whether the responses were "right" or "wrong" and that, instead of giving a bead immediately after every correct reponse, she would give the beads later. Following the completion of the 12th trial, the experimenter presented 10 beads irrespective of the child's performance. The X-Y relations (X1-Y1,X2-Y2) were tested in Step 2a and the Y-X relations (Y1-X1,Y2-X2) in Step 2b. In each substep, correct responding on at least 11/12 test trials was required.

Step 3: Training A1-R1 and A2-R2. This step was directed at teaching the children to clap (R1) when given A1 and to wave (R2) when given A2 (A1-R1,A2-R2). Each block consisted of two demonstration trials followed by 12 no-help trials. Each demonstration trial started with the experimenter presenting a stimulus card. Then she said, "Look what I do" and demonstrated the correct response: clapping

(three times in a row) when given A1 (first demonstration trial) or waving (left-right-left) when given A2 (second demonstration trial), each time inviting the child to do the same ("You do that too"). The no-help trials were presented without any instruction. Correct responding on at least 11/12 trials was required.

Step 4: Training B1-R1 and B2-R2. Same as Step 3 except that the children were trained to clap when given B1 and to wave when given B2 (B1-R1,B2-R2).

Step 5: First mixed training. Same as Steps 3 and 4 except that each block consisted of 12 no-help trials: 3 A1-R1 trials quasirandomly mixed with 3 A2-R2, 3 B1-R1, and 3 B2-R2 trials.

Step 6: Training A1-R3 and A2-R4. This step was directed at establishing novel responses to one member of each class, A1 and A2, but in a different context (three-quadrants card). The procedures were the same as in Step 3 except that the children had to move the stimulus cards from the center quadrant to one of the peripheral quadrants: A1 to the top left quadrant and A2 to the bottom right quadrant (A1-R3,A2-R4). Correct responding on at least 11/12 trials was required.

Step 7: Second mixed training. This step assessed whether A1-R1, A2-R2, B1-R1, and B2-R2 (trained in Steps 3–5) were still intact after the training of A1-R3 and A2-R4 (Step 6). A1-R1, A2-R2, B1-R1, and B2-R2 were presented in one block (three trials each) and A1-R3 and A2-R4 in the second block (six trials each). Correct responding on at least 11/12 trials of each block was required.

Step 8: Mixed testing. This step assessed the accurate performance on all trained tasks under testing conditions. Two blocks of 12 trials were used. The procedures were the same as in Step 7 but without immediate consequences (verbal feedback and beads). Instead, the children received 10 beads after each block irrespective of their performance. Children who did not demonstrate criterion performance (at least 11/12 correct in each block) but responded correctly on at least 9 trials (75%) of each block received Step 8 again. Those who responded correctly on less than 9 trials of each block returned to Step 7 (mixed training) before receiving Step 8 again.

Step 9: Testing B1-R3 and B2-R4. Only the stimulus-placement task (three-quadrants card) was used. The test consisted of 24 trials, 12 on the trained A1-R3 and A2-R4 tasks (6 trials each) quasirandomly mixed with 12 on the nontrained B1-R3 and B2-R4 tasks (6 trials each). Criterion was met when a child responded correctly on at least 11/12 trials (92%) on the trained tasks and on at least 10/12 trials (83%) on the nontrained tasks.[1] Children who demonstrated criterion performance on the trained but not on the nontrained tasks received Step 8 once more. Failure to demonstrate criterion performance on the nontrained *and* on the trained tasks resulted in a return to Step 7 (mixed training) and a repeat of Step 9. At that point, the children proceeded to the next steps irrespective of their performance.

Step 10: Testing Y-X (review). This step assessed whether the previously trained matching tasks (Y1-X1,Y2-X2) were still in tact and served as a "warming up" for Step 11. The procedures were the same as in Step 2b.

---

[1] This somewhat more relaxed criterion was based on our experience that young children tend to make more errors on novel probe tasks.

TABLE 2
Basic Training and Test Sequences

Train or test	Tasks				
			Experiments 1 and 5		
	Many-To-One	One-To-Many-1	One-To-Many-2	One-To-One-1	One-To-One-2
Train	X1-Y1,X2-Y2	Same	Same	Same	Same
Test	X1-Y1,X2-Y2	Same	Same	Same	Same
	Y1-X1,Y2-X2	Same	Same	Same	Same
Train	A1-R1,A2-R2	A1-R1,A2-R2	A1-R3,A2-R4	A1-R3,A2-R4	B1-R1,B2-R2
Train	B1-R1,B2-R2	A1-R3,A2-R4	A1-R1,A2-R2	B1-R1,B2-R2	A1-R3,A2-R4
Train	A1-R3,A2-R4	B1-R1,B2-R2	B1-R1,B2-R2	A1-R1,A2-R2	A1-R1,A2-R2
Test	B1-R3,B2-R4	Same	Same	Same	Same
Test	A1-B1,A2-B2	Same	Same	Same	Same
	B1-A1,B2-A2	Same	Same	Same	Same
			Experiments 2, 3, and 5		
	Control				
Train	X1-Y1,X2-Y2				
Test	X1-Y1,X2-Y2				
	Y1-X1,Y2-X2				
Train	A1-R1,A2-R2				
Train	C1-R1,C2-R2				
Train	A1-R3,A2-R4				
Test	B1-R3,B2-R4				
Test	A1-B1,A2-B2				
	B1-A1,B2-A2				
			Experiment 4		
	Many-To-One	One-To-Many-1	One-To-Many-2	One-To-One-1	One-To-One-2
Train	X1-Y1,X2-Y2	Same	Same	Same	Same
Test	X1-Y1,X2-Y2	Same	Same	Same	Same
	Y1-X1,Y2-X2	Same	Same	Same	Same
Train	B1-R1,B2-R2	B1-R1,B2-R2	B1-R3,B2-R4	B1-R3,B2-R4	A1-R1,A2-R2
Train	A1-R1,A2-R2	B1-R3,B2-R4	B1-R1,B2-R2	A1-R1,A2-R2	B1-R3,B2-R4
Train	B1-R3,B2-R4	A1-R1,A2-R2	A1-R1,A2-R2	B1-R1,B2-R2	B1-R1,B2-R2
Test	A1-R3,A2-R4	Same	Same	Same	Same
Test	A1-B1,A2-B2	Same	Same	Same	Same
	B1-A1,B2-A2	Same	Same	Same	Same

Step 11: Testing A-B and B-A. This step assessed the conditional relations between same class A and B stimuli. Two blocks of symbolic match-to-sample trials were used. One block consisted of 12 trials with B1 and B2 as comparisons. On 6 of these trials A1 served as sample (measuring A1-B1). These trials were quasirandomly mixed with 6 other trials in which A2 served as sample (measuring A2-B2). The other block was the same except that the B stimuli served as samples and the A stimuli as comparisons (measuring B1-A1,B2-A2). The procedures

were the same as in Step 2. The test was conducted twice with an interval of 24 h between the first and second presentation. Criterion was set at at least 10/12 correct responses (83%) in each block during the second test presentation.

One-To-Many and One-To-One Conditions

The procedures for the One-To-Many and One-To-One conditions were the same as for Many-To-One except for the order in which the training tasks were introduced in Steps 3, 4, and 6. All children received a mixed training of the initially trained relations in Step 5 and mixed training and testing of all trained relations in Steps 7 and 8, respectively. Table 2 shows the basic training and testing sequence for each condition.

Interobserver Agreement

The observers monitored 2136 training trials (32.4%) and 1884 test trials (38.4%). The experimenter and observers agreed on all but 4 training trials (99.8%) and on all but 11 test trials (99.4%).

Results

All children learned the X-Y matching tasks without major difficulties and continued to respond accurately on these tasks and on the symmetry tasks (Y-X) under testing conditions (Steps 2 and 10). The numbers of required baseline training trials (Steps 3–7), however, varied markedly between and within conditions. One-to-Many-1 required the least number of training trials (Mean: 122; Range: 90–226), followed by One-to-Many-2 (Mean: 146; Range: 90–258), One-To-One (Mean: 169; Range: 90–240), and Many-To-One (Mean: 202; Range: 114–408). Table 3 shows (a) the age (years and months) and sex of each child, (b) the numbers of required baseline training trials in Steps 3–7, and (c) the percentages of correct responses on the stimulus–response and stimulus–stimulus relation tests.

The trained A1-R3,A2-R4 performances of two children, one from Many-To-One (3) and one from One-To-One (26), deteriorated when interspersed among the B-R probes. Therefore, these children's performances during the stimulus–response (B-R) and stimulus–stimulus probes (A-B,B-A) were excluded from further analyses. Of the remaining 30 children, 16 (53%) evidenced class-consistent B-R relations (B1-R3,B2-R4): 5/7 (71%) in Many-To-One, 5/8 (63%) in One-To-Many-1, 5/8 (63%) in One-To-Many-2, and only 1/7 (14%) in One-To-One. Thirteen of these children (81%) also matched same class A and B stimuli with one another (e.g., A1-B1,A2-B2).

Seven other children (23%) evidenced class-reversed B-R relations (B1-R4, B2-R3): 1/7 (14%) in Many-To-One, 2/8 (25%) in One-To-Many-2, and 4/7 (57%) in One-To-One. All these children demonstrated A-B and B-A relations (e.g., A1-B2,A2-B1) that were consistent with the trained A1-R3,A2-R4 and tested B1-R4,B2-R3 relations.

Seven children (23%) responded unsystematically during the B-R probes. Three of these children matched the A and B stimuli, one (5) class consistently

TABLE 3
Age (Years;Months) Sex, Numbers of Trials for Training the Stimulus–Response
Relations in Steps 3–7, and Percentages of Correct Responses on
Each Test for EachSubject in Experiment 1

Subject	Age	Sex	Train trials	S-R tests 1	S-R tests 2	S-S tests 1	S-S tests 2
			Many-To-Many				
1	5;1	M	114	100[a]		100	92[a]
2	5;2	M	166	100[a]		100	96[a]
3	5;4	M	216	75	50	63	88
4	5;5	M	156	92[a]		100	100[a]
5	5;7	M	128	75	42	71	96[a]
6	5;2	F	284	58	92[a]	46	71
7	5;5	F	142	100[a]		96	100[a]
8	5;7	F	408	0	0[b]	0	4[b]
M	5;4		202				
			One-To-Many-1				
9	5;3	M	226	25	67	17	0[b]
10	5;5	M	178	50	58	63	58
11	5;6	M	104	100[a]		100	100[a]
12	5;7	M	90	100[a]		100	100[a]
13	5;4	F	90	92[a]		100	96[a]
14	5;5	F	90	58	58	29	42
15	5;6	F	104	100[a]		100	100[a]
16	5;7	F	90	100[a]		100	100[a]
M	5;5		122				
			One-To-Many-2				
17	5;3	M	90	100[a]		88	96[a]
18	5;4	M	222	0	0[b]	21	0[b]
19	5;5	M	102	100[a]		92	100[a]
20	5;6	M	104	92[a]		100	96[a]
21	5;2	F	176	75	25	46	25
22	5;4	F	114	25	0[b]	0	0[b]
23	5;5	F	258	83[a]		33	25
24	5;5	F	104	92[a]		54	46
M	5;4		146				
			One-To-One-1				
25	5;4	M	190	100[a]		92	100[a]
26	5;5	M	194	50	58	88	96[a]
27	5;4	F	138	42	8[b]	0	4[b]
28	5;7	F	212	0	0[b]	0	0[b]
M	5;5		184				
			One-To-One-2				
29	5;4	M	116	50	50	50	17[b]
30	5;4	M	240	25	0[b]	0	0[b]
31	5;1	F	90	0	0[b]	4	0[b]
32	5;7	F	174	67	75	63	54
M	5;4		155				

[a] 10/12 B-R trials correct (Step 9) or 10/12 A-B *and* 10/12 B-A trials correct (Step 11).
[b] 10/12 B-R trials incorrect (Step 9) or 10/12 A-B *and* 10/12 B-A trials incorrect (Step 11).

(e.g., A1-B1,A2-B2) and two (9 and 29) class reversed (A1-B2,A2-B1). The other four children (10, 14, 21, and 32) did not match A and B stimuli.

The derived stimulus–response relations were also related to the required numbers of trials for learning the baseline tasks. Excluding One-To-One, in Many-To-One, One-To-Many-1, and One-To-Many-2, children who derived the designated stimulus–response relations required less trials for learning the baseline tasks (Mean Many-To-One: 172; Mean One-To-Many-1: 96; Mean One-To-Many-2: 132) than those who did not derive these relations (Mean Many-To-One: 268; Mean One-To-Many-1: 165; Mean One-To-Many-2: 171).

Discussion

Clearly, the Many-To-One and both One-To-Many protocols produced the class-consistent B-R relations more effectively than One-To-One. Although the discrepancy between One-To-One and the other three conditions is modest, the outcomes of the Many-To-One and both One-To-Many conditions are consistent with those reported in several previous studies (Astley & Wasserman, 1996; Wasserman & DeVolder, 1993; Wasserman, DeVolder, & Coppage, 1992). Thus, the obtained findings are in support of the equivalence account.

This conclusion, however, could be challenged. Because the proportions of class consistent (16/23 or 70%) and class reversed (7/23 or 30%) approximated chance level, the B-R performances could simply be a demonstration of generalized conditional responding. Previous studies have shown that if humans are given training with multiple two-choice conditional discriminations, some percentage will respond conditionally to new conditional discriminations even when no feedback is provided (Saunders et al., 1999; Saunders & Spradlin, 1990; Saunders, Saunders, Kirby, & Spradlin, 1988). Thus, if no other factors were involved, the chances of the conditionally responding children showing B1-R3,B2-R4 or B1-R4,B2-R3 should be about 50%. Therefore, the class-consistent B-R performances should not be seen as evidence for derived relations.

The class-reversed B-R performances, however, were not evenly distributed across conditions and occurred more often in One-To-One than in all three other conditions combined. Perhaps, therefore, the class-reversed B-R performances resulted from some form of interfering stimulus control. If for some reason, the children had a bias toward demonstrating B1-R4,B2-R3 relations, most likely this bias would be evident in the least effective condition (One-To-One), thereby making the present data far more convincing. This issue was addressed in Experiment 2.

EXPERIMENT 2

This experiment examined whether in Experiment 1 the class-reversed B-R (and A-B,B-A) performances should be seen as a demonstration of generalized conditional responding or as a product of spurious stimulus control. What proportions of B1-R3,B2-R4 and/or B1-R4,B2-R3 performances would be obtained if the B stimuli were prevented from participating in the experimenter-designated classes? If the class-reversed performances indicated generalized conditional

TABLE 4
Age (Years;Age) Sex, Numbers of Trials for Training the Stimulus–Response
Relations in Steps 3–7, and Percentages of Correct Responses on
Each Test for Each Subject in Experiments 2 and 3

Subjects	Age	Sex	Train trials	S-R tests		S-S tests	
				1	2	1	2
Experiment 2							
1	5;2	M	116	17	0[b]	4	0[b]
2	5;2	M	104	17	17[b]	13	4[b]
3	5;3	M	142	75	67	46	42
4	5;4	M	114	58	67	67	63
5	5;4	F	178	50	42	71	58
6	5;4	F	142	25	8[b]	4	0[b]
7	5;5	F	166	33	17[b]	25	8[b]
8	5;6	F	114	25	0[b]	21	33
M	5;4		135				
Experiment 3							
1	5;1	M	154	33	58	58	54
2	5;4	M	114	75	42	67	58
3	5;5	M	194	50	75	79	46
4	5;2	F	116	100[a]		96	100[a]
5	5;3	F	90	83[a]		75	79
6	5;3	F	104	75	92[a]	100	96[a]
7	5;6	F	178	92[a]		100	100[a]
8	5;7	F	272	67	58	67	21
M	5;4		153				

[a]10/12 B-R trials correct (Step 9) or 10/12 A-B *and* 10/12 B-A trials correct (Step 11).
[b]10/12 B-R trials incorrect (Step 9) or 10/12 A-B *and* 10/12 B-A trials incorrect (Step 11).

responding, the proportions of "class-consistent" performances (B1-R3,B2-R4) and of "class-reversed" performances (B1-R4,B2-R3) should be about the same as in Experiment 1 (about 2 to 1). Alternatively, if the class-reversed performances resulted from extraneous control, a bias for B1-R4,B2-R3 performances should be evident.

Method

Eight Dutch children participated. Their ages and sex are listed in Table 4. All children received the same (control) condition. The training and testing sequence was the same as in Many-To-One in Experiment 1 (see Table 1) except that the B1-R1,B2-R2 training was replaced by C1-R1,C2-R2 training (see Table 2). Although the present sequence did not permit transfer from A to B, consistent with Experiment 1, B1-R3,B2-R4 and A1-B1,A2-B2 performances were recorded "correct" or "class consistent" and B1-R4,B2-R3 and A1-B2,A2-B1 performances as "incorrect" or "class reversed."

The observer monitored a total of 468 training trials (26.7%) and 424 test trials (26.1%). The experimenter and observer agreed on all but 2 training trials (99.6%) and on all but 4 test trials (99.0%).

Results and Discussion

All children completed the program. They required 116 to 178 trials ($M = 134$) to learn the training tasks (Steps 3–7) and continued to respond accurately on the A1-R3 and A2-R4 trials when tested among the B1-R3,B2-R4 trials. Five children responded conditionally on the B-R trials, all "class reversed" (B1-R4,B2-R3). For four of these children, the A-B and B-A matching probes were consistent with the trained A1-R3,A2-R4 and tested B1-R4,B2-R3 relations (e.g., A1-B2,A2-B1). The other three children responded unsystematically during the B-R probes *and* during the A-B and B-A matching probes. These findings suggested that the B1-R4,B2-R3 relations in Experiment 1 were probably not based on chance or lack of control, but on interfering stimulus control.

Observations and interviews conducted with two children (2 and 6) suggested that the B-R relations were controlled by the stimulus configuration of B1 (see Fig. 2). During the B-R probe, Child 6 occasionally tapped card B1 on the bottom right quadrant before placing the card on it (B1-R4). When asked why he tapped, he indicated that B1 looked like the head of a bird (eyes and beak) and that he made the bird eat from the food container on the floor (bottom right quadrant). The other child (2) indicated that B1 was a "thing" pointing down.

EXPERIMENT 3

This experiment examined whether, in Experiment 2, the B1-R4,B2-R3 performances were indeed affected by the configuration of B1. The X-Y match-to-sample training, together with the A-R and C-R (Experiment 2) or A-R and B-R training (Experiment 1) could have been sufficient to place B1 at the bottom right quadrant (B1-R4) and, as a result of that, place B2 at the upper left square of the quadrant (B2-R3). If correct, the B1-R4,R2-R3 bias should be absent or perhaps even reversed when presenting B1 upside down (bird's beak or "thing" pointing up).

Method and Results

Eight new Dutch children participated (see Table 4). The procedures were the same as in Experiment 2 except that B1 was presented upside down (i.e., pointing up). Two observers (one at a time) monitored 412 training trials (24.8%) and 436 test trials (25.3%). The experimenter and observers agreed on all training trials (100%) and on all but 2 test trials (99.5%).

All children completed the program, learned the trained A-R and C-R relations within 90–272 trials ($M = 153$) and, except for Child 2, continued to respond accurately on the A1-R3,A2-R4 trials when interspersed with the B-R probe trials. Of the remaining seven children, four responded conditionally during the B-R probes, all showing B1-R3,B2-R4. Three of these children also matched the A and B stimuli consistent with the trained A-R and tested B-R relations (e.g., A1-B1,A2-B2). Three other children responded at chance level during the B-R and during the A-B and B-A probes. These findings suggested that the class-reversed

B1-R4,B2-R3 performances in Experiments 1 and 2 probably resulted from inadvertent control by the configuration of B1.

EXPERIMENT 4

This experiment examined whether the training order effects of Experiment 1 could be replicated when using test stimuli that presumably did not induce a response bias. Would Many-To-One and One-To-Many be again superior to One-To-One and, irrespective thereof, would the proportion of class-reversed B-R (and A-B; B-A) performances be substantially lower than in Experiment 1?

Method

Twenty-four Dutch children participated, six in each condition. The stimuli (including B1), procedures, and protocols were the same as in Experiment 1 except that now the B stimuli were used for training (B1-R1,B2-R2 and B1-R3,B2-R4) and the A stimuli for training (A1-R1,A2-R2) and testing (A1-R3,A2-R4) (see Table 2).

Three observers monitored 1240 training trials (26.4%) and 822 test trials (24.0%). The experimenter and observers agreed on all but 3 training trials (99.8%) and 7 test trials (99.1%).

Results and Discussion

Table 5 shows the age and sex of each child together with the major results (training and testing). All children learned the X-Y matching tasks and continued to respond accurately on these tasks and on the symmetry tasks (Y-X) under testing conditions. The numbers of required baseline training trials (Steps 3–7) varied from 90 to 222 ($M = 143$), with no obvious differences between conditions.

One One-To-Many-1 child (8) reversed the trained B-R relations (i.e., demonstrated B1-R4,B2-R3) when these trials were interspersed with the untrained A-R trials. His performances during the A-R probes (A1-R4,A2-R3) and subsequent matching probes (A1-B2,A2-B1 and B1-A2,B2-A1) were consistent with the trained A1&B1-R1,A2&B2-R2 and the tested B1-R4 and B2-R3 performances.

Of the other 23 children, 17 (74%) responded conditionally during the stimulus–response (A-R) probes, 16 of whom (94%) responded class consistently (A1-R3,A2-R4): 5/6 (83%) in Many-To-One, 5/5 (100%) in One-To-Many-1, 4/6 (67%) in One-To-Many-2, and 2/6 (33%) in One-To-One. Fourteen of these children (88%) matched the A and B stimuli class consistently (e.g., A1-B1,A2-B2). Only 1 child (13) evidenced class-reversed A-R relations (A1-R4,A2-R3). His performances during the A-B and B-A matching tasks (e.g., A1-B2,A2-B1) were consistent with the trained B1-R3,B2-R4 and tested A1-R4,A2-R3 relations. The remaining 6 children responded at chance level during the A-R probes and did not match the A and B stimuli.

As in Experiment 1, in Many-To-One and both One-To-Many conditions, children who derived the designated stimulus–response relations required less trials for learning the baseline tasks (Mean Many-To-One: 133; Mean One-To-Many-1:

TABLE 5
Age (Years;Age) Sex, Numbers of Trials for Training the Stimulus–Response Relations
in Steps 3–7, and Percentages of Correct Responses on Each Test
for Each Subject in Experiment 4

Subjects	Age	Sex	Train trials	S-R tests		S-S tests	
				1	2	1	2
Many-To-One							
1	5;2	M	90	100[a]		92	100[a]
2	5;4	M	190	75	50	63	54
3	5;5	M	128	92[a]		100	100[a]
4	5;1	F	104	75	100[a]	100	96[a]
5	5;5	F	168	100[a]		71	92[a]
6	5;6	F	174	100[a]		100	100[a]
M	5;4		142				
One-To-Many-1							
7	5;3	M	138	100[a]		100	100[a]
8	5;4	M	216	0	0[b]	0	0[b]
9	5;5	M	112	75	100[a]	67	58
10	5;8	M	112	92[a]		100	100[a]
11	5;3	F	166	100[a]		88	92[a]
12	5;4	F	114	100[a]		100	100[a]
M	5;5		143				
One-To-Many-2							
13	5;1	M	174	17	0[b]	13	4[b]
14	5;2	M	196	33	58	54	58
15	5;4	M	124	100[a]		79	100[a]
16	5;6	M	156	83[a]		75	67
17	5;4	F	102	75	92[a]	100	100[a]
18	5;5	F	104	100[a]		100	100[a]
M	5;4		143				
One-To-One-1							
19	5;3	M	156	50	42	79	25
20	5;4	F	174	100[a]		100	100[a]
21	5;5	F	112	42	58	54	42
M	5;4		147				
One-To-One-2							
22	5;4	M	222	50	50	45	54
23	5;7	M	102	100[a]		88	100[a]
24	5;2	F	90	75	67	67	58
M	5;4		138				

Note. R = revised.
[a]10/12 B-R trials correct (Step 9) or 10/12 A-B *and* 10/12 B-A trials correct (Step 11).
[b]10/12 B-R trials incorrect (Step 9) or 10/12 A-B *and* 10/12 B-A trials incorrect (Step 11).

118; Mean One-To-Many-2: 122) than those who did not derive these relations (one child in Many-To-One: 190; Mean One-To-Many-2: 185).

Thus, as in Experiment 1, the emergent A-R relations were clearly affected by the order in which the baseline relations were trained. Again, Many-To-One, One-To-Many-1, and One-To-Many-2 produced the derived A-R relations more

TABLE 6
Age (Years;Age) Sex, Numbers of Trials for Training the Stimulus–Response
Relations in Steps 3–7, and Percentages of Correct Responses on
Each Test for Each Subject in Experiment 5

Subject	Age	Sex	Train trials	S-R tests		S-S tests	
				1	2	1	2
Many-To-One							
1	18;6	M	92	100[a]		100	96[a]
2	20;2	M	78	100[a]		100	92[a]
3	19;9	F	92	92[a]		100	100[b]
4	21;10	F	78	100[a]		96	100[a]
M	20;1		85				
One-To-Many-1							
5	17;11	M	78	100[a]		92	100[a]
6	19;6	M	92	92[a]		100	92[a]
7	19;3	F	92	92[a]		100	100[a]
8	22;1	F	78	100[a]		96	96[a]
M	19;8		85				
One-To-Many-2							
9	20;2	M	92	100[a]		92	100[a]
10	21;6	M	92	92[a]		96	100[a]
11	18;7	F	102	100[a]		100	96[a]
12	18;9	F	78	100[a]		100	100[a]
M	19;9		91				
One-To-One-1							
13	19;0	M	92	100[a]		100	100[a]
14	20;3	M	102	100[a]		96	100[a]
M	19;8		97				
One-To-One-2							
15	18;8	F	102	100[a]		92	100[a]
16	21;0	F	78	100[a]		100	100[a]
M	19;10		90				
Control							
17	19;3	M	78	58	50	46	54
18	21;8	M	92	42	58	63	50
19	18;3	F	92	67	42	50	58
20	22;1	F	102	50	50	63	42
M	20;4		91				

[a] 10/12 B-R trials correct (Step 9) or 10/12 A-B *and* 10/12 B-A trials correct (Step 11).

reliably than One-To-One. The class-reversed A-R performances, however, were reduced to a minimum and comparable to those reported in previous research (Smeets et al., 1997; Wasserman et al., 1993). As in Experiments 1 and 2, the derived A-B and B-A relations were highly consistent with the tested B-R and A-R performances.

EXPERIMENT 5

This experiment was a replication of Experiments 1 and 2 but with adults. Experiment 5 examined whether the data obtained in Experiments 1 and 2 were

age related or, in more general terms, related to the participants' competence in learning and deriving new conditional relations. Specifically, the following two questions were addressed. First, would adults show the same training order effects as children? Second, when no class formation was possible (replication of Experiment 2), would adults respond conditionally during the B-R probes and, like the children, show a B1-R4,B2-R3 response bias?

Method

Five conditions were used. Four of these conditions involved the same three protocols (same stimuli and same procedures) as in Experiment 1: Many-To-One, One-To-Many-1, One-To-Many 2, and One-To-One (two versions). The fifth condition served as a control and involved the same training and testing sequence as Experiment 2: Training of A1-R1,A2-R2 (first), C1-R1,C2-R2 (second), and A1-R3,A2-R4 (third), followed by B-R and A-B,B-A probes (see Table 2).

Twenty 1st-year undergraduate students from University College, Cork, were recuited through faculty notice board advertisements. None had participated in similar research before, and none of them were psychology majors. Their age and sex are listed in Table 6. The participants were quasirandomly assigned to the experimental conditions, four per condition.

The sessions were conducted in the department laboratory. The sessions were arranged such that the participants would not meet each other in the vicinity of the laboratory. Each participant completed the experiment in one session (duration between 1 and 3 h). The experimental arrangements were the same as in Experiment 1, except that the participants were paid 2 Irish pounds each time the beads in the tube reached the mark (50 beads).

A graduate female served as experimenter. Two other graduate students served as observers. The observers monitored 442 training trials and 600 test trials. The experimenter and observers agreed on 441 training trials (99.8%) and on all test trials (100%).

Results

Table 6 shows the age, sex, and results for each participant. All 16 participants receiving the experimental conditions (Many-To-One, One-To-Many-1, One-To-Many-2, and One-To-One) quickly learned the baseline tasks, continued to respond accurately on the A-R tasks (A1-R3,A2-R4) when interspersed with B-R probe trials, and responded class consistently during the stimulus–response and stimulus–stimulus relations probes. The performances of the four control participants were the same except that none of them responded consistently during the probes (i.e., either close to 100% or 0% correct). Instead, their probe performances were consistently about 50% correct. Verbal reports indicated that they simply did one thing on one trial and the other thing on the next so that at least some of their responses were correct. In a sense, therefore, they responded consistently (i.e., in accordance with a rule) but without evidencing generalized conditional responding or any form of extraneous stimulus control.

TABLE 7
Percentages of Participants Showing Class-Consistent, Class-Reversed,
or No Novel Conditional Stimulus–Response Relations

Novel S-R relations	MTO	OTM		OTO
		1	2	
Experiment 1—Children				
Class Con	72	63	63	14
Class Rev	14	0	25	58
None	14	37	12	28
Experiment 4—Children				
Class Con	83	100	67	33
Class Rev	0	0	16	0
None	17	0	16	67
Experiments 1 and 4 combined				
Class Con	77	77	64	23
Class Rev	8	0	21	31
None	15	23	14	46
Experiment 5—Adults				
Class Con	100	100	100	100

Note. MTO = Many-To-One; OTM = One-To-Many; OTO = One-To-One.

GENERAL DISCUSSION

The major purpose of the study was to determine whether the derived stimulus–response relations are sensitive to the order in which the baseline relations are trained and, if so, whether the differential effects are consistent with the response-mediated transfer account (Urcuioli, 1996; Urcuioli, Zentall, & deMarse, 1995) or with Sidman's equivalence account (Sidman, 1994). Experiments 1, 4, and 5 showed that the training protocols affected class formation in children but not in adults (see Table 7).

All 16 adults in Experiment 5 demonstrated the derived stimulus relations immediately, irrespective of the training protocol that was used (Many-To-One, One-To-Many, and One-To-One). This finding is clearly consistent with the equivalence account. So was the evidence of child Experiments 1 and 4, though the data are perhaps not as compelling. Of the 53 children who maintained accurate responding on the trained stimulus-placement tasks (e.g., A1-R3,A2-R4) during testing, 32 (60%) demonstrated novel class-consistent stimulus–response relations (e.g., B1-R3,B2-R4): 10/13 (77%) in Many-To-One, 10/13 (77%) in One-To-Many-1, 9/14 (64%) in One-To-Many-2, and only 3/13 (23%) in One-To-One.[2] Although the evidence in Experiments 1 and 4 may not seem equally convincing, One-To-One clearly was the least effective protocol in each experiment, probably because it did not permit the four-term equivalence classes

[2] Fisher exact tests revealed that the proportions of children with class-consistent and other (i.e., class-reversed and no conditional) performances in One-To-One were statistically different from those in Many-To-One (*p* = .008), One-To-Many-1 (*p* = .008), and One-To-Many-2 (*p* = 0.033).

(e.g., A1-R1-R3-B1) to be formed in a stepwise fashion. This conclusion gains additional strength when considering the derived performances of only those children who learned the baseline tasks as fast as the adults and hence may have been more advanced than the other children. In Experiment 5, the adults required a mean of 89 trials (Range: 78–102) for learning the baseline tasks. In Experiments 1 and 4, 17 children learned the baseline tasks in 90–104 trials (Mean: 97), 14 Many-To-One and One-To-Many children, and 3 One-To-One children. Thirteen of these Many-To-One and One-To-Many children (93%) and 1 One-To-One child (33%) showed transfer (Fisher exact test, $p = .06$). Thus, while the derived performances of the fast-learning Many-To-One and One-To-Many children closely approximated those of the adults, those of the fast-learning One-To-One children did not. These findings indicate that the One-To-One protocol clearly is inadequate for generating transfer in most young children. This conclusion does not imply, of course, that young children are incapable of benefiting from this protocol per se. Given the conceptualizations and successes reported in previous research (Barnes-Holmes & Barnes-Holmes, 2000; Barnes-Holmes, Barnes-Holmes, & Roche, in press; Stokes & Baer, 1977), some form of multiple exemplar training might be sufficient for making the One-To-One protocol effective with 5-year-old or perhaps even younger children.

The present study also generated some unexpected findings. First, the high proportion of class-reversed performances (30%) during the stimulus–response (B-R) probes in Experiment 1. Because this proportion approximated chance level, it might be tempting to consider these performances as a case of generalized conditional responding (Saunders et al., 1988, 1999; Saunders & Spradlin, 1990; 1993); thus, the class consistent performances could also be seen in this way (i.e., no derived relations). Experiments 2 and 3 indicated, however, that the class-reversed performances resulted from inadvertent control by the configurations of the B stimuli, notably B1. When B1 "pointed down" (Experiment 2), all conditional B-R performances were class reversed. When B1 "pointed up" (Experiment 3), all conditional B-R performances were class consistent. Although these findings also permitted the class-consistent performances to be the product of extraneous control, these performances could have been derived despite the interfering control. Experiment 4 proved this latter assumption to be correct. In fact, the results of Experiments 4 and 5 indicate that chance-level conditional responding probably may not (or not exclusively) be the product of being tested on a task similar to previous training tasks. In both these experiments, the participants (children and adults) had the same history of training and testing as in Experiment 1. Yet, class-reversed conditional responding occurred seldomly in Experiment 4, and no conditional responding occurred in the control condition of Experiment 5. These findings are consistent with those reported by Pérez-González (1994; Experiment 3). After being trained on A-B, P-Q, and AB-X match-to-sample tasks, and showing transfer from AB to PQ (PQ-X), none of his three participants responded conditionally during the subsequent E-F test. Clearly, the phenomenon of "generalized conditional responding" deserves experimental analysis.

TABLE 8
Numbers of Participants with Class-Consistent (CC), Class-Reversed (CR),
or No Conditional Performances (NC)

S-R test			S-S test		
CC	CR	NC	CC	CR	NC
Experiments 1 and 4					
32	—	—	27	—	5
—	8	—	—	8	—
—	—	13	1	2	10
Experiments 2 and 3					
4	—	—	3	—	1
—	5	—	—	4	1
Experiment 5					
12	—	—	12	—	—
—	—	4	—	—	4

Although one could easily see how the "bird's beak" or "thing pointing down" configuration of B1 could have led to a position bias (Experiments 1–3), it is not clear why this configuration affected the children's B-R performances in the present study and not in the previous study (Smeets et al., 1997), in which the same (X, Y, A, and B) stimuli were used. This discrepancy could be related to the nature of the trained A1&B1-R1 and A2&B2-R2 tasks. In the present study, these tasks involved clapping (R1) and waving (R2). In the previous study, training of A1&B1-R1 and A2B2-R2 involved another three-quadrants card similar to the one used for testing (see Fig. 2) but with periferal quadrants at different locations (one in the upper right corner and one in the bottom left corner). Thus, even if a child might have seen B1 as a "thing pointing down," he or she was trained to put it in the *upper* right corner, thereby overriding any inadvertent control this configuration may have had.

Another unexpected finding in the current study was the impressive consistency between the performances during the stimulus–stimulus tests (A-B,B-A) and those during the stimulus–response tests (A-R,B-R), irrespective of whether the tested B-R relations were consistent or inconsistent with the trained stimulus–response (A-R,B-R) relations. Table 8 shows the numbers of participants with class-consistent, class-reversed, and no conditional performances during the stimulus–response and stimulus–stimulus tests. For example, Table 8 shows that of the 32 children who in Experiments 1 and 4 responded class consistently during the stimulus–response probe (e.g., A1&B1-R3,A2&B2-R4), 27 also responded class consistently during the stimulus–stimulus test (A1-B1,A2-B2 and B1-A1,B2-A2). The other 5 children did not respond conditionally.

The predominance of the tested relations over the trained relations (A1&B1-R1,A2&B2-R2) for the formation of equivalence relations was most evident in the eight participants in Experiments 1 and 4 with class-reversed stimulus–response probe performances. The stimulus–stimulus test permitted

these participants to relate the A and B stimuli on the basis of the trained relations (e.g., A1&B1-R1,A2&B2-R2) *or* on the tested stimulus–response relations (A1&B2-R3,A2&B1-R4). Yet, all eight participants demonstrated A-B and B-A performances (e.g., A1-B2,A2-B1) that were consistent with the tested stimulus–response relations. This finding may be related to the fact that the format and response requirements of the matching task (pointing to one of two comparisons) resembled those of the stimulus-placement task more (placing a card at one of two locations) than those of the trained clap/wave tasks. If correct, this finding may indicate that not all stimulus–response relations are equally effective in producing class-consistent match-to-sample performances.

Although the present and previously reported findings (Smeets et al., 1997) permit us to view the derived stimulus–response relations (e.g., Astley & Wasserman, 1996; Wasserman & DeVolder, 1993) from an equivalence perspective, Wasserman's findings with pigeons demonstrate that one cannot dismiss the response-mediated generalization account altogether (Sidman, 1994). Perhaps nonverbal organisms do derive novel relations solely through mediated response generalization and only through Many-To-One training protocols. If correct, this would show that animals are less versatile in acquiring "nonsimilarity based categorizations" (Astley & Wasserman, 1996; Wasserman, & DeVolder, 1993) than children who, as the present findings indicated, are less versatile in acquiring these categorizations than adults.

REFERENCES

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record*, **47**, 309–320.

Astley, S. L., & Wasserman, E. A. (1996). Mediating associations, essentialism, and nonsimilarity-based categorization. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 111–133). Amsterdam: Elsevier.

Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record*, **44**, 91–124.

Barnes, D., Browne, M., Smeets, P. M., & 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-Holmes, D., & Barnes-Holmes, Y. (2000). Explaining complex behavior: Two perspectives on the concept of generalized operant classes. *The Psychological Record,* **50**, 251–265.

Barnes-Holmes, Y., Barnes-Holmes, D., & Roche, B. (in press). Exemplar training and the development of relational framing in young children. In S. C. Hayes & D. Barnes-Holmes (Eds.), *Relational frame theory: Creating a new behavioral agenda in language and cognition*. Reno, NV: Context Press.

Dube, W. V., McIlvane, W. J., Callahan, T. D., & Stoddard, L. T. (1993). The search for stimulus equivalence in nonverbal organisms. *The Psychological Record*, **43**, 761–778.

Fields, L., Adams, B. J., & Verhave, T. (1993). The effects of equivalence class structure on test performances. *The Psychological Record*, **43**, 697–712.

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

Fields, L., Landon-Jimenez, D. V., Buffington, D. M., & Adams, B. J. (1995). Maintained nodal distance effects in equivalence classes. *Journal of the Experimental Analysis of Behavior*, **64**, 129–145.

Hayes, L. J. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, **51**, 391–395.

Lipkins, R., Kop, P. F. M., & Matthijs, W. (1988). A test of symmetry and transitivity in the conditional discriminations of pigeons. *Journal of the Experimental Analysis of Behavior*, **49**, 395–409.

Pérez-González, L. A. (1994). Transfer of relational stimulus control in conditional discriminations. *Journal of the Experimental Analysis of Behavior*, **61**, 487–503.

Pilgrim, C., Chambers, L., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: II. Children. *Journal of the Experimental Analysis of Behavior*, **63**, 239–254.

Rodewald, H. K. (1974). Symbolic matching-to-sample by pigeons. *Psychological Reports*, **34**, 987–990.

Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of the Experimental Analysis of Behavior*, **71**, 195–214.

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior*, **50**, 145–162.

Saunders, K. J., & Spradlin, J. E. (1990). Conditional discrimination in mentally retarded adults: The development of generalized skills. *Journal of the Experimental Analysis of Behavior*, **54**, 239–250.

Saunders, R. R., Wachter, J., & 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.

Saunders, K. J., Williams, D. C., & Spradlin, J. E. (1996). Derived stimulus control: Are there differences among procedures and processes? In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 93–109). Amsterdam: Elsevier.

Schenk, J. J. (1994). Emergent relations of equivalence generated by outcome-specific consequences in conditional discrimination. *The Psychological Record*, **44**, 537–558.

Schusterman, R. J., & Kastak, D. (1993). A California sea lion (*Zalophus californianus*) is capable of forming equivalence classes. *The Psychological Record*, **43**, 823–839.

Sidman, M (1992). Equivalence relations: Some basic considerations. In S. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 15–27). Reno, NV: Context Press.

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

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

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs matching to sample. *Journal of the Experimental Analysis of Behavior*, **37**, 5–22.

Smeets, P. M., Barnes, D., & Roche, B. (1997). Functional equivalence in children: Derived stimulus–response and stimulus–stimulus relations. *Journal of Experimental Child Psychology*, **66**, 1–17.

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

Smeets, P. M., & Striefel, S. (1994). A revised blocked-trial procedure for establsihing arbitrary matching in children. *The Quarterly Journal of Experimental Psychology*, **47B**, 241–261.

Spradlin, J. E., & Saunders, R. R. (1986). The development of stimulus classes: Sample classification vs comparison classification. *Analysis and Intervention in Developmental Disabilities*, **6**, 41–58.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generatlization. *Journal of Applied Behavior Analysis*, **10**, 349–367.

Urcuioli, P. (1996). Acquired equivalences and mediated generalization in pigeon's matching-to-sample. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals*, (pp. 55–70). Amsterdam: Elsevier.

Urcuioli, P. J., Zentall, T. R., & DeMarse, T. (1995). Transfer to derived sample-comparison relations by pigeons following many-to-one versus one-to-many matching with identical training relations. *The Quarterly Journal of Experimental Psychology*, **48B**, 158–178.

Wasserman, E. A., & DeVolder, C. L. (1993). Similarity- and nonsimilarity-based conceptualization in children and pigeons. *The Psychological Record*, **43**, 779–793.

Wasserman, E. A., DeVolder, C. L., & Coppage, D. J. (1992). Nonsimilarity based conceptualization in pigeons via secondary order or mediated generalization. *Psychological Science*, **3**, 374–379.

Wetherby, B., Karlan, G. R., & Spradlin, J. F. (1983). The development of derived stimulus relations through training in arbitrary-matching sequences. *Journal of the Experimental Analysis of Behavior*, **40**, 69–78.

Zentall, T. R., & Smeets, P. M. (1996). *Stimulus class formation in humans and animals*. Amsterdam: Elsevier.