JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2005, **84,** 417–433 NUMBER 3 (NOVEMBER)

DERIVED STIMULUS RELATIONS, SEMANTIC PRIMING, AND EVENT-RELATED POTENTIALS: TESTING A BEHAVIORAL THEORY OF SEMANTIC NETWORKS

DERMOT BARNES-HOLMES, CARMEL STAUNTON, ROBERT WHELAN, YVONNE BARNES-HOLMES, SEAN COMMINS, DEREK WALSH, IAN STEWART, PAUL M. SMEETS, AND SIMON DYMOND

NATIONAL UNIVERSITY OF IRELAND, MAYNOOTH;
UNIVERSITY COLLEGE, DUBLIN;
NATIONAL UNIVERSITY OF IRELAND, GALWAY;
LEIDEN UNIVERSITY; AND UNIVERSITY OF WALES, SWANSEA

Derived equivalence relations, it has been argued, provide a behavioral model of semantic or symbolic meaning in natural language, and thus equivalence relations should possess properties that are typically associated with semantic relations. The present study sought to test this basic postulate using semantic priming. Across three experiments, participants were trained and tested in two 4-member equivalence relations using word-like nonsense words. Participants also were exposed to a single- or two-word lexical decision task, and both direct (Experiment 1) and mediated (Experiments 2 and 3) priming effects for reaction times and event-related potentials were observed within but not across equivalence relations. The findings support the argument that derived equivalence relations provides a useful preliminary model of semantic relations.

Key words: derived equivalence relations, semantic priming, event-related potentials, humans, adults

---

The study of human language and cognition has attracted increasing attention within behavioral psychology in recent years (e.g., Barnes & Holmes, 1991), and one group of researchers recently has offered a behavioral account of language and cognition known as Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001). One of the core assumptions of RFT is that the behavioral units of human language and thought may be defined in terms of derived stimulus relations and relational networks. One of the simplest examples of a derived stimulus relation is the equivalence relation, which some have argued provides the basis for semantic or symbolic meaning in natural language (e.g., Sidman, 1986, 1994). A number of behavioral researchers also have argued that traditional network theories of verbal or semantic meaning (e.g., Anderson, 1976, 1983; Collins & Loftus, 1975; McClelland & Rumelhart, 1988) share similarities with the concept of derived stimulus relations (Barnes & Hampson, 1993; Cullinan, Barnes, Hampson, & Lyddy, 1994; Fields, 1987; Hayes & Hayes, 1992; Reese, 1991). If this basic postulate is correct, it follows that the pattern of findings that have been observed using semantic stimuli should also be found when using stimuli from equivalence or other derived relations (Branch, 1994).

One very common finding that has been reported using semantic stimuli is the semantic priming effect (see Meyer & Schvaneveldt, 1971, for the seminal study in this area). If two words are presented (prime and target), and they are semantically related (e.g., tiger–lion), participants' average reaction times (RTs) in a recognition task are significantly shorter than if the words are semantically unrelated (e.g., tiger–house). The priming literature includes many variants such as semantic, associative, mediated, and episodic priming as well as numerous experimental preparations to demonstrate priming, such as lexical decision and pronunciation tasks (see Neely, 1991, for a review).

Only one published study has sought to determine if priming in a lexical decision task occurs for previously trained and tested

---

The original study upon which the current research is based was conducted by Carmel Staunton under the supervision of Dermot Barnes-Holmes. Carmel lost her life tragically in a road traffic accident in September 2003, and this article is dedicated to her memory.

Dermot Barnes-Holmes, Yvonne Barnes-Holmes, Sean Commins, and Derek Walsh are at the National University of Ireland, Maynooth; Robert Whelan is at University College, Dublin; Ian Stewart is at the National University of Ireland, Galway; Paul M. Smeets is at Leiden University; and Simon Dymond is at the University of Wales, Swansea.

Correspondence and requests for reprints should be sent to Dermot Barnes-Holmes at the Department of Psychology, National University of Ireland, Maynooth, Maynooth, Co. Kildare, Ireland (e-mail: Dermot.Barnes-Holmes@may.ie).

doi: 10.1901/jeab.2005.78-04

equivalence relations (Hayes & Bisset, 1998). Participants were asked to press a ''YES'' key if both words were from any of the previously learned equivalence relations, and to press a ''NO'' key if one or both words were previously unseen. Mean RTs to equivalently related word pairs were significantly faster than mean RTs to nonequivalently related word pairs. In effect, the equivalence relations appeared to generate priming effects not unlike those typically found when real words are used in cognitive research (e.g., Balota & Lorch, 1986; de Groot, 1983; McNamara & Altarriba, 1988; Meyer & Schvaneveldt, 1971).

Although the Hayes and Bisset (1998) study produced the predicted priming effects, a number of key features of their experimental work limit the extent to which strong conclusions may be drawn from the data. For example, Hayes and Bisset employed the two-word lexical decision task in which the participant is required to respond to both stimuli (i.e., YES if both are real words and NO if one or both stimuli are nonsense words). The most common procedure in modern priming studies, however, is the single-word priming paradigm in which the participant responds only to the target, not to the prime (see Neely, 1991). Given that reliable priming effects have been reported across numerous studies using the single-word paradigm (see Neely, Keefe, & Ross, 1989), it seems important to replicate Hayes and Bisset's data with this procedure if the generality of their results is to be sustained.

Another limitation to the Hayes and Bisset (1998) study concerns the fact that they presented participants with corrective feedback for correct and incorrect responses during the lexical decision task. Some priming studies in the cognitive literature have demonstrated semantic priming in the absence of differential feedback (e.g., Hill, Strube, Roesch-Ely, & Weisbrod, 2002; Holcomb and Anderson, 1993; Weisbrod et al., 1999), and thus it seems important also to replicate this effect with equivalence-based priming if the derived stimulus relations model of semantic meaning is to be upheld. The Hayes and Bisset study also suffered from another critical limitation, but this will be addressed in the context of Experiments 2 and 3 of the current report.

Experiment 1 of the present study involved training and testing participants in the necessary conditional discriminations for the formation of two 4-element equivalence relations, followed by exposure to a lexical decision task designed to test for priming effects among the stimuli participating in the equivalence relations. Unlike the Hayes and Bisset (1998) study, however, a single-word priming paradigm was employed, with no corrective feedback.

EXPERIMENT 1

Method

Participants

Participants were 27 right-handed undergraduate students (18 females and 9 males) attending the National University of Ireland, Maynooth who had not studied stimulus equivalence or related phenomena prior to their participation in the experiment. The participants were between 18 and 28 years old (mode 19).

Apparatus and Stimuli

Each participant completed the experiment in a small, dimly lit, sound-attenuated cubicle in the Experimental Psychology Laboratory in the Department of Psychology at the National University of Ireland, Maynooth. All experimental tasks were presented on Dell® personal computers equipped with Pentium 4® processors and standard keyboards and monitors. Given a viewing distance of 60 cm for the lexical decision task, each word subtended a visual angle of approximately 2° to 6° in width and 1° in height.

The stimuli employed across both the matching-to-sample (MTS) and lexical decision procedures were taken from Massaro, Venezky, and Taylor (1979). The stimuli were letter permutations derived from the most frequent 150 six-letter English words as listed in Kucera and Francis (1967). The letter strings met the following criteria: (a) they were orthographically regular; (b) they were pronounceable; (c) they contained common vowel and consonant spellings; (d) they had no more than three letters for a medial consonant cluster, if one occurred; and (e) none of the strings kept the first and last letters from the original English words in their original positions. Consistent with the procedures employed by Hayes and Bisset (1998), the participants were told that some of the letter strings were genuine foreign words and some were nonsense words (during postexperimental debriefing the participants were informed that all of the strings were in fact nonsense words).

Procedure

The experimental procedure consisted of MTS training, followed by a partial equivalence test, which was then followed by a single-word lexical decision task.

Phase 1: MTS training. On each trial, a sample stimulus appeared in a rectangular box in the center upper half of the computer screen. After 2,000 ms the sample was removed and 500 ms later two comparison stimuli appeared, one in a rectangular box in the lower-left corner of the screen and the other in a second rectangular box in the lower-right corner of the screen. The comparison stimuli remained visible until the participant clicked on one of the boxes. If the participant clicked on a comparison stimulus that was deemed to be correct by the experimenter, then the comparison stimuli were removed and the word ''correct'' was presented in the middle of the computer screen for 2,000 ms. A computer-generated chime (presented via headphones worn by the participants) also was presented with this visual feedback. If the participant clicked on a comparison stimulus that was deemed to be incorrect by the experimenter, then a similar sequence occurred except that the word ''wrong'' was presented and no computer-generated chime occurred. When the feedback message (correct or wrong) was removed, the screen remained blank for a 3,500 ms intertrial interval. Immediately thereafter the next trial was presented. At the beginning of the first trial, each participant was provided with instructions on the computer screen that described the task (adapted from Hayes & Bisset, 1998; all verbatim instructions used in this and subsequent experiments are available from the first author upon request).

All participants were trained using six MTS trial types designed to establish two 4-element equivalence relations. The training protocol was of a linear design (e.g., A1-B1, B1-C1, C1-D1) and a schematic representation of the six MTS trial types is shown in the upper half of Table 1. If, for example, the nonsense word A1 was presented as a sample, clicking the mouse on the comparison B1 produced the ''correct'' feedback, but clicking B2 produced the ''wrong'' feedback. Note that the participants never saw the alphanumeric labels, which are used here for ease of communication. The six MTS training trials were presented quasi-randomly in blocks of six trials, such that each trial type was presented once within each six-trial block. The position of the two comparison stimuli was randomized across trials, such that the correct comparison could appear with equal probability in the left- or right-hand rectangular box. Each participant was required to produce 24 consecutively correct trials in order to complete this training phase. When a participant successfully completed the MTS training, the computer presented the instructions for the equivalence test.

Phase 2: Equivalence testing. The second phase consisted of an equivalence test that was presented immediately after the participants had completed the MTS training. The MTS procedure for the equivalence test was similar to that of the MTS training, with two key differences. First, the six C-A, D-B, and D-A trial types listed in the lower section of Table 1 were used. Second, no feedback (i.e., the words correct and wrong) was presented on any trial, but instead the program simply progressed to the intertrial interval.

At the beginning of the first trial, instructions were presented to each participant on the computer screen indicating that no cor-

Table 1

A schematic representation of the trained conditional discriminations and tested equivalence relations.

	Comparison	
Sample	Correct	Incorrect
Trained conditional discriminations		
A1	B1	B2
B1	C1	C2
C1	D1	D2
A2	B2	B1
B2	C2	C1
C2	D2	D1
Tested equivalence relations		
D1	A1	A2
D1	B1	B2
C1	A1	A2
D2	A2	A1
D2	B2	B1
C2	A2	A1

rective feedback would be presented during this phase. The six MTS testing trials were presented quasi-randomly in blocks of six trials, such that each trial type was presented once within each six-trial block. Each participant was exposed to 24 test trials in order to complete the equivalence test. If a participant failed to produce a minimum of at least 22 correct responses in the equivalence test phase, the computer displayed the following message on the screen: ''You may now take a short break. Please click on the button below when you are ready to continue.''

When the participant clicked on the continue button, the program returned to the MTS training phase (starting with the instructions), which in turn was followed by a second exposure to the equivalence test (again starting with the instructions). This cycle of MTS training and testing continued until the participant successfully passed the equivalence test (i.e., produced at least 22 correct responses) or completed six cycles of training and testing. When a participant passed the equivalence test, the computer displayed a message asking the participant to report to the Experimenter.

Pilot research had shown that participants would sometimes fail an equivalence test immediately after completing the next phase of the experiment (see below), thus indicating that one successful exposure was insufficient to establish robust equivalence relations. To combat this problem, participants who passed the equivalence test returned to the laboratory within 48 hr for a second session, and were reexposed to the MTS training and equivalence testing procedures until they produced five consecutively successful equivalence test performances. When a participant had achieved this criterion, he or she was asked to take a brief break outside the laboratory (5 min), and the experimenter then proceeded to set the computer up for the third and final phase of the experiment, the lexical decision task, before asking the participant to return.

Phase 3: The lexical decision task. The lexical decision task was modeled on Holcomb and Anderson (1993). The background color of the computer screen throughout the lexical decision task was white. Each trial began with a warning stimulus, the presentation of a red ''X,'' in the middle of the screen. This X remained on the screen for 500 ms and was then replaced by the prime (e.g., the nonsense word A1), which remained on screen for 200 ms. When the prime was removed from the screen there was a 50 ms delay, during which the screen remained blank, and then the target stimulus (e.g., the nonsense word B1) was presented (a stimulus onset asynchrony [SOA] of 250 ms). After 1,500 ms the target was removed and a green ''X'' appeared in the middle of the screen. Finally, after 1,250 ms the green X was replaced by the red X and the next trial began. A total of 96 trials was presented, with a short break (of about 1 min) provided after the first 48 trials.

Participants were instructed to observe and ''mentally'' pronounce the first word that appeared on each trial (the prime). They also were told to rest their hands comfortably at the bottom of the keyboard and to place the left and right index fingers on the ''M'' and ''Z'' keys, respectively, and to press the M key (for yes) on the computer keyboard if the second word (the target) was a foreign word (i.e., a word that they had seen during the MTS training), and to press the Z key (for no) if the target was a nonsense word they had not seen before. The left–right position of the *yes* and *no* keys was counterbalanced across participants (the instructions were adapted accordingly). The instructions emphasized that the yes or no button-response should be made to the target word alone, not the prime. Finally, the participants were asked to respond as fast and as accurately as they could on every trial. If a response was made after the green X had been removed from the screen (i.e., the next trial had begun), it was recorded as ''too late,'' and this datum was excluded from subsequent analyses.

Following a brief practice session of 24 trials with English words, participants were presented with the instructions on the computer screen that explained the lexical decision task (adapted from Hayes & Bisset, 1998, for a single-word task). During the lexical decision task, each participant was presented with 24 pairs of stimuli that were from the same equivalence relation, and 24 pairs that were from different equivalence relations or contained one or two previously unseen nonsense stimuli. Table 2 contains a schematic representation of all 48 trial types divided into eight categories that were subsequently employed in the descriptive and inferential statistical analyses of the data.

Table 2

A schematic representation of the 48 trial types presented during the lexical decision procedure. Pm = Prime, Tg = Target, Rp = Correct Response, N = Nonsense Word.

Class–class			Class–nonclass			Class–nonsense			Nonsense–class			Nonsense–nonsense		
Pm	Tg	Rp	Pm	Tg	Rp	Pm	Tg	Rp	Pm	Tg	Rp[a]	Pm	Tg	Rp
Directly Trained														
A1	B1	Yes	B1	A2	Yes	A1	N1	No	N1	A1	Yes	N1	N6	No
B1	C1	Yes	A1	C2	Yes	B1	N2	No	N2	B1	Yes	N2	N5	No
C1	D1	Yes	D1	B2	Yes	C1	N3	No	N3	C1	Yes	N3	N4	No
A2	B2	Yes	B2	A1	Yes	A2	N4	No	N4	A2	Yes	N4	N3	No
B2	C2	Yes	A2	C1	Yes	B2	N5	No	N5	B2	Yes	N5	N2	No
C2	D2	Yes	D2	B1	Yes	C2	N6	No	N6	C2	Yes	N6	N1	No
Symmetry														
B1	A1	Yes												
C1	B1	Yes												
D1	C1	Yes												
B2	A2	Yes												
C2	B2	Yes												
D2	C2	Yes												
Transitivity														
A1	C1	Yes												
B1	D1	Yes												
A2	C2	Yes												
B2	D2	Yes												
A1	D1	Yes												
A2	D2	Yes												
Equivalence														
C1	A1	Yes												
D1	B1	Yes												
C2	A2	Yes												
D2	B2	Yes												
D1	A1	Yes												
D2	A2	Yes												

[a] In Experiment 3, the correct response was *No*.

The 48 trial types were presented in a quasi-random order, such that each trial type was presented twice across the 96 trials, with a short 1-min break after the first 48 trials. When the 96 trials had been presented, participants then were reexposed to the MTS equivalence test (but not the training) to determine if the equivalence relations were still in tact, and this marked the end of the experiment.

Results

The 27 participants required between six (the minimum) and 10 exposures to each cycle of MTS training and equivalence testing before successfully completing five successive equivalence tests (i.e., 22 of 24 correct responses) in the second session. All 27 participants passed the final equivalence test that was presented immediately after the lexical decision task, thus indicating that the equivalence relations were still in tact at the end of the experiment. The mean RTs (in ms) for correct responses and percentage of errors for each of the eight types of lexical decision task were calculated across the 27 participants. Table 3 shows the comparisons among trial types.

Reaction Times

Participants responded more rapidly to prime-target pairs that were directly or indirectly related (i.e., directly trained, symmetry, transitivity, or equivalence) than to unrelated prime-target pairs, or to pairs that contained one or two previously unseen stimuli (i.e., class–nonclass, class–nonsense, nonsense–class, nonsense–nonsense). Table 3 presents additional descriptive analyses of the RT and error data for each of the eight types of lexical decision task. The types of task are numbered 1 to 8 and these numbers are used subsequently in Table 4 to indicate which types of priming task were compared statistically during post hoc tests.

The RT data from the eight priming tasks were subjected to a repeated measures one-way analysis of variance (ANOVA), which proved to

Table 3

The eight types of priming task and related descriptive statistics for Experiment 1.

Stimuli presented		Correct response	Proportion of trials	Reaction times (ms)		% of errors	
Prime	Target			Mean	Std. error	Mean	Std. error
Equivalent pairs							
1. Directly trained	Directly trained	Yes	0.125	586.1	21.7	3.7	1.0
2. Symmetry	Symmetry	Yes	0.125	604.3	18.8	5.6	1.5
3. Transitivity	Transitivity	Yes	0.125	602.1	16.2	4.3	1.1
4. Equivalence	Equivalence	Yes	0.125	595.8	22.4	5.6	2.0
Nonequivalent Pairs							
5. Class	Nonclass	Yes	0.125	724.3	19.8	6.5	1.8
6. Class	Nonsense word	No	0.125	781.3	33.4	9.6	2.2
7. Nonsense word	Class	Yes	0.125	748.3	34.3	9.3	2.3
8. Nonsense word	Nonsense word	No	0.125	747.0	35.6	8.0	2.0

be significant, $F(7, 182) = 14.264$, $p < .0001$, $\eta_p^2 = .35$. A series of Scheffé post hoc tests also were conducted, comparing data for equivalent and nonequivalent stimulus pairs presented during the lexical decision tasks, the results of which are shown in Table 4. The specific post hoc tests that were conducted were chosen based on those reported by Hayes and Bissett (1998). Reaction times were significantly faster when both the prime and the target words were from the same equivalence class than when they were from different classes (Table 4: Equivalent vs. nonequivalent comparisons). No significant differences were obtained when comparisons were made between priming tasks in which the stimuli from both tasks were from the same equivalence class (Table 4: Equivalent vs. equivalent comparisons). Finally, no significant differences were obtained between the two YES-response conditions (5 and 7) or the two NO-response conditions (6 and 8) in which the stimuli were either from different equivalence classes or were previously unseen (Table 4: Nonequivalent vs. nonequivalent comparisons).

Table 4

Results of Scheffé post hoc tests comparing reaction times for the eight types of priming tasks in Experiment 1, with percentage of individual participants (in parentheses) conforming to the direction of the statistically significant effects.

Post hoc comparisons	
Equivalent versus nonequivalent comparisons	
Equivalent (direct) vs. nonequivalent (1 vs. 5)[a]	(88.9)
Equivalent (symmetry) vs. nonequivalent (2 vs. 5)[a]	(88.9)
Equivalent (transitivity) vs. nonequivalent (3 vs. 5)[a]	(96.3)
Equivalent (equivalence) vs. nonequivalent (4 vs. 5)[a]	(85.2)
Equivalent versus equivalent comparisons	
Direct vs. symmetry (1 vs. 2)	
Direct vs. transitivity (1 vs. 3)	
Direct vs. equivalence (1 vs. 4)	
Symmetry vs. transitivity (2 vs. 3)	
Symmetry vs. equivalence (2 vs. 4)	
Transitivity vs. equivalence (3 vs. 4)	
Nonequivalent versus nonequivalent comparisons	
Nonequivalent (5) vs. nonequivalent (7)	
Nonequivalent (6) vs. nonequivalent (8)	

[a] $p < .05$

These data indicate that priming is facilitated within equivalence classes, but not between them. Furthermore, there was no evidence to suggest that the priming effect was differentially sensitive to the effects of direct training versus symmetry, transitivity, or equivalence. Finally, stimulus pairs from different equivalence classes did not prime each other more rapidly than stimulus pairs incorporating previously unseen stimuli.

Error Data

Unlike the RT data, the errors were relatively uniform across the eight conditions (see Table 3). The error data for the eight lexical decision tasks were subjected to a repeated measures one-way ANOVA, but this proved to be nonsignificant, $F(7, 182) = 1.703$, $p = .1108$.

Discussion

The current data replicate the RT effects reported by Hayes and Bisset (1998), but not their error data (the failure to replicate the error data will be revisited in the context of Experiment 3). The results of Experiment 1 therefore support the view that derived stimulus relations are functionally similar to semantic relations, to the extent that priming is a semantic process. Furthermore, insofar as priming is an associative process, derived stimulus relations appear to function like direct associations. On balance, however, we should be cautious in drawing this final conclusion. In Experiment 1, all participants were required to pass an equivalence test *before* exposure to the lexical-decision task, and thus the four stimuli contained within each of the two equivalence classes had been matched repeatedly (i.e., directly associated) during the test, albeit without differential reinforcement. This fact limits the extent to which the priming effects observed here can be defined as *mediated* rather than *direct* priming.

Mediated priming refers to the priming effect that is sometimes obtained when the prime and target are *indirectly* semantically related via a mediating word or concept (e.g., Balota & Lorch, 1986; Weisbrod et al., 1999). For example, the word ''stripes'' may prime recognition of ''lion'' based on the mediating concept ''tiger.'' In Experiment 1, priming was clearly demonstrated across combined symmetry and transitivity relations (e.g., A1 primed C1 and vice versa), and this was taken as evidence for mediated priming because the A and C stimuli were indirectly related via the B stimuli. As was the case in the Hayes and Bisset (1998) study, however, all of the participants had successfully completed an equivalence test prior to the lexical decision task and thus the indirectly related stimuli had in fact been directly related during this test. Consequently, the priming among indirectly related stimuli observed in Experiment 1 and in the Hayes and Bisset study may have simply reflected direct rather then mediated priming. Ideally, therefore, an equivalence test should not be presented until *after* the lexical decision task if unequivocal mediated priming is to be observed across indirectly related elements of an equivalence relation. Experiment 2 adopted this strategy.

EXPERIMENT 2

Method

Participants

Participants were 54 right-handed undergraduates (38 females and 16 males) attending the National University of Ireland, Maynooth. They had not studied stimulus equivalence or any related phenomena prior to their participation in the study. Participants were between 17 and 24 years of age (mode 19).

Apparatus and Stimuli

The same apparatus and stimuli that were employed in Experiment 1 were employed in Experiment 2.

Procedure

The procedure of this experiment was similar to that of the first experiment with only one major difference; participants were exposed to equivalence training first, followed by the lexical decision task, which was then followed by the equivalence test. Once again, the experiment was divided across two separate sessions. In Session 1, each participant was required to produce 24 consecutively correct trials in order to complete one training exposure. To ensure that the trained conditional discriminations were firmly established, participants were each required to complete 10 cycles of training across two separate sessions before proceeding to the lexical decision task.

Results and Discussion

Twenty-eight of the 54 participants did not pass the equivalence test (< 22 of 24 correct; mean correct = 12.72; range = 9 to 16) following the lexical decision task. The data (RTs and errors) were therefore divided into two separate sets, one for the participants who passed the equivalence test (the pass group) and one for those who failed (the fail group). The mean numbers of training trials completed by the pass and fail groups were 424 and 413, respectively; an independent *t* test indicated that this difference was nonsignificant at the .05 level.

Reaction Time

Table 5 presents descriptive analyses of the data, showing the mean RTs for correct responses averaged over participants for all eight lexical decision tasks, for both the pass and fail groups. The mean RTs indicate that participants who passed the equivalence test responded more rapidly to prime-target pairs that were directly or indirectly related than to unrelated prime-target pairs or to pairs that

Table 5

The eight types of priming task and related descriptive statistics for participants who passed and for participants who failed the equivalence test in Experiment 2.

Stimuli presented		Correct response	Proportion of trials	Reaction times (ms)				% of errors			
				Mean		Std. error		Mean		Std. error	
Prime	Target			Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail
Equivalent pairs											
1. Directly trained	Directly trained	Yes	0.125	581.8	592.3	21.2	25.1	4.2	6.5	1.5	1.9
2. Symmetry	Symmetry	Yes	0.125	637.0	707.2	14.5	21.4	5.4	6.2	1.9	1.7
3. Transitivity	Transitivity	Yes	0.125	631.0	698.7	14.8	19.1	5.4	6.8	1.8	1.4
4. Equivalence	Equivalence	Yes	0.125	654.7	705.0	18.5	21.3	5.8	7.1	1.6	1.6
Nonequivalent pairs											
5. Class	Nonclass	Yes	0.125	731.5	773.3	20.5	24.8	7.7	8.3	1.6	2.2
6. Class	Nonsense word	No	0.125	725.9	800.5	23.2	31.3	8.3	12.0	2.2	2.4
7. Nonsense word	Class	Yes	0.125	768.3	731.0	29.5	21.8	5.8	8.6	1.6	2.2
8. Nonsense word	Nonsense word	No	0.125	760.2	773.4	22.2	25.0	6.7	7.1	2.0	1.9

contained one or two previously unseen stimuli; these results are similar to Experiment 1. In contrast to the pass group, the participants who failed the equivalence test produced a relatively brief mean RT for the directly trained relations only.

The RT data were subjected to a 2 × 8 mixed repeated measures ANOVA, with performance on the equivalence test (pass or fail) as the between-participant variable and the eight-types of lexical decision task as the within-participant variable. The ANOVA indicated no significant difference between the pass and fail groups, $F(1, 52) = 1.920$, $p = .1718$, but a significant difference was obtained across priming trial types, $F(7, 364) = 41.969$, $p < .0001$, $\eta_p^2 = .45$, and a significant interaction between groups and priming trial types was also recorded, $F(7, 364) = 3.825$, $p = .0005$, $\eta_p^2 = .07$. These analyses indicate that RTs to the eight priming trial types were differentially effected across the pass and fail groups. Two 1-way repeated measures ANOVAs were conducted on the data from the pass and fail groups to investigate these differences.

The ANOVA for the pass group proved to be significant, $F(7, 175) = 27.109$, $p < .0001$, $\eta_p^2 = .52$. A series of Scheffé post hoc tests also were conducted, comparing data for equivalent and nonequivalent stimulus pairs presented during the lexical decision tasks, the results of which are shown in Table 6. Reaction times were significantly faster when both the prime and the target words were from the same equivalence class than when they were from different classes (Table 6: Equivalent vs. nonequivalent comparisons). Only one significant difference was obtained when comparisons were made between priming tasks in which the stimuli from both tasks were from the same equivalence relation (directly trained pairs generated greater priming than pairs related indirectly via equivalence; see Table 6: Equivalent vs. equivalent comparisons). Finally, no significant differences were obtained between the two YES-response conditions (5 and 7) or the two NO-response conditions (6 and 8) in which the stimuli were either from different equivalence classes or were previously unseen (Table 6: Nonequivalent vs. nonequivalent comparisons). In short, the RT data from the pass group in Experiment 2 were similar to those obtained in Experiment 1, except that a significant difference was obtained between directly trained and equivalence priming trial types.

The ANOVA for the fail group also proved to be significant, $F(7, 189) = 18.834$, $p < .0001$, $\eta_p^2 = .42$. Scheffé post hoc tests indicated that RTs were significantly faster for the directly trained primes and targets relative to the symmetry, transitivity, and equivalence priming trial types (Table 6: Equivalent vs. equivalent comparisons). Although the remaining effects were in the same direction as the pass group, the only statistically significant effect was obtained between the directly trained and nonequivalent trial types (Table 6: 1 vs. 5). Unlike the pass group, therefore, who showed significant priming among all four components of the equivalence relations, the fail group only showed this effect, at a statistically significant level, for the directly trained primes and targets.

Table 6

Results of Scheffé post hoc tests comparing reaction times for the eight types of priming tasks for participants who passed and for participants who failed the equivalence test in Experiment 2, with percentage of individual participants (in parentheses) conforming to the direction of the statistically significant effects.

Post hoc comparisons	Pass and fail groups: Pass	Fail
Equivalent versus nonequivalent comparisons		
Equivalent (direct) vs. nonequivalent (1 vs. 5)	[a](100)	[a](92.8)
Equivalent (symmetry) vs. nonequivalent (2 vs. 5)	[a](96.1)	
Equivalent (transitivity) vs. nonequivalent (3 vs 5)	[a](84.6)	
Equivalent (equivalence) vs. nonequivalent (4 vs. 5)	[a](88.5)	
Equivalent versus equivalent comparisons		
Direct vs. symmetry (1 vs. 2)		[a](92.8)
Direct vs. transitivity (1 vs. 3)		[a](85.7)
Direct vs. equivalence (1 vs. 4)	[a](96.1)	[a](85.7)
Symmetry vs. transitivity (2 vs 3)		
Symmetry vs. equivalence (2 vs. 4)		
Transitivity vs. equivalence (3 vs. 4)		
Nonequivalent versus nonequivalent comparisons		
Nonequivalent (5) vs. nonequivalent (7)		
Nonequivalent (6) vs. nonequivalent (8)		

[a] $p < .05$.

Error Data

The error data (see Table 5 for descriptive statistics) were subjected to a 2 × 8 mixed repeated measures ANOVA, and no significant difference between the pass and fail groups, $F(1, 51) = 1.776$, $p = .1886$, or among the eight priming trial types, $F(7, 357) = 1.471$, $p = .1763$, was obtained (note, error data for 1 participant from the fail group was lost due to experimenter error). Furthermore, no interaction between groups and trial type was recorded, $F(7, 357) = .211$, $p = .9829$. Like Experiment 1, therefore, the RT, but not the error data, indicated priming effects.

The results of Experiment 2 show that the RT priming effect was obtained only for the group that successfully passed the subsequent equivalence test, whereas the group that failed the test also failed to demonstrate priming, at a statistically significant level (except for the directly trained relations). The pass group therefore showed priming for stimulus pairs that were never directly associated (i.e., paired together), and thus the priming effects appear to parallel what cognitive researchers have termed ''mediated priming.''

EXPERIMENT 3

In both Experiments 1 and 2, the most common measures of semantic priming were used; RTs and error scores. However, there is a substantive body of research on semantic priming that also has employed electrophysiological measures of the priming effect (e.g., Kutas & Hillyard, 1980; Weisbrod et al., 1999). For example, one measure of brain activity, known as Event-Related Potentials (ERP), is particularly well suited to studying the effects of discrete stimulus presentations on human learning (see Holcomb, 1988; Holcomb & Neville, 1991; Kutas, 1993). The technique involves placing electrodes at specified locations on the scalp of the head, from which it is possible to record electroencephalograms (EEGs) from each location. A marker is then placed in the continuous EEG signal each time a stimulus is presented to the participant. A specific segment of EEG following each marker is then averaged across stimulus presentations and across participants to produce the ERP waveform for that particular stimulus.

There are numerous waveforms associated with ERP measures. For example, some ERPs are thought to be associated with cognitive functions such as understanding words or being able to distinguish one type of visual or auditory stimulus from another. These ERPs occur at around 300 or 400 ms after the stimulus onset. The ERP measure that is most relevant in the context of the current research is a late negative waveform, known as the N400 (see Holcomb & Anderson, 1993; Kounios & Holcomb, 1992). This waveform is typically produced when participants are asked to respond to words that are semantically unrelated. In contrast, when the words are from the same semantic categories, the N400 is greatly reduced or completely absent.

In a recent study on direct and mediated semantic priming (Weisbrod et al., 1999), ERP measures were taken during a lexical decision task. The results of the study showed that the N400 waveform was more pronounced for the unrelated word pairs than for either the directly or indirectly related word pairs, which indicates the standard priming effect. The primary purpose of Experiment 3 was to determine if the N400 waveform also would differentiate between nonequivalent and directly trained and equivalent stimulus relations on a lexical decision task. Insofar as the N400 is more sensitive to semantic associations than RT (e.g., Kounios & Holcomb, 1992), demonstrating N400 sensitivity to equivalence relations will provide important additional evidence for the functional overlap between semantic and derived relations. In Experiment 3, therefore, ERPs were collected for all participants while they completed the lexical decision task. Experiment 3 also employed a two-word lexical decision task similar to that used by Hayes and Bisset (1998) to determine if it would produce a significant difference among the error scores.

Method

Participants

Twenty-one of the participants who passed the equivalence test in Experiment 2 also completed Experiment 3 (15 females and 6 males, 18 to 24 years of age; mode 19).

Apparatus and Stimuli

The experiment was conducted in a sound-attenuated, electrically shielded cubicle in the electrophysiology laboratory in the Department of Psychology at the National University of Ireland, Maynooth. The apparatus used in Experiments 1 and 2 were also employed in Experiment 3. A novel set of stimuli was employed for the MTS procedures and lexical decision task, and these were selected based on the same five criteria as stipulated in the *Apparatus and Stimulus* section of Experiment 1. In order to record the EEG measures during the lexical decision task, a Brain Amp MR (Class IIa, Type BF) with approved control software (Brain Vision Recorder Version 1.0) and an approved electrode cap (BrainCap/MR) were employed. The Brain Amp was controlled by a Dell® personal computer with a Pentium 4® processor. Finally, the ERP data were analyzed using approved analysis software (Brain Vision Analyzer Version 1.0). All of the hardware and software were manufactured and supplied by Brain Products GmbH, Munich, Germany.

Procedure

The procedure was similar to that employed in Experiment 2, except that following the ninth successful cycle of MTS training (which always occurred in a second session) participants were fitted with the electrode cap in preparation for recording the EEG measures that were to be taken during the lexical decision task. Following a final 10th cycle of MTS training, the Brain Amp was activated, and the participants then were exposed to the two-word lexical decision task (i.e., there was an approximate break of 60 s between completing the MTS testing and commencing the latter task).

The two-word lexical decision task was similar to the one-word task used in Experiments 1 and 2, except that on each trial the prime and the target word remained on screen together. The sequence of stimulus presentations was as follows. The red X appeared for 500 ms, followed immediately by the prime stimulus, which was then followed 100 ms later by the target (a 100 ms SOA); the prime and target were positioned in the center of the screen with the former located 3 cm above the latter. The prime and target remained on screen for 1,750 ms; the green X was then presented for 1,000 ms followed by the red X (for 500 ms), which marked the beginning of the next trial. Responses made after the green X had been removed from the screen were recorded by the computer program as ''too late'' and this datum was excluded from subsequent analyses. The instructions employed in Experiment 3 were similar to those employed in Experiment 1, except that the instructions for the lexical decision task specified that participants must respond to both the prime and the target words, and not simply to the target.

Recordings. Ten sites were chosen to monitor EEG signals, and these were taken from the standard International 10-20 system locations over the left and right hemispheres at the central (C3 and C4), parietal (P3 and P4) temporal (T3 and T4, T5 and T6), and occipital (O1 and O2) sites. These sites were chosen because they were employed in a recent study that demonstrated that the N400 waveform is sensitive to direct and indirect semantic priming using words from a natural language (Weisbrod et al., 1999; see also Holcomb, Coffey, & Neville, 1992; Holcomb & Neville, 1990; 1991). Four frontal sites (Fp1, Fp2, F3, F4) also were used but the N400 was not reliably detectable—in accordance with Weisbrod et al., these data are not reported.

Evoked potentials were recorded and analyzed from 10 sintered AG/AG-CI scalp electrodes positioned as outlined above. Linked mastoids were used as reference sites and FPz as ground. The amplifier resolution was 0.1 μV (range, ± 3.2768 mV), and the bandwidth was set between 0.5 and 62.5 Hz with a sampling rate of 250 Hz. All electrode impedances were at or below 5 kΩ. The EEG was continuously collected and edited off-line. Epochs of 900 ms with a pretarget stimulus time of 200 ms were analyzed. The electro-oculogram (EOG) was recorded using the same specifications as the EEG. The horizontal and vertical EOG were monitored with two electrodes each. Special care was taken to reject segments with eye movements or other artifacts. Trials on which EEG or EOG activity exceeded ± 75 μV were rejected (Silva-Pereyra et al., 1999). The EEG data from one participant was removed from subsequent analyses due to an excessive number of artifacts (less than nine artifact-free trials per condition). Baseline correction was performed in relation to a pretarget stimulus time of 200 ms (i.e., 100 ms before the presentation of the prime). Average ERPs to primes and targets that were directly trained (e.g., A1-B1), related through equivalence (e.g., C1-A1, D2-A2), or were unrelated through equivalence (class–nonclass; e.g., B1-A2, A1-C2) were obtained from correct response trials for each participant. In accordance with Weisbrod et al. (1999), these individual averages were digitally low-pass filtered at 16 Hz (24 dB/octave roll-off), and the N400 waveform was measured between 350 and 550 ms after the target onset. The directly trained, equivalent, and nonequivalent prime and target pairs were selected for ERP recording because they appear to provide the derived-stimulus-relations analogue of directly related, indirectly related, and unrelated word pairs examined in previous ERP studies of semantic priming (e.g., Weisbrod et al., 1999).

Results and Discussion

The mean number of MTS training trials completed by the 21 participants was 398.

Reaction Time

The mean RTs shown in Table 7 indicate that participants responded more quickly to prime-target pairs that were directly or indirectly related than to unrelated prime-target pairs or to pairs that contained one or two previously unseen stimuli. Table 7 presents additional descriptive analyses of the RT data for each of the eight types of lexical decision task. The types of task are numbered 1 to 8, and these numbers are used subsequently in Table 8 to indicate which types of priming task were compared statistically during post hoc tests.

The RT data from the eight priming tasks were subjected to a repeated measures one-way ANOVA, which proved to be significant, $F(7, 140) = 32.007$, $p < .0001$, $\eta_p^2 = .61$. A series of Scheffé post hoc tests compared data for equivalent and nonequivalent stimulus pairs presented during the lexical decision tasks, the results of which are shown in Table 8. Reaction times were significantly faster when both the prime and the target words were from the same equivalence class than when they were from different classes (Table 8: Equivalent vs. nonequivalent comparisons). Similar to Experiment 2 (pass group) only one significant difference was obtained when comparisons were made between priming tasks in which the stimuli from both tasks were from the same equivalence relation (directly trained pairs generated greater priming than pairs related indirectly via equivalence; see Table 8: Equivalent vs. equivalent comparisons). Finally, no significant differences were obtained among any of the three NO-response conditions (6 vs. 7, 6 vs. 8, and 7 vs. 8) in which one or both of the stimuli were previously unseen (Table 8: Nonequivalent vs. nonequivalent comparisons; bottom three rows).

These data indicate, therefore, that priming, as measured by RT, is facilitated within equivalence classes, but not between them. Furthermore, consistent with Experiment 2 (but not Experiment 1) there was evidence to suggest that the observed priming effects were differentially sensitive to direct training versus equivalence relations. Finally, differential priming effects did not emerge among stimulus pairs incorporating previously unseen stimuli.

Error Data

The mean percentage errors shown in Table 7 show that this measure was slightly lower for the four equivalent trial types relative to the four nonequivalent trial types (see

Table 7

The eight types of priming task and related descriptive statistics for Experiment 3.

Stimuli presented		Correct response	Proportion of trials	Reaction times (ms)		% of errors	
Prime	Target			Mean	Std. error	Mean	Std. error
Equivalent pairs							
1. Directly trained	Directly trained	Yes	0.125	614.8	18.8	8.7	2.1
2. Symmetry	Symmetry	Yes	0.125	684.4	22.4	13.5	2.9
3. Transitivity	Transitivity	Yes	0.125	670.3	19.7	11.1	2.4
4. Equivalence	Equivalence	Yes	0.125	731.9	23.3	15.1	2.9
Nonequivalent pairs							
5. Class	Nonclass	Yes	0.125	848.6	29.9	18.2	2.9
6. Class	Nonsense word	No	0.125	887.2	38.9	16.2	2.9
7. Nonsense word	Class	No	0.125	903.8	25.6	15.4	2.6
8. Nonsense word	Nonsense word	No	0.125	868.1	28.0	20.2	2.7

Table 7 for additional descriptive statistics). The error data for the eight lexical decision tasks were subjected to a repeated measures one-way ANOVA, and this proved to be significant, $F(7, 140) = 2.111$, $p = .0462$, $\eta_p^2 = .09$. A series of Scheffé post hoc tests were then conducted, comparing data for equivalent and nonequivalent stimulus pairs presented during the lexical decision tasks, but none of these tests proved to be significant at the .05 level. In summary, therefore, the error data in Experiment 3 provided some evidence for greater priming among members of the two equivalence classes, but the statistical effects were far weaker than for the RT data.

Table 8

Results of Scheffé post hoc tests comparing reaction times for the eight types of priming tasks in Experiment 3, with percentage of individual participants (in parentheses) conforming to the direction of the statistically significant effects.

Post hoc comparisons	
Equivalent versus nonequivalent comparisons	
Equivalent (direct) vs. nonequivalent (1 vs. 5)[a]	(95.2)
Equivalent (symmetry) vs. nonequivalent (2 vs. 5)[a]	(90.5)
Equivalent (transitivity) vs. nonequivalent (3 vs. 5)[a]	(100)
Equivalent (equivalence) vs. nonequivalent (4 vs. 5)[a]	(95.2)
Equivalent versus equivalent comparisons	
Direct vs. symmetry (1 vs. 2)	
Direct vs. transitivity (1 vs. 3)	
Direct vs. equivalence (1 vs. 4)[a]	(90.5)
Symmetry vs. transitivity (2 vs. 3)	
Symmetry vs. equivalence (2 vs. 4)	
Transitivity vs. equivalence (3 vs. 4)	
Nonequivalent versus nonequivalent comparisons	
Nonequivalent (6) vs. nonequivalent (7)	
Nonequivalent (6) vs. nonequivalent (8)	
Nonequivalent (7) vs. nonequivalent (8)	

[a] $p < .05$.

ERP Measures

The averaged EEG data were calculated across 20 participants. The grand average waveforms for each of the 10 electrode sites for the directly trained, equivalent, and nonequivalent lexical decision trial types are shown in Figure 1. Visual inspection of these waveforms indicates that the most pronounced negative deflections occurred for the nonequivalent priming trial types (thick black lines) commencing at around 400 ms, thus approximating the classic N400 waveform. For all 10 sites, the post-400-ms waveforms for the directly trained prime-target pairs (thin dark lines) were clearly more positive than for the nonequivalent pairs. Usually, the equivalent prime-target pairs (thick light lines) produced post-400-ms waveforms that fell somewhere between the directly trained and nonequivalent waveforms. In comparing the grand averages between 350 and 550 ms, the differences between the equivalent versus nonequivalent priming trial types appear to be greater for the left hemisphere relative to the right (left and right panels of Figure 1, respectively); this difference is most pronounced across the parietal and temporal sites.

The area dimensions ($\mu V \times ms$) for each waveform for each participant recorded between 350 and 550 ms were calculated, yielding either positive or negative values with respect to the 0 μV level. These data were subjected to a three-way ANOVA with priming trial type (directly trained, equivalent, and nonequivalent), scalp location (central, parie-

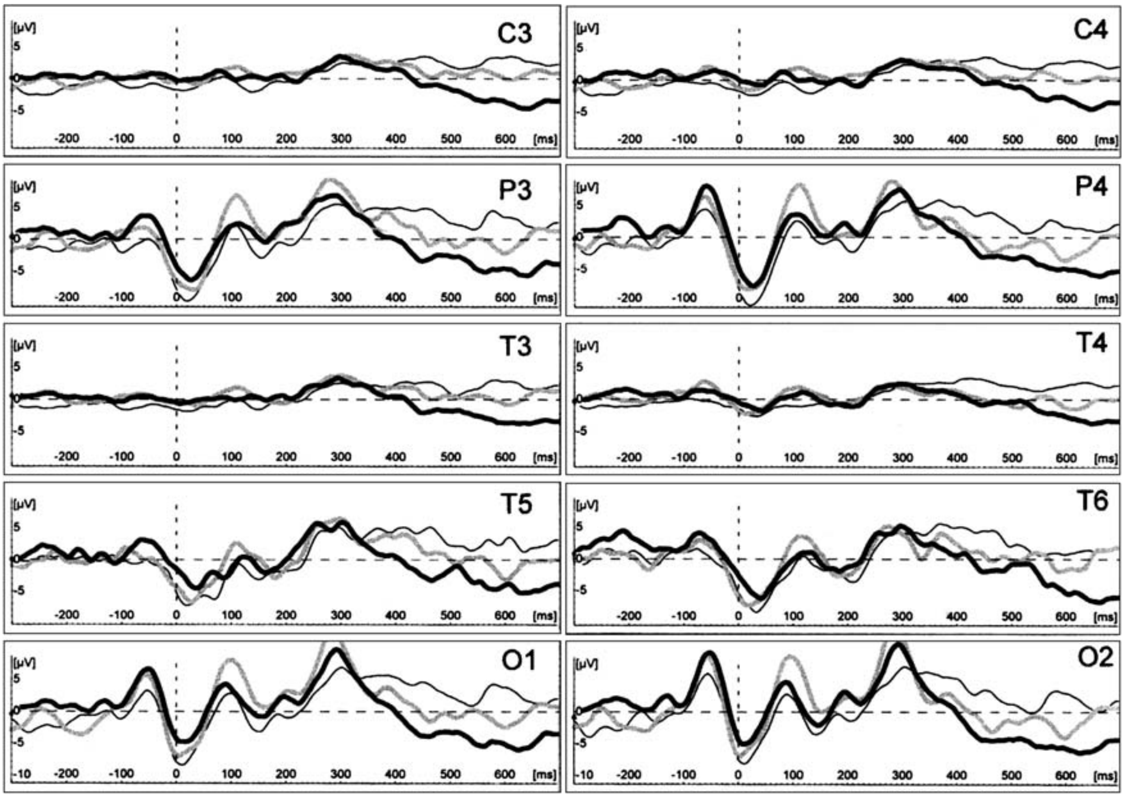

Fig. 1. Grand average waveforms from 20 participants for prime-target stimulus pairs that were directly trained (thin black lines), equivalent (thick gray lines) and nonequivalent (thick black lines) at electrode sites C3, C4 (top panel), P3, P4 (second panel), T3, T4 (third panel), T5, T6 (fourth panel), and O1, O2 (bottom panel). Note that the prime was presented 100 ms (−100) prior to the target stimulus (0 ms).

tal, temporal 3/4, temporal 5/6, and occipital), and laterality (left/right) as repeated measures factors. The ANOVA revealed main effects for laterality, $F(1, 19) = 7.039$, $p = 0.0157$, $\eta_p^2 = .27$; scalp location, $F(4, 76) = 4.781$, $p = 0.0017$, $\eta_p^2 = .20$; and priming trial type, $F(2, 38) = 25.336$, $p < 0.0001$, $\eta_p^2 = .57$. All four possible interaction effects also were found to be significant: laterality by scalp location, $F(4, 76) = 16.798$, $p < 0.0001$, $\eta_p^2 = .47$; laterality by priming trial type, $F(2, 38) = 10.050$, $p = 0.0003$, $\eta_p^2 = .34$; scalp location by priming trial type, $F(8, 152) = 19.078$, $p < 0.0001$, $\eta_p^2 = .50$; and laterality by scalp location by priming trial type, $F(8, 152) = 3.406$, $p = 0.0013$, $\eta_p^2 = .15$.

Given that significant differences were obtained for all main effects and all interactions, the data were subjected to 10 individual one-way ANOVAs, one for each electrode site. All 10 ANOVAs proved to be significant, all Fs$(2, 19) > 10.1$, all ps < 0.003, all $\eta_p^2 > .35$. Scheffé post hoc tests conducted for each ANOVA (see Table 9) indicated that the waveforms (between 350 and 550 ms) at the five left-hemisphere sites were significantly more negative for the nonequivalent priming trial types relative to the directly trained and equivalent trial types. The post hoc tests also revealed that the equivalent trial types were significantly more negative than the directly trained trial types for all left-hemisphere sites except for C3. The right-hemisphere sites also produced significant differences between the directly trained trial types and both the equivalent and nonequivalent trial types, but unlike the left sites none of the post hoc comparisons between the equivalent and nonequivalent trial types proved to be significant, although T6 and O2 approached significance ($p = 0.058$ and $p = 0.052$, respectively).

In summary, the ERP measures indicate that the N400 waveform is sensitive to the differences among directly trained, derived equiva-

Table 9

Results of Scheffé post hoc tests comparing area dimensions (μV × ms), 350 to 550 ms, for the directly trained, equivalent, and nonequivalent trial types for each of the 10 electrode sites, with percentage of individual participants (in parentheses) conforming to the direction of the statistically significant effects.

	Statistical significance									
	Area dimensions (μV × ms)									
Post hoc comparisons	C3	P3	T3	T5	O1	C4	P4	T4	T6	O2
	Left hemisphere sites					Right hemisphere sites				
Equivalent versus nonequivalent comparisons										
Equivalent (direct) vs. nonequivalent (1 vs. 5)	[a] (85)	[a] (90)	[a] (85)	[a] (85)	[a] (95)	[a] (80)	[a] (85)	[a] (85)	[a] (90)	[a] (95)
Equivalent (equivalence) vs. nonequivalent (4 vs. 5)	[a] (70)	[a] (80)	[a] (75)	[a] (80)	[a] (90)				[b]	[b]
Equivalent versus equivalent comparisons										
Direct vs equivalence (1 vs 4)		[a] (85)	[a] (70)	[a] (75)	[a] (85)		[a] (85)	[a] (70)	[a] (70)	[a] (85)

[a] $p < .05$.
[b] Approaching significance ($p < .059$).

lent, and nonequivalent stimulus pairs that were established using a behavioral model of semantic relations. The results also indicated a significant effect for laterality, with greater negativity observed for the left, relative to the right sites. The post hoc analyses suggest that the left and right hemispheres may be differentially sensitive to the mediated priming effect. That is, all five of the left-hemisphere sites but none of the right-hemisphere sites produced significant differences between equivalent and nonequivalent prime-target pairs, although two of the latter sites were only marginally nonsignificant. Overall, Experiment 3 provides evidence for both direct and mediated priming among derived stimulus relations using the three most common measures employed in semantic priming research; RTs, error scores, and the N400 waveform.

GENERAL DISCUSSION

The current study demonstrates priming effects with derived stimulus relations using a single-word lexical decision task both after (Experiment 1) and before (Experiment 2) a formal MTS equivalence test. The results of Experiment 2 are particularly compelling in that priming was not observed for those participants who subsequently failed the equivalence test. This indicates that training in a set of interrelated conditional discriminations is not sufficient to produce the priming effect normally observed with semantic relations in natural language. Rather, the conditional discrimination training must give rise to derived equivalence relations if semantic-like effects are to be obtained. This result certainly supports the argument that derived relations, rather than directly reinforced stimulus relations alone, provide a behavioral model of what cognitive researchers refer to as semantic processes (Barnes-Holmes, Hayes & Dymond, 2001; Barsalou, 1999; Deacon, 1997; Hayes & Bisset, 1998).

The results of Experiment 3 provide additional evidence in favor of a functional overlap between semantic and derived stimulus relations, in that the N400 waveform was shown to be sensitive to the directly trained and equivalent stimulus pairs versus the nonequivalent pairs. Furthermore, a significant effect for laterality was observed, with greater negativity observed for the left relative to the right sites, which is consistent with previous research using natural language stimuli (Weisbrod et al., 1999; unfortunately, individual site analyses were not provided, and thus a direct comparison with this earlier natural-language research is not possible).

The current data are important in that only one published study has undertaken an investigation of the neural correlates of derived relational responding (Dickins et al., 2001, who used fMRI). The current research employed

ERP measures rather than fMRI, but the present findings are broadly consistent with that earlier work in that both measures indicate that derived stimulus relations produce neural effects that are typically observed when humans are engaged in activities that cognitive psychologists call semantic processing.

In comparing the results of Experiment 1 with Experiments 2 and 3, it appears that mediated priming effects can be observed or assessed only when an equivalence test is *not* presented before the lexical decision task. That is, significant differences in RT measures between directly trained and equivalent prime-target pairs were not observed in Experiment 1, but were observed in both Experiments 2 and 3. These results suggest that a formally tested equivalence class does not provide an accurate model of an indirect semantic relation. During a successful MTS equivalence test the participant matches all of the indirectly related stimuli, and although no explicit reinforcement or feedback is provided, the stimuli are nonetheless directly associated by each matching response. This type of repeated pairing seems highly unlikely for the indirectly related word-pairs typically employed in semantic priming studies (e.g., lion and stripes), and thus the apparent absence of a mediated priming effect in Experiment 1 might well have been predicted. Indeed, Hayes and Bisset (1998) required that all of their participants successfully pass an equivalence test before the lexical decision task, and this study also failed to find significant within-class RT differences.

The fact that mediated priming could not be shown in Experiment 1, but was shown in Experiments 2 and 3, demonstrates how a derived relations model may provide a level of precision in the analysis of semantic relations that is difficult to attain using stimuli from natural language (because the exact histories of association among real words remain unknown). In the current study, the relations were created ab initio in each experiment, and the structure of those relations, in terms of direct versus indirect relatedness, and whether or not the relations had been tested, could be manipulated. An example of the precision that this methodology provides is highlighted in Experiments 2 and 3, in which significant differences were obtained between the directly trained and equivalent relations, but not between the directly trained and transitive relations. This result suggests that the relative strength of the mediated priming effect observed in the context of the current model depends, in part, on the direction of the relation, not simply the presence of a mediating node (if direction was irrelevant, then both transitive and equivalent prime-target pairs would have produced significantly longer RTs relative to directly trained relations).

At the present time it remains unclear why this difference in the priming effect emerged between transitive and equivalence relations. It is worth noting, however, that previous research with nonhumans has provided clear evidence for transitive relations (see Zentall & Smeets, 1996, for a review), but a reliable demonstration of equivalence relations, which requires a combination of both symmetry and transitivity, has remained elusive (e.g., Dugdale & Lowe, 2000; Lionello-DeNolf & Urcuioli, 2002; but see Shusterman & Kastak, 1993). It is possible, therefore, that transitive and equivalence relations may involve different psychological processes. For example, transitivity may be based on some form of mediated generalization, for which there is considerable evidence in the nonhuman literature (see Hall, 1996; Urcuioli, 1996), and thus this relation alone may not provide solid evidence for a semantic relation. In contrast, perhaps the equivalence relations in the current study (which cannot be readily explained in terms of mediated generalization; Hall, 1996; Urcuioli, 1996) could be considered genuinely semantic. Insofar as this turns out to be the case, the current work has served to highlight some of the possibly important but subtle properties of semantic relations as studied in an experimental context.

Although Experiments 1 and 2 of the current study produced a priming effect for RTs on the lexical decision task, the error scores did not show a priming effect similar to that of Hayes and Bisset (1998). The current procedures differed in a number of ways from those employed in the earlier research (e.g., no feedback was presented during the lexical decision task, two rather than three equivalence classes were employed), and thus the source of the different pattern of error scores across the two studies is obscured. Nevertheless, it is important to note that when one compares the error data from both the current study and that of Hayes and Bisset with the results from typical priming research, the present findings more closely resemble those reported in the priming literature. For example, some researchers have reported the absence of statistically significant differences among error scores, although RTs were significant (e.g., Balota & Lorch, 1986, Experiment 1; McNamara & Altarriba, 1988, Experiment 1). Furthermore, the significant difference that was found among error scores in Experiment 3 of the current study suggests that the derived relations model is not completely insensitive to what may be considered the weakest measure of semantic priming (e.g., Kounios & Holcomb, 1992).

A great deal of research and considerable debate surrounds the topic of semantic priming in the cognitive literature. In particular, the theoretical debate has focused on the relation between what are called automatic (or spreading activation) versus controlled processes (Neely, 1977, 1991; Posner & Snyder, 1975). It has been suggested (Neely, 1977) that automatic spread of activation is the relevant process when the interval between prime and target (SOA) is short, whereas controlled processes (expectancy-induced priming) are engaged with long SOAs (longer than 400 ms). Given that relatively short SOAs were employed in each of the three experiments reported in the current study (250 ms in Experiments 1 and 2, and 100 ms in Experiment 3), it is unlikely that the observed priming effects were expectancy-induced. This conclusion is important for the derived relations model of semantic relations because it has recently been shown that expectancy-induced priming may not be directly involved in language processing (Brown, Hagoort, & Chwilla, 2000).

Although expectancy-induced priming seems unlikely in the current study, alternatives to spreading activation have been suggested, including a range of integration processes, which do not depend upon the SOA duration (Hess, Foss, & Carroll, 1995; McKoon & Ratcliffe, 1992; Rugg & Doyle, 1994). At the present time, therefore, the current priming effects might be explained, *cognitively*, in terms of spreading activation, or an integration process, or perhaps some combination of both (Hill et al., 2002). In any case, drawing simultaneously on the theories and methodologies of both the cognitive and behavioral psychologies, as we have done in the current study, may serve not only the behavioral understanding of human language and cognition, but also contribute to the research endeavors of cognitive psychologists and neuroscientists.

REFERENCES

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Balota, D. A., & Lorch, R. F., Jr. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 336–345.

Barnes, D., & Hampson, P. J. (1993). Stimulus equivalence and connectionism: Implications for behavior analysis and cognitive science. *The Psychological Record, 43*, 617–638.

Barnes, D., & Holmes, Y. (1991). Radical behaviorism, stimulus equivalence, and human cognition. *The Psychological Record, 41*, 19–31.

Barnes-Holmes, D., Hayes, S. C., & Dymond, S. (2001). Self and self-directed rules. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational Frame Theory: A post-Skinnerian account of human language and cognition* (pp. 119–139). New York: Plenum.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–660.

Branch, M. N. (1994). Stimulus generalization, stimulus equivalence, and response hierarchies. In S. C. Hayes, L. J. Hayes, M. Sato, & K. Ono (Eds.), *Behavior analysis of language and cognition* (pp. 51–70). Reno, NV: Context Press.

Brown, C., Hagoort, P., & Chwilla, D. (2000). An event-related brain potential analysis of visual word-priming effects. *Brain and Language, 72*, 158–190.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Cullinan, V. A., Barnes, D., Hampson, P. J., & Lyddy, F. (1994). A transfer of explicitly and non-explicitly trained sequence responses through equivalence relations: An experimental demonstration and connectionist model. *The Psychological Record, 44*, 559–585.

Deacon, T. (1997). *The symbolic species.* London: Penguin.

de Groot, A. M. B. (1983). The range of automatic spreading activation in word priming. *Journal of Verbal Learning and Verbal Behavior, 22*, 417–436.

Dickins, D. W., Singh, K. D., Roberts, N., Burns, P., Downes, J., Jimmieson, P., & Bentall, R. P. (2001). An fMRI study of stimulus equivalence. *Neuroreport, 12*, 2–7.

Dugdale, N., & Lowe, C. F. (2000). Testing for symmetry in the conditional discrimination of language trained chimpanzees. *Journal of the Experimental Analysis of Behavior, 73*, 5–22.

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

Hall, G. (1996). Learning about associatively activated stimulus representations: Implications for acquired equivalence and perceptual learning. *Animal Learning and Behavior, 24*, 233–255.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational Frame Theory: A post-Skinnerian account of human language and cognition.* New York: Plenum.
Hayes, S. C., & Bisset, R. T. (1998). Derived stimulus relations produce mediated and episodic priming. *The Psychological Record, 48*, 617–630.
Hayes, S. C., & Hayes, L. J. (1992). Verbal relations and the evaluation of behavioral analysis. *American Psychologist, 47*, 1383–1395.
Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General, 124*, 62–82.
Hill, R., Strube, M., Roesch-Ely, D., & Weisbrod, M. (2002). Automatic vs. controlled processing in semantic priming – differentiation by event-related potentials. *International Journal of Psychophysiology, 44*, 197–218.
Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language, 35*, 66–85.
Holcomb, P. J., & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related potentials. *Language and Cognitive Processes, 8*, 379–411.
Holcomb, P. J., Coffey, S. A., & Neville, H. J. (1992). Visual and auditory sentences processing: A developmental analysis using event-related potentials. *Developmental Neuropsychology, 8*, 203–241.
Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes, 5*, 281–312.
Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19*, 286–300.
Kounios, S. A., & Holcomb, P. J. (1992). Structure and process in semantic memory: Evidence from event-related potentials and reaction times. *Journal of Experimental Psychology General, 121*, 460–480.
Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.
Kutas, M. (1993). In the company of other words: Electrophysiological evidence for simple-word and sentence context effects. *Language and Cognitive Processes, 8*, 533–578.
Kutas, M., & Hillyard, S. A. (1980). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 1161–1163.
Lionello-DeNolf, K. M., & Urcuioli, P. J. (2002). Stimulus control topographies and tests of symmetry in pigeons. *Journal of the Experimental Analysis of Behavior, 78*, 467–495.
Massaro, D. W., Venezky, R. L., & Taylor, G. A. (1979). Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General, 108*, 107–124.
McClelland, J. L., & Rumelhart, D. E. (1988). *Explorations in parallel distributed processing: A handbook of models, programs, and exercises.* Cambridge, MA: MIT Press.
McKoon, G., & Ratcliffe, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1155–1172.
McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language, 27*, 545–559.
Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognition pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227–234.
Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology, 106*, 226–254.
Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processing in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.
Neely, J. H., Keefe, D. F., & Ross, F. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectations and retrospection semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1003–1019.
Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.
Reese, H. W. (1991). Mentalistic approaches to verbal behavior. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 151–177). Reno, NV: Context Press.
Rugg, M. D., & Doyle, M. C. (1994). Event-related potentials and stimulus repetition in indirect and direct tests of memory. In H. Heinze, T. Munte, & G. R. Mangun (Eds.), *Cognitive electrophysiology* (pp. 124–148). Boston: Birkhauser.
Shusterman, 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. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213–245). Hillsdale, NJ: Erlbaum.
Sidman, M. (1994). *Stimulus equivalence: A research story.* Boston: Authors Cooperative.
Silva-Pereyra, J., Harmony, T., Villanuera, G., Fernandez, T., Rodriguez, M., Galan, L., et al. (1999). The N400 and lexical decisions: Automatic or controlled processing? *Clinical Neurophysiology, 40*, 813–824.
Urcuioli, P. J. (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.
Weisbrod, M., Kiefer, M., Winkler, S., Maier, S., Hill, R., Roesch-Ely, D., & Spitzer, M. (1999). Electrophysiological correlates of direct versus indirect semantic priming in normal volunteers. *Cognitive Brain Research, 8*, 289–298.
Zentall, T. R., & Smeets, P. M. (Eds.) (1996), *Stimulus class formation in humans and animals.* Amsterdam: Elsevier.

Received August 18, 2004
Final acceptance March 21, 2005