6-1996

Comparison of Selection-Based vs. Topography-Based Verbal Behavior

William F. Potter
Western Michigan University

COMPARISON OF SELECTION-BASED VS. TOPOGRAPHY-BASED VERBAL BEHAVIOR

by

William F. Potter

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Psychology

Western Michigan University
Kalamazoo, Michigan
June 1996

COMPARISON OF SELECTION-BASED VS. TOPOGRAPHY-BASED VERBAL BEHAVIOR

William F. Potter, Ph.D.

Western Michigan University, 1996

Michael (1985) distinguished between selection-based and topography-based verbal behavior. Some researchers have started to examine this distinction, as does this research. This study examined the contribution that response-produced kinesthetic stimulation has on the acquisition of conditional discriminations and equivalence relations by college students. To accomplish this, a special computerized nonidentity matching-to-sample task (a selection-based task) was created, which arranged for each participant to perform under two conditions. The first condition arranged for participants to make a stereotypical response to each choice stimulus selected. The second condition arranged for a unique response to be made to each choice selected. The number of incorrect responses and latencies were recorded.

Initial findings indicated little difference between the two conditions. Exit interview data indicated that all participants used vocal verbal behavior (overt or covert) as an aid in performing the arranged task. A final session was conducted in which less discriminable sample stimuli were used. In addition the vocal verbal behavior of participants was examined while they engaged in the arranged task (a protocol analysis, Ericcson & Simon, 1993).

The results of the protocol analysis indicate that specific types of comments typically preceded correct choices, lending support to the possibility that some conditional discrimination tasks, and emergent equivalence relations, are mediated by topography-based responding.

ACKNOWLEDGMENTS

It is a long journey to obtain a doctorate, one fraught with numerous distractions, hurdles and sometimes frustrations. When the journey is completed, one has the time to reflect on all the people who contributed to the accomplishment. Those who have aided my efforts are listed below.

My parents, Thomas and Carol Potter, gave me the skills to stand on my own, along with the good judgment to do so without compromising my ideals. They did this with a wonderful mixture of love and firmness.

The chair of my committee, Jack Michael, brought my world view into focus. He did this with his keen insights, enthusiasm and great skill in teaching. The rest of my dissertation committee, Dick Malott, Lisa Baker and Phil Chase, each contributed in their unique way. I am forever beholden to each of you.

My friends' intellectual and emotional support made this journey stimulating and fruitful. Matthew Mason, Asiah Mayang, Judy Agnew, Bruce Hesse, Satoru Shimamune, Mike Hixson, Mary Slominas, Mike Makepeace, John Hampel, Mark Lesage, Yoshi Hayashi, Roland Johnson, Tara De La O, Mike Schmandt, Todd Nelson, Lisa Leppert, Maureen Zeff, and Jennifer O'Donnell. Special thanks to Shawn Huber and Deb Brown for their efforts on the dissertation research.

Finally, thanks to the California State University, Stanislaus, department of psychology for making me feel part of the "family" as well as providing support throughout the dissertation process.

William F. Potter

UMI Number: 9640075

TABLE OF CONTENTS

Table of Contents-Continued

Table of Contents-Continued

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Selection-based and Topography-based Verbal Behavior

Michael (1985) pointed out the distinction between two types of verbal behavior: selection-based (SB) and topography-based (TB). Selection-based verbal behavior consists of pointing to a verbal stimulus, for example a bliss symbol (Bliss, 1965; McNaughton, 1978) or some similar type of symbol. The symbol pointed to presumably affects the listener in an appropriate manner. This form of communication has been used with apes (Savage-Rumbaugh, 1984) for which vocalizations are difficult or impossible (Donahoe & Palmer, 1994) and is currently used extensively as a communication method for the developmentally disadvantaged (Shafer, 1993). Common examples of selection-based verbal behavior include a child requesting food by pointing to various symbols on a communication board or a student selecting the correct answers on a multiple-choice test.

Topography-based verbal behavior, as its name implies, consists of making a response with a unique response form or topography, the resulting stimulus affecting the listener in an appropriate manner. Common examples of topography-based verbal behavior include vocal verbal behavior, sign language and writing.

Contrasting selection-based verbal behavior to topography-based verbal behavior is useful in illustrating the differences pointed out by Michael (1985). Consider a situation in which a teacher is training a developmentally disadvantaged student to use a communication board. A communication board usually consists of a flat board-like device upon which a number of symbols are displayed. The student may point to a particular symbol on the device and the teacher will provide feedback about the correctness of the response. For example, the teacher may present a picture of a dog and the student may then point to the symbol in the upper-left square. If the selected stimulus is the correct symbol for a dog, the teacher will respond in a positive fashion. The teacher may then hold up a picture of a cow. Again, an appropriate consequence follows depending on the student's selection. Note that the topography of the response, pointing, is very similar whether the student is pointing at the symbol for a dog or for a cow. Except for minor positional differences, the responses may appear identical.

In the case of a teacher presenting a picture of a cow to a deaf student learning sign language, the response evoked (if accurate) will be the particular sign for cow. The teacher judges the accuracy of the response based on the topography of the response. When the teacher presents a picture of a dog, the sign emitted is different from that of the sign for cow, or in other words, the topography of the signed response "cow" differs significantly from that of the signed "dog" response. This is not the case with SB verbal behavior. Other differences between the two types of verbal behavior are tabulated in Table 1.

Table 1 (derived from Cresson, 1994; Michael, 1985; and Stratton, 1992) compares the differences between SB and TB tasks in four general categories: the stimuli controlling the SB or TB response (stimulus control), the response itself, the nature of the response product (response-produced stimulation), and the general environmental arrangement required for each of the task types.

As noted in the part of Table 1 addressing stimulus control, an SB task is controlled by a relation between two stimuli, namely the sample stimulus (or an establishing operation, although this is generally not referred to as a stimulus), which affects the participant in such a manner as to increase the evocative strength of one of the choice stimuli. This choice stimulus evokes the pointing response from the participant. Such a situation is often called a conditional discrimination. Contrast this type of stimulus control with that of a TB task, in which no choice array is present. In this case, the response is evoked directly by the antecedent stimulus.

Several differences also exist between the types of responses emitted in SB and TB verbal behavior. All final SB responses tend to be similar (e.g., pointing to a choice stimulus), while TB responses are all necessarily different (as the form of the response is the criterion by which the response is deemed appropriate or not in any given situation) (Michael, 1993). Prior to the final SB response, a scanning response must occur; that is, the speaker must come into contact with the appropriate sample stimulus

Table 1

A Comparison of Selection-based and Topography-based Verbal Behavior

Comparison Item	Selection-based	Topography-based
Stimulus Control	•Conditional. The pointing response is controlled by a stimulus or establishing operation which alters the organism in such a manner that the evocative strength of a particular choice stimulus is increased (from an array). •All choice stimuli are present allowing the student to react to them	•A direct relation. The antecedent stimulus condition evokes the response (no second stimulus is involved.) •Not all choice stimuli are present; thus, if the response is not readily evoked, no other environmental support is available. Even if the student starts to recite all responses involved in the training set, it is not the same level of environmental support.
Response	•Nearly indistinguishable from other selection-based responses. •Requires an additional "scanning" repertoire to come into contact with the choice stimulus effective in evoking the pointing response. •It is likely that the pointing response already exists in the student's repertoire (no new topography needs to be learned)	•Clearly distinguishable from other topography-based responses. •Requires no "scanning" repertoire. •The topography of the response must be learned before it can be emitted.

Table 1-Continued

Comparison Item	Selection-based	Topography-based
Response-produced Kinesthetic Stimulation	•Nearly indistinguishable from other selection-based response-produced kinesthetic stimuli (see text for more detail).	•Clearly distinguishable from other topography-based response-produced kinesthetic stimuli.
Correspondence Between Response and the Antecedent Stimulation for the Listener	•No point-to-point correspondence between the response and the response-produced stimulus functioning as an antecedent stimulus for the listener's behavior.	•Point-to-point correspondence exists between the response and the response-produced stimulus functioning as an antecedent stimulus for the listener's behavior.
Environmental Arrangement	•Necessitates additional apparatus (the stimulus array). •In most such communicative activity, if it can be assumed that the correct symbol is on the board (as with any training exercise), finding it even if it is not well known is made easier by being able to eliminate the known symbols for other objects or events. •If the number of symbols is large enough to require a considerable search time, there will be a loss of control by the variable that initiated the search, even if the relevant symbol is well-known.	•No additional apparatus required. •There is no sense in which the entire relevant repertoire can be examined and incorrect responses eliminated, other than by emitting the responses which are strong, which may not include the relevant one. •If the response is strong, there is no time delay between the presentation of the controlling variable and the response occurrence.
Notes	Derived from Cresson, 1994; Michael, 1985; Stratton, 1992.	

before that stimulus can evoke the pointing response. This is not the case with a TB response. When acquiring new relations in an SB framework, no new response needs to be added to the existing repertoire (i.e., the pointing and scanning repertoires having already been established). In a TB framework however, acquiring a new relation generally involves adding a new response to the speaker's existing repertoire (i.e., learning a new sign). However, in both TB and SB types, the response must be brought under appropriate stimulus control. Most of Table 2 describes research which has focused on examining gross differences between SB and TB relations in terms of stimulus control and response type (see the review of those studies below).

An area which has not been researched in terms of the differences between SB and TB verbal behavior is the different response-produced stimulation provided under each task (Michael, 1993). Clear differences exist between such tasks in terms of kinesthetic feedback. SB tasks generally result in little or no difference in the response-produced stimulation as each physical movement is nearly identical regardless of the selection (e.g., point and touch or point and press). Only slight kinesthetic differences will be experienced between pointing to a stimulus on the left versus the right. TB tasks, however, require a unique topography in order to provide stimulation which is distinct enough to function in the same manner as the choice stimuli in the SB task (i.e., in a verbal exchange, have the appropriate effect on the listener). This necessarily results in unique kinesthetic feedback for different TB responses. This study will investigate the effects such differences might produce.

Finally, certain logistical elements differ between the two types of verbal behavior, which have some interesting practical implications as noted in Table 1 (Cresson, 1994; Michael, 1993). For example SB verbal behavior has been called an "aided" (Romski, Sevcik & Pate, 1988; Sigafoos & Iacono, 1993) communication system, in that it necessitates the use of additional apparatus such as a communication board or a computer.

There appear to be at least four general areas in which research in this area might be considered valuable. First, as some researchers have recently noted, there appears to be an increase in the use of selection-based techniques (e.g., communication

Table 2

Research Examining Differences in TB and SB Verbal Behavior

Article Specification	Bristow & Fristoe, 1984	Hodges & Schwethelm, 1984	Sundberg & Sundberg, 1990	Wraikat, 1990	Wraikat, 1990	Stratton, 1992	Wallender
Participants	20 children av.age age 8.2 YO	52 retarded children (Av. age 12.21)	4 DD adults	5 DD adults	7 DD adults	28 college students	20 college
Pre-training	TB: saw picture then sign was modeled. Participant required to do same sign SB: said word and picture was pointed out. Participant. was required to pick correct picture from array. Same procedure for picking symbols, but pictures served as sample.	TB: Trained out behavioral problems (crying etc.) then trained help up object and asked them to sign (used prompting and fading). Given object signed (mand training) SB: Trained out behavioral problems (crying etc.) then trained matching to sample.	TB: modeled 5 times with antecedent Participant imitated SB: modeled 5 times with antecedent Participant imitated	TB: modeled 5 times with antecedent. Participant imitated SB: modeled 5 times with antecedent. Participant imitated	TB: modeled 5 times with antecedent Participant imitated SB: modeled 5 times with antecedent. Participant imitated	TB: shown English word and experimenter said Japanese word. Participants repeated word until correct. SB: shown English word and choice. Participants then picked choice from array (repeats until correct). All stimuli and choice were presented once, random order.	SB: show English w choice. Participan picks choi array (repe correct). A stimuli an were prese once, rand order.
Maintenance Training	None provided	None provided	5 times each relation at start of each phase 2 times each relation at start of each session	5 times each relation at start of each phase 2 times each relation at start of each session	2 times each relation at start of each session Interspersal training: 1/2 trials were assigned to old relations (with 3 relations the # of trials per session increased from 48 to 54).	None provided	None prov

Table 2

Research Examining Differences in TB and SB Verbal Behavior

& 1990	Wraikat, 1990	Wraikat, 1990	Stratton, 1992	Wallender, 1993	Cresson , 1994	Tan et. al, 1995
s	5 DD adults	7 DD adults	28 college students	20 college students	16 college students	8 college students
d 5 d 5 dent	TB: modeled 5 times with antecedent. Participant imitated SB: modeled 5 times with antecedent. Participant imitated	TB: modeled 5 times with antecedent Participant imitated SB: modeled 5 times with antecedent. Participant imitated	TB: shown English word and experimenter said Japanese word. Participants repeated word until correct. SB: shown English word and choice. Participants then picked choice from array (repeats until correct). All stimuli and choice were presented once, random order.	SB: shown English word and choice. Participants then picks choice from array (repeats until correct). All stimuli and choice were presented once, random order.	TB: 1. Shown pattern, part. pointed to match; 2. Hear sound, part. said sound; 3. shown symbol, part. wrote it (then wrote all 8, random order, no sample). SB: 1. Shown visual pattern, Participant point to match; 2. Hear sound, Participant said sound; 3. shown symbol, Participant pointed to match.	Trained response in either SB or TB (typing) then tested for symmetry in either the same or opposite paradigm. Displayed on a computer participant. either selected or typed in choice.
h tart of h tart of 1	5 times each relation at start of each phase 2 times each relation at start of each session	2 times each relation at start of each session Interspersal training: 1/2 trials were assigned to old relations (with 3 relations the # of trials per session increased from 48 to 54).	None provided	None provided	Reviews at start of session II when still in same paradigm. Review of both training paradigms was given before a full test was given of all relations, equivalence and generalization.	Interspersal training - every 6th trial was symmetry test

Table 2 - Continued

Article Specification	Bristow & Fristoe, 1984	Hodges & Schwethelm, 1984	Sundberg & Sundberg, 1990	Wraikat, 1990	Wraikat, 1990	Stratton, 1992
Correction Training	Correct -said "Correct" or something similar Incorrect- said incorrect of similar. Participant. required to imitate correct response.	Fading and prompting were used, but no correction training was cited (e.g. remedial)	Non answer = Demo with verbal prompt Incorrect = Told incorrect, then model with verbal prompt, correct.	Non answer = Demo with verbal prompt Incorrect = Told incorrect, then model with verbal prompt, correct.	Non answer = Demo with verbal prompt Incorrect = Told incorrect, then model with verbal prompt, correct.	Correct answer an ask Participant to do response Incorrect = presen correct answer anc ask Participant to do response
Reinforcers used	Correct/Incorrect feedback. No others reported.	Various foods, objects and praise.	Pennies & Praise. Weekly outing.	Praise, stickers, etc. No money. Weekly outing.	Praise, stickers, etc. No money. Weekly outing.	Praise (via computer) and feedback "Correct"
Mastery Criterion	100% correct on 2 consecutive presentations of all 12 relations.	80% percent correct responding on 8-10 signs or 80% on 9 sample-comparison (SB) relations.	For all relations (including partial equivalence), 9 out of 10 correct responses.	For all relations (including partial equivalence), 9 out of 10 correct responses.	With 2 objects = 11 of 12. With 3 objects = 7 of 8 (successive correct. responses. Interspersal not used in calculation.)	Three consecutive blocks of no errors. A block consisted of each stimulus being presented once. (either set of 5 or of 15)
Relations Trained or tested (•):	TB Tact SB Tact Transitivity	TB Mand SB Tact	TB Tact TB Intraverbal SB Tact SB Intraverbal Transitivity	TB Tact TB Intraverbal SB Tact SB Intraverbal	TB Tact TB Intraverbal SB Tact SB Intraverbal Transitivity	TB Intraverbal SB Intraverbal Symmetry
Results Acquisition:	Learned signs (TB) slightly quicker than symbols (SB).	Learned signs (TB) quicker than symbols (SB Tact)	TB faster for tact and IV and equivalence	TB Faster for tact and IV	TB faster for tact and IV and equivalence	For 5 stimulus set SB and TB about equal For 20 stim. set SB was faster (by aver. of 6 blocks)

& 1990	Wraikat, 1990	Wraikat, 1990	Stratton, 1992	Wallender, 1993	Cresson , 1994	Tan et. al, 1995
er = h verbal = Told then h verbal orrect.	Non answer = Demo with verbal prompt Incorrect = Told incorrect, then model with verbal prompt, correct.	Non answer = Demo with verbal prompt Incorrect = Told incorrect, then model with verbal prompt, correct.	Correct answer and ask Participant to do response Incorrect = present correct answer and ask Participant to do response	Non answer = present correct answer and ask participant to do response Incorrect = present correct answer and participant did it.	Incorrect = correction: repeat trial with correct answer displayed (for both paradigms)	Correct answer resulted in a high tone. Not reported if repeated or not. Incorrect answer resulted in a low tone. Not reported if repeated or not.
: Praise. uting.	Praise, stickers, etc. No money. Weekly outing.	Praise, stickers, etc. No money. Weekly outing.	Praise (via computer) and feedback "Correct"	Praise (Via computer))and feedback "Correct"	Praise; $5 per completed session	Tones for correct/incorrect. Others not reported
ations partial e), 9 out ct	For all relations (including partial equivalence), 9 out of 10 correct responses.	With 2 objects = 11 of 12. With 3 objects = 7 of 8 (successive correct. responses. Interspersal not used in calculation.)	Three consecutive blocks of no errors. A block consisted of each stimulus being presented once. (either set of 5 or of 15)	Two consecutive blocks (20 trials) with no errors. A block consisted of each stimulus being presented once.	None taken	———
'act verbal 'act verbal tivity	TB Tact TB Intraverbal SB Tact SB Intraverbal	TB Tact TB Intraverbal SB Tact SB Intraverbal Transitivity	TB Intraverbal SB Intraverbal Symmetry	SB Intraverbal	TB Tact TB Intraverbal SB Tact SB Intraverbal Equivalence Generalization*	SB IV TB IV
for tact e	TB Faster for tact and IV	TB faster for tact and IV and equivalence	For 5 stimulus set, SB and TB about equal For 20 stim. set SB was faster (by aver. of 6 blocks)	English words as sample stimuli were learned much faster (half as many blocks to mastery) than Katakana as sample stimuli	Not reported	All learned equally well.

Table 2 - Continued

Article Specification	Bristow & Fristoe, 1984	Hodges & Schwethelm, 1984	Sundberg & Sundberg, 1990	Wraikat, 1990	Wraikat, 1990	Stratton, 1992
<u>Results/Accuracy:</u>	Not reported	More TB Mand relations were acquired then SB Tact relations.	TB Higher (tact) TB Higher (IV) TB Higher - Equivalence	TB Better (tact) TB Higher (IV)	TB Better (Tact) TB Better (IV) TB Better - Equivalence	Not reported (tac nor IV) All SB and TB groups easily mastered symme
<u>Results/Relation Type:</u> within paradigm Tact Intraverbal	———	———	The tact was acquired faster for both paradigms (3 of 4 participants). Lower VB skill seemed to be a factor.	The tact was acquired faster for both paradigms (TB all participants SB 3 out of 5)	In TB, Tact acquired faster than intraverbal, and better % correct. In SB, slightly better Tact vs. IV performance in both acquisition and % correct.	5 stimulus set took less blocks master than the stimulus set, in both paradigms (expected)
Results Other:	None reported	None reported	None reported	None reported	None reported	Within SB & TB react. time faster for the 5 stimulu group (slight diff Between SB & T react. time averaged 3' more for SB
Antecedents Used	TB: Pictures SB: Pictures Nonsense word	TB & SB: Objects (e.g. candy, ball)	TB: Show Object Say nonsense syll. SB: Show Object Say nonsense syll. Asked to pick Obj.	TB & SB: Show Object Say nonsense syll.	TB: Show Object Auditory nonsense syllable SB: Show Object Say nonsense syll. Asked to pick Obj	TB & SB: Written English word
Comparison's Used (Selection-Based)	Symbols or pictures (ordering not reported)	Symbols (4 used) (ordering not reported)	Symbols - 3 used Pseudo random Order	Symbols - 3 used Pseudo random order each trial	Symbols - 2 used for 3 Participants. 3 used for other 4 part.. Pseudo random order each trial	Symbol (Japanese Kanji characters) Random order eac block

& 1990	Wraikat, 1990	Wraikat, 1990	Stratton, 1992	Wallender, 1993	Cresson , 1994	Tan et. al, 1995
(tact) (IV) - e	TB Better (tact) TB Higher (IV)	TB Better (Tact) TB Better (IV) TB Better - Equivalence	Not reported (tact nor IV) All SB and TB groups easily mastered symmetry	Not reported (IV)	TB better - tact TB better - IV TB better - trans. TB better for generalization test (diff. were slight)	Near perfect symm. performance when response was SB; generally poor symm. when response was TB.
s ter for gms (3 pants). skill e a	The tact was acquired faster for both paradigms (TB all participants SB 3 out of 5)	In TB, Tact acquired faster than intraverbal, and better % correct. In SB, slightly better Tact vs. IV performance in both acquisition and % correct.	5 stimulus set took less blocks to master than the 20 stimulus set, in both paradigms (as expected)	———	Intraverbal was generally learned better in both paradigms (fewer errors). Probably due to nature of sample stimuli (quilt pattern vs. trigram)	———
ed	None reported	None reported	Within SB & TB, react. time faster for the 5 stimulus group (slight diff.) Between SB & TB, react. time averaged 3' more for SB	No difference in reaction times between the two sample stimulus groups (either English words or Japanese Katakana characters)	Auditory sample stimulus resulted in better performance than visual (complex quilt patterns)	None reported
ct se syll. ct se syll. ck Obj.	TB & SB: Show Object Say nonsense syll.	TB: Show Object Auditory nonsense syllable SB: Show Object Say nonsense syll. Asked to pick Obj	TB & SB: Written English word	TB & SB: Written English or Japanese words	TB & SB Visual Pattern Auditory nonsense trigram	TB & SB: All stimuli used were either French or English written words
3 used lom	Symbols - 3 used Pseudo random order each trial	Symbols - 2 used for 3 Participants. 3 used for other 4 part.. Pseudo random order each trial	Symbol (Japanese Kanji characters) Random order each block	Symbol (Japanese Kanji characters) Random order each block	Symbols (Katakana SB Intraverbal) and tact) pseudo randomized	French or English word (4 at each time) (ordering not reported)

Table 2 - Continued

Article Specification	Bristow & Fristoe, 1984	Hodges & Schwethelm, 1984	Sundberg & Sundberg, 1990	Wraikat, 1990	Wraikat, 1990	Stratton. 1992
TB Response Form	Signs	Signs	Manual Signs	Manual Signs (but none touching body)	Manual Signs (but none touching body)	Vocal (Japanese 2 syll. word)
Limited Hold for Response	30"	Not reported	10"	20"	20"	20"
Additional Comments	Found a tendency for participants to do better on relations which were practiced more.	Participants appeared more attentive in mand condition.	Several participants demonstrated overt mediating responses. These participants performed better than others	Participants appeared more attentive and positive when engaging in the TB task	Participants appeared more attentive and positive when engaging in the TB task	Noted that some participants foun similarities between samples and choices (engages in TB responding durin SB task)
Notes	* In this study generalization referred to a TB version of a trained SB task. Or in reverse, a SB version of a TB trained ** IT should be noted that this may be partially SB and TB, depending on the proficiency with which each participant					

/raikat, 1990	Wraikat, 1990	Stratton, 1992	Wallender, 1993	Cresson , 1994	Tan et. al, 1995
lanual Signs ut none touching ody)	Manual Signs (but none touching body)	Vocal (Japanese 2 syll. word)	None	Write Katakana symbol	Typing**
20"	20"	20"	20"	No Time limit (all part. but one given instruction "Go quickly")	Not Reported
rticipants peared more tentive and sitive when gaging in the TB sk	Participants appeared more attentive and positive when engaging in the TB task	Noted that some participants found similarities between samples and choices (engages in TB responding during SB task)		Found students with high errors did better on TB tasks, those with lower errors did better on SB	Symmetry performance improved over repeated exposures (non reinforced)

ned SB task. Or in reverse, a SB version of a TB trained task.
pending on the proficiency with which each participant types

boards) with developmentally-delayed and disabled populations (Shafer, 1993; Shane & Bashir, 1980; Sundberg, 1993; Sundberg & Sundberg, 1990). This increase has not been accompanied by much empirical research (Shafer, 1993; Sundberg & Sundberg, 1990), and these studies that have been noted often are case studies (Romski et al., 1988).

Second, most studies examining the acquisition of verbal behavior by nonhumans have not considered the distinction between SB and TB verbal behavior, although there has been much research using communication boards with nonhumans. The aims of these researchers vary considerably, but each, at some level, recognize the importance of nonhuman research in understanding the nature of language (Epstein, Lanza & Skinner, 1980; Patterson & Linden, 1981; Pepperberg 1988; Rumbaugh, 1977; Savage-Rumbaugh, 1990; Terrace, 1979). It would be valuable to determine if differences exist between these types of verbal behavior, to aid in interpreting the results of such studies.

Third, an area in which selection-based procedures are quite frequently used (see Sidman, 1994), and with strong implications for the acquisition of verbal behavior, is research on stimulus equivalence (discussed below). Much of this research involves the use of the so called "matching to sample" technique (although see Sidman, 1994, for a somewhat different definition of matching-to-sample). This technique involves presenting some kind of antecedent stimulus (e.g., a picture of a car, or the auditory stimulus "car") along with several choice stimuli from which the participant is required to select the correct relation (e.g., pick the word "car" in the presence of a picture of a car). It also should be noted though that research is conducted in which selection-based and topography-based responses are intermixed in a single equivalence experiment. For example, one of the earliest studies (Sidman, 1971) involved training a severely developmentally disabled participant to match pictures of a cat to the printed word cat (selection-based response), as well as to say the word "cat" upon seeing a picture of a cat (topography-based response).

Some researchers have suggested that SB responding may consist of both TB and SB responding (Dugdale & Lowe, 1990; Horne & Lowe, 1996; Lowenkron, 1991, 1996; Lowenkron & Colvin, 1995). Lowenkron (1991) suggests that selection-based responding may be responding under what he has termed "joint control." More specifically, he suggests that a topography-based response is emitted in the presence of the sample stimulus and then repeated while the participant scans the choice stimuli. When a response similar to the repeated response is evoked by a choice stimulus, the selection response is emitted. Dugdale and Lowe (1990) and Horne and Lowe (1996) have proposed that a combination of SB and TB verbal behavior, trained at similar times, could be called the "naming" relation, and might be responsible for the emergence of stimulus equivalence (see below). Some evidence exists for the role of mediating TB responses in the development of stimulus equivalence (Dugdale & Lowe, 1990; Lowenkron & Colvin, 1995; Wulfert, Dougher & Greenway, 1991) but most of this research is necessarily correlational.

Finally, as implied in the preceding three paragraphs, furthering our knowledge of the distinctions between selection-based and topography-based verbal behavior may be useful in designing experiments which take these distinctions into consideration (Hall & Chase, 1991). As noted above, much research in stimulus equivalence either mixes the two types of verbal behavior, or focuses on only one (generally the selection-based type). Chase (personal communication, 1995), pointed out a previous study that examined both rate and accuracy of responding using a selection-based (identifying examples of a concept) and topography-based (defining or exemplifying a concept) task (Chase et al., 1985) . This study showed a clear difference between the SB and TB rates of responding and accuracy (the SB/TB distinction had not been developed at this point). These researchers found that the SB task produced a high rate of responding, but with many errors. The TB tasks produced lower rates of responding, but with more accuracy. Given the differences which appear to exist between these two types of verbal behavior (see Table 1), it would seem important to investigate these differences to aid in future research.

Interestingly, in more cognitive areas of psychology, some researchers are focusing on problems which also encompass the SB/TB distinction. Research on automaticity (and signal detection) tend to focus on SB responding, for example asking a participant to find several types of objects from an array of choices (Anderson, 1980; Strayer & Kramer, 1990). Behaviorally, these participants are required to engage in scanning behavior, then emit the appropriate response, based on a particular contextual arrangement (e.g., find all the "B's" in this list). Although these researchers tend to measure different dependent variables and interpret these data differently, methodologically there are many similarities between the behavioral and cognitive approaches. Although this literature is not reviewed here, it is important to note the pervasiveness of the SB/TB distinction, especially when it may be a variable affecting research results.

Stimulus Equivalence and Language

Prior to reviewing studies which have directly examined the SB/TB division, a brief review of stimulus equivalence and its relevance to language is beneficial. Basically, stimulus equivalence involves training certain relations, for example, after hearing the nonsense word "Ork", the participant is required to select a picture of some object (e.g., a lump of coal). Next, the participant is required to select the correct nonsense symbol (e.g., "¬") from an array when presented with a picture (e.g., the lump of coal). Once these relations are trained, then equivalence testing can be conducted. In this case, one could test to see if reversing the sample and choice stimuli results in accurate responding (e.g., present "¬" and require the participant to select the picture of the lump of coal from an array of pictures). This type of relation is called symmetry, and composes one of three relations which Sidman and Tailby (1982) used to define stimulus equivalence. Another relation, called reflexivity, consists of presenting a particular stimulus both as the sample and the correct comparison in an array of similar stimuli. The last relation, called transitivity, would consist of presenting a nonsense word (e.g., "Ork") and having the symbol "¬" as the correct comparison stimulus in an array. In symbolic terms, these relations can be summed as: A = B and B = C (training), then test A = A (and all other reflexive relations), B = A, C = B (symmetrical relations) and A = C (transitive relation). When all three emergent relations, transitivity, reflexivity, and symmetry have been demonstrated, then stimulus equivalence is said to exist for the set of stimuli used (Sidman & Tailby, 1982). The relevance of these emergent relations to verbal behavior has been noted by many researchers (Barnes, 1994; Fields, Verhave, & Fath, 1984; Hall & Chase, 1991; Sidman, 1986; Wulfert & Hayes, 1988;) although some have raised questions as to the applicability of stimulus equivalence to topography-based verbal behavior (Hall & Chase, 1991; Tan, Bredin, Polson, Grabavac & Parsons, 1995).

While most researchers have focused on demonstrating reflexivity, symmetry and transitivity, others have explored various combinations and arrangements of these relations. For example, Sidman and Tailby (1982) examined the emergence of both a symmetrical and transitive relation. To do this they trained the relations B = A, C = A, then tested B = C and C = B. In summary, these researchers examined whether or not equivalence was demonstrated between two sample stimuli which had been paired only with a common choice stimulus (equivalence was demonstrated). This study will arrange for a similar training and testing. For the purpose of this research, this type of emergent relation will be called "equivalence" as done so by Lazar, Davis-lang and Sanchez (1984).

There is much debate on the relevance of the stimulus equivalence research to verbal behavior. Sidman (1994) generally believes that the capacity to demonstrate stimulus equivalence is an innate feature of humans, one essential to the development of verbal behavior. Hayes and Hayes (1986) also believe stimulus equivalence to be essential to verbal behavior, although the development of this capability they believe to be a function of a specific history. Others have claimed that stimulus equivalence is a function of verbal behavior; that is, without verbal behavior, stimulus equivalence would not emerge (Dugdale & Lowe, 1990; Hall & Chase, 1991; Mandell & Sheen, 1994). To date no study has demonstrated the development of stimulus equivalence as defined by Sidman and Tailby (1982) in nonhumans, nor in nonverbal humans, which has been taken as evidence that language is a prerequisite for demonstrations of stimulus equivalence (Horne & Lowe, 1996). As noted, others have offered an alternative to this (e.g., Hayes & Hayes, 1989).

Topography-based and Selection-based Studies

A number of studies have directly investigated the differences between these two types of language. Many of these studies are summarized in Table 2. Nearly all the studies conducted to date have focused on the acquisition of tact and intraverbal relations. A brief overview is provided of these two types of relations prior to examining the results of SB and TB studies.

The tact relation as defined by Skinner (1957) is "a verbal operant in which a response of given form is evoked (or at least strengthened) by a particular object or event or property of an event. We account for the strength by showing that in the presence of the object or event a response of that form is characteristically reinforced in a given verbal community" (p. 82). Thus, in the tact relation, the controlling variable is nonverbal, that is, is not the product of another's verbal behavior, and the response is typically reinforced by other members of the verbal community. An example of this would be a child saying "red" in the presence of a red apple. Note that the aspect of the environment controlling the response is the "red" property of the apple, which of course might be shared by other items (e.g., a barn, a sunset). Skinner classified this particular type of tact as abstract. An example in which an auditory stimulus controls the tact response would be someone saying "Airplane" in the presence of the sound of an airplane overhead. Notice that the definition does not specify that a particular type of reinforcer is required for this relation: it is not an important distinction (Peterson, 1978). Some researchers have stressed that generalized conditioned reinforcers probably play a rather large role in the acquisition of tact responses as well as the other relation discussed below (Michael, 1993; Skinner, 1957; Sundberg, 1990).

The intraverbal (Skinner, 1957) is a verbal response controlled by a verbal stimulus, but little similarity exists between the controlling verbal stimulus and the product of the evoked response. For example, saying "The Herald" after hearing someone say "newspaper" would be considered an intraverbal relation. The intraverbal relation is also not limited to any sense modality. For example, signing "Hello" after reading the lips of someone saying "Good Morning" would be considered an intraverbal response. As with other relations controlled by verbal antecedents, the intraverbal tends to be consequated generalized conditioned reinforcers (as noted above). Differences found between the acquisition of tact and intraverbal relations will be discussed later in this section.

For the seven studies in Table 2 which investigated ease of acquisition across language type, there was an advantage for topography-based verbal behavior demonstrated in all but two studies. Tan et. al, (1995) found no difference in acquisition between the two language types, but this may have been due to the small number of relations trained in each condition (four), which was the same result obtained by Stratton (1992). In one part of his experiment, Stratton used only five relations in the SB and TB conditions (see Table 2), and found little difference between SB and TB performance. He attributed these results to a floor effect. However, when Stratton increased the number of relations to be learned in both the TB and SB conditions to twenty, the SB relations were acquired on the average of six blocks (i.e., a single presentation of all 20 relations) faster than the TB relations. This is the only study reviewed which demonstrated an advantage for the SB paradigm. This was also the first study to test these differences with normal adults - most other researchers focused on developmentally disadvantaged children or adults. Bristow and Fristoe (1984) and Cresson (1994) used normally functioning children and college undergraduate students, respectively. Both of these researchers found only small differences between the SB and TB paradigms (Cresson did not record the number of blocks to acquisition; however, to the extent to which accuracy is related to acquisition, this holds true). Bristow and Fristoe (1984) showed a slight advantage for the TB paradigm. Preliminary results would seem to indicate that clearer differences emerge between the two paradigms when participants with poor initial verbal skills are tested, and only slight differences appear when participants with well-developed verbal skills are tested. However, more investigation is needed to clarify this issue.

If one were to adopt the hypothesis that SB verbal behavior may be composed of TB components (Dugdale & Lowe, 1990; Lowenkron, 1991 - see above), one would expect to see the results described above. Given no verbal skills with which to use in learning SB relations, it would be necessary for a participant to learn a conditional discrimination to select correctly, which has been demonstrated in nonhuman research, but often takes a considerable amount of time (see discussion section). An alternative to this is that a participant might learn a topographical response and then have it function as an aide in responding correctly in the SB condition. This of course would require an extra learning step, one not programmed by the researcher, and would also require an extended training period.

This hypothesis is partially supported by a follow-up study to Stratton (1992) by Wallender (1993). Wallender suggested that Stratton's results were due to the fact that English (very familiar) words were used as sample stimuli (Stratton used the English names of various animals). According to Wallender, this familiarity most probably allowed participants to readily emit verbal responses to the sample stimuli and also to the choice stimuli, possibly introducing a TB response as a mediator in the SB task (i.e., identifying common features, emitting the TB response while scanning, etc.). Wallender had two groups of participants engage in a SB task similar to that used by Stratton. With one group, however, unfamiliar Japanese characters served as samples and for the other group English animal names served as sample stimuli. In both groups, unfamiliar Japanese characters functioned as choice stimuli. Thus the "familiar" group saw English animal names then selected the appropriate Japanese character and the "unfamiliar" group saw Japanese characters then selected the appropriate (but dissimiliar) Japanese character. A clear difference was demonstrated between the two groups: the group receiving the English words as samples learned the relations nearly twice as fast as the other group (Wallender, 1993). Other researchers have examined this same phenomenon, drawing similar conclusions. Mandell and Sheen (1994) examined the effects of pronounceable (e.g., "FLODG") and nonpronounceable (e.g., "NSJBM" or "+]*^!") sample stimuli on the acquisition of conditional discriminations (SB tasks) and on equivalence classes (transitivity, symmetry and reflexivity). In both cases, acquisition was better with more pronounceable sample stimuli, leading the researchers to note that stimulus equivalence (and conditional discriminations) are likely to be mediated by verbal behavior. It seems clear however, that more investigation is needed to clarify this issue.

Of course, if one were to adopt the hypothesis that TB verbal behavior mediates SB verbal responding, one is still left with the task of explaining Stratton's (1992) results in which SB performance was better then TB performance. Stratton noted:

> The topography-based task is essentially learning 20 new foreign words for 20 English words. There are clearly two aspects to this task: learning to say the foreign words as units, and learning to say the correct one when the English word is shown ..... whereas no new topographies are required in the selection-based task (pp. 23-24).

Stratton also noted that the SB task offers the highly verbal participant the opportunity to engage in verbal behavior which may eliminate some of the choice stimuli (e.g., "it's not the one with the grid pattern") as comparisons on any given trial. Others have echoed this sentiment (Cresson, 1994; McIlvane et al., 1987; Wallender, 1993). McIlvane et al. (1987) showed that responding (using normal adults) in a matching-to-sample task could be controlled by the relation between the sample stimulus and the incorrect choice stimulus. These researchers found similar results when three and four comparisons were used, versus only two. These authors also note that previous research with children seems to indicate that exclusion responses "... have been reported to emerge in the second year, coinciding with initial acquisition of verbal behavior" (p. 206).

It is important to examine the results of Stratton's study in light of the different procedures adopted in each study. Three seem particularly relevant: the number of relations trained (which is usually the number of choice stimuli used for these studies); the nature of the TB response (e.g., a sign, writing response, vocal, etc.); and finally, the extent to which the TB response was pretrained, thus necessitating acquisition of the response during regular training or not, in addition to coming under the appropriate stimulus control for that response.

Stratton (1992) was the only researcher to require more than 12 relations be learned (Cresson, 1994 combined 2 groups of 8 for a generalization test, but not in training), arranging for 20 relations. As noted, Cresson required 8 relations be learned in any given condition and Bristow and Fristoe (1984) required 12 SB relations be learned, but only a maximum of 6 TB relations. No other researcher required more than four relations be learned in any condition. Stratton was also the only researcher incorporating a vocal verbal TB response. All others used sign (Bristow & Fristoe, 1984; Hodges & Schwethelm, 1984; Sundberg & Sundberg, 1990; Wraikat, 1990, 1991), typing (Tan et. al, 1995), or writing (Cresson, 1994). Stratton's verbal response consisted of two-syllable Japanese words. It is possible that some of those words contained phonemes not in the English language as Japanese contains some phonemes which English does not. In comparison, the writing, typing and sign responses required of participants in all other studies had no similar subcomponent of the final response that would be unfamiliar to them. For example, all would have had experience with drawing lines, with characters in the English alphabet - (used when typing French words in the Tan et al., 1995 study), and it is likely all participants had previously emitted all subparts of the sign responses included in these studies.

Stratton (1992) also only provided minimal pretraining of the TB response, namely one preexposure to the sample stimulus and the correct response, at which point the participant would echo the response. Cresson (1994) on the other hand thoroughly trained the TB writing response used, requiring the participants to go through a seven step process starting from copying the Katakana symbols used to generate a series of them, in any nonrepeating order, quickly. Pretraining for the TB response varied across the remaining studies, but could be classified as being more intensive than Stratton's, but less intensive than in Cresson's study (e.g., Sundberg & Sundberg, 1990 demonstrated and required imitation for each relation five times). Bristow and Fristoe (1984) and Hodges and Schwethelm (1984) did not report the extent of pretraining.

In summary then, it is unclear whether Stratton's results, as compared to other researcher's results, are due to true differences in SB and TB responding, procedural differences, or to both. However, at the very least, the preceding analysis does illustrate the number of variables which might influence obtained results.

Again examining Table 2, it is noted that of the five studies reporting accuracy data between SB and TB tasks, all reported more accurate responding in the TB condition. These performances paralleled those cited above (as would be expected): relatively large differences were recorded in studies using participants with poor verbal skills and relatively small differences were demonstrated in studies using participants with well-established verbal skills (Bristow & Fristoe, 1984; Cresson, 1994). Although Hodges and Schwethelm (1984) showed clear differences between TB and SB accuracy (and number of relations acquired), it should be noted that these researchers used nonspecific reinforcers when training the SB task (praise and food) and specific reinforcers for the TB task. These researchers actually conducted training which ultimately resulted in mand responses. Skinner (1957) defined the mand as "a verbal operant in which the response is reinforced by a characteristic consequence and is under the functional control of relevant conditions of deprivation or aversive stimulation" (p. 36). For example, participants in Hodges and Schwethelm's study were required to sign "Candy" when the researcher held up a candy bar. If the emitted sign was correct, the candy was given to the child. Prompting (molding the child's hands) and fading of these prompts were also incorporated in this training. Eventually, the child was only asked "What do you want?", at which point if the child emitted an appropriate sign, the object requested was given to the child. Skinner (1957) and Sundberg (1990) suggest that the mand is probably the first type of verbal relation acquired by humans. They also note that it is likely to be an easily acquired response, given the powerful nature of the reinforcers used. It is possible that the results of this study were actually due to the type of relation trained and not a true difference between SB and TB verbal behavior. Overall however, TB performance appeared to be more accurate than SB performance across studies.

For those studies which incorporated some aspect of stimulus equivalence (transitivity, symmetry and reflexivity - see Results/Accuracy sections of Table 2), TB performance was generally better. Again, the magnitude of the differences recorded seem related to the type of participants tested. All studies using DD participants had a relatively large difference between the TB and SB conditions (primarily transitivity was tested). With highly verbal participants (Bristow & Fristoe, 1984; Cresson, 1994; Stratton, 1992), TB performance was only slightly higher for two studies (Bristow & Fristoe; 1992; Cresson, 1994) and showed no difference in one study (symmetrical relations - Stratton, 1992). Tan et. al (1995) examined the difference between symmetry performances under four conditions: (1) SB training/SB Testing, (2) SB Training/TB Testing, (3) TB Training/SB Testing " and (4) TB Training/TB Testing. These researchers used either French words (unfamiliar) as samples and English words as the choice stimuli (either typed out - TB, or selected - SB). In testing for symmetry, when the response was selection-based, regardless of the training, accuracy was generally high. However, when the response was topography-based in symmetry testing, accuracy was generally low. This was an expected outcome as the authors' note: "... seeing the stimulus word, either during selection-training or topography training, no matter how often, does not guarantee that one will later be able to produce it, especially if the word is unfamiliar" (p. 2).

In those studies which trained both tact and intraverbal relations in the SB and TB conditions, overall the tact relation appears to have been acquired more readily than the intraverbal relation (see Table 2). Cresson (1994) demonstrated an exception to this, with the intraverbal relation acquired more readily. It is difficult to draw conclusions from these results for several reasons. First, different participants were used in those studies (Cresson used normal adults, others used DD adults and children). Second, within each study, the stimuli used as samples are in different modalities (i.e., hearing a sound for the intraverbal and seeing a graphic for the tact), thus making it difficult to equate the two in terms of complexity. Since the researchers in each study used different choice stimuli, it becomes even more difficult to interpret these data. More research is needed to clarify this issue.

Finally, some studies incorporated reaction time measures (Stratton, 1992; Wallender, 1993). As would be expected, reaction times were higher for the SB group in Stratton's study (participants needed to scan the choice stimuli). Wallender trained and tested only in the SB paradigm (as he was examining the effect of familiar and nonfamiliar stimuli on acquisition), and little differences existed between these groups in terms of reaction times.

Aside from these studies, no other studies have attempted to directly determine what particular aspects of the two types of verbal behavior (see Table 1) might be responsible for the differential results obtained. The present study will attempt to do so, focusing on the differences in response-produced stimulation.

Experiment Overview

As noted in Table 1 and Table 2, there are many differences between selection-based and topography-based verbal behavior which need to be examined more closely. For example, some research might investigate the extent to which a conditional discrimination contributes to the differences between the types of verbal behavior, the impact that a scanning repertoire has on acquisition or accuracy, or the effects of imposing time delays between the presentation of the antecedent stimulus and the emittance of the verbal response. This study focused on the "Response" and "Response-produced Kinesthetic Stimulation" comparison items of Table 1. Specifically, it focused on the role of a stereotypical (pure selection-based) response versus a unique response emitted when selecting a choice stimulus from an array of choices. By keeping the task within the SB paradigm, it allowed for the relations learned to be equated, except for the nature of the actual selection response. This allowed for the comparison of a SB task providing little or no distinctive response-produced kinesthetic stimulation with a SB task that provided quite distinct response-produced kinesthetic stimulation. This was accomplished by arranging for one condition in which each comparison was selected in a stereotypical manner (e.g., click four times in a specified location) and a second condition in which a unique response was required for each comparison selected (see the method section for a complete description). A test for equivalence was also conducted, which examined the effect that a unique topography had in demonstrating equivalence.

In addition, a protocol analysis was conducted in session 5. Protocol analyses investigate the verbal behavior of participants under a variety of conditions. Ericsson and Simon (1993) summarize numerous research articles which have investigated the use of protocol analyses. Based on this literature review, the researchers have classified such analyses into three categories. The first, "Talk Aloud" will be the only one used in this research. Basically, this involves asking the participant to simply speak aloud any covert responses which might occur during a task (concurrent verbalizations as compared to post session or post trial verbalizations). Efforts are taken in these analyses to stop the participant from emitting self-observation responses. For example, instructions given to the participants often state that participants should talk aloud as they would do when working on a difficult problem or working alone, and that participants should not attempt to explain what they are doing (engage in self-observing responses) (Ericsson & Simon, 1993). There is a relatively large body of evidence which indicates that such a procedure increases the amount of time required to complete various tasks, but does not affect accuracy. The basic approach of this research is to require participants or groups of participants to perform some task in a "talk-aloud" condition or silently. Little or no difference in performance is demonstrated between these two conditions (Ericsson & Simon, 1993). These types of analyses are compatible with a behavioral approach. As Skinner (1957) wrote: "... covert behavior evokes the same response as the overt behavior because it is essentially the same stimulus except for magnitude" (p. 142).

Recently, several behaviorally oriented researchers have either proposed the use of talk-aloud protocol analyses (Hayes, 1986), or have incorporated them in research (Wulfert et al., 1991). Interestingly, Watson (1920) used talk-aloud procedures in studying problem-solving. The common assumptions underlying this type of research is that covert verbal responses, no different from overt verbal responses, are made overt without loss or confounding of the research at hand. Wulfert et al. (1991) recorded participant's vocal verbal responses and then transcribed them, encoding the various statements into several categories. It is interesting to note that the researchers did not examine the statements in terms of Skinner's elementary verbal operants. Instead, the researchers categorized participant's verbal responses as "relational responses" which included some kind of statement regarding the relationship between two stimuli: "common physical features" which included statements regarding similarities between the comparison and choice, "Stimulus compounds" which included statements which appeared to integrate the sample and choice (e.g., "put together they look like an elephant"), and finally an "other" category for all remaining responses. These researchers found that those participants who generally emitted relational responses demonstrated stimulus equivalence. Participants who did not emit relational responses generally did not demonstrate stimulus equivalence. Wulfert et al. then manipulated the types of training given to participants (either taught to name stimulus compounds or to name relations between stimuli). In general, those in the "relations" group demonstrated stimulus equivalence and the other group did not.

This study will also use a protocol analysis, but will examine the resulting transcripts in terms of the elementary verbal operants. In addition, this study will examine the results in terms of the distinctions between SB and TB responding, especially in light of the possibly that SB responding may in some cases be composed of TB and SB components.

METHODS

Overview

The overall question this research addresses is whether or not response-produced kinesthetic stimulation affects the acquisition of conditional discriminations and the emergence of equivalence. Keeping the task performed within the SB paradigm allowed for the relations to be equated in all ways except for the nature of the actual selection response, the variable which is manipulated in this study. This was accomplished by arranging for one condition in which each choice in a matching-to-sample condition is selected in a stereotypical manner (e.g. click four times in a specified area of the choice stimulus). A second condition was arranged in which a unique response was required for each choice selected, also in a matching-to-sample task. Details are explained below.

Participants

Six participants were recruited by the use of posters. The participants consisted of students attending California State University, Stanislaus. Two males and four females participated in this experiment, with ages ranging from 21 to 41 years old. Applicants were screened to ensure they met the following criteria: (a) over 18 years old, (b) could read and write, (c) had the available time to participate in this study, (d) were unfamiliar with the selection-based and topography-based research area, and (e) were interested in participating in the study once they read the informed consent form. No special effort was made to recruit within a specific age range beyond 18 years, nor was gender a participant selection factor. Only one session per day was conducted with the participants, and they were asked at the start of each session if they felt fit to participate (e.g. "did you get enough sleep?") to determine if the session should be conducted. Successive sessions were not conducted more than two days apart (with the exception of session 5).

Setting

Experimental sessions were conducted in a quiet office in the Classroom Building of California State University, Stanislaus. The office was not being used and contained only a desk, an empty file cabinet and an empty bookshelf. The room had no windows, but was well-ventilated via an air conditioning vent which also functioned to provide masking noise.

A Macintosh Quadra 610 was situated on the desk. The participant sat in front of the computer which arranged all experimental conditions and gathered all data. The participant was alone in the room, with the experimenter waiting just outside the office door.

Apparatus and Materials

All experimental conditions were presented and arranged by a Macintosh Quadra 610 personal computer, programmed by the experimenter using HyperCard 2.2 (Apple Computer, 1989). Participants were taught relations between visual patterns (described as flag-like from this point on) and squares of dot patterns (the tact relation) and relations between nonsense sounds (two syllables) and squares of dot patterns (the intraverbal[1] relation). See Appendix A for an illustration of the patterns used. Nonsense

---

[1]Strictly speaking, this is also a tact relation, because the nonsense sounds are not really verbal stimuli. It would be more accurate to say that two kinds of tact relation are being studied, one based on visual nonverbal stimuli and one based on auditory nonverbal stimuli. However, starting with one of the first studies (Sundberg &

syllables and visual/dot patterns (also displayed in Appendix A) were used to reduce the influence of existing verbal repertoires on the results of this study. All nonsense sounds used were checked by the experimenter and his two assistants for similarity to English words. The flaglike patterns underwent a similar process, with the three reviewers excluding any easily recognizable/namable patterns. Finally, the dot patterns used were created using a HyperCard program which created patterns that were dissimilar (mathematically calculated) and eliminated patterns using the following criteria:

1. No two patterns shared four dots that were in the same positions or differed by only one position (each pattern contained four dots).

2. No two patterns had three dots in the same position with the last dot off by only one position.

Finally, all dot patterns were reviewed by the experimenter and his two assistants to ensure that no pattern was easily namable nor closely resembled any other dot patterns.

Finally in session 5, stimuli similar to the flag-like patterns were constructed, but with the aim of making them less discriminable than those used in sessions 1 to 4. As illustrated in Appendix A, these stimuli consisted of four dimensions: background pattern (with three variations, horizontal, vertical or diagonal lines; shape of inset item (two variations, rectangle or semicircle); location of inset item (two variations, left or right side); and the shading within the inset item (four variations, white, light gray, dark gray and black). All 48 possible patterns (3 x 2 x 2 x 4) were created using these dimensions. Used in session 5, these sample stimuli provided for many similarities among the sample stimuli, and thus were less discriminable. The created patterns were

---

Sundberg, 1990) such relations have been referred to as intraverbal, possibly because vocally produced stimuli are most commonly verbal stimuli, or because in this language training context, such stimuli become verbal if the training is successful. For ease of comparison with other studies, the relation will be called intraverbal in this study as well.

randomly assigned to one of the four conditions used in session 5 (see description below).

Response Definitions

Each participant interacted with the computer program using the computer mouse. Two general responses were required: a selection response with an added topographical component, namely clicking on each dot of the selected square, and a pure selection response, namely clicking twice on the top-left corner and lower-right corner of the square selected. This provided for two different conditions: one in which unique response-produced stimulation accompanied a selection response, and one in which no unique-response produced stimulation accompanied the selection response.

Click on Each Dot

This response type will also be referred to as the point-to-point (PTP) condition in the remainder of this paper. This is the added topographical response to the selection-based task. As each dot pattern has been constructed to be different from every other dot pattern, the form of the response is necessarily different for each pattern (this could be compared to the response used when dialing one's home telephone number as compared to one's office number). Note that in this condition, selecting a stimulus (clicking on each of the dots that makes the stimulus distinctive from the other dot patterns) provides unique response-produced stimulation for each stimulus chosen. The computer determined that each dot was clicked on, and only one time. A 10x10 pixel black square flashed at the mouse-click location. If the black square at least partially covered one of the dots (and not the same one twice), then the computer judged the response correct. If the black square fell outside of a displayed dot, a recording played, requesting that the participant click on each dot. A click outside of a dot required that the sequence of four clicks be started over. A click outside of a dot was not considered an incorrect response. The participants were able to click on the dots in any order they chose. Pilot data indicated that the pattern varied little over repeated exposure to the same dot pattern. Both a tact and intraverbal relation were trained in this condition.

Click on the Corners Response

This response type will also be referred to as the nonpoint-to-point (NPTP) condition in the remainder of this paper. This was the pure selection-based response. All choices in this condition involved this same topography. Four clicks were required to equalize the number of clicks and the amount of time a participant spent focused on a choice stimulus (the amount of exposure). The nature of this response was determined by pilot data - clicking on each of the corners required approximately the same amount of time it took pilot participants to click on each of the dots (the PTP condition). Note that in this condition, selecting the stimulus does not provide unique kinesthetic response-produced stimulation. After each click, the computer determined the accuracy of the click location. A 10x10 pixel black square flashed at the center of each mouse-click location - if the center of the black square (the mouse-click location) was within a 15x15 pixel square defined as the corner of the square, then the response was considered correct. If the response fell outside of the defined corner, then a recording played, reminding the participant to click on the corners (but not tallied as an incorrect response). A click outside of a corner also required that the sequence of four clicks be started over. The sequence of clicks was the same for all participants: click twice on the top-right corner and twice on the bottom-left corner. In addition to ensuring equal exposure time to the selected sample stimulus, this pattern of clicking ensured that the participant was exposed to all the dots in the square as they cut diagonally across the square to accomplish the task. Both a tact and intraverbal relation were trained in this condition.

Intraverbal and Tact Relations

Figure 1 illustrates the tact and intraverbal relations which were trained in this study. The intraverbal consisted of the participant hearing a nonsense word, repeated twice, through the computer speakers. After hearing the sound, the participant was required to select the dot pattern paired with the particular sound. The tact relation consisted of seeing a flag-like pattern displayed (similar to the quilt patterns used by Cresson, 1994) in the center of the comparison stimuli arranged on the computer screen. The participant was then required to select the dot pattern paired with the particular flag-like pattern.

Protocol Analyses

In session 5, participants were asked to talk aloud while performing in both the PTP and NPTP conditions. These vocalizations were recorded, transcribed and encoded. The frequency of various types of statements was calculated. The experimenter and one assistant established criteria, prior to encoding, for classifying these vocalizations into the elementary verbal operants listed below. Once the transcripts were completed (see Appendix B for an example), the participants were recalled to clarify unclear parts of the tape recordings and to indicate what aspect of the situation controlled the emitted response. Responses were labeled as tacts (T) under the control of either a sample or choice stimulus if the participant indicated which stimulus was "referred to" and it was clearly some kind of referential statement. Responses were labeled as repeated intraverbals (RI) if they appeared to be tacts from a previous trial (e.g. when the choice stimulus was made apparent to the participant - in tutorial or

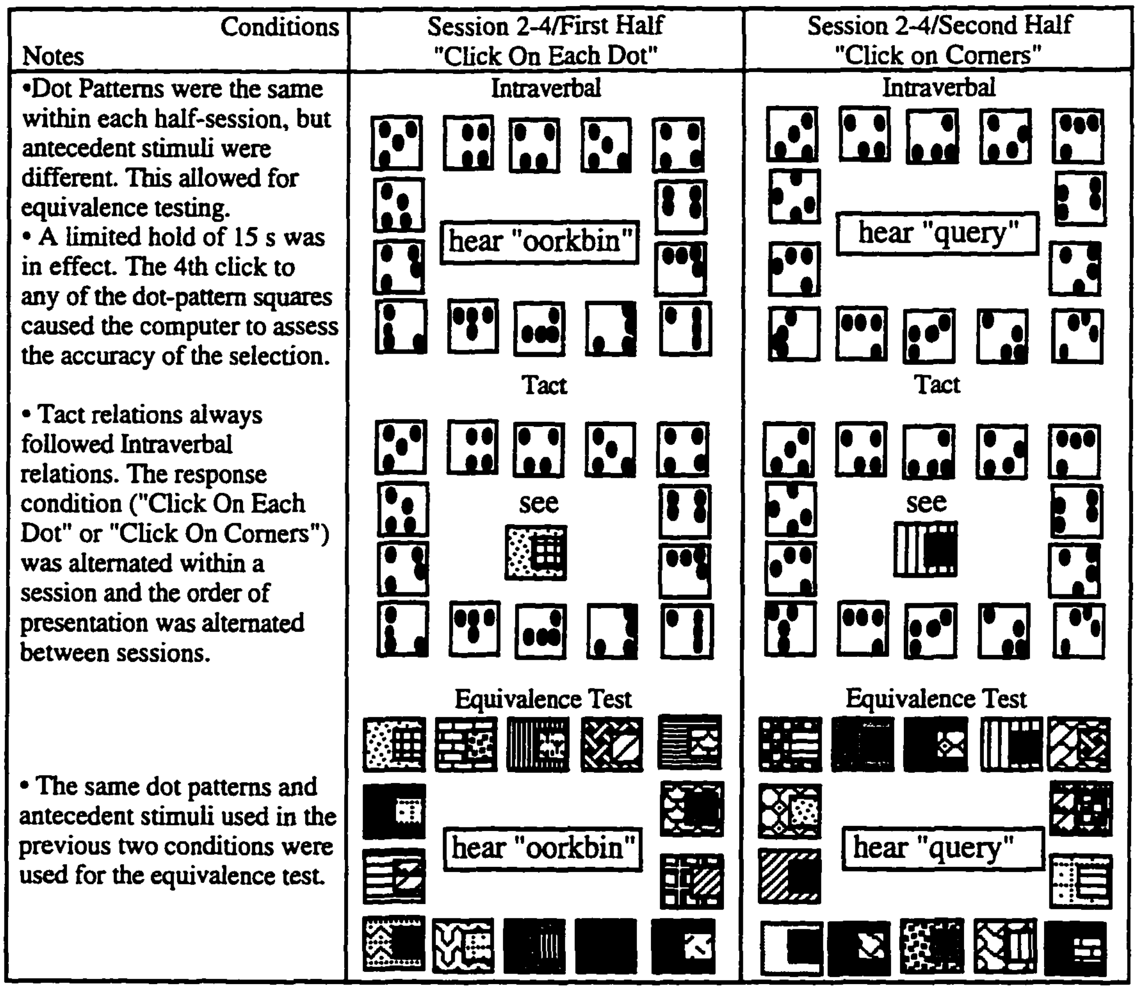


Figure 1. Sequence for Sessions 2- 4. See text for details.

remedial trials), repeated in the trial to be encoded. Repeated tacts or intraverbals were defined as sharing common words within a statement, and referring to the same stimulus as in the original statements (this was confirmed by having each participant review the transcript of his/her protocol session). See the results section for interobserver agreements.

Vocalizations were encoded (operationalized) as follows:

1. "T": Tact of the sample or choice stimulus (when apparent, e.g. tutorial and remedial trials. See description of each of these below).

2. "-": No tact or intraverbal.

3. "RT": Repeated tact to sample stimulus from previous tutorial or remedial trial within the same relation.

4. "RI": Repeated tact to sample stimulus from previous tutorial or remedial trial within the same relation.

5. "NI": New intraverbal. Not a previous tact for this relation.

6. "RET": Repeated "Exam" Tact ("Exam" is used as a synonym for "Test", the proper term, which would be confused with the T used for "Tact"). Tact to sample stimulus in the trial that was encoded, which was repeated from an earlier test trial, correct or incorrect (within same relation).

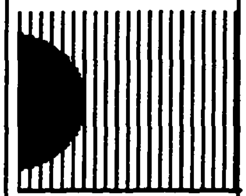
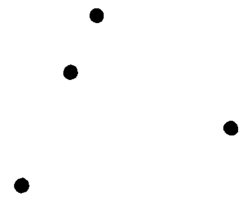
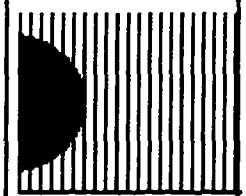
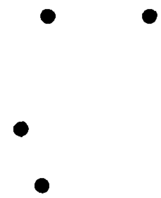
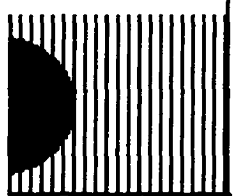
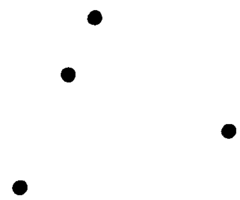
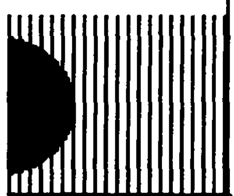
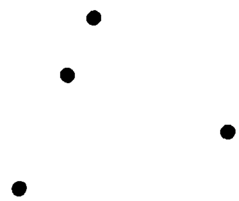
7. "REI": Repeated "Exam" intraverbal. Intraverbal emitted in encoded trial, which was emitted as an intraverbal in an earlier test trial, correct or incorrect (within same relation).

The following transcript from participant B was encoded as indicated to the right of each statement (e.g. T/T and RT/RI). The "/" indicates responses to the sample stimuli (left side) and choice stimuli (right side) - or as an intraverbal in test conditions, when the choice is not indicated.

1. Tutorial: "Gray model T" was encoded as "T/T". The participant noted that Gray referred to the sample stimulus (a tact) and the "Model T" was the choice (a tact)

2. Test 1: "Gray model T, T, T, T, T" was encoded as "RT/RI". As above, this time the "Model T" was not emitted upon seeing the choice stimulus, but upon saying "Gray", thus an intraverbal that was a repeat of an earlier tact (RI).

A more complete example follows in Figure 2. This is a transcript of part of C's performance on a particular relation. This figure also shows the situation in effect when the participant aided in clarifying the transcripts.

TUTORIAL TRIAL "Dark gray semicircle would go with the... mountain shape" T/T T (dark gray semicircle - Tact) /T (mountain shape - Tact)			C
TEST TRIAL "Dark semicircle with the vertical lines ahhh.... " RT/- RT (dark semicircle - repeated tact from tut. trial) /- (.... no comment made)			I
REMEDIAL TRIAL "With the T like pattern" -/T - (no comment made) /T (T like pattern - new tact for this relation as on remedial trials the correct choice is shown to the participant)			C
TEST TRIAL "Vertical lines with the gray semicircle would be the mountain-like pattern" RT/RI RT (Vertical lines with the gray semicircle) /RI (mountain-like pattern). This is classified as an intraverbal as the choice stimulus was not made apparent to the participant			C

Note: The coding is listed in the left most box under the actual transcribed comments. To the right is the sample stimulus present for each trial followed by the comparison stimulus which the participant selected. The column furthest to the right denotes that the trial was C = Correct, I = Incorrect, or T = Timed out. This is a section of the actual form used when participants aided in clarifying the transcripts and to specify what controlled each part of the response. Note, however, that all relations were presented in random order thus, these trials appeared in this order, but were separated by many other relations. This transcript is part of participant C's data.

Figure 2. A Sample Coding of One Representative Relation.

Independent Variables

The tact and intraverbal relations were trained under two different conditions: the "Click on Corners" (NPTP) condition (no unique response-produced stimulation) and the "Click on Each Dot" (PTP) condition (unique response-produced stimulation). Within a session these two conditions were alternated (see Table 3).

Dependent Variables

For each trial (composed of the training or testing of a single intraverbal or tact relation) the latency to the first mouse click, the duration of the four mouse-click response and the correctness of the response were recorded. A tally of the total number correct and incorrect within a block (consisting of exposure to all tact or intraverbal relations within a condition), and within a set (consisting of all blocks participants were exposed to within a condition) were also recorded. In addition, block and set averages were obtained for latencies and durations. The correct choice and the participant's choice were recorded for each trial, allowing for an item analysis (to determine if particular relations were easier/harder to acquire and if the same errors reoccurred).

A response was recorded as correct or incorrect immediately after the participant emitted the fourth correct click to a comparison stimulus. If the choice selected was appropriate for the antecedent presented, a correct choice was registered; otherwise, the response was considered incorrect.

In the PTP condition, the pattern in which participants clicked on the four dots was also recorded, to determine if consistent patterns were being used (that a consistent topographical response occurred).

In session 5, additional measures were implemented. In the NPTP condition, the movement of each participant's mouse was recorded to determine if topographical responses were occurring to choice stimuli in this condition. These movements were classified as a topographical response to the choice stimulus if the mouse pattern was seen to connect at least three dots together or a stereotypical movement appeared to occur. Also in this session, a protocol analysis was conducted. Participants were asked to "talk aloud" while engaged in both the PTP and NPTP conditions of this experiment. These vocalizations were recorded and encoded as noted above.

Procedure

Overview

In general, participants were exposed to alternating PTP and NPTP conditions in each session. The order of presentation of these conditions alternated with each subsequent session. Session 1 was used primarily to familiarize participants with the computer task and to gather data with no equivalence testing in effect. Participants were exposed to all conditions of the experiment in this session, with the exception of the equivalence testing (described below). In sessions 2 to 3, acquisition was examined for both the PTP and NPTP conditions, in addition to testing for the emergence of equivalence. In these sessions, sample stimuli (nonsense sounds or flag-like patterns) were individually matched to the same set of dot patterns. In equivalence testing, the participant was required to match the appropriate flag-like pattern to a sample nonsense sound (that is, to determine which flag-like pattern had been paired with the same dot-pattern that the sound presented had been paired with). Session 4 was similar to sessions 2 and 3, with the exception that a mastery criterion was added to the PTP and NPTP training received prior to equivalence testing. Finally in session 5, more difficult (harder to name) sample stimuli were used in an attempt to reduce this naming behavior by participants. A protocol analysis was also conducted to determine the extent and nature of this vocal verbal behavior.

Each participant attended four or five sessions, as described below. Each session lasted between one and two hours. All contingencies that were in effect are outlined below. All experimental conditions are summarized in Table 3.

Table 3

A Summary of All Experimental Conditions by Session

Session	Conditions
1. Pre-training and Placement *	A. Informed Consent B. Introductory Training: 2 sets (4 relations-2 blocks each) Set 1: "Click On Corners"; Intraverbal Set 2: "Click On Each Dot"; Tact C. Pre-training/placement: 4 sets (14 relations - 1 tutorial and 3 test blocks each) Set 1: "Click On Corners"; Intraverbal Set 2: "Click On Each Dot"; Tact Break - 15 minutes Set 3: "Click On Corners"; Tact Set 4: "Click On Each Dot"; Intraverbal
2. Training and Equivalence *	A. Training #1 (all training sets consist of 1 tutorial and 4 test blocks) Set 5. "Click On Each Dot"; Intraverbal; ** relations Set 6. "Click On Each Dot"; Tact; ** relations B. Equivalence Tests using stimuli from set 5 and 6 above Break - 15 minutes C. Training #2 Set 7. "Click On Corners"; Intraverbal; ** relations Set 8. "Click On Corners"; Tact; ** relations D. Equivalence Test using stimuli from set 7 and 8 above
3. Training and Equivalence *	A. Training #3 (all training sets consist of 1 tutorial and 4 test blocks) Set 9. "Click On Corners"; Intraverbal; ** relations Set 10. "Click On Corners"; Tact ** relations. B. Equivalence Tests using stimuli from set 9 and 10 above Break - 15 minutes C. Training #4 Set 11. "Click On Each Dot"; Intraverbal; ** relations Set 12. "Click On Each Dot"; Tact; ** relations D. Equivalence Test using stimuli from set 11 and 12 above

Table 3 - Continued

Session	Conditions
4. Training to Mastery and Equivalence Testing *	A. Training #5 Set 13. "Click On Each Dot"; Intraverbal; 10 relations-*** blocks Set 14. "Click On Each Dot"; Tact; 10 relations-*** blocks B. Equivalence Tests using stimuli from set 13 and 14 above Break - 15 minutes C. Training #6 Set 15. "Click On Corners"; Intraverbal; 10 relations-*** blocks Set 16. "Click On Corners"; Tact; 10 relations - *** blocks D. Equivalence Test using stimuli from set 15 and 16 above E, Exit interview
5. Training tact relations only, using harder to discriminate sample stimuli. Four participants only.	A. Training #7 Set 17. "Click on Corners"; Tact; 12 relations - 5 blocks Set 18. "Click On Each Dot"; Tact; 12 relations - 5 blocks Break 15 minutes Set 19. "Click on Corners"; Tact; 12 relations - 5 blocks Set 20. "Click On Each Dot"; Tact; 12 relations - 5 blocks B. Exit interview #2 and Protocol Analysis
Notes	* Participants 4 - 6 had the alternate response condition first. ** The number of relations trained was determined by the participant's performance in the pre-training/placement condition. See text for details. *** The number of blocks varied on the participants' ability to reach mastery (3 successive blocks with no incorrects).

Session 1 - Pre-training and Placement

The participants were introduced to the study by asking them to read the informed consent form (see Appendix B). Once they read it and demonstrated that they understood it by answering several questions, pre-training occurred. With the experimenter present, the participant was exposed to the task on the computer, but with only four comparison stimuli (the dot patterns) arranged on the screen. This introductory training allowed the participant to become accustomed to the nature of the task. Prior to starting each of the tasks, a set of instructions was provided by the computer. For the "Click on Each Dot" (PTP), condition "Click on each dot in the square of your choice" was heard. For the "Click On Corners" (NPTP) condition, "Click on the top-left corner twice, then click on the bottom-right corner twice, of the square of your choice" was heard. These same instructions were used throughout the experiment at the start of each set and each block within a set. A set is composed of blocks, and blocks are composed of one trial for each of the relations trained (i.e. 10, 12, or 14 relations). A set could also be called a condition as each new set resulted in a condition change.

For three participants, the introductory training consisted of the following (the other three participants received an identical arrangement except "Click On Corners" was arranged first):

1. One block - tutorial intraverbal relation ("Click on each Dot", four relations only).

2. One block - test intraverbal relation ("Click on each Dot", four relations only).

3. One block - tutorial tact relation ("Click on Corners", four relations only).

4. One block - test tact relation ("Click on Corners", four relations only).

"Tutorial" consisted of the computer presenting the antecedent stimulus, (e.g. in block 1, an auditory nonsense word), then highlighting the correct match for that sound. "Test" consisted of the computer presenting the antecedent stimulus, but without highlighting the correct stimulus. In either case, the participant was required to emit the correct response to the correct comparison stimulus (e.g. in block 1, clicking on each dot in the square). Once the participant completed this response, the computer provided the auditory feedback "Correct" if the correct stimulus was selected. As in all conditions mentioned from this point on, the participant had 15 s in which to respond. If 15 s elapsed without a response, the computer provided the auditory feedback

"Incorrect, try again." At this point, the computer repeated the nonsense word or re-presented the antecedent visual pattern (flag-like pattern) and highlighted the correct dot pattern (a remedial trial). Once the participant emitted the correct response to the highlighted comparison stimulus, the next relation in the block was presented. A mouse click within one of the dot patterns prior to the 15 s time-out period extended the time-out period by 5 s to reduce the number of time-outs occurring while participants were in the process of emitting a response to a comparison stimulus. A response to an incorrect comparison stimulus had the same effect as a time-out: the computer presented the auditory "Incorrect" and then provided remediation training as described above.

Finally, each correct response resulted in a brief intertrial interval (less than 1 s) in which all stimuli were removed from the screen and re-presented in a scrambled order on the screen . This "scrambling" was randomly determined by the computer. The rectangular shape the dot patterns made on the screen was retained: only the positions of each dot pattern within the rectangle changed (see Appendix A). This procedure was implemented to prevent positional cues from aiding discriminations, and was used after a correct response occurred and at the start of each new block.

Finally, at the start of each block of trials, the order of the stimuli used as antecedents was randomized, thus providing for a unique presentation order for those stimuli for each block the participant completed.

During this condition, the experimenter only provided the following information:

1. A 15 s limited hold is in effect.

2. Switching between comparison stimuli is allowed as long as it occurs before the fourth click (the computer monitored this).

3. Given the monetary contingencies in effect (see below), it is best to work quickly and accurately.

4. Each sound is only related to one dot or flag-like pattern.

5. If a mouse-click location error occurs, the four-click sequence must be repeated.

6. Clicking outside of a pattern on the screen has no effect.

Following this introduction the following conditions were in effect for three participants:

1. Intraverbal, Click On Each Dot, one set composed of one tutorial block and three test blocks. Fourteen relations were trained.

2. Tact, Click On Corners, one set composed of one tutorial block and three test blocks. Fourteen relations were trained.

3. A break period of approximately 15 minutes.

4. Tact, Click On Each Dot, one set composed of one tutorial block and three test blocks. Fourteen relations were trained.

5. Intraverbal, Click On Corners, one set composed of one tutorial block and three test blocks. Fourteen relations were trained.

Different dot patterns were used for both the intraverbal and tact relations, unlike sessions 2, 3, and 4 in which equivalence tests were conducted. The other three participants received the "Click On Corners" condition first, then alternated with the "Click On Each Dot" condition, with all other arrangements staying the same. This allowed for half of the participants to receive training with unique response-produced stimuli first and the other half to receive this type of training second. Although the experimental question was addressed using a within-subjects design, this arrangement provided some useful between-subject information.

The pre-training session was designed to familiarize the participants with the computer task. Data from pilot participants indicated that the condition received first seemed to be the one performed most poorly. It was hoped that this initial session would reduce this effect by providing a relatively intensive training - however, as noted in the results section, this was not the case. The pre-training session was also designed to determine the number of relations to be used for each individual participant. Again, data from pilot participants showed a broad range of performances, resulting in a range of responding from nearly perfect (totally correct) to nearly totally incorrect responding. For participants who scored poorly in this initial session, a reduction in the number of relations to be learned made the task easier. For those participants who scored quite high in this initial session, an increase in the number of relations to be learned resulted in more errors and thus more useful data for this experiment. As it turned out, only one participant required a reduction in the number of relations used (participant S). The following criteria was used to determine the number of relations used - based on the last two sets (#4 and #5 above) described in Table 3:

1. An average of 12, 13, or 14 incorrects for both sets resulted in using 10 relations.

2. An average of 9, 10, or 11 incorrects for both sets resulted in using 12 relations.

3. An average of 6, 7, or 8 incorrects for both sets resulted in using 14 relations.

4. An average of 3, 4, or 5 incorrects for both sets resulted in using 16 relations.

5. An average of 0, 1, or 2 incorrects for both sets resulted in using 18 relations.

Regardless of the number of relations used, the appearance of the comparison stimuli on the screen was similar; that is, they formed a rectangle with the tact antecedent stimuli appearing in the center of the rectangle. Once the number of relations was established for a participant, this number remained in effect for the next two sessions. During the final session, all participants were only required to learn 10 relations, but were required to learn them to a mastery level (see session 4).

A summary of conditions in effect throughout the remaining four sessions is provided here:

1. All choices must have occurred within 15" or (20" if an initial response was made) or the trial was considered incorrect.

2. An incorrect response was followed by the sound "Incorrect, try again" and a “remedial” trial occurred in which the sample stimulus was re-presented and the participant was required to emit the correct response (the correct comparison stimulus was highlighted).

3. Correct responses (except in the equivalence condition - see below) were followed by the sound "Correct." The positions of dot patterns within the rectangular shape on the screen were then scrambled prior to the start of the next trial.

4. Clicks which did not fall on a dot (for the PTP condition) or in the defined corner (NPTP condition) resulted in each sound respectively: "Click on each dot" and "Click on the top left and bottom right corners."

5. "Tutorial" referred to the first block of a set where the relations were shown to the participant (the correct dot pattern for each antecedent was highlighted), and the participant was required to emit the correct response.

6. "Test" refers to an antecedent stimulus being presented with no highlighting of the correct answer.

7. At the start of each block of trials, the order of the stimuli used as antecedents were scrambled, thus providing a unique presentation order for those stimuli for each block the participant completed.

8. The incentives provided to the participant were as follows (including session 1): (a) $.02 for each correct choice, excluding tutorial and remedial choices, (b) $.10

for each minute under 90 that the participant completed the session, and (c) $5 for completing the session. This incentive was designed to foster speed and accuracy. Participants earned in a range of $8 to $14 per session. All participants were paid immediately following each session.

Session 2 - Training and Equivalence Testing

This session was conducted as illustrated in Figure 1. Unlike session 1, in the next three sessions, intraverbal training always preceded tact training. Since a comparison of tact and intraverbal ease of learning was not an issue in this study, the relations were kept in this fixed order to simplify the design . The number of relations used in this session was determined as discussed above. The first condition the participants were exposed to was determined by alternating it with the first condition they received in the previous session (this will also hold for the remaining two sessions). Thus, if a participant received the "Click On Corners" condition in the first block of the previous session (which was determined by random assignment), he/she started session 2 with the "Click on Each Dot" condition. See Table 3 and Figure 1 for an arrangement of conditions presented in session 2.

The first two trials of each half of the session (composed of intraverbal and tact relations) were conducted using the same dot patterns for each. This was necessary for testing the equivalence relation. Thus, after one tutorial trial and four test trials on the intraverbal relation, clicking on each dot, a second set was composed of one tutorial trial and four test trials of the tact relation, clicking on each dot, and using the same dot patterns as used in the intraverbal relation. The order of presenting the intraverbal first and the tact second was partially determined by pilot data indicating that the tact relation was somewhat more difficult to acquire. Since the same dot patterns were used across the different relations, requiring the tact to be second increased accuracy, probably due to the participants prior exposure to the dot patterns during the intraverbal relation.

Once intraverbal and tact tutorial and testing trials were complete, participants were administered an equivalence test. This condition started with the instructions "Click on the pattern which seems most appropriate for the sound you hear." In addition, prior to the start of session 2, participants were informed to expect that some relations would be presented in which no feedback as to the correctness of their responses would be provided. They were all informed, however, that performance in this condition contributed to their overall number of corrects and was included in the incentive contingency. Once the instructions for the equivalence condition were delivered, all the flag-like patterns used in the tact relation were arranged in a rectangle similar to the one used for the dot patterns (see Figure 1). One of the nonsense sounds used for the preceding intraverbal relation training was presented, and the participant had 20" to respond by clicking on one of the flag-like patterns. A selection was registered by the computer by sounding a brief tone, but no feedback was given as to the correctness of the response. After a response, a brief intertrial interval occurred in which the flag-like stimuli were randomly ordered within the rectangular shape on the screen and at which time they were re-presented and a new nonsense word also presented. If a trial timed out, the next nonsense sound was presented. This condition was completed when all nonsense sounds had been presented once. The order of presentation for the nonsense sounds was arranged by a random function of the computer at the start of the condition.

As indicated in Table 3, this session consisted of a first half, composed of intraverbal and tact conditions, (both either "Click On Corners"/NPTP or "Click On Each Dot"/PTP), followed by a equivalence test. A fifteen minute break followed in which the participant could stretch his/her legs, get a drink of water, etc. The second half of the session was identical to the first, except the conditions both consisted of whatever response type was not used in the first half (e.g. "Click On Corners"/NPTP or "Click On Each Dot"/PTP).

The following summary lists additional conditions that were in effect throughout the remaining three sessions:

1. Intraverbal relations preceded tact relations and both used the same dot patterns for comparison stimuli. Once both were complete, an equivalence test followed (except for session 5 which consisted of only tact relations).

2. The equivalence test consisted of arranging the flag-like patterns used during the immediately preceding tact condition in a rectangle on the screen, then presenting one of the nonsense syllables from the immediately preceding intraverbal condition. A response within 20 s was followed by a brief tone, but no feedback as to whether or not the selection was correct. The sample stimuli used in the preceding two trials were used for these tests.

3. Each half of the session was dedicated to either the "Click On Corners" (NPTP) or "Click On Each Dot" (PTP) conditions, with the condition presented first alternating with the condition received first in the previous session.

Session 3 - Training and Equivalence Testing

This session was identical to session 2, with the exception that the order of the response-type conditions was reversed. Thus, if in session 2 "Click on Corners" (NPTP) was first, then in this session, "Click on Each Dot" (PTP) was first.

Session 4 - Training to Mastery and Equivalence Testing

All conditions in effect during this session were identical to sessions 2 and 3, with the exception that only 10 relations were trained. This reduction in relations used was to allow for mastery to occur within a reasonable time period. Mastery was defined as three successive blocks in which no incorrect responses occurred. This mastery requirement was adopted from Stratton (1992). The mastery requirement was imposed to replicate similar procedures used in stimulus equivalence research (Sidman, 1994). Note that in sessions 2 and 3, mastery was not required before equivalence was tested.

Session 5 - Training Using Less Discriminable Sample Stimuli

After sessions 1 to 4, it became apparent that all participants were relying heavily on topography-based (overt or covert vocal-verbal behavior) verbal behavior to perform correctly on test trials. This use of verbal behavior was believed to obscure any differences which might exist between the PTP and NPTP conditions. Thus, more difficult sample stimuli (less discriminable) were introduced to try and reduce the vocalizations which might occur. Session 5 was very similar to previous sessions, with the following exceptions:

1. Only tact relations were used. For a description of the stimuli used in these relations, see the section "Materials and Apparatus".

2. New patterns were used in each condition (see table 3), thus no equivalence testing was conducted.

3. The number of blocks was extended to a total of six: one tutorial block and five test blocks for each condition.

4. The number of stimuli was reduced to 12 to allow for completion of the session in a timely manner (based on pilot data).

Finally, this session ended with each participant repeating the last two conditions (PTP and NPTP, in reverse order) of this session, but this time talking aloud as they worked - the protocol analysis. The order of the sessions were reversed so that the participant's most recent vocalizations were examined first, before engaging in more relations. Only one tutorial and two test trials were arranged for this analysis. All vocalizations were recorded and transcribed. Participants were then asked to clarify parts of the transcript.

Experimental Design

This study utilized an alternating-treatments design (Kazdin, 1982), in which the "Click on Corners" (no unique response-produced stimulation) condition was alternated with the "Click on Each Dot" condition (unique response-produced stimulation). Unlike most alternating-treatment designs in the applied literature, the dependent variable was a learning rather than a performance variable, thus the design is also a form of repeated acquisition.

RESULTS

Individual Data

The number of incorrects in the tutorial conditions are not reported in this section as only a few errors were made overall. Recall that in the tutorial condition, the participants were shown the correct choice stimulus for each sample stimulus. In addition, the duration each participant spent on the choice stimuli (from first click to last click in both the PTP and NPTP conditions) are not reported as those data were very stable across all conditions and participants. Averaged across sessions 1 to 5, the difference in averaged durations between the PTP and NPTP conditions was never greater than 0.5 s for any participant. Averaged across all participants, the difference between the means of the PTP and NPTP conditions showed little difference (across sessions 1 to 4 the PTP mean was 2.9 s and 3.1 s for the NPTP condition). These data provide verification that exposure to the choice stimuli did not differ between the PTP and NPTP conditions.

For each participant a table is included which summarizes block by block data for each session. Considering the similarity of results across participants, only participant B's data will be examined in detail. For the remaining participants, results common to all participants will only be briefly noted, while results unique to each participant will be examined more thoroughly.

Participant B

The main interest of this research was to examine the difference in performance between the PTP and NPTP conditions. B performed best overall in the PTP condition (137 errors) versus the NPTP condition (147 errors), across all sessions. However, an examination of Table 4 reveals that this difference is probably not due to condition differences, but rather to which condition was in effect second. For each session B participated in, B performed best in the condition which occurred second within a session. Thus, B performed best in the PTP condition in sessions 1, 3, 5 and best in the NPTP conditions in sessions 2 and 4. This same pattern was shown with all participants. For this participant, most of the difference between the PTP and NPTP conditions was contributed by the very first situation encountered in a session, namely the PTP or NPTP intraverbal (IV) condition. Particularly illustrative of this are sessions 2 and 3, in which most of the errors in the first condition are contributed by the IV condition (see Table 4). One probable contributor to this effect is that the same choice stimuli are used in the immediately following tact condition (thus the participants were more familiar with those patterns).

Another possible reason is that the participants may have needed to "warm-up", most likely by generating mediating verbal responses to the sample and choice stimuli (see the protocol analysis later in this section). In this context, "warm-up" refers to adjusting to the experimental conditions. An analogy might be drawn to the initial minutes a pigeon is placed in an experimental chamber, at which time responding may be inaccurate or the pigeon may not be fully involved with the arranged task. Although entirely new sample and comparison stimuli were introduced in the second condition of sessions 2 to 4, the number of errors did not increase to the level observed in the first condition for that session (the first IV condition, whether PTP or NPTP). This would seem to give some support to the "warm-up" effect, although it could also be a function of improving performance over time (see Table 4). Performance in sessions 1 and 5 do between the PTP and NPTP conditions. B performed best overall in the PTP condition (137 errors) versus the NPTP condition (147 errors), across all sessions. However, an examination of Table 4 reveals that this difference is probably not due to condition differences, but rather to which condition was in effect second. For each session B participated in, B performed best in the condition which occurred second within a session. Thus, B performed best in the PTP condition in sessions 1, 3, 5 and best in the NPTP conditions in sessions 2 and 4. This same pattern was shown with all participants. For this participant, most of the difference between the PTP and NPTP conditions was contributed by the very first situation encountered in a session, namely the PTP or NPTP intraverbal (IV) condition. Particularly illustrative of this are sessions 2 and 3, in which most of the errors in the first condition are contributed by the IV condition (see Table 4). One probable contributor to this effect is that the same choice stimuli are used in the immediately following tact condition (thus the participants were more familiar with those patterns).

Another possible reason is that the participants may have needed to "warm-up", most likely by generating mediating verbal responses to the sample and choice stimuli (see the protocol analysis later in this section). In this context, "warm-up" refers to adjusting to the experimental conditions. An analogy might be drawn to the initial minutes a pigeon is placed in an experimental chamber, at which time responding may be inaccurate or the pigeon may not be fully involved with the arranged task. Although entirely new sample and comparison stimuli were introduced in the second condition of sessions 2 to 4, the number of errors did not increase to the level observed in the first condition for that session (the first IV condition, whether PTP or NPTP). This would seem to give some support to the "warm-up" effect, although it could also be a function of improving performance over time (see Table 4). Performance in sessions 1 and 5 do

Table 4

Number of Incorrects per Block for Participant B

Session Number	Relations Trained	Condition in Order of Presentation	# of Incorrects per Test Block Block = one presentation of each relation 1	2	3	4	5	6	7	Totals	
1	14	IV-NPTP	11	14	8					33	NPTP
		Tact-PTP	10	10	9					29	67
		Tact-NPTP	13	11	10					34	PTP
		IV-PTP	9	4	2					15	44
2	14	IV-PTP	11	7	5	1				24	PTP
		Tact-PTP	7	5	1	0				13	37
		Equivalence	11								
		IV-NPTP	4	3	1	0				8	NPTP
		Tact-NPTP	1	1	1	0				3	11
		Equivalence	1								
3	14	IV-NPTP	11	6	5	0				22	NPTP
		Tact-NPTP	3	0	0	0				3	25
		Equivalence	0								
		IV-PTP	3	3	2	1				9	PTP
		Tact-PTP	5	1	0	3				9	18
		Equivalence	3								
4	10	IV-PTP	2	0	0	1	0	0	0	3	PTP
		Tact-PTP	1	1	0	0	0			2	5
		Equivalence	0								
		IV-NPTP	0	0	0					0	NPTP
		Tact-NPTP	0	0	1	0	0	0		1	1
		Equivalence	0								
5	12	Tact-NPTP	8	6	5	2	1			22	NPTP
		Tact-PTP	9	5	1	3	0			18	43
		Tact-NPTP	9	5	3	3	0			20	PTP
		Tact-PTP	8	5	2	0	0			15	33

not support this hypothesis however, as in those sessions, the number of errors increased near to the number observed in the first IV condition, after a short break occurred. Recall that in sessions 1 and 5, new sample and comparison stimuli were introduced for each condition (as no equivalence testing took place). It is unclear why these two different types of "warm-up" (with and without the same choice stimuli) would differentially affect performance, but they appear to have done so, although only to a slight extent.

When comparing sessions 2 and 3 only (the two most similar sessions), B performed better overall in the NPTP condition. The difference between the PTP and NPTP conditions was relatively large in session 2 (26 errors) and relatively small, and in the opposite direction, in session 3 (7 errors). However, when one takes into consideration the general improvement across sessions, it would appear that the difference is likely to be due to the improving performance across sessions in combination with the "warm-up" effect. In both exit interviews (after session 4 and session 5), B stated a preference for the PTP condition: "I preferred clicking the dots in the pattern (PTP). It made it easier to quickly identify the pattern under stress of time" (parenthetical text added).

Within conditions, B showed a general improvement across blocks (see Table 4). By session 2, B was reaching nearly perfect performance by the last test block. This last block performance is interesting when one considers B's equivalence performance. The first time B encountered the equivalence condition (session 2), performance was poor, 11 out of 14 incorrect responses. Yet, as noted already, the number of errors in the last block for the previous two sessions was very low, only 1 in the IV condition and none in the tact condition, suggesting mastery had occurred. In the next equivalence condition, B's performance improved dramatically (only 1 error), which would seem to indicate that B employed some strategy in the second equivalence condition. This was supported by B's exit interview in which B was asked to recall strategies/techniques that were used in the equivalence condition. B wrote: "Try to remember the relations between the first and second parts. If you identified the dot patterns and continued it into the flag series (the tact condition), it was easier" (parenthetical statements added). B clarified this later, indicating that the dot patterns were given a "name" and this name was carried over into the tact condition, which allowed B to select correct choices in the equivalence sessions. It is important to note that B (and any other participant) was not informed of the nature of the equivalence tests, yet each were easily able to describe the type of relationship tested in that condition when asked in exit interviews.

Across sessions, B's performance generally improved, as indicated by Table 4. As expected, in session 5, B's performance worsened. In this session, stimuli that were harder to discriminate were introduced in an attempt to prevent participants from naming the stimuli (a strategy all participants claimed to use when asked in the exit interviews). B performed the best in session 5, as compared to all other participants, and also had the most consistent type of verbal statements preceding correct choices (see the protocol analysis).

B performed best in the tact condition, as compared to the IV condition, for sessions 2 and 3, and had equal tact and IV performance in session 4. In session 1, B performed better in the IV condition. If B was using the strategy described above of carrying stimulus names over from the IV condition, one would expect better performance in the tact condition. However this better performance could also be due to the shared choice stimuli between the IV and tact conditions. In both exit interviews, this participant specified the tact condition was the one in which the best performance probably occurred.

For all participants, averaged test and tutorial latencies per session, for each condition, were reported. For participant B these data are summarized in Table 5. Block by block data are not shown as these demonstrated only slight variation (generally a gradual, but very slight, decrease across blocks).

Table 5

Summary Data for Participant B

		Sess. 1	Sess. 2*	Sess. 3	Sess. 4*	Sess. 1-4	Sess. 2-3	Sess. 5
Lat. PTP	IV	4.2/2.4	3.9/2.4	4.2/3.0	2.5/2.3	3.6/2.6	4.1/2.7	
	Tact	6.1/2.2	5.8/2.5	5.2/3.1	3.7/1.9	5.2/2.7	5.5/2.9	6.1/3.0
	Tot PTP	5.1/2.5	4.9/2.7	4.7/3.1	3.0/2.2	4.3/2.8	4.8/2.9	
Lat. NPTP	IV	4.3/2.4	4.4/3.2	5.3/3.0	2.6/2.2	4.3/2.9	4.8/3.1	
	Tact	6.9/2.3	5.1/2.5	5.7/3.1	3.3/1.5	5.1/2.7	5.4/2.8	7.0/3.2
	Tot NPTP	5.6/2.7	4.7/2.9	5.5/3.0	3.1/1.8	4.8/2.9	5.1/3.0	
Tut. Lat PTP	IV	0.6/0.3	0.6/0.4	1.0/0.8	1.0/0.5	0.7/0.6	0.7/0.6	
	Tact	1.4/0.8	1.4/0.6	1.8/1.1	2.2/0.8	1.5/0.9	1.5/0.9	1.7/1.2
	Tot PTP	1.0/0.8	0.9/0.7	1.4/1.1	1.6/0.9	1.1/0.9	1.1/0.9	
Tut. Lat NPTP	IV	0.7/0.5	0.7/0.4	1.1/0.7	1.0/0.5	0.8/0.6	0.9/0.6	
	Tact	1.0/0.6	2.3/1.1	2.5/0.9	2.4/0.9	1.7/1.1	2.4/1.0	1.8/0.9
	Tot NPTP	0.8/0.6	1.4/1.1	1.5/1.0	1.7/1.0	1.2/0.9	1.5/1.1	
Eq. Lat	PTP	5.5/2.0	5.1/3.7	1.4/1.8	4.2/3.2	5.3/3.0		
	NPTP	3.3/2.7	3.7/3.1	1.7/1.3	3.0/2.7	3.5/3.0		
Notes	* signifies the PTP condition occurred first. Lat: Latency; Tut: tutorial condition; Dur: duration, Eq: equivalence condition. Numbers to the left of the "/" indicate mean latencies in seconds per block. Numbers to the right of the "/" indicate the SD.							

As can be seen in the "Latency IV" and "Latency Tact" rows in Table 5 (in both the PTP and NPTP conditions), relatively large latency differences existed between these conditions. It is likely, however, that these differences are artifacts of the procedure. In the IV condition, the participant heard the nonsense word twice and then the latency timer started. In the tact condition, the flag-like pattern appeared on the screen and then the latency timer started. It can be seen that the IV latencies were likely to be shorter than tact latencies, as the participants were able to start their search, untimed, as soon as they heard the first utterance of the nonsense word.

For participant B, there was only a slight difference in test latencies between the PTP and NPTP conditions, with the PTP latencies less than the NPTP latencies (averaged across both sessions 1 to 4 and 2 to 3).

While little difference existed between the tutorial latencies in the PTP (1.1 s) and NPTP (1.5 s) conditions, B's latencies in this condition over subsequent sessions actually increased, as noted in Table 5 in the "Tutorial Latency" rows toward the bottom of the table. The largest increases occurred between sessions 2 and 3 for the PTP conditions and between sessions 1 and 2 for the NPTP condition.

Equivalence latencies for participant B decreased with subsequent sessions, with the exception of session 3, which showed a slight increase. Over sessions 1 to 4, the PTP equivalence latency was considerably higher than the NPTP condition. As noted in the group report (later in this section), B's performance follows the general trend of having longer latencies when equivalence incorrects were relatively high.

Participant's C, M, P, S and W

Tables are provided for each participant, summarizing block, condition and session number of incorrects (see Tables 6 to 14). As noted previously B performed best in the PTP condition, as did participants C and W (see Tables 4, 6, and 14 respectively). Participants M, P and S performed best in the NPTP condition (see Tables 8, 10 and 12 respectively). It should be noted that S showed an odd session 4 PTP performance (48 errors, see Table 12). This participant took an inordinate number of blocks to complete the mastery portion of sessions 4 (a total of 35 blocks). Recall that in this session mastery was required (defined as three consecutive blocks with no errors). S apparently was unaware of this criterion as the number of blocks required for S to reach mastery was 19 and 16 respectively in the PTP IV and PTP tact conditions (all other participants reached mastery in a maximum of 8 blocks for all conditions). In the exit interview, S reported that up to the second part of the 4th session (the NPTP condition), performing was "like reading the Sunday newspaper, casual not focused". It seems likely that S's PTP and NPTP performances would have been nearly equal if the misunderstanding of the instructions had not occurred. W also showed an odd performance, in this case in session 2 (see Table 14). W performed very poorly in the NPTP condition, which was presented first in this session. It is not clear why this was the case. Overall, none of the tot PTP/NPTP differences (summed over all sessions for each participant) were large, however, with differences ranging from 3 to 49 incorrects (for participants P and M, respectively). P's total errors (PTP and NPTP) amounted to 573, and M's to 449. It can be seen that little difference existed between these two conditions. As noted previously, all participants demonstrated the highest number of errors in the condition presented first (PTP or NPTP).

C and W were the only participants who did not show a resurgence of errors in sessions 1 or 5. The others, B, M, P and S, showed this increase shortly after the arranged 15 minute break during sessions 1 or 5 (for participants B, M, and P this

Table 6

Number of Incorrects per Block for Participant C

Sess. #	Rel. Trained	Condition in Order of Presentation	Number of Incorrects per Test Block Block = one presentation of each relation							Totals	
			1	2	3	4	5	6	7		
1	14	IV-NPTP	10	13	9					32	NPTP
		Tact-PTP	9	10	8					27	57
		Tact-NPTP	9	8	8					25	PTP
		IV-PTP	9	6	4					19	47
2	14	IV-PTP	5	7	3	4				19	PTP
		Tact-PTP	9	6	0	2				17	36
		Equivalence	9								
		IV-NPTP	6	4	0	3				13	NPTP
		Tact-NPTP	6	4	2	2				14	27
		Equivalence	6								
3	14	IV-NPTP	7	4	0	3				14	NPTP
		Tact-NPTP	4	0	0	0				4	18
		Equivalence	2								
		IV-PTP	6	1	1	2				10	PTP
		Tact-PTP	4	2	2	1				9	19
		Equivalence	3								
4	10	IV-PTP	4	2	0	0	0			6	PTP
		Tact-PTP	0	1	0	0	0			1	7
		Equivalence	1								
		IV-NPTP	3	2	0	0	0			5	NPTP
		Tact-NPTP	3	2	0	0	0			5	10
		Equivalence	0								
5	12	Tact-NPTP	10	9	7	6	7			39	NPTP
		Tact-PTP	10	11	7	5	7			40	78
		Tact-NPTP	10	11	10	5	3			39	PTP
		Tact-PTP	10	7	3	5	1			26	66

Table 7

Summary Data for Participant C

		Sess. 1	Sess. 2*	Sess. 3	Sess. 4*	Sess. 1-4	Sess. 2-3	Sess. 5
Lat. PTP	IV	5.3/2.9	5.1/3.0	4.6/3.4	3.1/3.1	4.5/3.3	4.9/3.2	
	Tact	6.7/2.1	6.9/3.4	4.9/2.4	3.6/2.4	5.6/3.0	5.9/3.1	5.5/2.6
	Tot PTP	6.0/2.6	6.0/3.3	4.8/3.0	3.3/2.8	5.0/3.2	5.4/3.2	
Lat. NPTP	IV	5.9/2.4	6.2/3.7	5.4/3.7	2.9/2.5	5.1/3.4	5.8/3.7	
	Tact	8.1/2.8	6.0/2.7	4.7/2.2	3.9/2.5	5.6/3.0	5.4/2.6	6.3/2.5
	Tot NPTP	7.0/2.8	6.1/3.2	5.1/3.1	3.4/2.5	5.4/3.2	5.6/3.2	
Tut. Lat PTP	IV	1.0/0.7	0.6/0.4	0.9/0.5	0.9/0.3	0.8/0.6	0.7/0.5	
	Tact	1.7/0.4	1.5/0.6	1.5/0.8	1.1/0.6	1.5/0.6	1.5/0.7	1.3/0.4
	Tot PTP	1.4/0.7	1.1/0.7	1.2/0.8	1.0/0.5	1.2/0.7	1.1/0.7	
Tut. Lat NPTP	IV	1.5/0.5	1.0/0.5	1.2/0.6	1.1/0.8	1.3/0.6	1.1/0.6	
	Tact	2.2/0.6	1.7/0.7	1.6/0.5	1.7/0.6	1.9/0.7	1.7/0.6	1.5/0.4
	Tot NPTP	1.8/0.6	1.4/0.7	1.3/0.6	1.4/0.8	1.6/0.7	1.4/0.7	
Eq. Lat	PTP		8.8/4.7	4.7/3.9	3.8/1.7	6.0/4.4	6.8/4.8	
	NPTP		4.0/2.1	3.8/1.9	2.9/1.9	3.6/2.0	3.9/2.0	
Notes	* signifies the PTP condition occurred first. Lat: Latency; Tut: tutorial condition; Dur: duration, Eq: equivalence condition. Numbers to the left of the "/" indicate mean latencies in seconds per block. Numbers to the right of the "/" indicate the SD.							

occurred in sessions 1 and 5). Recall that these sessions used different sample and choice stimuli in all conditions, thus possibly indicating a warm-up effect. It is conjectured that this warm-up may be related to the vocal-verbal behavior which each participant claimed to use (when asked in exit interviews) to aid performance in both the PTP and NPTP tasks.

Table 8

Number of Incorrects per Block for Participant M

Sess. #	Rel. Trained	Condition in Order of Presentation	Number of Incorrects per Test Block Block = one presentation of each relation								Totals	
			1	2	3	4	5	6	7	8		
1	14	IV-PTP	9	10	9						28	PTP
		Tact-NPTP	7	7	5						19	64
		Tact-PTP	11	12	13						36	NPTP
		IV-NPTP	11	9	7						27	46
2	14	IV-NPTP	11	9	10	8					38	NPTP
		Tact-NPTP	9	9	6	5					29	67
		Equivalence	13									
		IV-PTP	12	10	7	8					37	PTP
		Tact-PTP	7	6	5	5					23	60
		Equivalence	12									
3	14	IV-PTP	7	7	8	4					26	PTP
		Tact-PTP	7	2	1	1					11	37
		Equivalence	5									
		IV-NPTP	6	3	0	0					9	NPTP
		Tact-NPTP	1	0	0	0					1	10
		Equivalence	1									
4	10	IV-NPTP	3	1	1	0	0	0			5	NPTP
		Tact-NPTP	1	0	0	0					1	6
		Equivalence	1									
		IV-PTP	0	0	0						0	PTP
		Tact-PTP	1	0	0	0					1	1
		Equivalence	1									
5	12	Tact-PTP	6	9	8	8	8				39	PTP
		Tact-NPTP	9	9	6	6	1				31	87
		Tact-PTP	11	10	10	9	8				48	NPTP
		Tact-NPTP	9	9	7	7	8				40	71

Table 9

Summary Data for Participant M

		Sess. 1*	Sess. 2	Sess. 3*	Sess. 4	Sess. 1-4	Sess. 2-3	Sess. 5*
Lat. PTP	IV	5.8/3.0	4.5/2.7	5.7/3.0	1.9/2.2	4.7/3.1	5.1/2.9	
	Tact	5.8/2.5	5.9/2.9	5.9/3.1	3.2/2.1	5.3/2.9	5.9/3.0	7.5/2.9
	Tot PTP	5.8/2.7	5.2/2.9	5.8/3.0	2.7/2.3	5.0/3.0	5.5/3.0	
Lat. NPTP	IV	4.6/2.1	5.1/2.5	4.9/3.8	2.3/2.2	4.1/3.0	5.0/3.2	
	Tact	7.8/3.0	7.0/2.8	5.0/2.8	3.2/1.2	5.8/3.1	6.0/3.0	7.1/2.8
	Tot NPTP	6.2/3.0	6.0/2.8	4.9/3.3	2.6/2.0	4.9/3.2	5.5/3.1	
Tut. Lat PTP	IV	0.9/1.1	0.6/0.5	0.6/0.4	1.8/1.4	0.8/0.9	0.6/0.5	
	Tact	0.9/0.7	1.8/1.5	1.6/1.0	2.1/1.3	1.4/1.2	1.7/1.3	1.3/0.5
	Tot PTP	0.9/0.9	1.1/1.2	1.0/0.9	2.0/1.3	1.1/1.1	1.1/1.1	
Tut. Lat NPTP	IV	0.5/0.3	0.6/0.5	0.9/0.6	1.6/0.9	0.7/0.6	0.7/0.6	
	Tact	2.2/1.0	1.4/0.8	2.3/1.2	2.1/0.8	1.9/1.0	1.6/1.0	1.3/0.4
	Tot NPTP	1.3/1.1	0.9/0.8	1.5/1.1	1.8/0.9	1.2/1.0	1.1/0.9	
Eq. Lat	PTP		5.3/2.8	7.4/4.2	2.5/2.6	5.3/3.8	6.4/3.7	
	NPTP		6.2/3.6	3.6/3.0	2.6/2.9	4.3/3.6	4.9/3.6	
Notes	* signifies the PTP condition occurred first. Lat: Latency; Tut: tutorial condition; Dur: duration, Eq: equivalence condition. Numbers to the left of the "/" indicate mean latencies in seconds per block. Numbers to the right of the "/" indicate the SD.							

Across blocks, conditions and sessions, all participants (including participant B) tended to demonstrate improved performance (see Tables 4, 6, 8, 10,12, and 14). When considering last block PTP and NPTP performance, no other participant reached the level of mastery in session 2 that participant B reached. C reached the highest level of all these other participants, and also demonstrated the best equivalence performance as illustrated in Table 6. W nearly reached mastery in the PTP condition of session 2

Table 10

Number of Incorrects per Block for Participant P

Sess. #	Rel. Trained	Condition in Order of Presentation	Number of Incorrects per Test Block Block = one presentation of each relation 1	2	3	4	5	6	7	8	Totals	
1	14	IV-PTP	13	12	10						35	PTP
		Tact-NPTP	13	14	13						40	76
		Tact-PTP	14	14	13						41	NPTP
		IV-NPTP	9	6	2						17	57
2	14	IV-NPTP	9	7	6	5					27	NPTP
		Tact-NPTP	12	14	12	11					49	76
		Equivalence	12									
		IV-PTP	8	8	6	4					26	PTP
		Tact-PTP	12	9	5	7					33	59
		Equivalence	12									
3	14	IV-PTP	10	8	4	5					27	PTP
		Tact-PTP	10	7	4	4					25	52
		Equivalence	11									
		IV-NPTP	12	9	4	7					32	NPTP
		Tact-NPTP	6	5	5	5					21	53
		Equivalence	11									
4	10	IV-NPTP	4	2	2	1	1	0	0	0	10	NPTP
		Tact-NPTP	1	2	0	0	0				3	13
		Equivalence	1									
		IV-PTP	2	0	0	0					2	PTP
		Tact-PTP	3	2	0	0	0				5	7
		Equivalence	0									
5	12	Tact-PTP	11	11	8	8	8				46	PTP
		Tact-NPTP	8	9	8	6	6				37	94
		Tact-PTP	8	11	11	9	9				48	NPTP
		Tact-NPTP	9	11	9	10	10				49	86

Table 11

Summary Data for Participant P

		Sess. 1*	Sess. 2	Sess. 3*	Sess. 4	Sess. 1-4	Sess. 2-3	Sess. 5 *
Lat. PTP	IV	3.8/1.6	4.0/3.0	4.2/2.5	2.6/2.5	3.7/2.6	4.1/2.8	
	Tact	1.8/1.2	5.1/2.7	4.9/1.9	4.3/1.8	4.2/2.4	5.0/2.3	3.0/1.4
	Tot PTP	2.8/1.8	4.6/2.9	4.6/2.2	3.5/2.3	4.0/2.5	4.6/2.6	
Lat. NPTP	IV	5.5/2.7	4.9/3.0	4.7/3.1	3.0/2.4	4.3/3.0	4.8/3.1	
	Tact	3.4/2.1	2.2/1.4	5.5/2.5	4.9/2.0	4.0/2.4	3.9/2.6	3.9/1.8
	Tot NPTP	4.5/2.6	3.6/2.7	5.1/2.8	3.8/2.4	4.2/2.7	4.3/2.9	
Tut. Lat PTP	IV	1.0/1.4	0.6/0.4	0.7/0.6	0.6/0.4	0.8/0.9	0.7/0.5	
	Tact	1.2/0.7	0.9/0.4	1.3/0.5	1.6/0.4	1.2/0.6	1.1/0.5	1.5/0.5
	Tot PTP	1.1/1.1	0.8/0.4	1.0/0.6	1.1/0.7	1.0/0.8	0.9/0.6	
Tut. Lat NPTP	IV	0.5/0.3	0.5/0.3	0.5/0.3	0.8/0.5	0.5/0.3	0.5/0.3	
	Tact	1.7/0.6	1.2/0.5	1.3/0.5	1.7/0.7	1.4/0.6	1.2/0.5	1.5/0.5
	Tot NPTP	1.2/0.8	0.9/0.5	0.8/0.5	1.2/0.7	1.0/0.7	0.9/0.5	
Eq. Lat	PTP		2.3/1.5	2.8/1.5	3.2/1.9	2.7/1.6	2.5/1.5	
	NPTP		3.5/1.7	1.3/0.8	2.3/2.9	2.4/2.1	2.4/1.8	
Notes	* signifies the PTP condition occurred first. Lat: Latency; Tut: tutorial condition; Dur: duration, Eq: equivalence condition. Numbers to the left of the "/" indicate mean latencies in seconds per block. Numbers to the right of the "/" indicate the SD.							

(see Table 14), at which point equivalence performance improved dramatically. Participants M, P and S showed little mastery, and subsequently demonstrated poor equivalence performance. In session 3, participants C, M and W (Tables 6, 8 and 14 respectively) showed good last block performances, and following these performances, relatively good performances in the equivalence conditions. W performed poorly in the equivalence condition following the PTP conditions (12 out of 14 incorrect) due to an error: the same equivalence sets were used for this condition as were used following

Table 12

Number of Incorrects per Block for Participant S

Sess. #	Rel. Trained	Condition in Order of Presentation	Number of Incorrects per Test Block Block = one presentation of each relation 1	2	3	4	5	6	7	8	Totals	
1	14	IV-NPTP	14	10	8						32	NPTP
		Tact-PTP	11	10	9						34	73
		Tact-NPTP	13	14	14						41	PTP
		IV-PTP	12	13	11						36	70
2	12	IV-PTP	8	9	6	5					28	PTP
		Tact-PTP	9	9	7	6					31	59
		Equivalence	12									
		IV-NPTP	8	4	4	5					21	NPTP
		Tact-NPTP	8	6	7	5					26	47
		Equivalence	8									
3	12	IV-NPTP	6	8	6	7					27	NPTP
		Tact-NPTP	9	6	5	6					26	53
		Equivalence	11									
		IV-PTP	4	2	5	5					16	PTP
		Tact-PTP	8	3	3	5					19	35
		Equivalence	10									
4	10	IV-PTP	4	4	4	3	1	4	1	0 *	27	PTP
		Tact-PTP	5	7	3	0	1	2	0	1 *	21	48
		Equivalence	3									
		IV-NPTP	1	2	0	0	0				3	NPTP
		Tact-NPTP	7	2	1	0	0	0			10	13
		Equivalence	2									

* For this participant the number of blocks to mastery in the IV condition was 19 and in the Tact condition was 16. The remaining blocks not displayed here consisted of either 1 error or no errors.

the NPTP condition of the previous session. However, W's equivalence performance was relatively good (4 out of 14 incorrect) in the next presentation. In session 4, all

Table 13

Summary Data for Participant S

		Sess. 1	Sess. 2 *	Sess. 3	Sess. 4*	Sess. 1-4	Sess. 2-3
Lat. PTP	IV	3.6/2.1	4.8/2.9	3.9/3.3	2.9/2.5	3.4/2.8	4.4/3.1
	Tact	6.6/2.7	4.9/2.3	6.2/3.2	4.4/2.6	5.1/2.8	5.6/2.9
	Tot PTP	5.1/2.8	4.9/2.6	5.1/3.4	3.6/2.7	4.2/2.9	5.0/3.1
Lat. NPTP	IV	3.4/2.0	3.6/2.7	4.3/2.8	2.3/1.8	3.4/2.5	4.0/2.7
	Tact	5.0/2.1	6.3/2.9	5.9/2.6	4.2/2.2	5.3/2.6	6.1/2.7
	Tot NPTP	4.2/2.2	4.9/3.1	5.1/2.8	3.4/2.2	4.4/2.7	5.0/2.9
Tut. Lat PTP	IV	0.6/0.5	0.5/0.5	0.5/0.4	0.5/0.3	0.5/0.4	0.5/0.4
	Tact	1.9/0.9	1.2/0.5	1.0/0.5	1.0/0.5	1.3/0.8	1.1/0.5
	Tot PTP	1.2/1.0	0.9/0.6	0.8/0.5	0.8/0.5	0.9/0.7	0.8/0.5
Tut. Lat NPTP	IV	0.6/0.5	0.4/0.4	0.6/1.0	1.0/0.6	0.6/0.7	0.5/0.8
	Tact	1.4/0.5	1.1/0.5	1.0/0.6	1.4/0.7	1.2/0.6	1.0/0.5
	Tot NPTP	1.0/0.6	0.8/0.5	0.8/0.8	1.2/0.7	0.9/0.7	0.8/0.7
Eq. Lat	PTP		1.7/1.9	3.6/2.2	7.4/3.6	4.0/3.5	2.7/2.3
	NPTP		2.4/2.1	2.5/2.2	5.9/4.4	3.5/3.4	2.5/2.2
Notes	* signifies the PTP condition occurred first. Lat: Latency; Tut: tutorial condition; Dur: duration, Eq: equivalence condition. Numbers to the left of the "/" indicate mean latencies in seconds per block. Numbers to the right of the "/" indicate the SD.						

participants were required to meet a mastery criterion (three consecutive blocks with no incorrects) and all demonstrated good equivalence performances. All but one participant showed overall better equivalence performance following the NPTP condition than the PTP condition. However, when the faulty data from W (incorrect stimuli presented) and S (the misunderstood directions) are taken into account, the overall difference in the number of incorrects between the two conditions narrows to a small margin.

Table 14

Number of Incorrects per Block for Participant W

Sess. #	Rel. Trained	Condition in Order of Presentation	Number of Incorrects per Test Block Block = one presentation of each relation 1	2	3	4	5	6	7	8	Totals	
1	14	IV-PTP	11	9	8						28	PTP
		Tact-NPTP	13	11	12						36	62
		Tact-PTP	13	11	10						34	NPTP
		IV-NPTP	10	9	5						24	60
2	14	IV-NPTP	14	13	10	9					46	NPTP
		Tact-NPTP	11	12	8	6					37	83
		Equivalence	10									
		IV-PTP	5	2	0	0					7	PTP
		Tact-PTP	10	6	2	3					21	28
		Equivalence	4									
3	14	IV-PTP	7	7	5	4					23	PTP
		Tact-PTP	11	7	5	4					27	50
		Equivalence*	12									
		IV-NPTP	5	2	4	1					12	NPTP
		Tact-NPTP	8	6	4	2					20	32
		Equivalence	4									
4	10	IV-NPTP	1	1	0	0	0				2	NPTP
		Tact-NPTP	2	0	0	1	0	0	0		3	5
		Equivalence	1									
		IV-PTP	0	1	0	0	0				1	PTP
		Tact-PTP	1	0	0	0					1	2
		Equivalence	1									

* This was the condition in which the incorrect stimulus set was arranged.

Participants B, C, M and P participated in session 5. As expected, the less discriminable sample stimuli caused more errorful responding to be demonstrated. B performed best in this session, reaching mastery by the final block of each condition.

Table 15

Summary Data for Participant W

		Sess. 1*	Sess. 2	Sess. 3*	Sess. 4	Sess. 1-4	Sess. 2-3
Lat. PTP	IV	4.6/2.4	3.2/3.0	4.3/2.6	1.7/2.1	3.4/2.8	3.7/2.9
	Tact	4.6/2.4	5.1/2.6	5.3/2.2	3.2/1.4	4.7/2.4	5.2/2.4
	Tot PTP	4.6/2.4	4.2/3.0	4.8/2.5	2.4/2.0	4.0/2.7	4.5/2.7
Lat. NPTP	IV	4.8/2.6	4.7/3.0	4.1/3.1	2.4/2.3	4.0/2.9	4.4/3.1
	Tact	4.8/2.4	5.0/2.2	6.0/2.9	3.6/2.1	4.8/2.5	5.5/2.6
	Tot NPTP	4.8/2.5	4.9/2.6	5.1/3.1	3.1/2.3	4.4/2.8	5.0/2.9
Tut. Lat PTP	IV	0.6/0.4	0.8/0.4	0.6/0.6	0.4/0.3	0.6/0.5	0.7/0.5
	Tact	1.3/0.7	1.0/0.3	1.0/0.5	0.9/0.1	1.1/0.5	1.0/0.4
	Tot PTP	1.0/0.7	0.9/0.3	0.8/0.6	0.7/0.3	0.9/0.6	0.9/0.5
Tut. Lat NPTP	IV	0.8/0.8	0.4/0.4	0.4/0.2	0.5/0.3	0.5/0.6	0.4/0.4
	Tact	2.0/1.1	0.9/0.4	1.2/0.5	1.0/0.6	1.3/0.9	1.0/0.5
	Tot NPTP	1.4/1.2	0.6/0.5	0.9/0.6	0.8/0.5	1.0/0.9	0.7/0.5
Eq. Lat	PTP		6.6/4.3	1.2/1.3	1.8/2.0	3.4/3.8	3.9/4.2
	NPTP		4.3/2.8	3.6/2.8	2.5/3.1	3.6/3.0	4.0/2.8
Notes	* signifies the PTP condition occurred first. Lat: Latency; Tut: tutorial condition; Dur: duration, Eq: equivalence condition. Numbers to the left of the "/" indicate mean latencies in seconds per block. Numbers to the right of the "/" indicate the SD.						

Participants C, M, and P, however, tended to not reach mastery, or in the case of participant's C and M, did so in only one condition (Tact-PTP for C, and Tact-NPTP for M). As shown in the protocol analysis section later in this paper, B also had the most consistent performance in using vocal-verbal behavior in performing these tasks.

Participants C, M and P demonstrated a similar performance to B's in that they initially performed better (or had equal performances) in the IV versus tact conditions, generally for sessions 1 and 2. For session 3 and 4, these participants performed better in the tact conditions. This better tact performance was also correlated with better equivalence for participants C and M. This switch in better performances (IV to tact) supports exit interview statements in which each participant noted that they applied names in the IV condition, which were then carried over and applied in the tact condition (as the choice stimuli were shared between the two conditions). Participants S and W did not show this IV to tact reversal (each performed best in the IV condition).

Overall, the test, tutorial and equivalence latencies showed little difference between the PTP and NPTP conditions (see Tables 5, 7, 9, 11, 13 and 15). In addition test latencies tended to decrease across sessions, while tutorial latencies remained relatively stable. A interesting exception to this is participant P's tendency to have the shortest latency in conditions in which the highest errors occurred (see Tables 10 and 11). This may be indicative that P was "guessing", in that P would immediately select some choice stimulus, resulting in a shorter latency and a higher number incorrect. This hypothesis is partially supported by the results of P's protocol analysis: P tended not to make overt vocal-verbal responses to the sample and choice stimuli .

Participant C showed a rather large difference in equivalence latencies following PTP and NPTP performances (see Table 7). Most of this difference occurred in session 1: the very first time C encountered the equivalence condition, those latencies were very high (mean of 8.8 s vs. 4.0 s following the NPTP condition). As noted above, all other participants showed little difference between equivalence latencies following PTP and NPTP conditions. However, trends across sessions were mixed, with participants B, C, M, and W showing decreases in equivalence latencies (Tables 5, 7, 9 and 15), and participants P and S demonstrating increasing equivalence latencies (Tables 11 and 13). Participant W showed a dramatic decrease in equivalence latencies following the PTP condition, between sessions 2 and 3 (see Table 15). Recall, however, that the equivalence stimuli presented in the first part of session 3 were not the correct ones for this participant. It is possible that the dramatic decrease in the latencies was due to "guessing", (poor/inadequate stimulus control), thus resulting in any comparison stimuli being selected (possibly seen with participant P, as discussed previously).

Summary

Three participants performed better in the PTP condition and three in the NPTP condition. However, none of the differences between these conditions were very large, ranging from 3 (participant P) to 49 (participant M), with the total number of errors (PTP + NPTP) for P being 573 and for M 449.

All participants tended to improve in performance across subsequent blocks, conditions and sessions. An exception to this was the performance by four participants in which a resurgence in errors occurred following a 15 minute break in sessions 1 and 5. This would seem to indicate a warm-up effect was operating (see the analysis of this in the description of participant B's data).

Overall performance in the equivalence condition was less clear. At first viewing, performance in the equivalence condition that followed NPTP training appeared to be better. However, as noted in the group section, when several problems with these data are accounted for, the difference between these two conditions narrows to only a slight difference. Individually, all but one participant performed better following the NPTP condition (one participant had equal performance in both conditions). Participant B had a dramatic improvement in the second equivalence exposure, possibly indicating that this participant had "figured out" the task, which is supported by exit interview data. Three participants showed improved equivalence performance as the preceding PTP or NPTP performances improved, and the number of exposures to the equivalence task increased. The last two participants performed poorly in the equivalence condition until session 4, at which point performance improved dramatically.

Test, tutorial and equivalence latencies tended to decrease or remain relatively stable over subsequent sessions for all participants. Two notable exceptions to this (participant P in test latencies and W in equivalence latencies) showed decreased latencies in several conditions. Both of these instances seem to indicate that these participants were "guessing" or not engaging in vocal-verbal behavior that other participants were engaging in (see the analysis of each of these participants individual data for a more detailed description).

Finally, all participants reported in the exit interviews that they emitted some kind of vocal-verbal behavior, which aided their performance on these tasks.

Group Data

Overview

Table 16 summarizes the grouped data for all participants across blocks and sessions. As with the individual data, durations are not reported as they were very similar, and stable, across all conditions and sessions. Overall, a slightly better performance was found in the NPTP condition with 1185 errors overall for all participants, and 1203 errors overall for the PTP condition. When one considers the total number of relations each participant learned (approximately 256) and the total number of trials involved (approximately 1150 for each participant), this difference is quite small. Recall that the participants were balanced in terms of which condition was received first, thus, even with a worse first-condition phenomenon operating, one would expect to see little difference between the two groups overall if the conditions

Table 16

Number of Incorrects per Block for All Participants

Sess. #	# of Rel. Trained	Condition in Order of Presentation	Number of Incorrects per Test Block Block = one presentation of each relation 1	2	3	4	5	6	7	8	Totals	
1	84	IV-PTP	63	54	44						161	PTP
		IV-NPTP	65	61	39						165	359
		Tact-PTP	68	68	62						198	NPTP
		Tact-NPTP	68	65	62						195	360
2	82	IV-PTP	49	43	27	22					141	PTP
		Tact-PTP	54	41	20	23					138	279
		Equivalence	60									
		IV-NPTP	52	40	31	30					153	NPTP
		Tact-NPTP	47	46	36	29					158	311
		Equivalence	50									
3	82	IV-PTP	37	28	25	21					111	PTP
		Tact-PTP	45	22	15	18					100	211
		Equivalence	34									
		IV-NPTP	47	32	22	15					116	NPTP
		Tact-NPTP	31	17	14	13					75	191
		Equivalence	29									
4	60	IV-PTP	12	7	4	4	1	4	1	0*	33	PTP
		Tact-PTP	12	10	3	0	1	2	0	1*	29	62
		Equivalence	6									
		IV-NPTP	12	8	3	1	1	0	0	0	25	NPTP
		Tact-NPTP	14	6	2	1	0	0	0		23	48
		Equivalence	5									
5	48	Tact-PTP	36	36	24	24	23				143	PTP
		Tact-NPTP	35	33	26	20	15				129	280
		Tact-PTP	37	33	26	23	18				137	NPTP
		Tact-NPTP	37	36	29	25	21				148	277

* For participant S the number of blocks to mastery in the IV condition was 19 and in the Tact condition was 16. The remaining blocks not displayed here consisted of either 1 error or no errors.

had little effect on performance or if vocal-verbal behavior overshadowed such effects. A dependent t test revealed no significant difference between the two conditions at the 0.05 alpha level (two-tailed). Unless otherwise noted, all significance tests reported are dependent t-tests with an alpha level of 0.05. The average number of errors per block (or latencies, see below), for each participant was used to calculate all t scores.

A dependent t test was calculated for the different performances in the PTP and NPTP conditions for each of the sessions (from 1 to 4) as well as an overall dependent t test on the grouped data (for all participants, across sessions 1 to 4). For sessions 1 to 3, these tests were calculated using the total number of incorrects in each condition for each participant (six participants, five degrees of freedom). For session 4 and the overall t tests, the mean number of incorrect responses per block was used. This method was adopted as the number of blocks varied for each participant in session 4 (the session in which mastery was required). No t tests were calculated for session 5 as the number of participants in that session was low (a total of 4).

In addition, dependent t tests (six participants, five degrees of freedom) were calculated for the difference between the PTP and NPTP test latencies, durations and tutorial latencies. In each case the mean value per condition was used to calculate the t scores. For all tests, an alpha level of 0.05 (two tailed) was adopted.

As might be expected, few significant t values were found (as PTP and NPTP performances were so similar). IV and Tact latencies showed a significant difference, but as discussed earlier, this difference was an artifact of the procedure used. One other significant difference was also found, this being between the PTP and NPTP test latencies. This is discussed below in the section titled "Time Measures - Latencies, Durations and Tutorial Measures" below.

Point-to-Point and Non Point-to-Point Performances

As would be expected, when the data from all participants are grouped (Table 16) the first conditions effects disappear, or become very small. Over sessions 2 and 3, although slight, better performance shifted from the PTP condition (279 errors PTP vs. 311 errors NPTP) in session 2 to the NPTP conditions in session 3 (191 errors NPTP vs. 211 errors PTP). The larger difference in session two is most likely due to participant W's large number of incorrects in the NPTP conditions (this participant showed a very large difference between the PTP and NPTP condition in that session). Overall however, sessions 2 and 3 reflected the general trend of little overall difference between the two conditions.

As can be seen in Table 16, performance generally improved over successive blocks in all conditions. The exception to this are the first few blocks of several conditions in sessions 1 and 5. Recall in those sessions completely new stimuli were introduced for each new condition. It is interesting to note the remarkably small differences between the PTP and NPTP conditions, across all participants, for sessions 1 and 5. In session 1 the difference was only 1 error and in session 5, only 3 errors.

It is also interesting to note the difference in equivalence performance in the PTP conditions in sessions 2 and 3, as compared to the last block performances in those conditions as well. In session 2, last block performance in both the IV and tact relations in the PTP condition was very close to the last block performance in session 3 for the same relations under the PTP condition (see Table 16). To the extent that last block performance indicates mastery, one could conclude that the overall mastery level, across all participants, was nearly equal across these two conditions. Equivalence performance was not, however. In session 2, equivalence performance after the PTP conditions was very poor, 60 incorrect out of a possible 82 correct choices. The equivalence performance following the PTP condition in session 3, however, increased dramatically, nearly twice as good as the performance in session 2. Since it would appear that the two conditions were equal in terms of mastery levels reached, it is possible that some other factor may have increased performance, namely becoming more familiar with the equivalence task, or learning to use vocal-verbal strategies in performing this task (all participants indicated that they used such strategies in their exit interviews). Overall equivalence performance improved across sessions (see the equivalence section below).

Across sessions, performance improved in all conditions when considering the grouped data (see the last column of Table 16). An analysis of number of time-outs (counted as incorrects) did not reveal significantly more in the earlier sessions, thus signifying that it is unlikely that this trend was due to participants becoming more familiar with the computer (all professed skills with a mouse before starting the experiment). It seems likely that this trend across sessions was due to participants becoming more familiar with the task and with their development of vocal verbal strategies for increasing number corrects (see the exit interview and protocol analysis sections).

Overall, performance was better in the IV vs. tact conditions, but by only a slight margin (IV total errors was 905, tact total errors was 916) for sessions 1 to 4. It is interesting to note the general trend for these performances, however. In sessions 1 and 2, IV performance was better, with session 1 IV errors at 326 and session 1 tact errors at 393, a relatively large difference. In session 2 this difference narrowed greatly, with total IV errors only 2 less than total tact errors: IV, 294 errors; tact, 296 errors. By session 3 the tact was the condition in which better performance was found, with the total tact errors at 175 and the IV errors at 227, again a relatively large difference. In session 4 tact performance was still better, but by only six errors: tact performance was at 52 errors and IV performance was at 58 errors. If verbal mediation of the sort that participants claimed to have been using in exit interviews (carrying names generated in the IV condition into the tact condition) was occurring, one would expect to see such results.

Finally, a correlation was calculated for the relationship between the number of patterns which participants used in the PTP condition (the exact order in which they clicked on the dots in each dot pattern) and the number of incorrects for that condition. A relatively high positive correlation (Pearson) was found: 0.73. This correlation was conducted for sessions 1 to 5. This indicates that the more consistent the click pattern to a particular choice stimulus, the more accurate the responding.

Time Measures - Latencies, Durations and Tutorial Measures

A relatively small mean difference existed between the PTP and NPTP latencies during test conditions summarized across sessions 1 to 4 (4.4 s, SD=2.9 and 4.7 s, SD=2.9 respectively), as shown in the four test latency rows of Table 17. This difference was statistically significant ($p = .029$). This difference was smaller (4.9 s vs. 5.1 s) when summarized across sessions 2 and 3, and was not statistically significant. As noted earlier, this was the only statistically significant effect found between the PTP and NPTP conditions. From a practical point of view, this difference between the two conditions is very small (in no case did a difference between averaged latencies exceed 0.5 s).

The IV mean latency in seconds (for both PTP and NPTP conditions) was 4.6 s (SD = 3.1) and for the tact condition 5.4 s (SD = 2.8) (see Table 17). Recall that this difference is an artifact of the procedure (in the IV condition the participants could scan

Table 17

Summary of Data for All Participants

		Sess. 1	Sess. 2	Sess. 3	Sess. 4	Sess. 1-4	Sess. 2-3	Sess. 5
Lat. PTP	IV	4.6/2.6	4.3/2.9	4.5/3.0	2.6/2.5	3.8/2.9	4.4/3.0	
	Tact	5.3/2.8	5.6/2.8	5.4/2.7	3.9/2.3	5.0/2.7	5.5/2.8	
	Tot PTP	4.9/2.7	4.9/3.0	4.9/2.9	3.2/2.5	4.4/2.9	4.9/2.9	
Lat. NPTP	IV	4.7/2.5	4.8/3.1	4.8/3.3	2.6/2.3	4.2/3.0	4.8/3.2	
	Tact	6.0/3.0	5.2/2.9	5.5/2.7	3.9/2.1	5.1/2.8	5.4/2.8	
	Tot NPTP	5.4/2.8	5.0/3.0	5.1/3.0	3.2/2.3	4.7/2.9	5.1/3.0	
Tut. Lat. PTP	IV	0.8/0.9	0.6/0.5	0.7/0.6	0.8/0.7	0.7/0.7	0.6/0.5	
	Tact	1.4/0.8	1.3/0.8	1.3/0.8	1.4/0.8	1.3/0.8	1.3/0.8	
	Tot PTP	1.1/0.9	0.9/0.7	1.0/0.8	1.1/0.8	1.0/0.8	1.0/0.7	
Tut. Lat. NPTP	IV	0.8/0.6	0.6/0.5	0.8/0.7	1.0/0.7	0.7/0.6	0.7/0.6	
	Tact	1.7/0.9	1.3/0.7	1.4/0.8	1.7/0.8	1.5/0.8	1.4/0.8	
	Tot NPTP	1.3/0.9	0.9/0.7	1.1/0.8	1.3/0.8	1.1/0.8	1.0/0.8	
Eq. Lat.	PTP		5.1/4.0	4.1/3.6	3.3/3.1	4.3/3.7	4.6/3.8	
	NPTP		4.0/2.9	3.1/2.6	3.0/3.2	3.4/2.9	3.5/2.8	
# Inc / Eq. (r)	PTP		0.83	0.86		0.88	0.88	
	NPTP		0.90	0.94		0.93	0.90	
Notes		Eq. = equivalence condition, Inc = incorrects, IV = intraverbal condition, Lat = latency, r = Pearson correlation coefficient, Tut = tutorial condition. Numbers to the left of the "/" indicate mean latencies in seconds per block. Numbers to the right of the "/" indicate the SD.						

as soon as they hear the word over the computer speakers, not so in the tact condition). When summarized across sessions 2 and 3, with an IV mean latency of 4.6 s (SD = 3.1) and a tact mean latency of 5.4 s (SD = 2.8). A relatively sharp reduction in the averaged latency time in all conditions of sessions 4 is most likely due to the reduced number of relations used in that session (10 versus 14 or 12).

Tutorial and remedial latencies, as well as durations, were recorded and reported to determine if participants were using the tutorial time to study the relations presented (more studying would result in longer latencies). It was presumed that these latencies (and to some extent durations) would increase with subsequent sessions if such strategies were being employed by the majority of participants. Tutorial latencies and durations were very stable across sessions, thus not demonstrating such a trend.

Equivalence Measures

Summarized across sessions 1 to 4, a difference existed between the total number of errors recorded in the PTP condition (100) and the NPTP condition (84). This difference was not statistically significant. This difference is also skewed in that an error resulted in one participant receiving an incorrect equivalence set, resulting in an elevated number of errors (participant W, session 3, 12 errors) for that condition (PTP). Another possible contributing factor to the increased errors in the PTP conditions was the large number of blocks which participant S took to reach mastery in the PTP condition of session 4 (35 blocks). A total of three equivalence errors were recorded for this participant in the PTP condition, which was the most of all participants (all others had zero or one). Taking these into consideration, it would seem that there was a less extreme difference between the two conditions than initially indicated, but with a slightly better performance in the NPTP condition. However, individual performances were quite varied. See each participant's individual results for these analyses.

Equivalence conditions in sessions 2 and 3 were generally administered prior to a participant mastering (defined as errorless responding) the preceding IV and tact conditions (as a fixed number of blocks per condition was arranged). Thus, a Pearson correlation was conducted to clarify any relationship between mastering the preceding relations and performance on equivalence. As suspected, high positive correlations were found between the number of incorrects in the preceding IV and tact conditions and the equivalence conditions. With the erroneous data from participants S and W removed (as noted above), the correlation for session 2 PTP was 0.83 and 0.86 PTP for session 3. For the NPTP condition, the session 2 correlation was 0.90 and for session 3 the correlation was 0.93. Session 4 correlations are not reported due to the extremely low number of errors in that session for all participants (see Table 17 for these correlations).

It is interesting to note that equivalence latencies for all participants decreased over subsequent sessions. The decrease in session 4 is most likely due to the decreased number of stimuli used (10). Presumably, if improvements in scanning and other techniques were responsible for this decrease, then a similar effect would have been observed in the averaged latencies in normal (non-equivalence) conditions. There was a relatively high difference between the equivalence latencies in the PTP and NPTP conditions. Averaged across sessions 1 to 4, the PTP average equivalence latency was 4.3 s (SD = 3.7) and 3.4 s (SD = 2.9) for the NPTP condition. This difference was even more pronounced in the session 2 to 3 aggregate data. The PTP average latency was 4.6 s (SD = 3.8) and the NPTP average was 3.5 s (SD = 2.8). In general, the participants tended to have longer latencies when incurring more incorrects. However, two participants did not adhere to this pattern. P tended to have increased equivalence latencies for the condition which came first, except for session 4 (see Table 4). W actually demonstrated a relationship opposite to the general trend: latencies were short when equivalence corrects were higher, and longer when equivalence incorrects were higher (see Table 15).

First Condition Effects

From each of the tables for all participants (Tables 4 to 15), it can be clearly seen that the condition which they encountered first in each session was the one in which they made the most mistakes. While part of this effect must be the result of improved performance over time, this is probably not the only contributor to this effect. For participants B, M, P, incorrect responses increased in session 1 and session 5 after a 15 minute break was taken. Recall that sessions 1 and 5 required that the participant learn entirely new relations during each set, unlike sessions 2, 3 and 4 in which the same dot patterns were used for the IV and tact relations within the PTP and NPTP conditions. Participants M and P also demonstrated some resurgence in incorrects in sessions 2 and 3 respectively. However, in these sessions, the dot patterns remained the same for both sets (IV and tact) prior to the break. Thus, the resurgence in errors during the second half of these sessions could also be attributed to the introduction of entirely new relations. During sessions 1 and 5, however, if a simple practice effect was operating, it would be expected that a gradual decrease in errors would occur over subsequent sets, as was seen in Participant C's data (see Table 6) for session 1. Given that the participants claimed to use vocal-verbal behavior liberally in this task (see exit interview and protocol analysis results below) it is possible that a certain amount of time was required for participants to become fluent in using this technique, and quite possibly the break disrupted this "priming" resulting in a resurgence of errors. Until further empirical testing is conducted, however, this is conjecture.

Session 5 - Incorrects, Latencies and Durations

This session was included to decrease the vocal-verbal behavior generated by participants when involved in the arranged tasks. To do this, less discriminable sample

stimuli were used in the hope that if naming could not occur, differences between the PTP and NPTP conditions might become apparent. As can be seen in Table 16, there was only a slight difference in the PTP and NPTP performance of the four participants completing session 5 (a difference of only 3 errors, with the NPTP slightly better). In the exit interviews, all participants reported that the stimuli
were harder to name, but that it was still possible to do so (see the protocol analysis).

Interestingly, the average latency was somewhat shorter in the PTP condition, 5.5 s (SD = 3.4), versus in the NPTP condition 6.1 s (SD = 2.9) (See Table 17). This difference was not statistically significant. The tutorial latencies for both conditions were nearly identical at 1.4 s (SD = 0.7) for the PTP condition and 1.5 s (SD = 0.6) in the NPTP condition.

Session 5 - Mouse Movements

An added programming feature in session 5 allowed for the movements of each participant's mouse to be played back and analyzed in terms of whether or not they might be engaging in stereotypical mouse movements (potential topography-based responses) in the NPTP condition. These were recorded only when the mouse was over a dot pattern (the likely places where these patterns would occur). Of the 2,116 incidents of the mouse being within a field in the NPTP condition, only 34 were considered to be a possible topography-based response. Of those 34, 15 did not involve situations in which that pattern was selected (clicked on) and were thus less likely to have been a topography-based response for that stimulus. Interobserver agreement based on a review of half of the total recorded incidents (randomly chosen) by a second observer, was 99%. Based on this data, it was deemed very unlikely that the participants were using the mouse to create a topographical response in the NPTP condition.

Exit Interviews

All participants were administered an exit interview after session 4 (see Appendix B). However, only two participants were debriefed as to the nature of this study. The other four were asked to participate in session 5, after which they completed a second exit interview, which was a repeat of the first exit interview, and participated in the protocol analysis describe below.

Of the six participants, four indicated a preference between the PTP and NPTP conditions. All stated they preferred the PTP condition, not all of whom performed best overall in this condition (see individual results). Reasons for the indicated preferences were that the patterns were easier to memorize when each dot was clicked on (two participants), another stated clicking on the corners was "boring" and the last indicated that the computer "registered the clicks too slowly" in the NPTP condition, although this was not the case.

Four participants indicated that they preferred the IV condition, although it was not always the case that their performance was best in these conditions. Two stated a preference for the tact condition and overall both did perform best in this condition.

One of the more interesting parts of the exit interview were the answers given to the first three questions, namely, what strategies or techniques were used, if any, when doing various parts of the experiment. All participants indicated that vocal (overt or covert) verbal behavior played a key role in their performance. For example, M wrote "I would try and make a word or phrase to the nonsense symbol that related to the dot pattern. For example "Jumit" (a nonsense syllable) turned into "Jump it" and related to

 (M drew this in) because it looked like something jumping over something" (parenthetical text added). Participant C wrote "I would try to turn the nonsense symbol into a word I use and connect that with the dot pattern, which I would try to make into a familiar object." C also wrote regarding the tact conditions: "... pick out a particular pattern within the flag-like pattern and make it something familiar to me and then gave that a word and the dot pattern a word and paired the two together." All other comments were along a similar vein.

When asked to recall strategies used in the equivalence portion of the experiment, two participants had clear (and similar) strategies. M wrote: "For the last two sessions (sessions 3 and 4), I made it a point to bring the same 'nicknames' for the dot patterns (in the IV condition) over to the flag-like patterns. Then I would associate the 'nicknames' I had to the flag-like patterns" (parenthetical text added). B wrote a similar account. Two participants (P and W) simply stated that they combined their previous two techniques, as described above. The last two participants also had very similar statements. As S wrote: "I went back through the links I had formed with sounds, dots and flags to arrive at the correct flag-like pattern". C wrote a similar account. It is important to note that none of the participants were given feedback as to the correctness of responding in the equivalence conditions, nor were they informed of the nature of the relations between the stimuli. All wrote this information when answering question 3 of the exit interview.

Finally, all participants were asked to recall as many dot patterns or nonsense syllables that they could recall. While all recalled some dot patterns, it was not a very useful measure, as the drawn patterns could not be related easily to the actual pattern used in the computer program due to some of the fine distinctions between the dot patterns. Those that could be related were not limited to the fourth session (the session after which the first exit interview was administered). A more interesting finding was that when asked to recall as many nonsense words as possible, two participants (B and W) said they could not remember any of the actual nonsense sounds, but could recall their English versions of the sounds. The other four participants generated a relatively short list of both English and nonsense words.

Protocol Analyses

Overview

As noted earlier, during sessions 1 to 4 and in the exit interviews, all participants indicated they used vocal verbal (covert and overt) behavior to aid in improving their performances. The protocol analysis was conducted to determine the nature of these utterances. Participants were asked to talk aloud while they were exposed to three blocks (one tutorial and two test blocks) from each of the last two conditions (one PTP and one NPTP) encountered in session 5. All protocol analyses occurred within 15 minutes of completing session 5. Utterances were recorded and encoded to determine the frequency of various types of utterances (e.g., tacts, intraverbals, and repeats of each). See Figures 3 to 7 and the Method section for a complete description of the coding scheme.

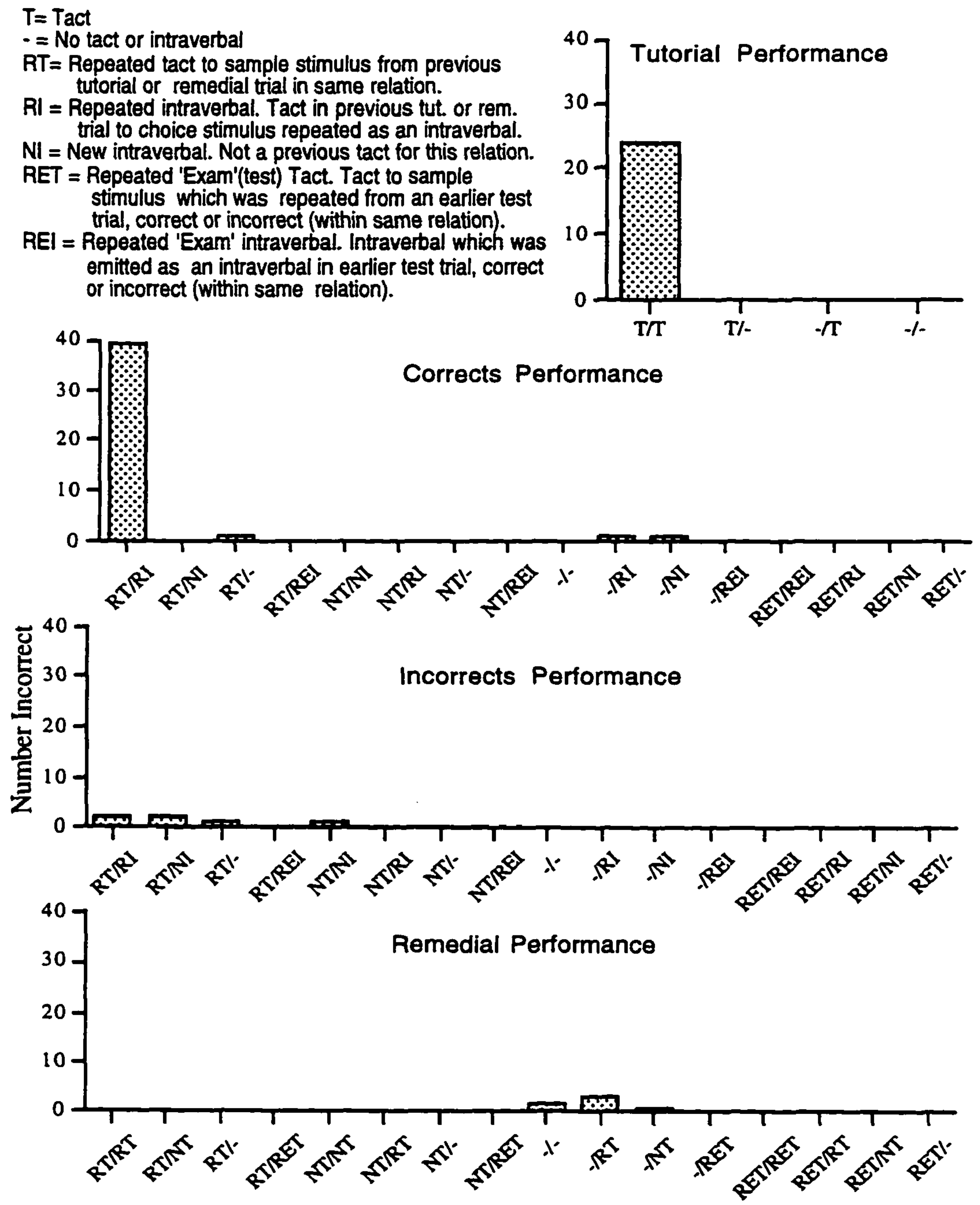
The percent of interobserver agreements (Page & Iwata, 1986) was calculated, based on this coding by two different observers (each scored all protocols). For all participants, the interobserver-agreement value was 94.5% for encoding each of the statement types. An interobserver-agreement value was also calculated for each of the individual participant's protocols that were encoded. All of these interobserver-agreement values exceeded 89%, ranging from 89.1% to 98.5%. An interobserver-agreement value was also calculated for whether or not each observer indicated that a relationship tact occurred (explained below) on any given trial. For all participants, the interobserver-agreement value was 89.6%, ranging from 81.1% to 97.0% for each individual participant.

Individual Data

Participant B

B performed the best of all participants in terms of the number of correct responses during this part of the experiment. As indicated by the graph in the upper right of Figure 3, B always emitted a tact response in the presence of the sample stimulus and the comparison, in the tutorial condition (the "T/T" category). B's verbal statements tended to reflect a relationship between the sample and choice stimuli (see the more thorough explanation in the group data section below). B's nearly perfect performance is illustrated in the second graph of Figure 3. Prior to selecting a correct answer, B tended to emit the same tact which was emitted to the sample stimulus in prior tutorial and remedial trials, and an intraverbal which was almost always the same as B emitted as a tact to the choice stimulus in previous tutorial and remedial trials (see the "RT/RI" category of the "Corrects Performance" of Figure 3). As noted earlier, these are labeled as IV responses, as the appropriate choice stimulus was not apparent and the response was most likely controlled by the ongoing verbal behavior of the participant (i.e., the tact of the sample stimulus). Of all statements preceding correct selections, B emitted this type of statement 92.8% of the time.

Three of the six statements which preceded incorrect selections involved a tact to the sample stimulus, as was done in previous tutorial and remedial trials, but were



Note: The key in the upper left-hand corner explains the abbreviations used. All frequencies indicate the number of specified statements (e.g., RT/RI) which preceded correct or incorrect selections. All selections in tutorial and remedial trials were correct (as the correct choice was revealed to the participant).

Figure 3. Protocol Data for Participant B.

accompanied by either a new intraverbal or none at all (see the "RT/NI" and "RT/-" category of the "Incorrects Performance" of Figure 3).

In remedial conditions, B tended to emit a tact to the choice stimulus only, and of all the remedial trials, half were repeats of previously used tacts in the tutorial or remedial conditions for the relation under study (see the "Remedial Performance" graph of Figure 3).

In one relation, it appeared that unnecessary parts of the verbal statement were dropped with each subsequent trial, perhaps an example of what Skinner (1957) noted: "Operant behavior tends to be executed in the easiest possible way" (p. 141):

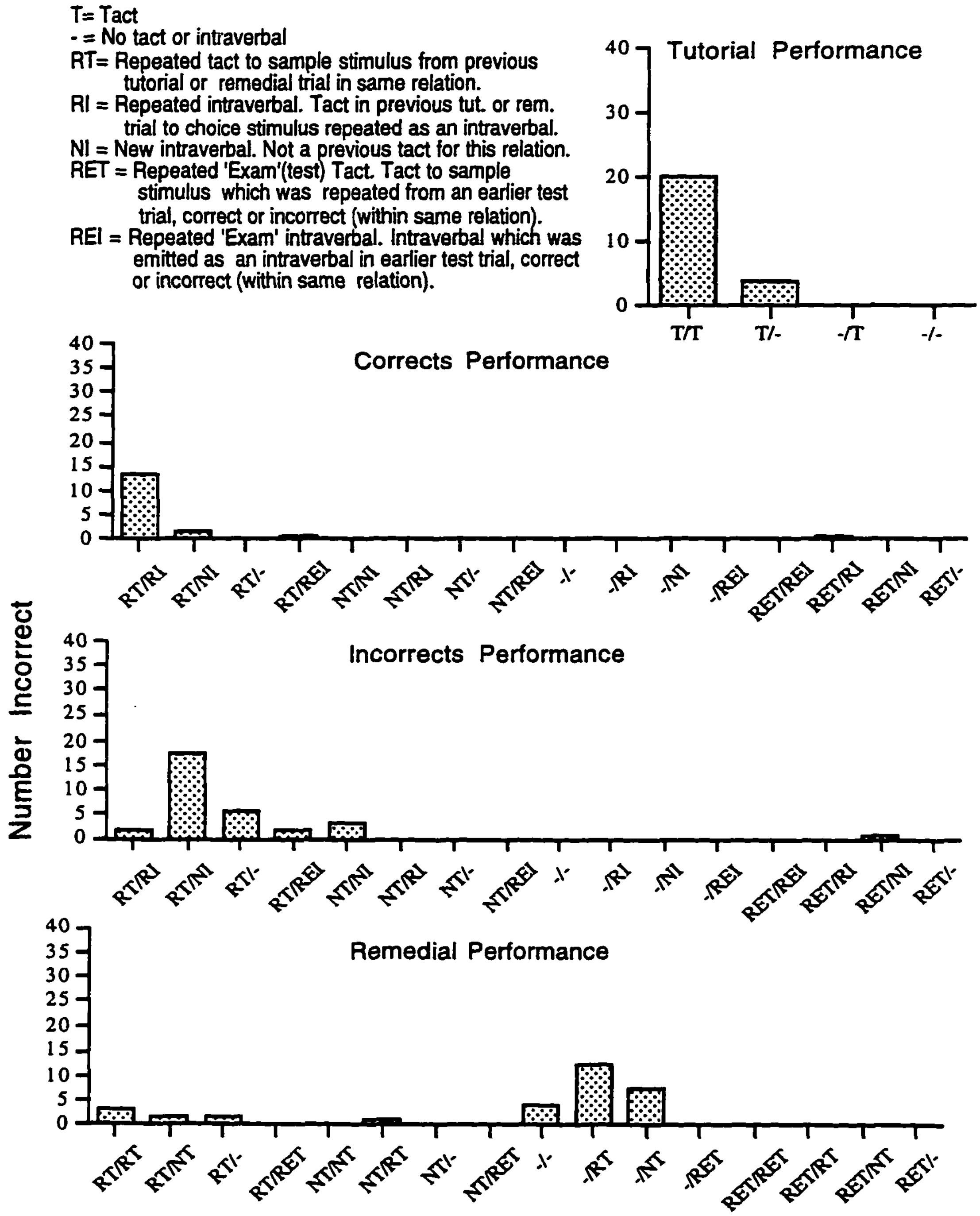
1. Let's see a white computer chip in a small basket (tutorial).
2. White computer chip in the basket (correct choice).
3. Ok white chip in a basket (correct choice).

The numbers indicate each subsequent exposure to the relation. The first exposure is the tutorial condition. Each subsequent exposure was a test trial, as B was correct in each trial for this relation (no remedial trials were required). The condition in effect, or the outcome of each trial is noted in the parentheses following each statement.

It is interesting to note that participant B seldom emitted a tact to the actual apparent characteristics of the sample and comparison stimuli (e.g., he would say "computer chip" and "basket" versus "dark diagonal" and "horizontal dots"). All the other participants had a much higher frequency of emitting more conventional tacts such as indicated above.

Participant C

In tutorial trials, C emitted tacts to the sample stimuli 100% of the time and 83.3% of the time to the choice stimuli (see the raw data displayed in the upper-right corner of Figure 4). Although C's overall number of correct responses was the next



Note: The key in the upper left-hand corner explains the abbreviations used. All frequencies indicate the number of specified statements (e.g., RT/RI) which preceded correct or incorrect selections. All selections in tutorial and remedial trials were correct (as the correct choice was revealed to the participant).

Figure 4. Protocol Data for Participant C

best of all participants (after B's), it was low, 35.4 % correct, compared to B's 87.5% correct. Of all C's statements preceding correct responses, 76.4% (a total of 13) consisted of repeats of tacts (RT) emitted in preceding tutorial and remedial trials for each relation and an intraverbal statement (RI) similar to the tact emitted in the presence of the choice stimuli during those same trials (see the "RT/RI" category of the second graph of Figure 4). The frequency of other types of statements preceding correct selections never exceeded two, also shown in the second graph of Figure 4.

Prior to selecting an incorrect choice on test trials, C generally repeated tacts to the sample stimulus emitted in previous tutorial and remedial trials (RT), but failed to emit intraverbals which were similar to the tacts emitted in previous tutorial and remedial trials. These type of statements ("RT/NI", "RT/-" and "RT/REI") composed 80.6% of all statements which preceded incorrect selections (see the third graph of Figure 4). In remedial trials, C generally emitted either a repeated tact from the previous tutorial or remedial trials, or emitted a new tact, in the presence of the choice stimulus (the "-/RT" and "-/NT" categories of the fourth graph of Figure 4). These composed 51.6% of all statements preceding remedial choice selections. The next most frequent response consisted of no tacts to either the sample or the choice stimuli the ("-/-" category). These responses composed 22.5% of all remedial responses. No other type of statement preceding remedial selections occurred more than three times, as shown in the last graph of Figure 4.

In watching C, it became apparent that this participant (and Participant P as well) would occasionally pause, emitting the second part of the vocal response only after scanning the choice stimuli and selecting it. While the following transcript illustrates this situation preceding a correct selection, this same pattern occurred with incorrect selections more frequently:

1. Ummm gray rectangle U shape (tutorial).

2. Gray rectangle with diagonal lines would go with the ... U shape (correct choice).

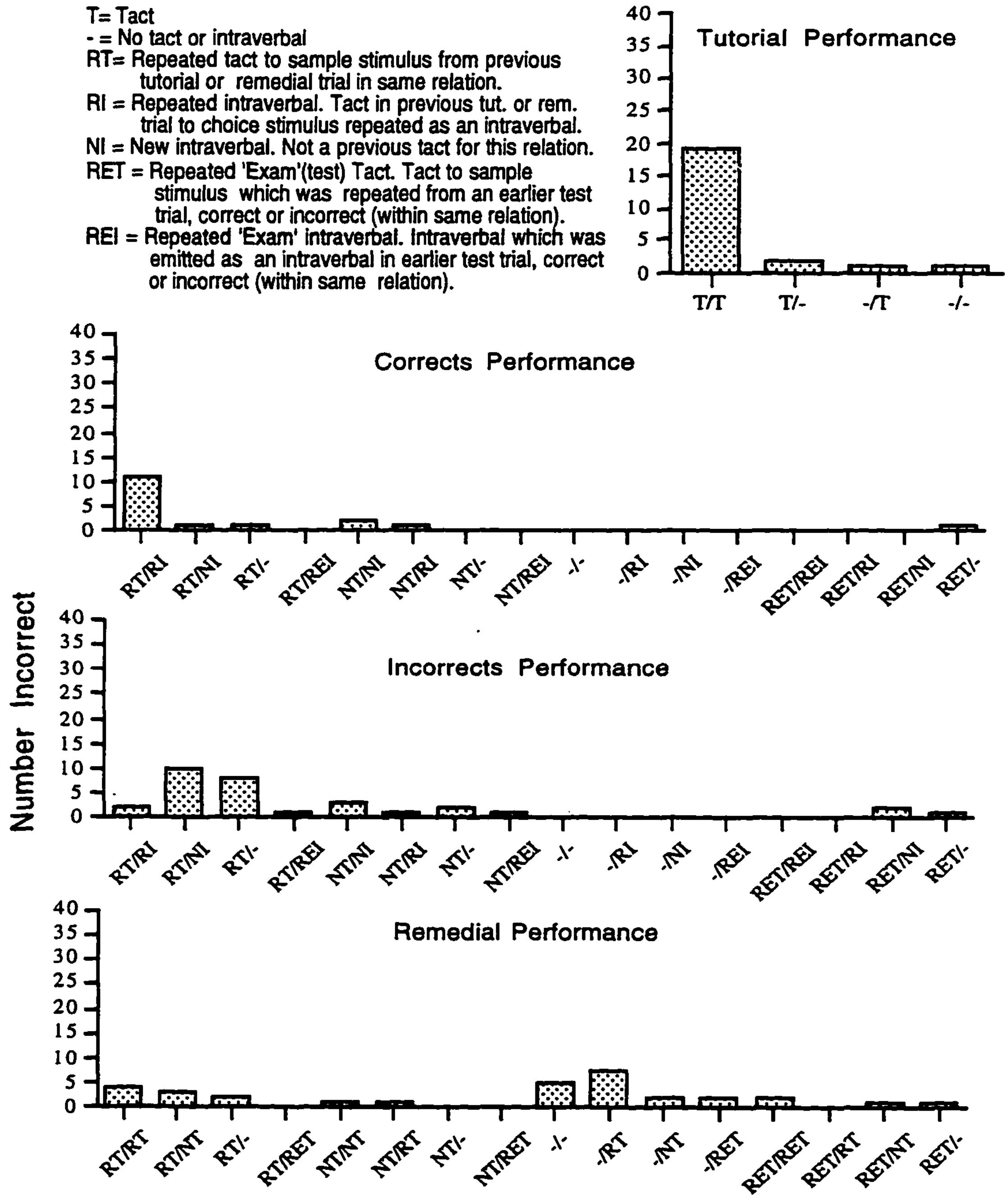
3. Gray rectangle with horizontal lines would go with the ... U shape (correct choice) (the "..." signify silence on the tape during which the participant is most likely scanning the choices).

Recall that the numbers indicate each subsequent exposure to this particular relation, and the parentheses immediately following each statement indicate the outcome or condition in effect. Also note that this participant would occasionally say "horizontal" for "diagonal", "vertical" for "horizontal", etc. Often C would correct these "mistacts" as they occurred. In the transcript labeled above, C did not, but a review of the stimuli used in this set indicated that this was the only stimulus which included a light gray rectangle. C did not tact any other stimulus in this set as "gray rectangle" and thus would not need to rely on the distinction of diagonal or horizontal lines to distinguish it from other similar sample stimuli.

Participant M

The majority of M's verbal statements during the tutorial conditions consisted of tacts to both the sample and comparison stimuli, the "T/T" conditions, which composed 82.6% of all verbal statements in this condition. Of the other types of statements encoded, only one ("T/-") reached a frequency of 2; the others had one incident each (see Figure 5).

M was the third highest scorer in terms of number of corrects (16 out of a possible 48), but was very close to C's score of 17. M's pattern of verbal statements also proved to be quite similar to participant C's, as illustrated in Figure 4. Of all



Note: The key in the upper left-hand corner explains the abbreviations used. All frequencies indicate the number of specified statements (e.g., RT/RI) which preceded correct or incorrect selections. All selections in tutorial and remedial trials were correct (as the correct choice was revealed to the participant).

Figure 5. Protocol Data for Participant M.

statements which preceded correct selections by M, 68.7 % were repeats of tacts to sample stimuli in the tutorial and remedial trials (RT), and intraverbals similar to tacts of choice stimuli in the same conditions (RI). No other type of relationship exceeded a frequency of two, as illustrated by the second graph of Figure 5.

Of all the statements preceding incorrect responses, 67.8% were repeated tacts of the sample stimuli emitted in previous tutorial or remedial trials, but not an intraverbal similar to tacts emitted in the same previous trials (namely the "RT/NI", "RT/-", and "RT/REI" conditions in the third graph of Figure 5). This was a similar pattern found with Participant C. No other type of statement preceded more than three incorrect selections.

M's statements during remedial trials were also similar in type to C's, but a bit more diversified (see last graph of Figures 4 and 5). Remedial trials in which tacts to sample stimuli were not emitted presided over all other types (53.1% of all statements, composed of "-/-" and "-/RT" statements). While the variety of M's types of statements was greater than in C's case, no statement type exceeded a frequency of 4 (see last graph of Figure 5).

M also demonstrated several relations in which the verbal statements were made more efficient with each subsequent trial:

1. It looks like a jail, hmmm, that looked like somebody head butting somebody (Tutorial).

2. That's gonna be jail, that's gonna be the head butt (correct).

3. That's jail, where's the head butt guy? There's the head butt guy. (correct).

M also explicitly demonstrated the use of elimination techniques as illustrated below. Note that several words were changed in this account, as Participant M used some graphic wording. However, the essence of M's verbal account is held intact:

1. Ok the black that goes straight across (Tutorial).

2. The black sun, can't be cup, done cup. it's gonna be dump truck hmmm... this one (Incorrect).

3. Ahh that's gonna be black sun that shoots straight across (remedial).

4. Ok, there's a sun it's black there is no blinds it's not cup, it's... it's dump truck, it's dump truck. (incorrect).

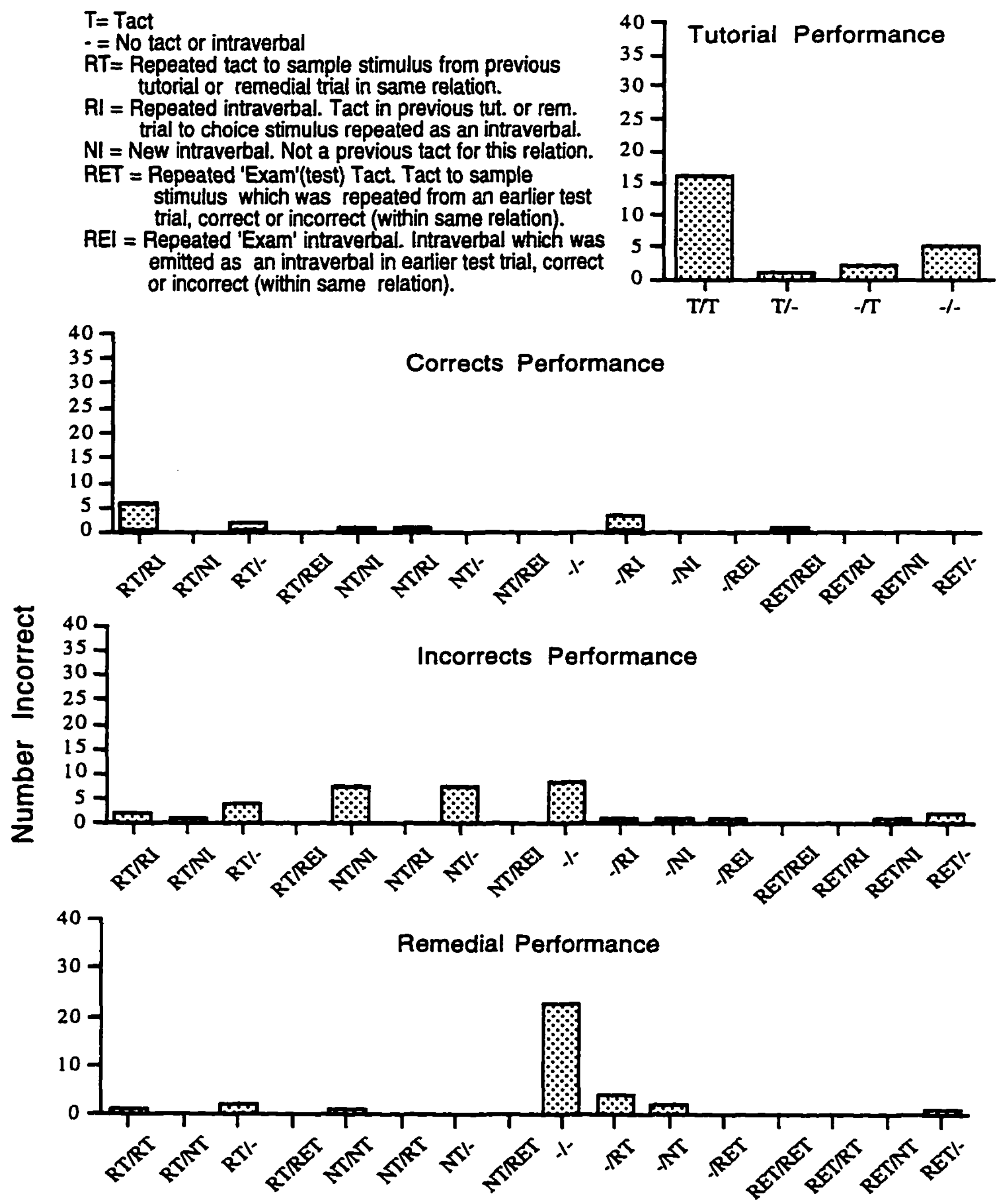
5. No it's not at all that, that's the one where it's just horizontal (remedial).

Note the elimination of choice stimuli with the verbal statement "can't be cup, done cup" and "it's not cup". It is interesting to note that M made the same mistake and type of statements in the next exposure to this relation, yet in each remedial easily made a similar tact to the choice comparison as was made in the tutorial.

Participant P

P performed the worst of all participants, but was very close in performance to both C and M, with a total of 14 correct responses out of a possible 48. As with the other participants, P generally emitted tact responses to both the sample and choice stimuli in the tutorial trials, the "T/T" category, composing 66.7% of all statements made in this condition. P had the highest number of tutorial (and remedial) trials in which neither the sample or choice stimuli evoked overt tacts (the "-/-" category, see Figure 6). In the tutorial condition, these types of statements composed 20% of the total number of statements made (a frequency of 5). The only other participant who made such a statement type in the tutorial condition was M, but only once.

Of all statements preceding correct selections, P's performance was similar to all other participants by having the highest number in the "RT/RI" condition (see the second graph of Figure 6). This type of statement preceded 42.8% of all correct



Note: The key in the upper left-hand corner explains the abbreviations used. All frequencies indicate the number of specified statements (e.g., RT/RI) which preceded correct or incorrect selections. All selections in tutorial and remedial trials were correct (as the correct choice was revealed to the participant).

Figure 6. Protocol Data for Participant P.

selections. The variety of statements preceding P's correct selections was greater than any other participant; however, none of these other statements occurred more than three times.

The type of statements preceding incorrect responses was also more diversified with participant P than any other participant. The highest number of statements was composed of neither tacts nor intraverbal statements of preceding tutorial or remedial trials (see condition "-/-" in graph 3 of Figure 6). This type of statement composed 23.5% of all statements preceding incorrect selections. Statements in which a new tact was made to the sample stimulus and either no intraverbal response occurred, or a new one for the relations under consideration occurred (the "NT/-" and "NT/NI" conditions in the third graph of Figure 6), accounted for 41.1% of all statements preceding incorrects. Finally, the statement type "RT/-" accounted for 11.7% (four incidents) of the statements preceding the incorrect selections. No other statement type exceeded a frequency of two incidents.

P had a remarkably high number of statements in the remedial condition in which neither the sample nor the comparison stimulus were overtly tacted. This type of statement composed 67.6% (23 incidents) of all remedial statement types. No other statement type exceeded 4 incidents (see last graph of Figure 6).

P also showed some use of elimination as indicated below in a transcript of one particular relation:

1. (prompted) I'm trying to think how I'm going to remember this. I don't know (Tutorial).

2. Ok the square, ummm ... this one? (Incorrect).

3. This one..... I have no idea it doesn't look like anything (remedial).

4. Ok... horizontal lightly shaded, was what, backwards flag? Wasn't that one, no, try, no, cuz that's a semicircle, lets try this one (Incorrect).

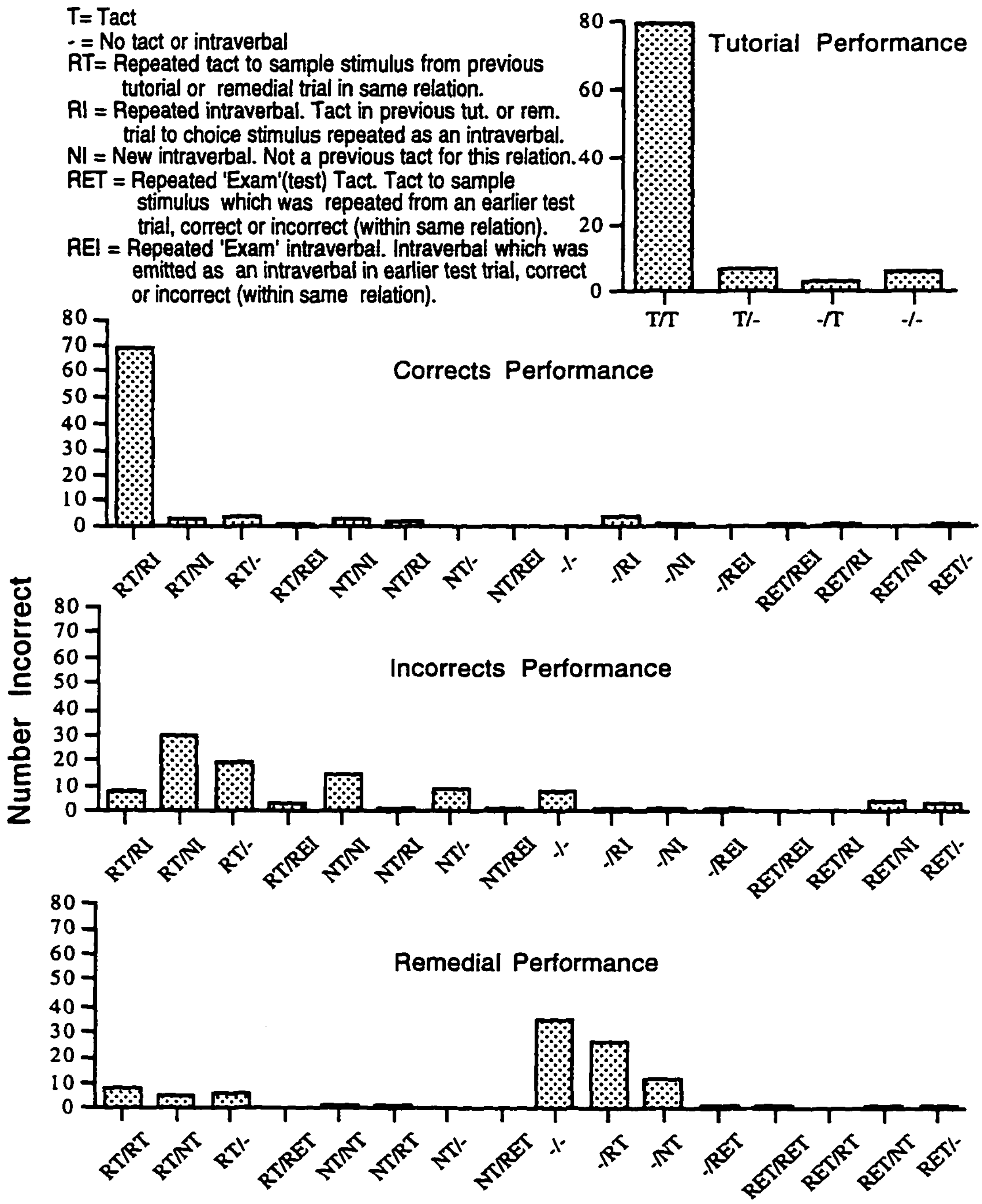
5. Don't know what that looks like, doesn't look like anything, backwards flag (remedial).

Note in number 4 above that P eliminates backwards flag as the "name" for the sample.

Group Data

Figure 7 shows the aggregate data for all participants in session 5 (a total of 4 participants) on the protocol analysis. In the tutorial condition, all participants made some tact to the sample and choice stimuli (both stimuli were apparent to the participant). In a few instances, some participants emitted a tact response to only the sample or the choice stimulus, as noted in the "T/-" and "-/T" bars of the upper left hand graph of Figure 7. Note that participant P was responsible for nearly all of the incidents in which tact responses were not emitted to both the sample and the choice stimuli ("-/-" bar). M was the only other participant to emit such a response, and did so only one time. It is clear from the graph, however, that, in nearly all cases, a tact response was emitted to both stimuli in the tutorial condition - in fact, of all tutorial trials, 83.1% of the time a tact response was made to both the sample and choice.

The second graph in Figure 7 shows the frequency of correct choices which occurred after each type of verbal statement emitted. Nearly all corrects occurred after a tact response to the sample stimulus and an intraverbal response occurred, modeled after that initial tact (the "RT/RI" condition). These tact and intraverbal responses were nearly always the same or similar to the original tact responses which occurred in the



Note: The key in the upper left-hand corner explains the abbreviations used. All frequencies indicate the number of specified statements (e.g., "RT/RI") which preceded correct or incorrect selections. All selections in tutorial and remedial trials were correct (as the correct choice was revealed to the participant).

Figure 7. Protocol Data for All participants.

preceding tutorial or remedial conditions. In fact, of all statements preceding correct trials, 76.6% fell in this category. No other type of verbal statement, by itself, preceded more than 4.4% of the total correct responses. The third graph in Figure 7 shows the frequency of incorrect choices which occurred after each type of verbal statement emitted. The highest number of incorrects occurred after the participants emitted the same tact response to the sample stimulus as they did in the tutorial or remedial conditions, but then either emitted a different intraverbal response than previously emitted ("RT/NI" verbal statement) or did not emit an intraverbal response at all ("RT/-" verbal statement). These two types of statements preceded 47.5% of the incorrects which occurred. The next type of verbal statement which preceded the most number of errors was when the participant emitted a new tact to the sample stimulus and a new intraverbal (that is, neither had been emitted in that relation's previous tutorial or remedial trials). This type of statement ("NT/NI") preceded 13.5 % of the total incorrects emitted. In all, these three types of statements composed 61.1% of the total statements preceding incorrect selections. No other type of verbal statement preceded more than 8.7% of the total incorrect selections emitted. Finally, during the remedial condition, the participants generally did not make a response controlled by the sample stimulus, and either did not emit a tact response to the choice (the "-/-" verbal statement type), emitted the same response as in a previous tutorial or remedial trial for that relation ("-/RT") or emitted a new tact to the choice stimulus ("-/NT"). These three types of statements preceded 70.8% of all remedial selections.

In addition to the previously discussed measures, the protocol transcripts were analyzed for several other types of verbal statements not shown in Figure 7. These included tacts of relationships between the sample and choice. For example, part of participant B's transcript read:

1. Descending black planet ricocheting (tutorial).

2. Moving gray planet, ahhh ricocheting (correct choice).

3. Descending black planet, ok ricochet, ricochet, ricochet, ricochet... hmmm... that worked, a good guess (correct choice).

This verbal statement was considered a tact of a relationship between the sample stimulus and the choice stimulus. Compare this to the following part of B's transcript which was not counted as including a relationship tact:

1. Gray chip little peak (tutorial).

2. Gray chip little peak (correct choice).

3. Gray chips small peak (correct choice).

B's statements reflected relationships 38 times in test conditions. Of these 38 times, only 5 preceded incorrect responses. It is interesting to note that B's statements also tended to be shorter than the statements of the other participants, which might have also contributed to B's high success rate (B had 42 corrects out of a possible 48 in the protocol condition, more than double the number of correct responses of any other participant).

As can be seen in the first example from B's transcript, participants occasionally repeated a phrase (echoic responses) while searching for the "matching" choice. This technique constitutes an efficient method for increasing the effectiveness of a scanning repertoire. Michael (1985) notes "... if the scanning takes much time, the effectiveness of the nonverbal stimulus will be lost by the time the appropriate verbal stimulus is encountered." (p. 4). The number of incidents of this type of repetition was counted for each participant. Interestingly, B had the highest number of such incidents, with a total of 10 and all were intraverbals which had been onetime tact responses (in the tutorial and remedial conditions) to the choice stimuli. M's transcripts revealed 3 such repeats and P's showed 2. Participant C did not overtly repeat phrases.

DISCUSSION

Overall, there was no reliable difference between the PTP and NPTP performances, across all measures. While individual participants tended to perform better in one condition than in the other (e.g., three performed better in the PTP condition, and three performed better in the NPTP condition), these differences tended to be quite small and most likely do not represent a true difference between the two conditions. This same pattern was evident in terms of IV and tact performances: three participants performed better in the IV condition and three in the tact condition. These data are similar to those obtained by Cresson (1994), Tan et al. (1995) and Bristow and Fristoe, (1984) in that the participants in this study were normally functioning and had strong verbal repertoires, and showed only slight differences between SB and TB (PTP) performances.

Some commonalties existed in all participants' performances and reports, namely the tendency to have improved performance over sessions, the tendency to have poorer performance in the first component of each session, and the fact that each reported to have used some form of TB verbal strategy to improve their performance (which led to including session 5 and the protocol analysis).

It is not clear what caused the improvement over sessions. Intuitively, it would seem not to be due to becoming more familiar with the task, as by the end of the second session each participant had a minimum of 420 individual trials with the task (this is assuming 100% accuracy, which of course did not occur) spread over two sessions. From the exit interviews and analysis of the protocol tapes, it seemed more likely that participants were developing skills in generating TB responses to the sample and choice stimuli. As noted in the protocol analyses, typical verbal statements consistently preceded selection of correct and incorrect choice stimuli (the evidence is stronger for those preceding correct choices, however). This hypothesis was partially supported by introducing new sample stimuli in session 5, ones which were presumed to be harder to "name" distinctively. Performance decreased as expected. It could be argued, however, that this decrease was simply a function of less discriminable controlling stimuli, and thus required more trials to acquisition.

Comparing nonhuman and nonverbal human performances to human verbal performances, it would appear that differences exist which are more than simple innate discrimination abilities. Researchers studying nonhuman conditional discrimination acquisition report that many trials are needed for acquisition, many more than participants in this study required. For example, Cumming and Berryman (1961) examined matching-to-sample performance in pigeons, using only two choice stimuli. Acquisition at the 90% accuracy level occurred after a minimum of approximately 350 trials for each relation for one bird, and after a maximum of approximately 560 trials for each relation for two others birds. Of course, the larger number of required trials could be due to a species difference, thus the comparison with human nonverbal performance. For example in Sundberg & Sunberg's (1990) study, four mild to moderately mentally retarded adults required an average of approximately 295 trials (excluding remedial and pretraining exposures) to acquire (90% accuracy) three SB tact relations, an average of approximately 195 trials to acquire three SB intraverbal relations, 60 trials for three TB tact relations and 150 trials for three TB intraverbal relations. Dividing these figures by the approximate number of trials per relation results in the following number of trials for each individual relation: SB tact, 98; SB intraverbal, 65; TB tact, 20; TB intraverbal, 50. Similar results were obtained in

Wraikat (1990, 1991), although the number of relations trained were somewhat different (three for some participants, six for others) and interspersal training was also included. Compare these data to the results of this study in which the best performance of session 4 reached 90% accuracy in as few as two exposures (trials) for participant B for ten relations in all conditions (i.e., PTP tact, PTP intraverbal, NPTP tact, NPTP intraverbal). The largest number of trials to acquire (90% criterion) all ten relations in all conditions during session 4 was eight (participant S, which as discussed previously, was likely due to a misunderstanding of the experimental instructions).

However, again one could argue that this difference simply reflects differences between normally functioning and DD humans, instead of differences in the application of verbal behavior. More research is needed in this area. Although unlikely, it may be possible to develop a task which precludes the use of verbal behavior in completing that task. The only other option is to find physically normal humans who have no verbal skills, again another unlikely possibility. Some researchers (Stephens & Hutchison, 1992) are taking a different tack, arguing that operant principles, modeled on a computer (adaptive network systems) may provide sufficiency arguments for the development of verbal behavior via operant conditioning. To the extent that this technique parallels organismic learning, it is useful in that all environment - behavior relations (overt or covert) are readily available for observation.

It is difficult to account for the poorer performance demonstrated by all participants in the first condition of each session. If it was simply a gradual improvement over subsequent trials, one would not expect to see the resurgence of errors in the second portions of sessions 1 and 5, which was demonstrated by participants B, M and P. It is possible that a "warm-up" was necessary for participants to perform well, and that this was disrupted by allowing participants to take a break midsession. It is interesting to note that some researchers have reported or controlled for "warm up" effects when conducting research with nonhumans (Cumming & Berryman, 1961). What this warm-up effect consists of and whether in this research it was a verbal or nonverbal phenomenon is unclear. Additional investigations might examine the effect of extended training (e.g., weeks of training) to determine if this phenomenon is reduced, eliminated or remains stable over time.

Finally, the fact that all participants reported using some verbal technique to aid their performance led to incorporating session 5 and the protocol analysis in this experiment. The results of the protocol analysis are clear, but must be taken as only correlational. That is, these data are only important to the extent that verbal statements influence the selections which follow such statements. As discussed above, it is not possible to clearly manipulate the verbal statements a participant makes, especially considering that humans are quite capable of making covert verbal responses. However this evidence does support the similarly found results of other researchers in this area (Wulfert, et al. 1991).

Potential Causes for the Lack of PTP and NPTP Differences

Several reasons exist as to why few differences were found between the PTP and NPTP conditions. First, it is possible that unique response-produced stimulation has no effect on the development of conditional discriminations or the development of emergent relations. Little if any research has been conducted in this area (i.e., distinctive responses to the comparison stimuli). Some researchers have investigated the variables affecting matching-to-sample or delayed matching-to-sample performance, but these accounts tend to arrange for differential responding to the sample stimulus only.

Second, it's possible that the response-produced stimulation arranged for in this study was not salient enough to affect acquisition. The differences between clicking on each of the dots versus on the corners is not a large one. This task was chosen for the number of variations that could be arranged, in addition to its ease of use in the computer medium. Initially, the experiment was arranged to consist of several levels of response-produced stimulation. In addition to the unique kinesthetic and visual feedback, different tones were arranged to sound after each click, unique for each dot position, thus resulting in a unique series of four tones for each stimulus. However, this was not incorporated in this study due to time constraints - sessions would have far exceeded the desired two-hour limit. A follow-up study could manipulate the level of stimulation provided, increasing the salience this variable might have on the acquisition and the emergence of equivalence. Given the results obtained in this study, however, it seems unlikely that such manipulations will have a very large effect, for the reasons described next.

Third, it is possible that existing verbal repertoires were brought to bear on the experimental task, and these repertoires obscured any effect that differences in response-produced stimulation might produce. It was interesting to note that virtually all experimental and pilot participants reported the same strategy in post-session interviews: they used verbal behavior to perform more accurately. In addition, all participants indicated that they employed a particular strategy in responding under the equivalence condition. This strategy was similar for all participants, namely carrying comparison names from the IV condition into the tact condition (that is, using the same names for tact comparison stimuli). In general, these findings are consistent with postsession reports and anecdotal observations obtained by other researchers (Cresson, 1994; Dugdale & Lowe, 1990; Stratton, 1992; Sundberg & Sundberg, 1991; Wulfert et al. 1991).

Sidman (1994) notes that participants in certain studies were given exit interviews to determine if they had developed names or used the names of objects to aid their performance. Little evidence of "naming" was demonstrated. The same procedure was followed in this study, and as Sidman found, the participants were not able to recall many of the actual nonsense syllables used as samples in this study. When pressed to recall, however, several of the participants were able to recall distorted English names for some of the choice stimuli. Several of the participants also expressed some hesitation in revealing those names, as they said they were "goofy" or something of that nature. Something akin to this was seen with two of the participants in the protocol analysis. Both women, they commented afterwards that in the actual session they were much more effective in generating names for the stimuli presented to them. However, the researcher's presence (a male) caused them to edit the names they supplied (several comments on the tape recording support these statements). The other two participants (both male) stated that they did not have any reactivity occur during the protocol analysis, which is partially supported by the somewhat graphic nature of one of the participant's labels - although it is possible that this too is indicative of reactivity. It is possible that postsession analyses are affected by too many extraneous variables to give insight into what actually occurred in the experiment. This is the reason why the concurrent "talk aloud" method (Ericsson & Simon, 1993) was used in this experiment, as was conducted in the Wulfert et al. (1991) study. Given the results of the protocol analysis (discussed below) and the supporting evidence from other researchers, it seems likely that verbal behavior was mediating the SB responses in this study. It should be noted, however, that no direct causal relationships can be inferred from protocol analyses; only a sufficiency argument can be made.

As noted earlier, as Sidman (1994) points out, and as the nonhuman research literature supports, it is clearly possible for conditional discriminations to develop without the capacity for verbal behavior. This is not as clear in regard to the demonstration of stimulus equivalence, however.

It seems likely that the number of choice stimuli used (or relations trained) may be relevant in the necessity of using existing verbal repertoires to aid in successful SB responding. In many of the equivalence and nonhuman studies, only two choice stimuli are arranged. This study used a minimum of ten relations, and a maximum of 14. It would seem likely that the more difficult the task, the more likely verbal mediation would occur. Something of the sort has been offered as an explanation for remembering (Donahoe & Palmer, 1994). These researchers drew a distinction between reminding and remembering. Reminding is classified as simple stimulus control, in which a stimulus controls a response (either in a respondent or operant manner). Remembering, however, occurs in a situation in which a response is scheduled to be reinforced, but for various reasons it can not immediately be evoked by the present stimulus conditions. The current stimulus conditions then evoke a series of responses (often verbal) which terminate in the production of the target response. This is one of several behavioral processes which have been labeled "problem solving". It is possible that such a process occurred in this study, given the difficulty level incurred when using such a relatively high number of relations to be learned. Stratton's (1992) manipulation showed clear differences when he manipulated the number of relations trained. However, no direct evidence (anecdotal data only was reported) was recorded for verbal mediation of the SB task in that experiment.

Protocol Analysis and Conclusions

Protocol analyses are not frequently used in behavior analysis, probably for both historical and practical reasons. Historically, such techniques were used primarily for gathering information on thoughts and to aid in uncovering inferred cognitive processes (Hergenhahn, 1986). Inferences and unsupported observations are not thought to be effective or necessary methods for uncovering functional relationships between environmental events and behavior (Skinner, 1974). However, in its current usage, and as used by Wulfert et al. (1991), the protocol analysis was used to make overt potentially important covert verbal behavior. This study extended those researchers work by encoding the protocol analysis in terms of Skinner 's (1957) elementary verbal operants. The results of this study supported those of Wulfert et al. (1991), showing relatively clear results in terms of verbal statements which accompany successful SB responses.

As noted earlier, some researchers have examined the utility of using verbal reports in behavior analysis (e.g., Hayes, 1986; Perone, 1988). Shimoff (1984) states (as cited in Perone, 1988): "an experimental analysis of behavior generally seeks causes of behavior in the environment, not in other behavior. Verbal behavior may serve as an intermediate cause, as when it is part of an extended chain preceding some nonverbal response, but an experimental analysis will trace the chain to its environmental origins" (p 74). It is believed that this study lives up to the spirit of this statement. The protocol analysis is used here only as a tool to clarify relations among stimuli and responses, whether they be response-produced (e.g., verbal behavior) or not.

In this study, the protocol analysis allowed direct observation of strategies used which were correlated with SB performance. The observation of exclusion responding, repeating phrases until the choice stimulus was selected, and completing a verbal statement after seeing the appropriate choice stimulus all seem to indicate that verbal strategies were used to increase accuracy in the SB task. Some researchers had hypothesized that these strategies do take place (Cresson, 1994; Michael, 1993; Stratton, 1992). Additional research might involve training participants to respond in these manners and to test accuracy and performance prior to and after such training. As a start, some researchers have examined the utility of teaching names to participants for various components of the conditional discrimination task. In general, such training facilitates acquisition (Dugdale & Lowe, 1990).

Behavior analysis can offer some improvements on the analysis of verbal reports. For example, by stressing the controlling variables of such reports, and the consequences following such reports, one obtains a more complete picture of the speaker and of the functionality of the responses. In addition, by adopting a behavior analytic approach, one is required to trace the chain of responses back to the controlling environmental stimuli, in addition to the histories required to establish such stimulus control. This would provide a very complete picture of the phenomena under study, one most likely complete enough that the behavior could be predicted and controlled. This has obvious implications for applied problems.

Considering the evidence presented here, and the high frequency of verbal behavior emitted by normally functioning humans, it seems likely that many SB tasks, at least those tasks of moderate difficulty, are in some manner facilitated by the use of that verbal behavior. It is unclear, however, how verbal behavior accomplishes this facilitation, although as noted previously, this is under investigation.

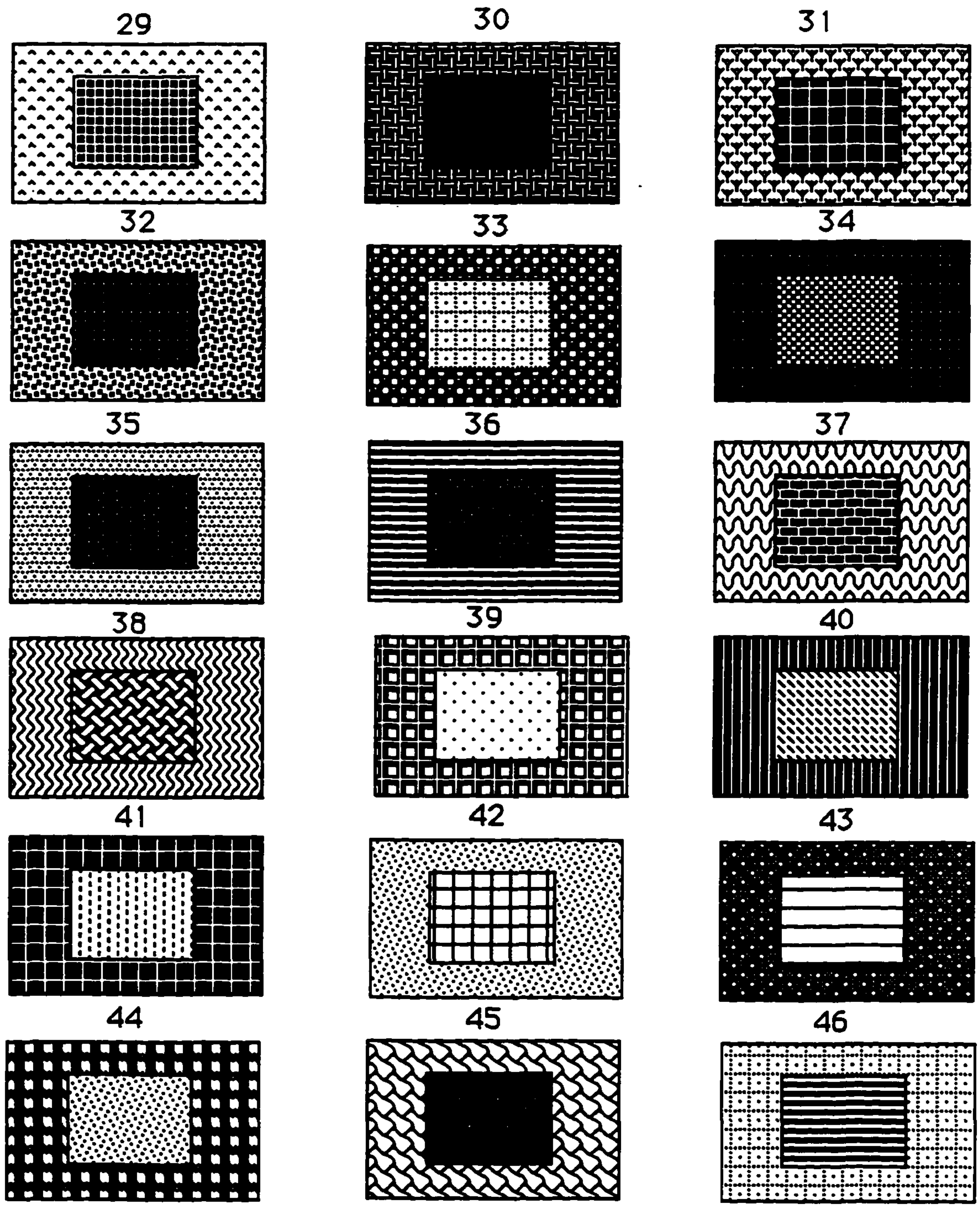
If TB responding does aid SB responding, then directly programming this into SB language training programs (e.g., communication board training) may facilitate the acquisition of those SB responses. Of course, this would only be useful in cases where the person being trained is incapable of learning extensive TB responses (this form of communication would appear to be more efficient). Some researchers have reported increases in vocal verbal behavior with the advent of SB training (Calculator & Luchko, 1983; McDonald & Schultz, 1973).

These findings, along with others indicate that more research and possibly new research techniques are needed to investigate this area. Protocol analyses offer a good starting point, but more rigorous accounts are necessary to provide for causal information. Some researchers have attempted to operationalize verbal reports in a more rigorous manner (Critchfield & Perone, 1990; Lane & Critchfield, 1996). These researchers arranged for structured self-reports to be taken immediately after conditional discrimination and equivalence trials. While these reports only indicate a participant's description of the accuracy of a immediately preceding response, they do so in an easily measured and observed manner, a start in a difficult area to research.

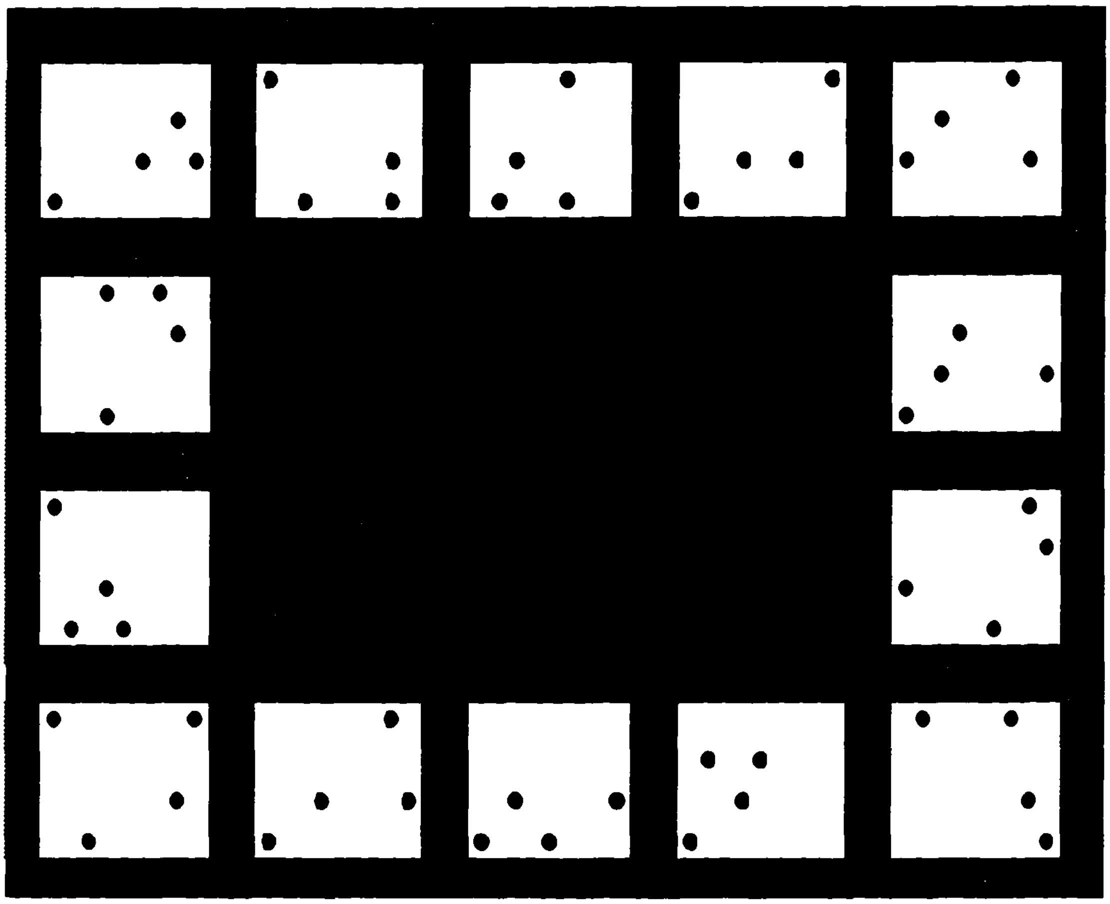
In regard to the differences between SB and TB verbal behavior, the results of this study proved inconclusive. Intuitively, it would seem that differential response-produced stimulation would aid in acquisition, but given the highly developed vocal-verbal skills of the participants used in this study, and the nature of the task, this was not demonstrated. It is hoped that the results of this study contribute to a better understanding of the types of variables operating in such research, and provide initial methodologies for analyzing these variables in light of existing verbal categories (the elementary verbal operants, Skinner, 1957).

Appendix A

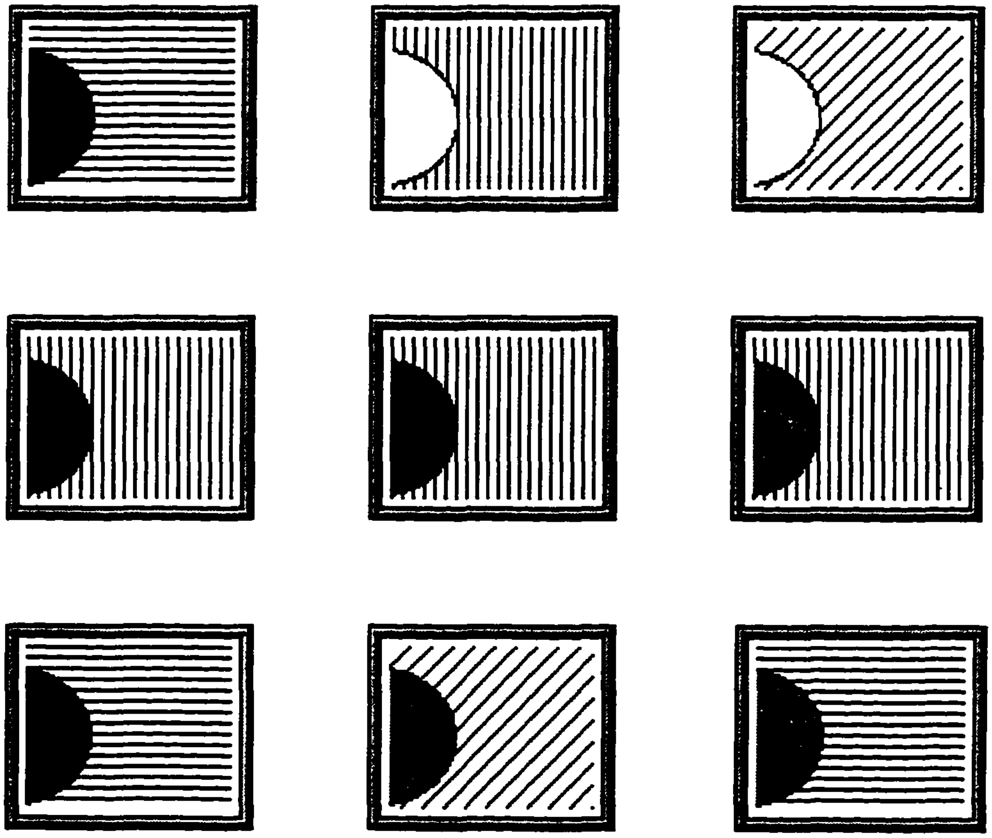
Stimulus Materials



Sample "flag-like" patterns.



Sample dot patterns in rectangular array.



Sample of the patterns used in session 5 (more difficult to discriminate).

nalwoz	Greby
yedtis	snellik
nofal	tigflub
reesvard	llamswam
Quoey	Errip
frolin	bermot
frebnib	Delkey
roetez	nalin
ugsorg	Subka
Tevib	chegor
diffluff	zibprok
umstip	Lishro
zekwone	tiestrome
Ibird	ebsnick
wikstead	nislar
hargiss	zebbult
razip	yilninj
quorgor	fasjolm
zebnor	wibflim
fraber	murte
vobel	zleleep
wopspu	Mokben
Etsor	Hutgin
Hebo	gliserk
Falto	yezdor
sotixed	Palird
hilzut	dohquib
Portuk	yislop
subtin	Galshay
igblik	Ribgin
ostmoog	brosork
xerlb	Imkul
twoarb	plonip
Fotib	Clible
Blicken	uncelib
Nasten	izlu
Dibsten	disblic
Taobe	ledjnor
duate	vaunee
pridsler	kestblu
quintor	pauksif
Jenkel	lotrab
Mishy	vineg

Sample nonsense words.

Appendix B

Forms

Informed Consent Form

Western Michigan University
Department of Psychology
Principal Investigator: William Potter
Co-Principal Investigator: Dr. Jack Michael

I have been invited to participate in a research project entitled "Comparison of Selection-based vs. Topography-based Verbal Behavior". I understand that this research is intended to examine the differences in acquiring language when a distinct response is required and when it is not. The overall purpose of this study is to examine which method leads to better or new language acquisition. I further understand that this project is William Potter's dissertation project.

My consent to participate in this project indicates that I will be asked to attend four private sessions with William Potter, each lasting between 1.5 and 2.5 hours. These sessions will take place in an office in the Classroom Building at California State University, Stanislaus. Prior to each session Mr. Potter will greet me, set up the computer which I will work on, then I will be alone for the remainder of the session, working on the tasks arranged for me by Mr. Potter on the computer. The tasks will consist of learning the relationship between either nonsense syllables or patterns, and symbols displayed on the screen of the computer. I will be using the computer mouse to click on the displayed symbols on the screen. At the end of each of these sessions I will be paid $5 and a bonus based on my performance. At the end of the final session Mr. Potter will ask me a series of questions about what I was thinking during the sessions. Mr. Potter will also explain the reasoning behind the experiment and answer any questions I might have concerning the study.

As in all research there may be unforeseen risks to the participant. If an accidental injury occurs, appropriate emergency measures will be taken; however, no compensation or treatment will be made available to me except as otherwise specified in this consent form. I understand that one potential risk of my participation in this project is that I may experience some stress when the computer presents the various choices from which I will choose a symbol to click on. This stress is expected to be no worse than that I experience when taking an exam. If I want to stop working during a session, I will simply have to tell Mr. Potter (who will be outside the office) that I do not wish to continue. I will be able to withdraw at any time during the experiment by simply telling Mr. Potter in person or telephoning him at 667-3255.

One way in which I may benefit from this activity is by having the chance to experience and talk about how psychologists study language. I also understand that others who study language may benefit from the knowledge gained from this research. In addition, I will gain some insight into the dissertation process, which Mr. Potter is in the process of completing.

I understand that all information collected from my participation in this study is confidential. This means that my name will not appear on any papers on which this information is recorded. The computer will keep track of the data from each session, with no information contained in that data identifying me. A code will be used to identify which data belongs to me. A single master list will be maintained linking the codes to the names of each participant. This list will remain in a locked box in the Classroom Building, with Mr. Potter being the only person with access to that box. Once the data are gathered and analyzed, the master list will be destroyed. In order to protect my confidentiality, any reference to the data from this study will incorporate only the codes assigned to each participant (this includes both written and oral presentations of the data from this study).

I understand that I may refuse to participate or quit at any time during the study without prejudice or penalty. If I have any questions or concerns about this study I may contact either William Potter 667-3255 or Jack Michael (616) 387-8325. I may also contact the chair of the Human Subjects Institutional Review Board at (616) 387-8293 or the Vice President for Research at (616) 387-8298 with any concerns that I have. My signature below indicates that I understand the purpose and requirements of the study and that I agree to participate.

____ ________________

Signature Date

Exit Interview

1. Do you remember any strategies/techniques that you used when doing the part of the program in which you heard a nonsense symbol and were required to select a particular dot pattern? Write your comments here:

   What did you think was occurring in this part of the program?

2. Do you remember any strategies/techniques that you used when doing the part of the program in which you saw a flag-like pattern and were required to select a particular dot pattern? Write your comments here:

   What did you think was occurring in this part of the program?

3. Do you remember any strategies/techniques that you used when doing the part of the program in which you heard a nonsense symbol and were asked to select a particular flag-like pattern? (this was the part in which you were not informed if your selection was correct or not)Write your comments here:

   What did you think was occurring in this part of the program?

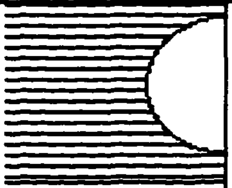
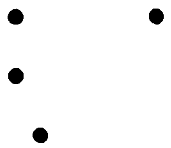
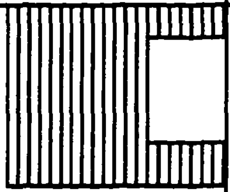
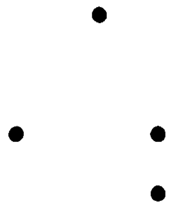
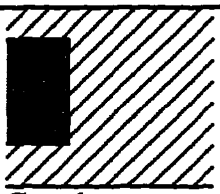
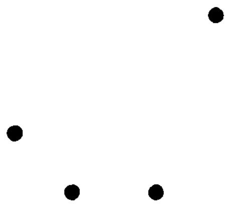
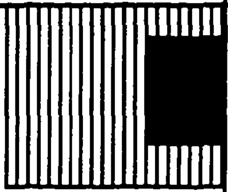
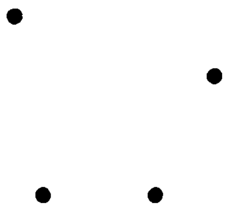
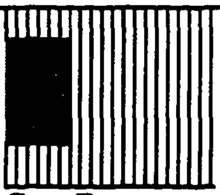
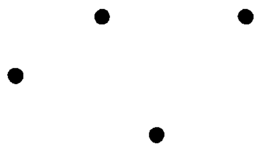
4. Did you prefer or dislike any part or aspect of the program over another? (specify which). Remember you had situations in which you:
   A. Heard a sound and selected a dot pattern,
   B. Saw a flag-like pattern and selected a dot pattern
   C. Clicked on the corners of the selected dot patterns
   D. Clicked on each dot of the selected dot patterns
   E. Heard a sound then selected a flag-like pattern (with no feedback)

If you preferred or disliked any of these, please explain:

5. Of the conditions listed in #4, which did you think you did best on?

6. Below, fill in as many of the dot patterns as you can remember: (use the next page if necessary):

C = Correct I = Incorrect T = Timeout

Tut	Talk Aloud (comments each trial)	Sample	Comparison Selected	
PTP	OK sun under the picture very bright	 Sun	 Sun (Picture)	C
	a * peak of white boxes	 White Box	 Peak	C
	mowing the gray lawn	 Gray lawn	 L.M.	C
	a basket full of black bars	 Black Bars	 Basket	C
	a mountain range of gray bars... or a Z	 Gray Bars	 Mountain Range or Z	C

Sample Transcript From Protocol Analysis of Participant B. These were used for participant clarifications.

Appendix C

Human Subjects Approval Form

Human Subjects Institutional Review Board

Kalamazoo, Michigan 49008-3899
616 387-8293

WESTERN MICHIGAN UNIVERSITY

Date: November 20, 1995

To: Wiliam Potter

From: Richard Wright, Chair

Richard A. Wright

Re: HSIRB Project Number 95-07-05

This letter will serve as confirmation that the change to your research project "Comparison of selection-based versus topography-based verbal behavior" requested in your memo dated November 17, 1995 has been approved by the Human Subjects Institutional Review Board. This change is:

1. audio taping during exit interviews.

The conditions and the duration of this approval are specified in the Policies of Western Michigan University.

You must seek reapproval for any changes in this design. You must also seek reapproval if the project extends beyond the termination date.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: October 4, 1996

xc: Jack Michael, PSY

Human Subjects Institutional Review Form.

BIBLIOGRAPHY

Anderson, J. R. (1980). Cognitive skills and their acquisition. Hillsdale N. J.: Lawrence Erlbaum Associates, Inc.

Apple Computer, Inc. (1989). HyperCard (version 2.2). Cupertino, CA: Author.

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

Bliss, C. K. (1965). Semantography. Sydney, Australia: Semantography Publications.

Bristow, D. & Fristoe, M. (1984). Learning of blissymbols and manual signs. Journal of Speech and Hearing Disorders, 49, 145-151.

Calculator, S., & Luchko, C. D. (1983). Evaluating the effectiveness of a communication board training program. Journal of Speech and Hearing Disorders, 48, 185-191.

Chase, P. N., Johnson, K. R., & Sulzer-Azaroff, B. (1985). Verbal relations within instruction: Are there subclasses of the intraverbal? Journal of the Experimental Analysis of Behavior, 59, 301-315.

Cresson, O. (1994). The writing response in studies of topography-based and selection-based verbal behavior. Unpublished doctoral dissertation, Western Michigan University, Kalamazoo, MI.

Critchfield, T. S., & Perone, M. (1990). Verbal self-reports of delayed matching to sample by humans. Journal of the Experimental Analysis of Behavior, 53, 321-344.

Critchfield, T. S., & Perone, M. (1993). Verbal self reports about matching to sample: Effects of the number of elements in a compound sample stimulus. Journal of the Experimental Analysis of Behavior, 59, 193-214.

Cumming, W. W., & Berryman, R. (1961). Some data on matching behavior in the pigeon. Journal of the Experimental Analysis of Behavior, 4, 218-284.

Donahoe, J. W., Palmer, D. C. (1994). Learning and complex behavior. Needham Heights, MA: Allyn and Bacon.

Dugdale, N., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.), Behaviour Analysis in Theory and Practice: Contributions and Controversies (pp. 115-138). Hillsdale, NJ: Erlbaum.

Epstein, R., Lanza, R. P., & Skinner, B. F. (1980). Symbolic communication between two pigeons (columba livia domestica). Science, 207, 543-545.

Ericsson, K. A., & Simon, H. A. (1993). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.

Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: A methodological analysis. Journal of the Experimental Analysis of Behavior, 42, 143-157.

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

Hayes, S. C. (1986). The case of the silent dog - Verbal reports and the analysis of rules: A review of Ericsson and Simon's *Protocol Analysis: Verbal Reports As Data*. Journal of the Experimental Analysis of Behavior, 45, 351-363.

Hayes, S. C., & Hayes, L. J. (1989). The verbal action of the listener as a basis for rule-governance. In S. C. Hayes (Ed.), Rule-governed behavior: Cognition, contingencies, and instructional control (pp. 153-190). New York: Plenum Press.

Hergenhahn, B. R. (1986). An introduction to the history of psychology. Pacific Grove, CA: Brooks/Cole Publishing.

Hodges, P., & Schwethelm, B. (1984). A comparison of the effectiveness of graphic symbol and manual sign training with profoundly retarded children. Applied Psycholinguistics, 5, 223-253.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. Journal of the Experimental Analysis of Behavior, 65, 185-243.

Kazdin, A. E. (1982). Single-Case Research Design. Oxford, NY: Oxford University Press.

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

Lazar, R. M., Davis-Lang, D., & Sanchez, L. (1984). The formation of visual stimulus equivalence in children. Journal of the Experimental Analysis of Behavior, 41, 251-266.

Lowenkron, B. (1988). Generalization of delayed identity matching in retarded children. Journal of the Experimental Analysis of Behavior, 42, 1-18.

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

Lowenkron, B. (1996). Joint control and word-object bidirectionality. Journal of the Experimental Analysis of Behavior, 65, 252-255.

Lowenkron, B. & Colvin, V. (1995) Generalized instructional control and the production of broadly applicable relational responding. The Analysis of Verbal Behavior, 12, 13-31.

McDonald, E. T., & Schultz, A. R. (1973). Communication boards for cerebral palsied children. Journal of Speech and Hearing Disorders, 38, 73-88.

McIlvane, W. J., Klederas, J. B., Munson, L.C., King, A. J., De Rose, J. C. , & Stoddard, L. T. (1987). Controlling relations in conditional discrimination and matching by exclusion. Journal of the Experimental Analysis of Behavior, 48, 187-209.

McNaughton, S. (1976). Bliss symbol - an alternative symbol system for the non-vocal pre-reading child. In G. Vanderheiden & K. Gillery (Eds.), Non-vocal communication techniques and aids for the severely physically handicapped. Baltimore: University Park Press.

Michael, J. (1985). Two kinds of verbal behavior and a possible third. The Analysis of Verbal Behavior, 3, 1-4.

Michael, J. (1993). Concepts and principles of behavior analysis. Kalamazoo, MI: Association of Behavior Analysis.

Page, T. J., & Iwata, B. A. (1986). Interobserver agreement: History, theory, and current methods. In Poling, A. & Fuqua, R. W. (Eds.), Research Methods in Applied Behavior Analysis (pp. 99-126). New York: Plenum Press.

Pepperberg, I. M. (1988). Comprehension of "absence" by an african grey parrot: Learning with respect to questions of same/different. Journal of the Experimental Analysis of Behavior, 50, 553-565.

Perone, M. (1988). Laboratory lore and research practices in the experimental analysis of human behavior: Use and abuse of subjects' verbal reports. The Behavior Analyst, 11, 71-75.

Peterson, N. (1978). An introduction to verbal behavior. Grand Rapids, MI: Behavior Associates.

Romski, M. A., Sevcik, R. A., & Pate, J. L. (1988). Establishment of symbolic communication in persons with severe retardation. Journal of Speech and Hearing Disorders, 53, 94-107.

Rumbaugh, D. M. (1977). Language learning by a chimpanzee: The Lana project. New York, NY: Academic Press.

Saunders, R. R., & Green, G. (1992). The nonequivalence of behavioral and mathematical equivalence. Journal of the Experimental Analysis of Behavior, 57, 227-241.

Savage-Rumbaugh, E. S. (1984). Verbal behavior at a procedural level in the chimpanzee. Journal of the Experimental Analysis of Behavior, 41, 223-250.

Savage-Rumbaugh, E. S. (1990). Language acquisition in a nonhuman species: Implications for the innateness debate. Developmental Psychobiology, 23(7), 599-620.

Shafer, E. (1992). Teaching topography-based and selection-based verbal behavior. The Analysis of Verbal Behavior, 10, 1-10.

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.

Shane, H. C., & Bashir, A. S. (1980). Elective criteria for the adaptation of an augmentative communication system. Journal of Speech and Hearing Disorders, 45, 408-414.

Shimoff, E. (1984). Post-session questionnaires. Experimental Analysis of Human Behavior Bulletin, 2, 1.

Sidman, M. (1971). Reading and auditory-visual equivalences. Journal of Speech and Hearing Research, 14, 5-13.

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

Sidman, M. (1994). Equivalence Relations and Behavior: A Research Story. Boston, MA: Authors Cooperative, Inc.

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

Sigafoos, J., & Iacone, T. (1993). Selecting augmentative communication devices for persons with severe disabilities: Some factors for educational teams to consider. Australia and New Zealand Journal of the Developmentally Disabilities, 18(3), 133-146.

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

Skinner, B. F. (1974). About behaviorism. New York: Alfred A Knopf.

Stephens, K. R., & Hutchison, W. R. (1992). Behavioral personal digital assistants: The seventh generation of computing. The Analysis of Verbal Behavior, 10, 149-156.

Stratton, M. A. (1992). Comparing selection-based and topography-based language systems with verbal adults learning Japanese words. Unpublished master's thesis, Western Michigan University, Kalamazoo, MI.

Strayer, D. L., & Kramer, A. F. (1990). An analysis of memory-based theories of automaticity. Journal of Experimental Psychology: Learning, Memory and Cognition, 16, 291-304.

Sundberg, C. T. (1990). A comparison of a topography-based language system and a selection-based language system. Unpublished master's thesis. Western Michigan University, Kalamazoo, MI.

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.

Sundberg, M. L. (1990). Teaching language to the developmentally disabled. Pleasant Hill, CA: Behavior Analysts, Inc.

Sundberg, M. L. (1993). Selecting a response form for nonverbal persons: Facilitated communication, pointing system, or sign language? The Analysis of Verbal Behavior, 11, 99-116.

Tan, K., Bredin, J., Polson, D., Grabavac, D., & Parsons, J. (1995). The effects of topography-and selection-based training and testing on the emergence of symmetry. Poster presented at the Association for Behavior Analysis, 21st Annual Convention, May 1995, Washington, DC.

Terrace, H. S. (1979). Nim: A chimpanzee who learned sign language. New York: Alfred A. Knopf.

Vaughan, W. (1988). Formation of equivalence sets in pigeons. Journal of Experimental Psychology: Animal Behavior Processes, 14, 36-42.

Wallander, Jr., R. J. (1993). Effects of familiarity with a sample stimulus in selection-based learning of verbal behavior. Unpublished master's thesis, Western Michigan University, Kalamazoo, MI.

Watson, J. B. (1920). Is thinking merely the action of language mechanisms? British Journal of Psychology, 11, 87-104.

Wraikat, R. M. (1990). Teaching tact and intraverbal behavior to developmentally disabled adults: A comparison of topography-based and selection-based paradigms. Unpublished project, Western Michigan University, Kalamazoo, Michigan.

Wraikat, R., Sundberg, C. T., & Michael, J. (1991). Topography-based and selection-based verbal behavior: A further comparison. The Analysis of Verbal Behavior, 9, 1-18.Wulfert, E., Dougher, M. J., & Greenway, D. E. (1991). Protocol analysis of the correspondence of verbal behavior and equivalence class formation. Journal of the Experimental Analysis of Behavior, 56, 489-504.

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