The Analysis of Verbal Behavior 1997, **14**, 19-40

Intraverbal Stimulus-Response Reversibility: Fluency, Familiarity Effects, and Implications for Stimulus Equivalence

David A. D. Polson, Diana M. Grabavac, and Joseph A. Parsons
University of Victoria

English-speaking subjects with little knowledge of the French language used a computerized flashcard program, *Think Fast*, to learn 16 English-French word pairs (intraverbals) by typing one word of the pair when presented with the other word as a textual stimulus. In Phase 1, half of the intraverbals were taught from French to English (FE-1) and half from English to French (EF-1). Then, in Phase 2, training continued with the stimulus and response items of each intraverbal reversed, i.e., cards previously in the FE-1 condition were trained from English to French (EF-2) and cards previously in the EF-1 condition were trained from French to English (FE-2). Feedback was provided throughout the experiment. Reversing the stimulus and response items in Phase 2 significantly reduced rate correct and accuracy scores for eight of the nine subjects. In Experiments 1 and 2, this effect was more pronounced for cards in the EF-2 condition; in Experiment 3, when the criterion for a "correct" response was more lenient, there was no consistent difference between cards in the EF-2 and FE-2 conditions. Symmetry, as indicated by accuracy scores on the first trial in Phase 2, was generally poor: eight of the nine subjects averaged only 29% correct when asked to respond to the reversed relations for the first time. We relate our paradigm and results to recent developments in fluency, verbal behavior, and stimulus equivalence, and provide directions for future research.

The acronym SAFMEDS was coined by precision teachers to refer to rules students should follow when using flashcards as a study method (Lindsley, 1996a, 1996b). Students are advised to *S*ay the answers to *A*ll the cards in a deck, work at a *F*ast pace during one *M*inute timings, practice *E*ach day, and *S*huffle the deck. The major goal is to obtain *fluency*, defined as a combination of accuracy and speed that characterizes competent performance and typically measured in terms of frequency per unit of time (Binder, 1993, 1996; Johnson & Layng, 1996). When fluent levels on basic tasks are obtained, three major outcomes have been reported: improved retention (e.g., Olander, Collins, McArthur, Watts, & McDade, 1986), greater resistance to fatigue and distraction (e.g., Binder, Haughton, & Van Eyk, 1990), and transfer to new situations and more complex behavior (e.g., Johnson & Layng, 1992). There have been some highly persuasive arguments that fluency is an educationally desirable goal (Binder, 1993; Binder & Watkins, 1990; Johnson & Layng, 1992; Lindsley, 1992). In light of these facts, Parsons (1989) developed a computerized flashcard program (*Think Fast*) designed to assist students in obtaining fluency with factual information (see also Polson, 1995). Preliminary research using *Think Fast* and a 20-card deck demonstrated that fluency could be influenced by card order, supporting the SAFMEDS recommendation to shuffle the deck (Polson, Wong, Parsons, & Grabavac, 1991).

Polson et al. (1991) also explored transfer effects. In that study, subjects first learned to type terms when presented with behavioral definitions. Then, subjects completed a pencil-and-paper transfer test, one part of which involved the reversed task of writing the definition when given the term. All

Portions of the data were presented at the 1992 convention of the Association for Behavior Analysis in San Francisco, CA. Requests for reprints of this paper and pilot reports cited in the footnotes should be sent to David Polson, Counselling Services, University of Victoria, P.O. Box 3025, Victoria, British Columbia, V8W 3P2, Canada. (e-mail: polson@islandnet.com)

four subjects performed very poorly on this part of the transfer test. Continued training in the forward direction and repeated testing of the reversed relations generally produced only minor improvements. In sum, being adept at seeing definitions and typing terms did not transfer to the reversed task of seeing terms and writing definitions. Among the variables that may have been responsible for this lack of transfer were the change from a typing to a writing response, size of the stimulus and response items, and practice histories (familiarity) with the terms and definitions.

The present study controlled for response modality and the size of the stimulus and response items. English-speaking subjects with little knowledge of the French language used *Think Fast* to learn 16 English-French word pairs. In Phase 1, half of the relations were taught from French to English (FE-1) and half from English to French (EF-1). Then, in Phase 2, training continued with the stimulus and response items of each word pair reversed. Thus, during Phase 2, the FE-1 cards were trained from English to French (EF-2) and the EF-1 cards were trained from French to English (FE-2). Our purpose was to explore how reversing the stimulus and response items affects performance for word pairs that are both accurate and fluent in the opposite direction, and how the familiarity of the stimulus and response items factors into the equation.

Skinner's (1957) analysis may be helpful here. The relations described above are intraverbals because a verbal stimulus (French or English word) occasions a verbal response (English or French word) without point-to-point correspondence. Each word pair contains two possible intraverbals, one taught in Phase 1 and the other in Phase 2. Consider the word pair *bottle – flacon*. In Phase 1, the subject sees *bottle* (S^D1) and types *flacon* (R1); then, in Phase 2 the subject sees *flacon* (S^D2) and types *bottle* (R2). These are separate intraverbals because they differ with respect to the topography of both the stimulus and the response. Is there any reason to expect transfer from S^D1-R1 to S^D2-R2? Hall & Chase (1991) suggest that there might be. While this particular S^D1-R1 relation may not be formally similar to any other previously trained intraverbal, it is the same on a higher level of abstraction because it shares the defining properties of intraverbals. Given an appropriate learning history, one of those defining properties may come to include an abstract relation between two relations, specifically, that intraverbals "go with" their reversed counterparts. Over time "subjects may come to respond in a consistent manner to this higher-order sameness," and "[a]fter each novel intraverbal is trained, the subject may produce the reversed relation without additional training" (p. 114).

Is there any reason to expect differential transfer based on word familiarity? For the intraverbal relations that require typing the French word, two repertoires seem to be involved: (1) learning the topography of the response, i.e., how to spell the French word, and (2) learning to emit that response in the presence of the corresponding English word. For the intraverbals that require typing the English word, only the latter repertoire seems relevant, since the English words selected for this study were all easy to spell for a native English speaker. We might therefore expect better initial performance for FE cards during both Phase 1 and Phase 2. The failure to show transfer from the FE-1 to EF-2 intraverbals "may result from the unavailability of the stimulus items as responses rather than from the reversibility of the associations" (Catania, 1992, pp. 287-288).

The bi-directional nature of training the French and English word relations in each phase suggests a close parallel to *symmetry* as defined and studied by stimulus equivalence researchers. Specifically, after learning to select Stimulus B in the presence of Stimulus A, the subject selects Stimulus A in the presence of Stimulus B with no additional training. Although this research has been confined to the match-to-sample (MTS) procedure (e.g., Sidman & Tailby, 1982; Sidman, 1986), recently there have been criticisms of that paradigm (Horne & Lowe, 1996) and calls for alternative methodologies (Saunders & Green, 1996). The MTS procedure is an example of what Michael (1985) calls *stimulus selection-based responding*, in which "[t]he unit of verbal behavior can be described as an increased control of the pointing response by a particular stimulus as the result of the presence of a different stimulus (or the strength of a particular establishing operation)" (p. 2). Had we employed a selection-based paradigm, subjects would have been trained to select the English (French) word from a list of choices following the presentation of an French (English) word sample. We chose instead to use a paradigm consistent with what Michael calls *topography-based responding*, in which "[t]he unit of verbal behavior can be described as an increased strength of distinguishable topography given some specific controlling variable" (p. 2). Training was set up so that each French (English) word stimulus occasioned an English (French) word response distinct in form from other English (French) word responses occasioned by other French (English) word stimuli. Hall and Chase (1991) suggest that it is possible to train and test for stimulus equivalence (including symmetry) using a topography-based paradigm and still remain consistent with the mathematical definition of equivalence that was originally applied to conditional discrimination performances by Sidman & Tailby (1982). According to Hall and Chase, a subject can demonstrate equivalence by selecting stimuli *or* producing them, so long as the nature of the task (e.g., typing) remains the same for each of the defining relations. Thus, if a subject is taught to produce (type) Stimulus B (*flacon*) in the presence of Stimulus A (*bottle*), and later produces (types) Stimulus A (*bottle*) in the presence of Stimulus B (*flacon*) with no additional training, then it seems appropriate to define this as *topography-based symmetry* and distinguish it from *selection-based symmetry* that is studied using the MTS procedure.

In the present study, the first trial in Phase 2 is of particular interest. The reason is that feedback was provided throughout Phase 2, and thus this trial was the only one in which feedback had yet to occur for the reversed relations. Correct performance on the first trial in Phase 2 would indicate *emergent* symmetrical responding because "the subject had never experienced the new contingencies before" (Sidman, 1986, p. 229). By considering these data, the present study represents a preliminary attempt to explore topography-based symmetry.

GENERAL METHOD

Subjects

All subjects were English-speaking university students. During the recruitment interview, they were told that the purpose of the research was to study language learning, and they were asked a series of questions that determined their eligibility. Subjects were allowed to continue only if they claimed to have (1) little or no knowledge of the French language and (2) good typing skills. As an additional screening devise, on the day of the experiment, subjects completed a paper-and-pencil pretest that asked for the English equivalent of each of the French words to be used in the experiment (see Table 1). Any subject who exceeded two correct responses was to be excluded from the experiment. None of nine participants in the three experiments scored more than one correct on this pretest.

Setting and Apparatus

All experiments were conducted in a sound-attenuating chamber that contained a workstation consisting of a table and two chairs. A Zenith Data Systems computer (Model 2F-158-42) was situated on the table. The computer ran a version of *Think Fast* that had been modified for experimental purposes. Stimulus and response events were timed and recorded to the nearest .05 s.

Each trial began with the program "shuffling" a deck of cards. Pressing the space bar started the timing of the trial and displayed the first card. Each card was comprised of a stimulus word and a response word. The stimulus word appeared in a

Table 1

English and French word pairs used throughout the study.

English	French
address	discours
harvest	moisson
hideout	cachette
freckles	taches
hammer	marteau
eyebrow	sourcil
peelings	pelures
puppet	pantin
bottle	flacon
hilltop	sommet
riddle	crible
rascal	coquin
cradle	berceau
bubble	bouillon
hallway	couloir
donkey	baudet

box in the middle of the screen. The task was to type the corresponding response word in another box centered near the bottom of the screen. A flashing cursor was displayed at the far left of this box; otherwise the box was empty. The last line of the screen prompted the subject to type the term, press the Tab key for a hint, or press the Enter key to skip that card. A correct keystroke or Tab key press displayed the appropriate character at the cursor position and then advanced the cursor to the position of the next letter in the response word. An incorrect keystroke produced only a low-sounding (200 Hz, 0.055 s) beep; however, five incorrect keystrokes at any one cursor position displayed the appropriate character and advanced the cursor. When the last letter of the response word was displayed, a response was deemed correct or incorrect. A correct response displayed the message "Good!" for 0.33 s in place of the prompt and sounded a high rising (940-1000 Hz) beep; an incorrect response displayed the message "Wrong, too many errors!" for 0.33 s in place of the prompt and sounded a low falling (160-100 Hz) beep. If the Tab key had been pressed at any cursor position, the response was considered to be incorrect, the message "Wrong, needed a hint!" was displayed for 0.33 s in place of the prompt, and the low falling beep sounded. Pressing the Enter key at any cursor position immediately displayed for 0.33 s the whole response word in the box along with the message "Skip!" in place of the prompt, and sounded the low falling beep; in such a case, the response was considered incorrect. Following each of these possibilities, the screen cleared and the next card was presented. When the last card in the deck was completed, the trial ended and a summary screen of performance scores for that trial was displayed, which included the trial duration, cards completed (always 16), correct cards per min, incorrect cards per min, and the percent correct (accuracy). The displayed rate measures were calculated by counting the cards scored correct (incorrect) in that trial, dividing by the total trial time, and then converting to correct (incorrect) cards per min.

With the *Strict Spelling* option "on" (Experiments 1 and 2), a response to a card was deemed correct only if no spelling error occurred. With the *Strict Spelling* option "off" (Experiment 3), a card was counted as correct if the first character was typed without error and no more than three errors occurred for the remaining characters.

All training trials were conducted using a 16-card deck consisting of English and French word pairs (see Table 1). English words consisted of 6-8 letter, two-syllable nouns that had a "meaningfulness score" of 80-95% for a native English speaker at the Grade 4 level of education (Dale & O'Rourke, 1976). The French words – a rough translation of their English counterparts – also consisted of 6-8 letters, and all but one (*taches*) had two syllables when pronounced correctly.

Procedure

To acquaint subjects with the program, the experimenter helped them work through a few trials with a four-card sample deck consisting of English-English word pairs (different from the English words in Table 1). Following this tutorial session, the experimenter explained that the subject's task was to learn a deck of 16 English and French word pairs by typing one item of the pair when presented with the other. Subjects were instructed to respond as accurately and as quickly as possible and to continue initiating 16-card trials until a beep sounded and the message "You have finished the session! Please notify the experimenter" appeared on the screen. If the scheduled session duration expired while a subject was responding within a trial, this signal was not presented until the end of that trial. The number of trials per session varied according to how quickly subjects worked through each 16-card trial and how long they paused between trials. Subjects completed all sessions within a 3-hour time block, with 5-min rest periods between sessions. While a subject was engaged with *Think Fast*, the experimenter waited in a room outside the chamber; the beep at the end of each session prompted her to enter the chamber and prepare the computer for the next session.

EXPERIMENT 1

Subjects

Three, fourth-year undergraduate students from the University of Victoria participated. All subjects claimed to have a minimum typing speed of 30 words per min and to have not advanced beyond the level of high school French.

Procedure

In Phase 1, half of the word pairs in Table 1 were practiced exclusively with the French word as the stimulus item and the English word as the response item (FE-1) and the other half with the English word as the stimulus item and the French word as the response item (EF-1). There were two training sessions, 40-min and 30-min, respectively. This was followed by Phase 2, consisting of a 20-min session in which stimulus and response items for each word pair were reversed. Thus, the FE-1 cards were trained with the English word as the stimulus item and the French word as the response item (EF-2), and EF-1 cards were trained with the French word as the stimulus item and the English word as the response item (FE-2). Subjects were not informed of this change. Throughout both phases, the *Strict Spelling* option was "on."

Independent and Dependent Variables

The independent variable was the direction of training (French to English or English to French) during Phase 1 (FE-1 and EF-1) and during Phase 2 when the stimulus and response items for each word pair in Phase 1 were reversed (EF-2 and FE-2). Thus, there were four conditions (card types), determined by the training mode and the phase. Two dependent variables were considered for each trial. First, as a measure of fluency, a rate correct score for each of the two card types in each of the two phases was calculated. This was determined by dividing the sum of the latencies for all eight cards in a condition into the number of cards scored correct in that condition, and then converting this figure into correct responses per min. (If no cards were scored correctly, then this datum was considered to be zero). Latency was defined as the time from the onset of the stimulus word to the time of the input of the last keystroke prior to the presentation of the next card or the end of the trial. The second dependent variable consisted of an accuracy score for each of the two card types in each of the two phases, calculated by dividing the number of cards scored correct in a condition by the number of cards in that condition (eight) and multiplying by 100.

RESULTS AND DISCUSSION

Figure 1 shows rate correct scores for the two cards types in each phase across trials for the three subjects. In this and all subsequent graphs to be presented, open triangles represent the card set trained from French to English in Phase 1 (FE-1) and closed triangles represent the same card set with the stimulus and response items reversed in Phase 2 (EF-2); closed rectangles represent the card set trained from English to French in Phase 1 (EF-1) and open rectangles represent that card set with the stimulus and response items reversed in Phase 2 (FE-2). For all subjects, rate correct scores improved across Phase 1, the scores being noticeably higher for the FE-1 cards during the early trials. For S1 and S2, rate correct scores leveled off, and comparable rates between the two card types were observed during later trials of Phase 1. For S3, rate correct continued to improve throughout Phase 1, with higher rates observed for the EF-1 cards during later trials. The effects of reversing the stimulus and response words in Phase 2 were dramatic for all subjects: rate correct plummeted for both card types, especially for the EF-2 cards. Rate correct improved for both card types across Phase 2, with sustained lower rates all through the phase for the EF-2 cards for two of the three subjects (S2 and S3).

Figure 2 shows accuracy scores for the two card types in each phase across trials for the three subjects. For S1 and S2, accuracy rapidly increased and scores of 100% correct for both card types were observed within the first session of Phase 1. Note that rate correct continued to increase past the point at which these perfect scores were first obtained (see Figure 1). Improvements in accuracy proceeded more slowly for S3: this subject scored 100% correct for the EF-1 cards, but only late in Phase 1, and never scored higher than 75% correct for the FE-1 cards. For all subjects, reversing the stimulus and response items in Phase 2 significantly reduced accuracy scores, especially for the EF-2 cards.

To reveal possible differences in initial acquisition of the four card types, Figure 3 displays cumulative correct responses across trials for the two card types in Session 1 of Phase 1 superimposed upon cumulative correct responses across trials for the two card types in the Phase 2 session. Note that the steeper the curve, the more rapid the acquisition. The dashed diagonal line shows the maximum acquisition rate (i.e., perfect performance across all trials). For each subject, the two reversed card types in Phase 2 (FE-2 and EF-2) were acquired more quickly than their counterparts in Phase 1 (EF-1 and FE-1). Within Phase 1, acquisition for S1 and S3 was more rapid for the FE-1 cards than for the EF-1 cards (although Figures 1 and 2 reveal superior performance for the EF-1 cards later in Phase 1 for S3); there was little difference between the two card types for S2. Within Phase 2, acquisition for all subjects was more rapid for the FE-2 cards than for the EF-2 cards.

Table 2 shows the subjects' accuracy scores for the two card types on the first trial of Phase 2, our measure of emergent topography-based symmetrical performance (symmetry test trial). Perfect symmetry would be indicated by scores of 100% correct. Accuracy was very low for the EF-2 cards, and while the scores were somewhat better for the FE-2 cards, no subject did better than five out of eight correct.

Overall, the data suggest that performance deteriorates when the stimulus and response items of intraverbals learned to accurate and fast (fluent) levels are reversed, the effect being greater when the reversed intraverbal involves an unfamiliar response. The results of the symmetry test trial contrast with the rapid emergence of selection-based symmetry frequently reported in the equivalence literature (see GENERAL DISCUSSION).

Table 2

Percent correct scores on first trial in Phase 2 for cards trained from French to English (FE-2) and cards trained from English to French (EF-2).

Subject	FE-2 Cards	EF-2 Cards	Average
S1	62.50	25.00	43.75
S2	62.50	0.00	31.25
S3	25.00	12.50	18.75

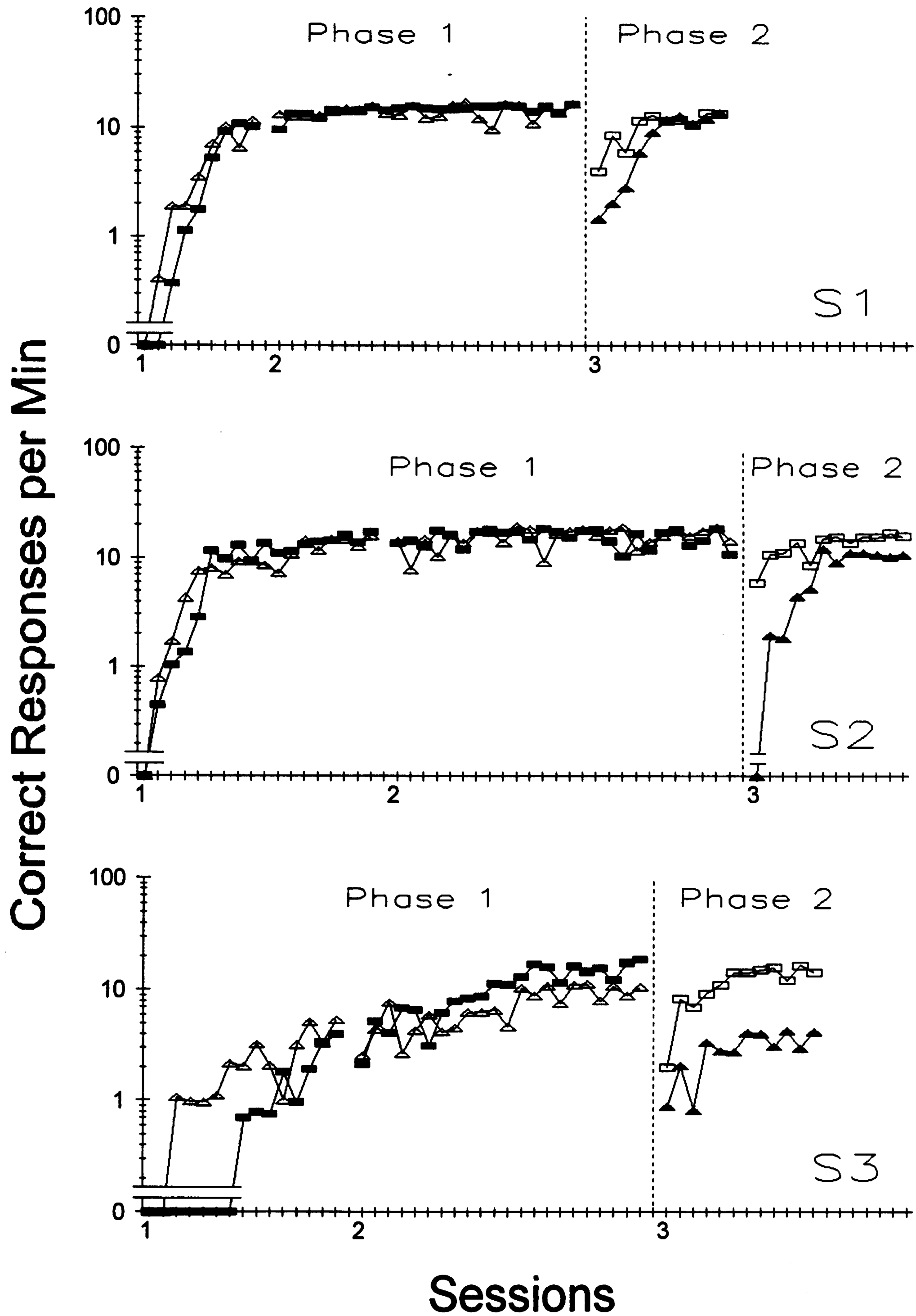

Fig. 1. Correct responses per min across trials for cards in Phase 1 trained from French to English (FE-1: open triangles) and from English to French (EF-1: filled rectangles) and for their reversed counterparts trained in Phase 2 (EF-2: closed triangles; FE-2: open rectangles). Graphs are presented for all subjects in Experiment 1. Sessions are numbered on the horizontal axis and session breaks are indicated by non-joined data points. The two phases are separated by the broken vertical line.

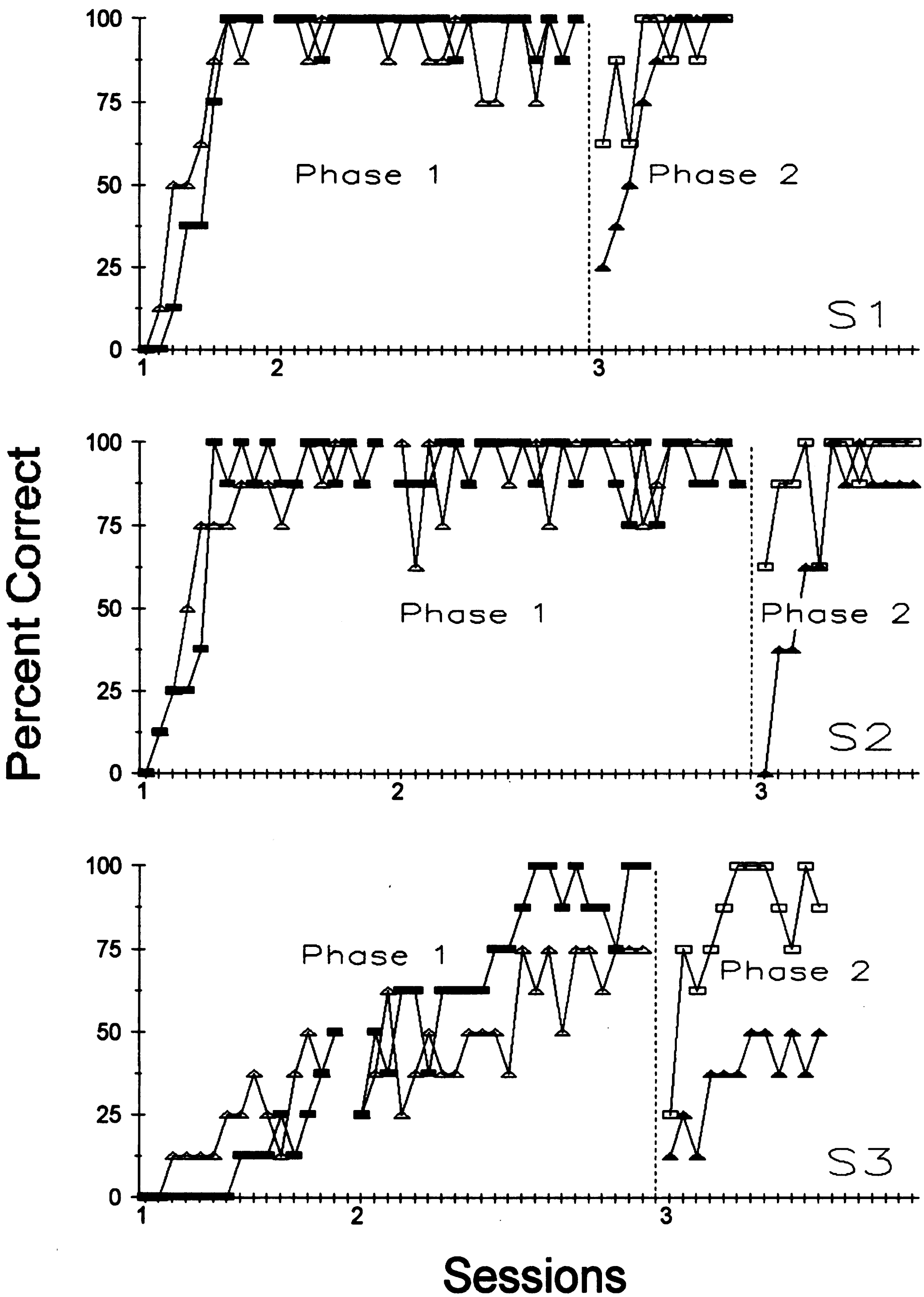

Fig. 2. Percent correct scores across trials for cards in Phase 1 trained from French to English (FE-1: open triangles) and from English to French (EF-1: filled rectangles) and for their reversed counterparts trained in Phase 2 (EF-2: closed triangles; FE-2: open rectangles). Graphs are presented for all subjects in Experiment 1. Other details as in Figure 1.

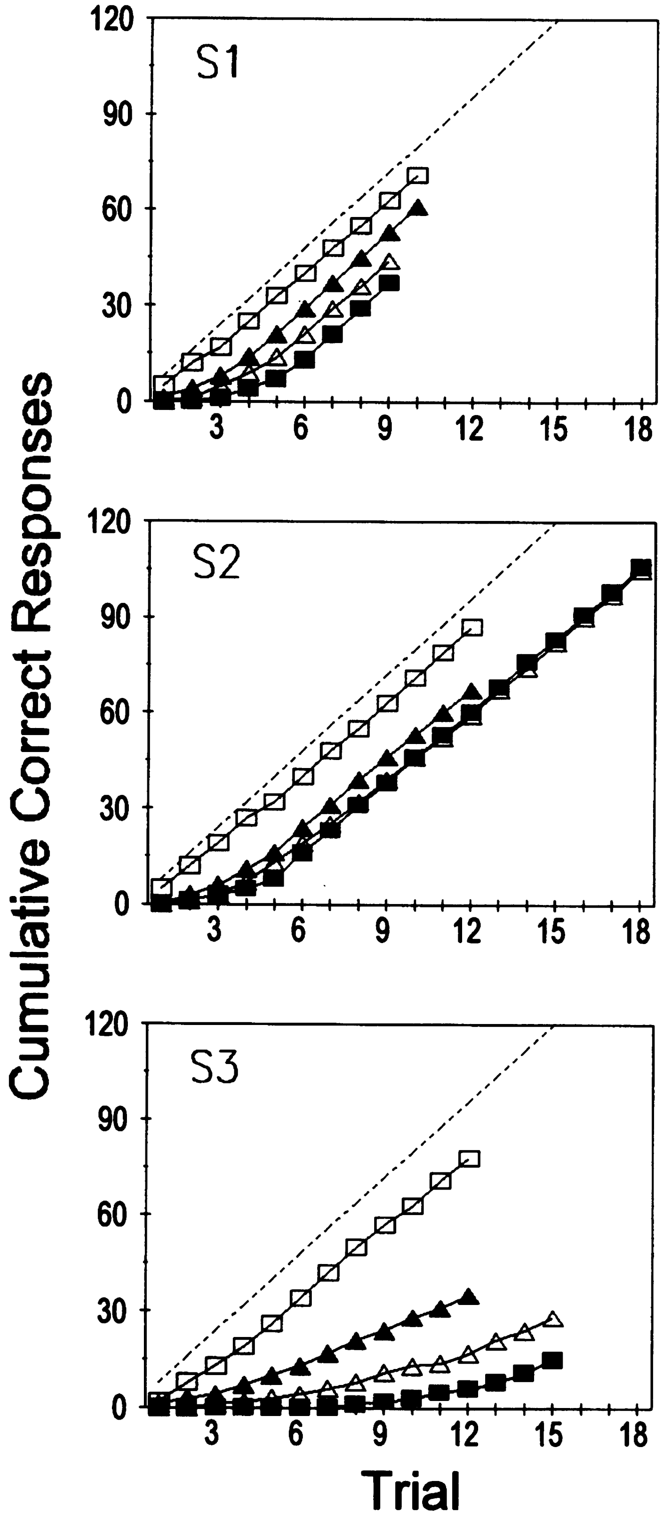

Fig. 3. Cumulative correct responses across trials in Session 1 of Phase 1 for cards trained from French to English (FE-1: open triangles) and from English to French (EF-1: filled rectangles) superimposed upon cumulative correct responses across trials in Phase 2 for their reversed counterparts (EF-2: closed triangles; FE-2: open rectangles). Graphs are presented for all subjects in Experiment 1. Broken diagonal line shows maximum rate, i.e., 100% correct for all trials within a session.

EXPERIMENT 2

One weakness of Experiment 1 was that the cards contained in each condition were the same for all subjects. Thus, it is possible that the unique properties of the cards in a particular condition, rather than how they were trained, were responsible for the results. In Experiment 2, cards were randomly assigned to the two card sets, the restriction being that there were eight cards in each set. Another procedural alteration concerned the length of the sessions: To reduce the possibility of within-session fatigue, the duration of each session was shortened.

Subjects

The subjects were three, first- and second-year undergraduate students from the University of Victoria. They were selected from the Psychology 100 subject pool and were compensated with course credit for three hours of participation. All subjects claimed to have a minimum typing speed of 30 words per min and to not have advanced beyond the level of high school French.

Procedure

The procedural details were identical to Experiment 1, except for the following changes. First, as stated above, the eight word pairs assigned to the two card sets varied among subjects. Second, in Phase 1, three, 25-min training sessions were scheduled; however, because fluctuations in accuracy were still apparent for all subjects after this period, a fourth, 15-min training session was added. This was followed by Phase 2, consisting of another 15-min session, this time with the stimulus and response items reversed.

RESULTS AND DISCUSSION

The rate correct scores shown in Figure 4 are comparable to Experiment 1. There was rapid improvement early in Phase 1. Rate correct scores were sometimes noticeably higher for the FE-1 cards during Phase 1 for S5, while there was little difference between the two cards types for S4 and S6. By the end of Phase 1, the scores leveled off for S4 and S6, and continued to improve for S5. Reversing the stimulus and response elements in Phase 2 produced an abrupt reduction for all subjects, which was more pronounced for the EF-2 cards. Rate correct scores improved across Phase 2, with the exception of the EF-2 cards for S4 and the FE-2 cards for S5. By the end of Phase 2, comparable scores between the two card types were observed for S5 and S6.

The accuracy scores presented in Figure 5 are also similar to Experiment 1. All subjects learned both card types to 100% correct, with minor fluctuations as training continued. To varying degrees, rate correct continued to improve past the point at which these perfect performances were first obtained (see Figure 4). When the stimulus and response words for each card were reversed in Phase 2, accuracy dropped to significantly lower levels, and this effect was clearly more evident for the EF-2 cards for S4 and S5.

The cumulative accuracy scores presented in Figure 6 reveal that the reversed FE-2 cards were acquired more quickly than their Phase 1 counterparts (EF-1) for all three subjects, but unlike Experiment 1, acquisition of the EF-2 cards was superior to the FE-1 cards for only one of the three subjects (S6). Within Phase 1, similar acquisition rates between the FE-1 cards and the EF-1 cards were observed for all subjects. Within Phase 2, all subjects demonstrated more rapid acquisition for the FE-2 cards than for the EF-2 cards, a replication of Experiment 1.

Table 3 shows the three subjects' accuracy scores on the symmetry test trial. Concordant with Experiment 1, when a difference was observed (S4 and S5), topography-based symmetry favored the card type that involved typing a familiar response item (FE-2); and, regardless of familiarity, no subject scored higher than five out of eight correct.

In sum, this experiment replicates the earlier one in revealing (1) large performance decrements when the stimulus and response items of intraverbals learned to accurate and fast (fluent) levels are

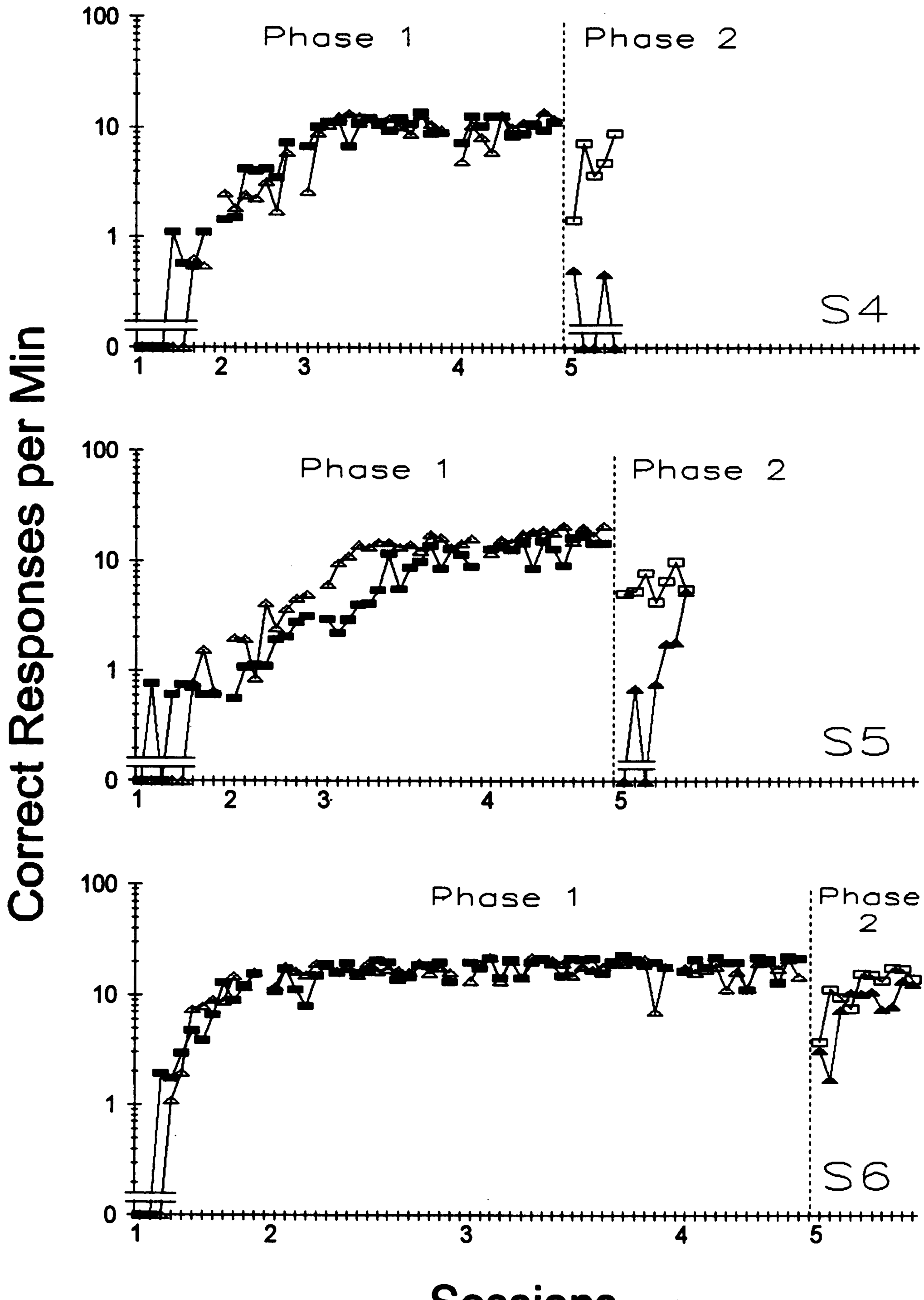

Fig. 4. Correct responses per min across trials for cards in Phase 1 trained from French to English (FE-1: open triangles) and from English to French (EF-1: filled rectangles) and for their reversed counterparts trained in Phase 2 (EF-2: closed triangles; FE-2: open rectangles). Graphs are presented for all subjects in Experiment 2. Other details as in Figure 1.

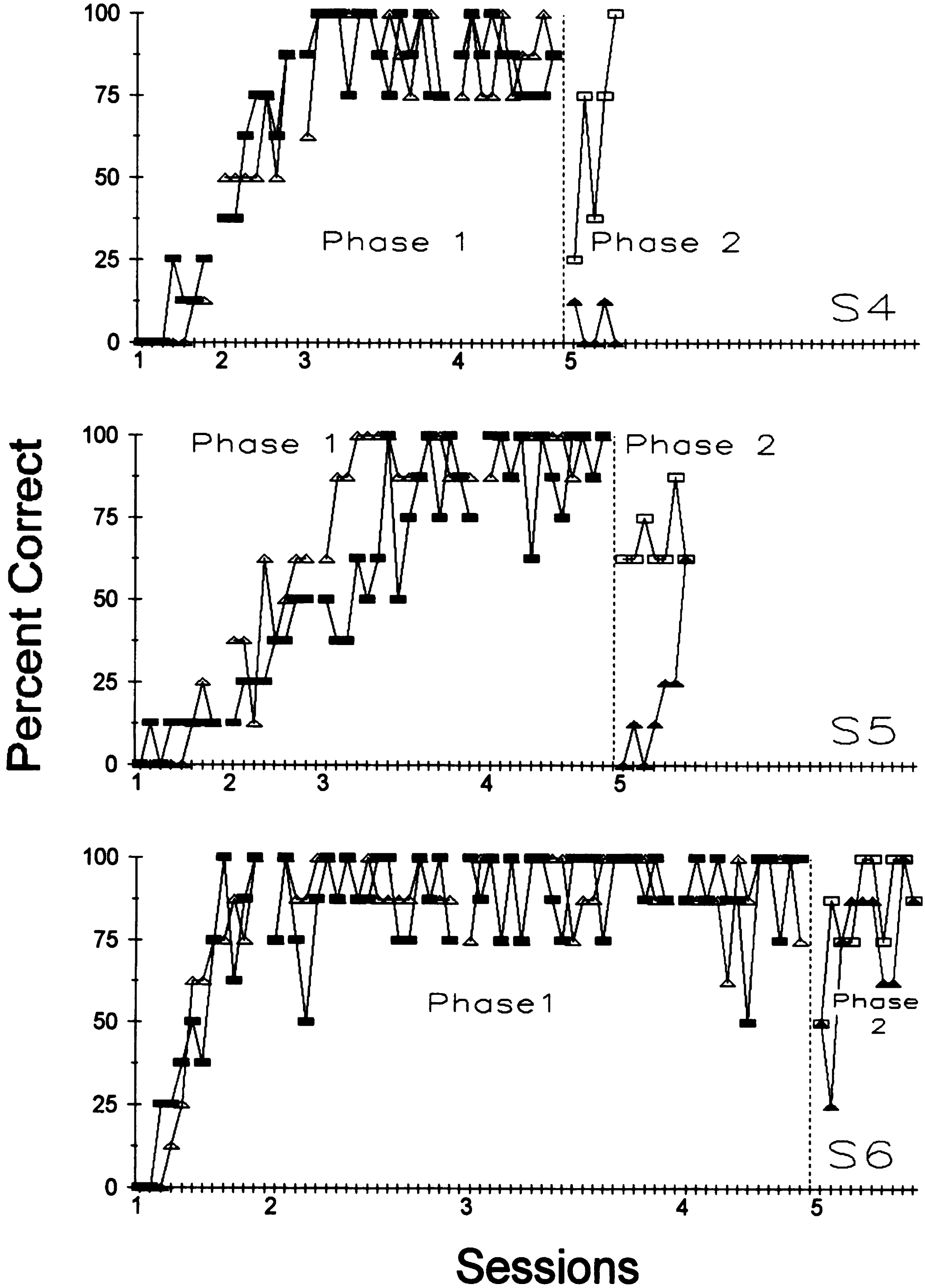

Fig. 5. Percent correct scores across trials for cards in Phase 1 trained from French to English (FE-1: open triangles) and from English to French (EF-1: filled rectangles) and for their reversed counterparts trained in Phase 2 (EF-2: closed triangles; FE-2: open rectangles). Graphs are presented for all subjects in Experiment 2. Other details as in Figure 1.

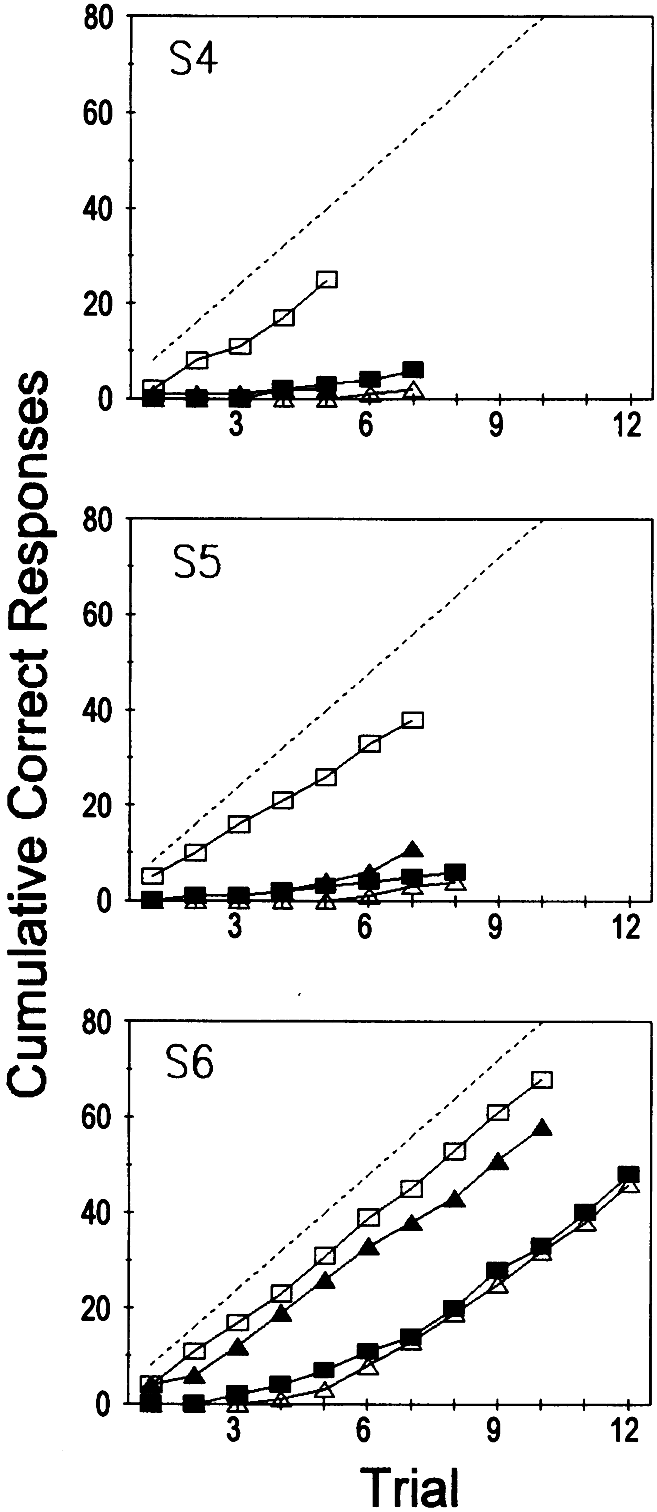

Fig. 6. Cumulative correct responses across trials in Session 1 of Phase 1 for cards trained from French to English (FE-1: open triangles) and from English to French (EF-1: filled rectangles) superimposed upon cumulative correct responses across trials in Phase 2 for their reversed counterparts (EF-2: closed triangles; FE-2: open rectangles). Graphs are presented for all subjects in Experiment 2. Other details as in Figure 3.

Table 3

Percent correct scores on first trial in Phase 2 for cards trained from French to English (FE-2) and cards trained from English to French (EF-2).

Subject	FE-2 Cards	EF-2 Cards	Average
S4	25.00	12.50	18.75
S5	62.50	0.00	31.25
S6	50.00	50.00	50.00

reversed, (2) greater decrements for intraverbals involving an unfamiliar response after the reversal, and (3) low scores on the symmetry test trial.

EXPERIMENT 3

In the first two experiments, the *Strict Spelling* option was "on." Thus, typing even one incorrect keystroke resulted in the response for that card being scored as incorrect. Consequently, approximations to correct responses were counted as mistakes. Consider the word pair *hallway – couloir*. A plausible sequence of events could be as follows. Upon seeing the stimulus *hallway*, a subject might covertly pronounce "*couloir*." This new verbal stimulus occasions a typing response, a relation known as taking dictation (Skinner, 1957). The subject then attempts to type *coloir*; when the cursor does not advance after the "l" keystroke, this may prompt typing "u," followed by the rest of the word *couloir* typed accurately. Perhaps subjects may have been able to correctly say and approximate the spelling of some of the response words immediately following the reversal, but our measure was insensitive to this fact. To eliminate the stringent requirements of the *Strict Spelling* option, we turned it off for this experiment.

Subjects

Three, first- and second-year undergraduate students from the Psychology 100 subject pool at the University of Victoria participated and were compensated with course credit for three hours of their time. All subjects claimed to have a minimum typing speed of 30 words per min and not to have advanced beyond high school French.

Procedure

The procedural details were identical to Experiment 2, except the *Strict Spelling* option was set to "off." Thus, a response word was scored correct if there was no error on the first character and three or fewer incorrect keystrokes on the remaining characters.

Results and Discussion

The data presented in Figures 7 and 8 are consistent with the previous two experiments in showing that reversing the stimulus and response items in Phase 2 dramatically reduced both rate correct and accuracy (S9 excluded). However, differences between the FE-2 cards and the EF-2 cards were not as readily apparent; improvements in both rate correct and accuracy proceeded rapidly for both card types. The data for S9 stand out from all other eight subjects: although his rate correct scores were reduced early in Phase 2, there was little evidence of this for his accuracy scores.

The cumulative accuracy scores displayed in Figure 9 show faster learning for the two reversed card types in Phase 2 (FE-2 and EF-2) than for their counterparts in Phase 1 (EF-1 and FE-1). Within Phase 1, acquisition was unaffected by card type for S7 and S9; it is, however, clearly more rapid for the EF-1 cards than for the FE-1 cards for S8, a result unseen for any subject in the previous two experiments. Within Phase 2, acquisition rates were similar for the FE-2 and EF-2 cards, in contrast to all six subjects in the previous two experiments who clearly showed more rapid acquisition for the FE-2 cards. The more lenient "correct" criterion appeared to facilitate accuracy scores for the EF-2 cards.

Table 4 shows the three subjects' accuracy scores on the symmetry test trial. Unlike the previous two experiments, topography-based symmetry was inconsistently affected by the familiarity of the

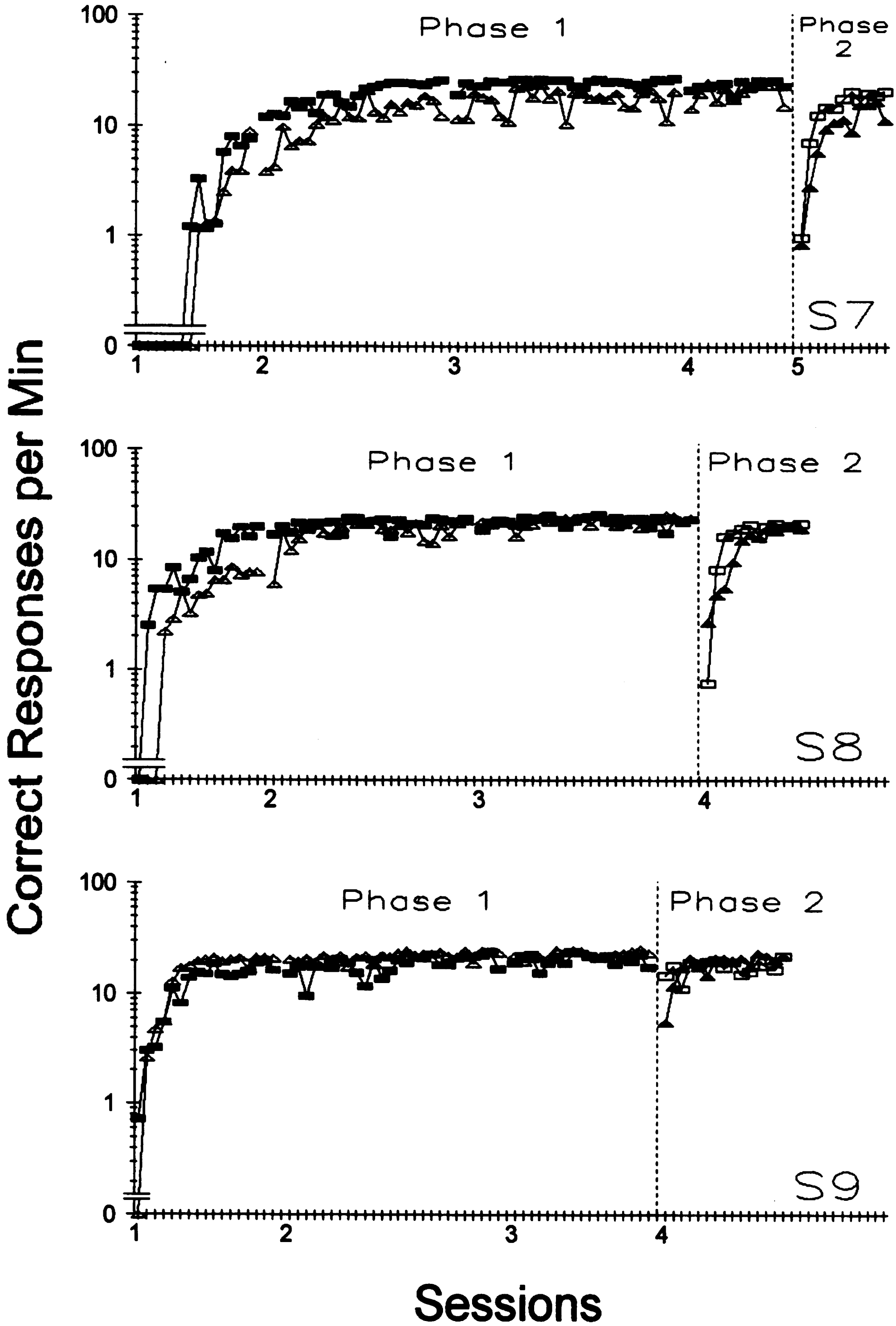

Fig. 7. Correct responses per min across trials for cards in Phase 1 trained from French to English (FE-1: open triangles) and from English to French (EF-1: filled rectangles) and for their reversed counterparts trained in Phase 2 (EF-2: closed triangles; FE-2: open rectangles). Graphs are presented for all subjects in Experiment 3. Other details as in Figure 1.

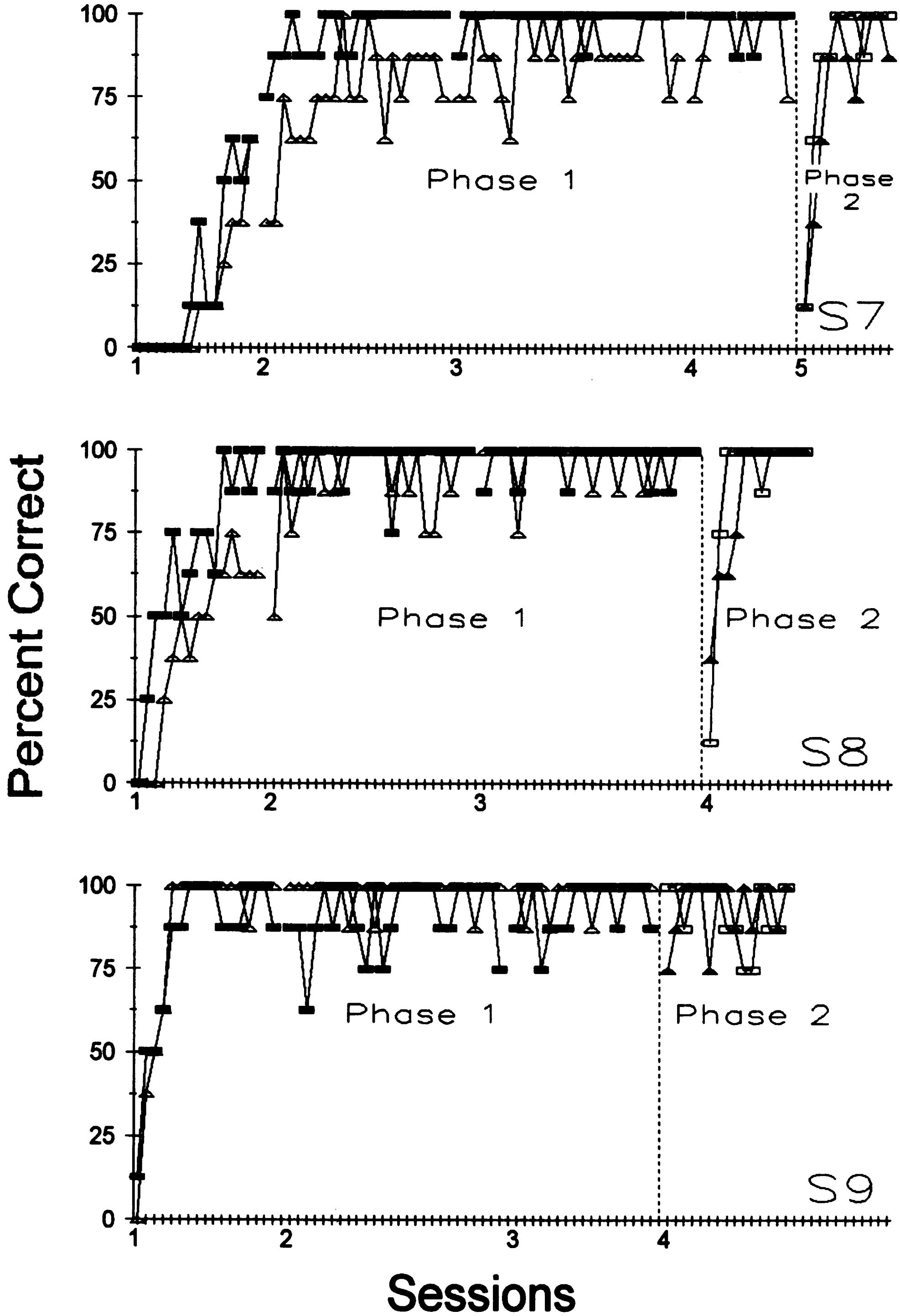

Fig. 8. Percent correct scores across trials for cards in Phase 1 trained from French to English (FE-1: open triangles) and from English to French (EF-1: filled rectangles) and for their reversed counterparts trained in Phase 2 (EF-2: closed triangles; FE-2: open rectangles). Graphs are presented for all subjects in Experiment 3. Other details as in Figure 1.

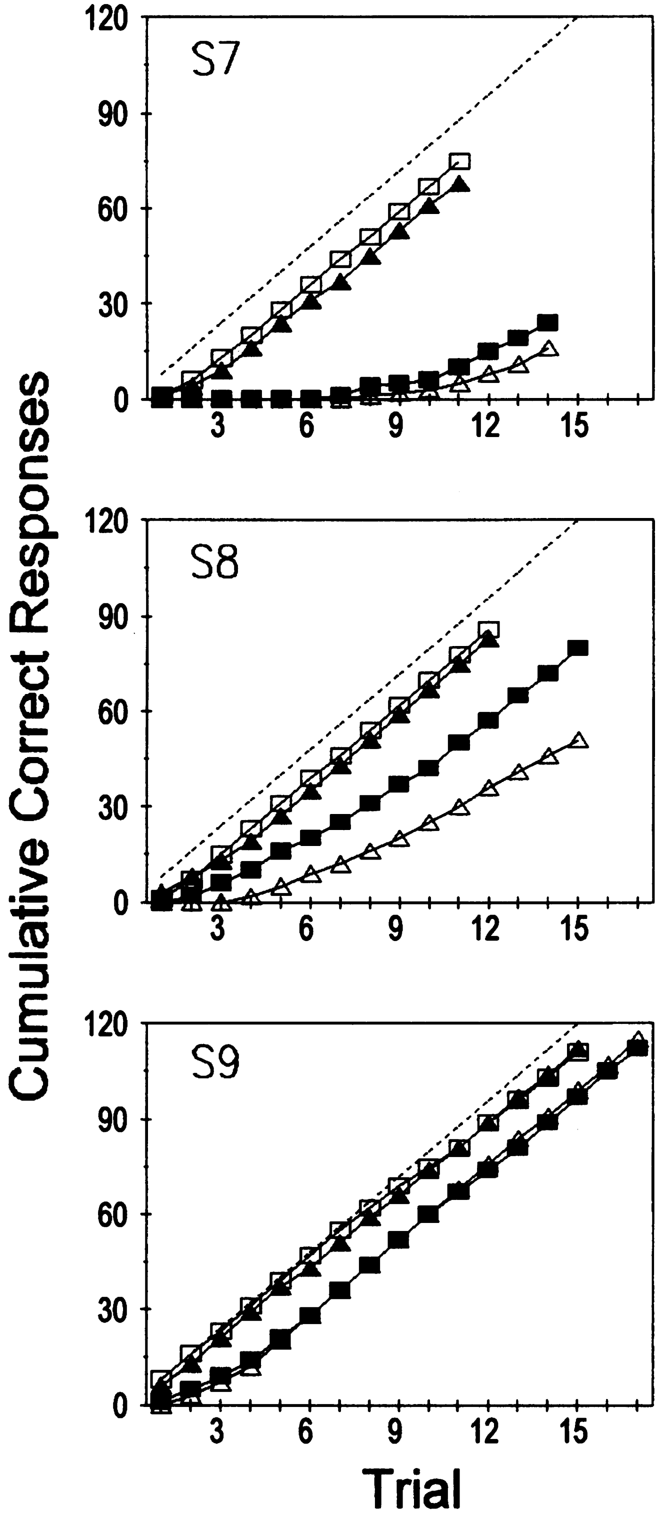

Fig. 9. Cumulative correct responses across trials in Session 1 of Phase 1 for cards trained from French to English (FE-1: open triangles) and from English to French (EF-1: filled rectangles) superimposed upon cumulative correct responses across trials in Phase 2 for their reversed counterparts (EF-2: closed triangles; FE-2: open rectangles). Graphs are presented for all subjects in Experiment 3. Other details as in Figure 3.

response item. Two of the three subjects (S7 and S8) performed poorly whether the response was typing the English word or the French word; the remaining subject scored considerably better, and demonstrated perfect symmetry for the FE-2 cards. Overall, with the exception of S9, these data replicate the previous two experiments in showing that topography-based symmetry occurred, but not reliably.

Table 4

Percent correct scores on first trial in Phase 2 for cards trained from French to English (FE-2) and cards trained from English to French (EF-2).

Subject	FE-2 Cards	EF-2 Cards	Average
S7	12.50	12.50	12.50
S8	12.50	37.50	25.00
S9	100.00	75.00	87.50

GENERAL DISCUSSION

Using a textual stimulus and a typing response, the first two experiments showed that reversing the stimulus and response items of intraverbals learned to highly accurate and fast (fluent) levels disrupted that performance. In addition, the amount of disruption was a function of the familiarity of the stimulus and response items: performance was poorer after the reversal when the response was a word unlikely to have ever been previously typed by the subject (EF-2 cards) than if the response was an easy-to-spell English word (FE-2 cards). In Experiment 3, when the *Strict Spelling* option was turned off, and thus the criterion for a "correct" response was more lenient, performance improved for the EF-2 cards and familiarity was no longer a factor.

This latter result suggests that subjects in the earlier two experiments may have been typing approximations to the response words in the EF-2 condition that were scored as incorrect. Unfortunately, *Think Fast* records only the outcome (i.e., "right" or "wrong"), not the behavior itself (i.e., what keys were pressed). Future research could develop and employ more sophisticated software that keeps track of all keystrokes, permitting a more detailed analysis of the "correctness" of subjects' responses. This result also suggests that subjects may have been able to say but not type some of the EF-2 response words. (For an explanation of how vocalizing may play a role, see the *Symmetry* section below.) To assess this possibility, subsequent research in this area could require subjects to both say and type responses following the reversal.

Transfer

The data showed that acquisition generally proceeded more rapidly for the two reversed card types in Phase 2 than for their counterparts in Phase 1. However, the extent to which this effect was due to practice *per se* is unknown without appropriate control conditions. To eliminate practice as a confounding variable, future research might expose a second group of subjects to the same Phase 1 conditions, but then use an entirely different set of English and French words in Phase 2. If our results were replicated for the experimental subjects and performances did not differ between Phase 1 and Phase 2 for control subjects, then concluding that prior training of intraverbals to fluency transferred and resulted in a savings of learning for their reversed counterparts would be justified.

Symmetry

Considering the first trial in Phase 2 to be a test of topography-based symmetry, our data demonstrated emergence for 32% [31/96] of the relations in Experiments 1 and 2; in addition, symmetry was differentially affected by the type of relation trained, with 48% [23/48] correct for FE-2 cards and 17% [8/48] correct for EF-2 cards. If we assume that the French words were less pronounceable than the English words, then this latter result is in accord with Mandell & Sheen (1994) who showed that equivalence class formation in a MTS paradigm was better when phonologically correct words were used as samples during training as opposed to non-phonological words. With the less stringent criterion for a "correct" response in Experiment 3, two of the three subjects still demonstrated low levels of symmetry (19% [6/32] correct), although the difference between the two card types was less apparent (13% [2/16] correct for the FE-2 cards, 25% [4/16] correct for the EF-2 cards). Perfect symmetry was shown by only one subject (S9) for only one card type (FE-2 cards). In addition to these accuracy measures, rate correct scores dropped for all subjects on the symmetry test trial, suggesting that fluency, like accuracy, was not symmetrical (see also Spencer & Chase, 1996).

Overall, symmetry was not as easily obtained in our topography-based paradigm as has been reported in the typical selection-based paradigm (e.g., Dougher, Augustson, Markam, Greenway, & Wulfert, 1994; Lane & Critchfield, 1996; Lynch & Cuvo, 1995; Mandell & Sheen, 1994). (It is, however, in line with reports of asymmetrical performances in the paired-associate literature [Nelson, 1972]). This observation is consistent with the growing evidence indicating that it is important to make a distinction between selection-based responding and topography-based responding (Lowenkron, 1991; Michael, 1985; Shafer, 1993; Sundberg & Sundberg, 1990; Wraikat, Sundberg, & Michael, 1991).

While performance was generally poor on the symmetry test trial, some correct responses did occur. Is it possible to make sense of this emergent performance in terms of known behavioral principles? A mechanism described by Horne & Lowe (1996, pp. 219-220) to account for selection-based symmetry hints at one possible explanation. Relevant to their account is a study by Lowe & Beasty (1987), in which children were taught to match a green comparison stimulus (B1) to a vertical line sample (A1) and a red comparison stimulus (B2) to a horizontal line sample (A2). During training, when presented with, say, the vertical line sample, children frequently self-echoically repeated phrases such as "up green up green up green," thereby establishing a bidirectional relation between the verbal responses of "up" and "green." Horne & Lowe suggest that upon hearing herself say "green," the child then engages in appropriate listener behavior by searching for and then selecting the green comparison stimulus. (The child might also search for the vertical line as a comparison after hearing herself say "up" as part of the self-echoic chain, but it is unavailable as a choice.) When the task is then reversed and the child sees the green stimulus as the sample for the first time, she says "green," thus occasioning the response "up," which in turn leads to the appropriate listener behavior of searching for and selecting the vertical comparison stimulus. (Here, the child might also search for the green stimulus as a comparison after saying "green," but it is not presented as a choice). Horne & Lowe speculate that the child named the sample and comparison, in this case "up" and "green" respectively, because "this behavior controls correct responding on the match-to-sample task (i.e., enables the child to "remember" what goes with what) and is thus reinforced through the experimental contingencies" (p. 220). In other words, these tact responses serve a precurrent function (cf., Polson & Parsons, 1994; Parsons, Taylor, & Joyce, 1981).

In the present study, a similar plausible sequence of events could be as follows. Consider the word pair *couloir – hallway*. Initially, upon seeing the stimulus *couloir*, the subject pronounces it (textual behavior: Skinner, 1957) and then types *hallway*; seeing the *hallway* stimulus briefly displayed on the screen, the subject engages in textual behavior with respect to it. With extended practice, the subject comes to say "*hallway*" even before typing it, saying becoming a more fluent response than typing. This would result in the subject typing the response *hallway* in the presence of her own self-generated stimulus "*hallway*," which would be considered taking dictation (Skinner, 1957). Eventually, upon seeing the stimulus *couloir*, the subject says *"couloir hallway,"* sometimes repeating it before typing *hallway*, thus generating a self-echoic chain. In this way, the subject learns the intraverbal response *"couloir"* in the presence of the stimulus *"hallway."* When the task is reversed, the subject sees the stimulus *hallway*, emits the textual response *"hallway,"* which in turn occasions the previously learned intraverbal response *"couloir."* At this point the subject engages in taking dictation and attempts to type the response *couloir*.

Interpretative accounts like the one described above have appeal because rather than assuming symmetry as a given, in which case we would expect all subjects to show minimal disruption when the stimulus and response items of fluent intraverbals are reversed, they suggest testable hypotheses of why some subjects (S9) performed considerably better after the reversal than others (S1 through S8). We would expect that those subjects with a history of reinforcement for self-echoic chaining (or other behaviors which serve the same precurrent function) in similar contexts will excel relative to subjects without that history. Future research could require self-echoic chaining or not to determine whether this does in fact facilitate performance (for affirmative evidence, see Horne & Lowe, 1996, p. 225). If the relation between two intraverbals with reversed stimulus and response items depends upon mediated responding, then this is yet another reason why we might look to even simpler experimental preparations involving nonverbal behavior to discover potentially unique properties of precurrent activities (Polson & Parsons, 1994).

It is tempting to speculate that the final stage of the process outlined above, taking dictation, is what differentiated the FE-2 and EF-2 performances. That is, following the reversal, subjects had to learn to spell the French word responses of the EF-2 cards, while spelling the English word responses of the FE-2 cards was already well-established in their repertories. However, if spelling unfamiliar French words was an impediment in Phase 2, then we would also expect it to be an impediment in Phase 1. This is not consistent with our results. Familiarity did not differentially affect learning prior to the reversal, at least not consistently, i.e., subjects varied as to whether the FE-1 cards were more easily acquired than the EF-1 cards.

Topography-Based Responding

One issue that deserves further comment is whether the current paradigm, which involved subjects typing words, is truly "topography-based." We argue that, at two ends of a continuum, good typists engage in topography-based responding and poor typists engage in selection-based responding.

The present study screened for good typists. Another possibility would have been to require subjects to say the answers, clearly topography-based responding (Michael, 1985). However, our pilot research revealed that manually recording and providing feedback for fluent vocal responses was not an easy task; thus, the automated typing mode was employed for convenience. Poor typists were excluded because they hunt and peck on the keyboard. In other words, given a certain visual stimulus (e.g., the stimulus word or the partially typed response word), they scan the array of letters on the keyboard, and only when the sight of a particular character is encountered do they then press it; this is consistent with Michael's (1985) description of selection-based responding. For each letter in a response word, there is a corresponding stimulus-response unit, the combination of which resembles stimulus-response chaining. For good typists, however, the operant appears to be different, larger, such that typing the entire response word given the stimulus word is a unit onto itself, one that is topographically distinct from responses typed to other stimulus words and one that does not require within-word intervening stimuli in order to be completed. As others have noted (Catania, 1992, pp. 123-125; Lindsley, 1996b), not all response sequences are instances of stimulus-response chaining, especially when the response components are emitted fluently. The rate data show that subjects were typically responding very quickly, too quickly to be scanning and pointing, suggesting that responding was, in fact, topography-based.

Fluency

By considering both rate and accuracy data, we were able to show that rate correct continued to increase after scores of 100% correct had been obtained. Thus, there was room for improvement beyond "perfect" performance, which may be an important consideration for facilitating transfer as well as retention and endurance (Binder, 1993, 1996; Dougherty & Johnson, 1996; Johnson & Layng, 1992). Continued research in this area might train intraverbals beyond 100% accuracy to various levels of fluency and explore potential differential effects when the stimulus and response items of those intraverbals are reversed.

It is not unreasonable to suggest that most persons can say words faster than they can type them. Employing a vocal response requirement may result in enhanced levels of fluency due to what Lindsley (1996a) refers to as the "freedom to form responses," resulting in better performance after the reversal than was observed in the present study.

Application

Understanding the conditions that promote transfer from one task to another has importance for educators trying to teach large and diverse repertoires. Consider again our pilot research (see Footnote 1). Subjects who learned to type terms when shown definitions could not write out the definitions when presented the terms. If we assume that writing out the definition was a less practiced (less familiar) response than typing the term, then the results of the current study would support a teaching strategy in which practice concentrates on emitting unfamiliar response components. For example, subjects might be given practice at saying or typing the whole or selected parts of definitions instead of supplying the terms when presented the definitions. Research in our laboratory using the "Type Keyword" mode of *Think Fast* confirms that this strategy enhances learning and transfer to the non-trained performance (Yuen, Polson, & Parsons, 1992).

REFERENCES

Binder, C. (1993). Behavioral fluency: A new paradigm. *Educational Technology, 33*, 8-14.

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst, 19*, 163-197.

Binder, C., Haughton, E., & Van Eyk, D. (1990). Precision teaching attention span. *Teaching Exceptional Children, 22*, 24-27.

Binder, C., & Watkins, C. L. (1990). Precision teaching and direct instruction: Measurably superior instructional technology in schools. *Performance Improvement Quarterly, 3*, 74-96.

Catania, A. C. (1992). *Learning* (3rd ed.) Englewood Cliffs, NJ: Prentice-Hall.

Dale, E., & O'Rourke, J. (1976). *The living world vocabulary: The words we know: A national vocabulary inventory*. Chicago: Field Enterprises Educational Corp.

Dougher, M. J., Augustson, E., Markam, M. R., Greenway, D. E., Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 62*, 331-351.

Dougherty, K. M., & Johnson, J. M. (1996). Overlearning, fluency, and automaticity. The *Behavior Analyst, 19*, 289-292.

Hall, G. A., & Chase, P. N. (1991). The relationship between stimulus equivalence and verbal behavior. *The Analysis of Verbal Behavior, 9*, 107-119.

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

Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist, 47*, 1475-1490.

Johnson, K. R., & Layng, T. V. J. (1996). On terms and procedures: Fluency. *The Behavior Analyst, 19*, 281-288.

Lane, S. D., & Critchfield, T. S. (1996). Verbal self-reports of emergent relations in a stimulus equivalence procedure. *Journal of the Experimental Analysis of Behavior, 65*, 355-374.

Lindsley, O. R. (1992). Precision teaching: Discoveries and effects. *Journal of Applied Behavior Analysis, 25*, 51-57.

Lindsely, O.R. (1996a). The four free-operant freedoms. *The Behavior Analyst, 19*, 199-210.

Lindsley, O.R. (1996b). Is fluency free-operant response-response chaining? *The Behavior Analyst, 19*, 211-224.

Lowe, C. F., & Beasty, A. (1987). Language and the emergence of equivalence relations: A developmental study. *Bulletin of the British Psychological Society, 40*, A49.

Lowenkron, B. (1991). Joint control and the generalization of selection-based verbal behavior. *The Analysis of Verbal Behavior, 9*, 121-126.

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

Mandell, C., & Sheen, V. (1994). Equivalence class formation as a function of the pronounceability of the sample stimulus. *Behavioural Process, 32*, 29-46.

Michael, J. (1985). Two kinds of verbal behavior plus a possible third. *The Analysis of Verbal Behavior, 3*, 2-5.

Nelson, S. E. (1972). In search of associative symmetry. In C. P. Duncan, L. Sechrest, & A W. Melton (Eds.), *Human memory: Festschrift for Benton J. Underwood* (pp. 133-153). New York: Appleton-Century-Crofts.

Olander, C. P., Collins, D. L., McArthur, B. L., Watts, W. O., & McDade, C. E. (1986). Retention among college students: A comparison of traditional versus precision teaching. *Journal of Precision Teaching, 4*, 80-82.

Parsons, J. A. (1989). *Think Fast*. [Computer program]. Victoria, BC: Author.

Parsons, J. A., Taylor, D. C., & Joyce, T. M. (1981). Precurrent self-prompting operants in children: "Remembering." *Journal of the Experimental Analysis of Behavior, 36*, 253-266.

Polson, D. A. D. (1995). Fostering multiple repertoires in undergraduate behavior analysis students. *The Behavior Analyst, 18*, 293-299.

Polson, D. A. D., & Parsons, J. A. (1994). Precurrent contingencies: Behavior reinforced by altering reinforcement probability for other behavior. *Journal of the Experimental Analysis of Behavior, 61*, 427-439.

Polson, D. A .D., Wong, W. K., & Parsons, J. A., & Grabavac, D. M. (1991, May). *Using Think Fast to examine variables affecting the efficacy of flash card instruction*. Poster presented at the Association for Behavior Analysis 17th Annual Convention, Atlanta, GA.

Saunders, R. R., & Green, G. (1996). Naming and equivalence relations. *Journal of the Experimental Analysis of Behavior, 65*, 312-314.

Shafer, E. (1993). Teaching topography-based and selection-based verbal behavior to developmentally disabled individuals: Some considerations. *The Analysis of Verbal Behavior, 11*, 117-133.

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: Lawrence Erlbaum Associates, Inc.

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

Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, NJ: Prentice-Hall.

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

Sundberg, C. T., & Sundberg, M. L. (1990). Comparing topography-based verbal behavior with stimulus selection-based verbal behavior. *The Analysis of Verbal Behavior, 8*, 31-41.

Wraikat, R., Sundberg, C. T., Michael, J. (1991). Topography-based and selection-based verbal behavior: A further comparison. *The Analysis of Verbal Behavior, 9*, 1-17.

Yuen, S., Polson, D. A. D., & Parsons, J. A. (1992, May). A study comparing the effectiveness of keywords and answers using Think Fast. In G. E. Yaber (Chair), *Think Fast: An authoring/delivery system for behavioral research and training*. Symposium conducted at the Association for Behavior Analysis 18th Annual Convention, San Francisco, CA.