Learning to Write without Writing: Using Conditional Discrimination Training to Establish an Expressive Repertoire

John Spear
Graduate Center, City University of New York
LEARNING TO WRITE WITHOUT WRITING:
USING CONDITIONAL DISCRIMINATION TRAINING TO ESTABLISH AN EXPRESSIVE REPertoire

by

JOHN R. SPEAR

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirement for the degree of Doctor of Philosophy, The City University of New York

2015
The manuscript has been read and accepted for the
Graduate Faculty in Psychology in satisfaction of the
Dissertation requirements for the degree of Doctor of Philosophy

Dr. Lanny Fields

November 6, 2014
Date
Chair of Examining Committee

Dr. Joshua Brumberg

November 11, 2014
Date
Executive Officer

Dr. Robert Lanson

Dr. Daniel Fienup

Dr. Erik Arntzen

Dr. Francis Mechner

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

LEARNING TO WRITE WITHOUT WRITING:
USING CONDITIONAL DISCRIMINATION TRAINING TO ESTABLISH AN EXPRESSIVE REPertoire

by

John R. Spear

Adviser: Professor Lanny Fields

Interpreting and describing complex information shown in graphs are essential skills to be mastered by students majoring in psychology. The dissertation describes conditional discrimination training procedures that induced students’ writing of descriptions of graphs. This training was designed to establish targeted conditional and joint stimulus control by elements of graphs and their printed descriptions. Thus, writing, a production-based performance, was induced by conditional discrimination training, a selection-based procedure.

Keywords: conditional discrimination, joint stimulus control, writing, college students
Dedication

This manuscript is dedicated to my teachers, friends, and family, who have been endlessly encouraging and have helped me to reach this point.
Acknowledgements

I cannot provide sufficient thanks to my advisor and mentor, Dr. Lanny Fields, for his unwavering commitment to such a challenging project. His quick, supportive, valuable, and seemingly endless feedback through this process was invaluable. My dissertation committee, Drs. Daniel Fienup and Robert Lanson, and outside readers, Drs. Erik Arntzen and Francis Mechener, also provided valuable insights and suggestions that could only come from fresh eyes. Friends and family have provided fun and relaxation over the years.
Table of Contents

Abstract ................................................................................................................................. iv

List of Tables ......................................................................................................................... viii

List of Figures ......................................................................................................................... ix

Introduction ............................................................................................................................. 1

Method................................................................................................................................. 11

Results and Discussion ......................................................................................................... 23

General Discussion ................................................................................................................. 42

Bibliography ........................................................................................................................... 71
List of Tables

Table 1. Glossary of terms and abbreviations of elements...............................50
Table 2. The value of each element presented during traditional training....................51
Table 3. All possible sources of stimulus control and those established by training........52
Table 4. Incorrect comparison stimuli used during Id training...............................53
Table 5. Incorrect comparison stimuli used during Ip training...............................54
Table 6. Incorrect comparison stimuli used during Is training...............................55
Table 7. Incorrect comparison stimuli used during non-relational training for no interaction........................................................................................................56
Table 8. Incorrect comparison stimuli used during non-relational training for crossover interaction..................................................................................................57
Table 9. Incorrect comparison stimuli used during non-relational training for divergent interaction.................................................................................................58
Table 10. Incorrect comparison stimuli used during non-relational training for synergistic interaction..........................................................................................59
Table 11. A symbolic representation of the ‘A’ training protocol.............................60
Table 12. A symbolic representation of the ‘B’ training protocol.............................61
Table 13. Inter-observer agreement........................................................................62
List of Figures

Figure 1. Graphs and their corresponding accurate descriptions…………………………...63
Figure 2. The grading rubric used to evaluate pre- and post-tests…………………………...64
Figure 3. The average number of trials required to reach master during training…………65
Figure 4. The percentage of participants achieving each of four different criteria during post-testing …………………………………………………………………………………..66
Figure 5. The average accuracy of each type of statement following each intervention…………………………………………………………………………………………..67
Figure 6. Accuracy scores of participants in the x-x-x group………………………………68
Figure 7. Accuracy of each of the three types of statement following each intervention…………………………………………………………………………………………..69
Figure 8. Accuracy of IV/DV and relational elements of each type of statement following each intervention……………………………………………………………………..70
Introduction

A high functioning individual can provide complete and accurate descriptions of complex situations or information. Some examples would be summarizing information presented in newspaper articles or newscasts, describing the symptoms of an illness, or explaining the problems of a car that is not running well. Similarly, a behavior analyst should be able to provide a complete and accurate description of the circumstances that precede the problem behavior of a client. Another example can be drawn from higher education: undergraduate students who major in Psychology should be able to write complete and accurate descriptions of graphs that depict statistical interaction (Garfield & Ahlgren 1988; Mulhern & Wylie, 2004; Behavior Analyst Certification Board, 2012). Unfortunately, this is a skill that is most difficult and time-consuming to establish (Garfield & Chance, 2000; Fields, Travis, Roy, Yadlovker, de Aguiar-Rocha, & Sturmey, 2009). For this reason, the development of interventions that expedite the teaching of this and similar skills—such as describing the composition of chemical molecules, summarizing written works, and listing characteristics of paintings—is a valuable endeavor that can have important societal effects.

The specific goal of the present experiment was to establish one of these complex repertoires, i.e., the writing of full and accurate descriptions of graphically presented information, by the prior establishment of conditional discrimination training between visually presented stimuli. Specifically, the present research determined how four intervention packages that contained combinations of two different training protocols influenced the emergence of accurate written descriptions of graphs depicting statistical interaction. Figure 1 shows graphs and descriptions involved in this process.

On the left side of Figure 1 are four graphs, each of which represents one form of the
interactive effects of two variables on behavior. From top to bottom, they illustrate no interaction (NO), a crossover interaction (XOVR), a divergent interaction (DVR) and a synergistic interaction (SYN). Each graph contains five non-relational elements and two functions. The non-relational elements are the two independent variables: one on the abscissa that is designated IV1, and the other in the legend that is designated IV2, the dependent variable that is designated DV, the directionality and slope of the function with the higher y-intercept that is designated DIR:HI, and the directionality and slope of the function with the lower y-intercept that is designated DIR:LO. All of these elements are considered to be non-relational because each describes a single element in a graph but not the interplay between the two independent variables.

There are also two types of elements considered to be relational: one is the difference in the slopes of the two functions and is designated as Interact because it refers changes in the magnitude and possibly the directionality of one independent variable’s effect based on the value of the other independent variable (Fields & Spear, 2012). The phrases used to characterize Interact consist of two sets of terms: one includes the terms produced and reversed, and will be referred to as Interact:p (abbreviated to Ip), and the other includes the terms constant or growing difference, and will be referred to as Interact:d (abbreviated to Id). The other relational element is the possible intersection of the two functions, which is designated as Intersect (abbreviated to Is), because it defines the point of contact of one function relative to the other. Each of the four types of interaction is characterized by a set of these non-relational and relational variables, and a glossary of the terms used to denote these variables is shown in Table 1.

Logically, the writing of an accurate description of a graph involves the specification of all of the above-mentioned variables, as shown on the printed descriptions listed on the right side of Figure 1. The production or writing of descriptions such as these implies that every element of
a graph is exerting discriminative control over the written words, phrases, or statements that accurately denote all of the relational and non-relational elements of the graphs. Thus, the writing of accurate descriptions of the interactive effects of two variables on behavior reflects the joint stimulus control (JSC) of responding, since it is governed by all of the elements in combination (Lowenkron, 1998, 2006).

Theoretically, joint stimulus control by all elements could be established by general-case or multiple-exemplar training (Hull, 1920; Keller & Schoenfeld, 1950; Tiemann & Markle, 1990; Cook, Cavoto, & Cavoto, 1995; Lazareva, Miner, Wasserman, & Young, 2008; Marzullo-Kerth, Reeve, & Reeve, 2011; Dib & Sturmey, 2011). While the suitability of these approaches for the establishment of joint control is an empirical question, such approaches would require an unknown number of exemplars to be trained, and this number would vary for each participant. In addition, the process would require a substantial amount of writing on the part of the student, and the provision of necessary grading and directed feedback on the part of the instructor. Although these procedures might be successful, they would also be quite time-consuming and labor-intensive for both student and instructor, particularly when done for a class of students (Sprague & Horner, 1984; Stokes & Baer, 1977; Marzullo-Kerth, et al., 2011; Reeve, Reeve, Townsend, & Poulson, 2007).

A more efficient and reliable alternative might involve the use of selection-based matching-to-sample training to establish conditional relations between graphs as sample stimuli with appropriate printed descriptive paragraphs as comparison stimuli. Once established, the graphs could be presented alone with instructions to write corresponding descriptions of the graphs. The production of correct written descriptions would constitute the transfer of function from the printed textual stimuli to the corresponding graphs. If successful, a selection-based
training procedure (conditional discrimination training) would induce a production-based repertoire (paragraph writing) without direct training. The use of this approach is supported by a rather extensive literature that has documented the transfer of function among the members of equivalence classes and between the members of conditional discriminations (Sidman, 1971; Sidman & Cresson, 1973; Cowley, Green, & Braunling-McMorrow, 1992; Eikeseth & Smith, 1992; Fields, Adams, Buffington, Yang, & Verhave, 1996; Belanich & Fields, 2003; Perez-Gonzalez, Herszlikowicz, & Williams, 2008; Groskreutz, Karsina, Miguel, & Groskreutz, 2010; Walker, Rehfeldt, & Ninness, 2010; Rosales, Rehfeldt, & Lovett, 2011; Walker & Rehfeldt, 2012).

For example, working with individuals with mental retardation who said the names of 18 commonly used objects when presented with their pictures but not with their printed names, Sidman (1971) established equivalence classes among the dictated names of the objects, their pictures, and their printed names. Thereafter, the responses evoked by the pictures were also evoked by the printed names of the objects without any direct training. Likewise, Cowley, et al. (1992) worked with three adults with traumatic brain injury who could not say the names of their three caregivers when shown pictures of their faces, even though they could say their names when their names were dictated to them. After forming equivalence classes of photographs of a caretaker’s face, the caretaker’s dictated name, and printed name, the patients said the names of their caretakers when shown pictures of the caretakers. This outcome constituted the transfer of function (the spoken name) from the dictated name to the pictorial stimulus.

Walker, et al. (2010) taught four college students 4-member equivalence classes consisting of medical diagnoses, each of which consisted of the name of the diagnosis, and printed statement of its definition, its causes, and its treatment. In a post training writing test,
participants wrote the names of the disorder when given the definitions, causes, or treatments, with accuracies that varied from 77% to 100% correct across participants. Thus, match-to-sample training did not consistently induce the writing of single word answers. The overall accuracy of function transfer in this study was lower than those reported in the prior experiments. Thus, accuracy of function transfer was positively correlated with stimulus complexity, or the number of features constituting the stimuli used in these experiments.

Walker and Rehfeldt (2012) used graduate student participants to study the transfer of function among the stimuli in equivalence classes. The experiment began with the establishment of equivalence classes that consisted of the names of the single-subject experimental designs (A), their printed definitional features of the design (B), graphs that exemplified each design (C), and short printed vignettes that described the application of a given single-subject design to the circumstance depicted in each vignette (D). Because these stimuli all consisted of several features, they can be considered complex stimuli. In a post-class-formation writing test, the participants accurately wrote the names of the single-subject design (single words or phrases) when presented with the printed definitions, graphs, or vignettes. In contrast, accuracy of written participants’ written answers was substantially lower when they were asked to write the defining features of a design (three or more separate words or phrases) given a printed vignette. This experiment, then, replicated on a within subject basis the difference in reliability of function transfer based on the complexity of the stimuli and responses. Walker and Rehfeldt (2012) opined that the production of descriptions that consisted of extended phrases required a participant to have attended to the many elements of the complex stimuli that served as models for the responses during conditional discrimination training. It follows from that assumption that function transfer failed because the training procedure did not establish control by, or attention
to, all of the elements in combination of the stimuli that were members of the conditional discriminations; i.e., the conditional discrimination training used in these experiments did not establish joint stimulus control.

Finally, Sella, Ribeiro, and White (2014) examined the transfer of function (complex writing responses) among members of an equivalence class consisting of names, definitions, examples, and notations of various research designs. During the writing post-tests, participants were instructed to produce descriptions of various designs, and generally did so with great accuracy. These results appear to contradict those of Walker and Rehfeldt (2012), since the written examples were equally complex. The difference, however, was probably driven by the fact that Sella et al. included prompts for each element of correct descriptions of designs during the writing tests.

When these experiments are considered together, they showed that accuracy of function transfer was inversely correlated with stimulus and response complexity, regardless of intellectual capacity. Also, those experiments did not include test-retest controls to determine whether test repetition alone would have enhanced accuracy of function transfer, and did not study transfer with complex responses in the absence of prompts. Further, Walker and Rehfeldt (2012) conducted training and testing over the internet, and the participants of Walker, et al. (2010) were enrolled in a course on medical diagnoses at the time of the study, which could have influenced their writing performance. Last, these experiments incorporated what can be called *traditional* conditional discrimination training procedures to establish conditional control by the complex stimuli used in these experiments: these procedures did not establish control by individual elements, in combination, of the stimuli, and so did not establish joint stimulus by all of the elements of those complex stimuli. The latter issue was addressed in part by Spear and
Fields (submitted) who used conditional discrimination training designed to establish joint stimulus control to improve students’ written descriptions of interaction-based graphs. Participants first received a pre-test in which they wrote descriptions of graphs that represented four types of interaction shown in Figure 1.

Participants then received match-to-sample training, in which each of the four graphs shown in Figure 1 was presented as a sample stimulus, along with different printed descriptions as comparisons. One group received traditional conditional discrimination training, in which a graph was presented with two printed descriptions. One of these referred accurately to all of the relational and non-relational elements of the graph; the other referred incorrectly to most of the relational and non-relational elements of the graph. This information is elaborated in Table 2. Each column designates one of the types of interaction, with a graph that exemplifies each. All of the relational and non-relational variables are enumerated in the second column of the Table. Across the four types of interaction, the printed descriptions differed with respect to the phrases used to characterize the interactions (Id, Ip, and Is), the independent and dependent variables (IV1, IV2, DV), and the directionalities of the two functions in each graph (DIR:LO and DIR:HI); i.e., they differed with respect to most of the elements shown in Table 2.

Before training, the accuracy of the pre-test writing performances was quite low (20% accurate on average, with little inter-subject variability). During training, which established conditional control by any, but not necessarily all, of the elements of graphs and their printed descriptions, all participants acquired the graph-printed text conditional relations. After training, the performances in the writing test showed no improvement in the accuracy of the written descriptions. The absence of improvement in their written descriptions can be explained by a consideration of the content of the comparison stimuli presented in all trials as depicted in Table
2. Because the correct and incorrect comparison stimuli differed with respect to many elements, the selection of the correct comparison on any given trial could be achieved by reading only one word of each description, such as the name of IV1, instead of all relevant words. Since all elements of the printed descriptions were not controlling behavior, it was not possible for the printed texts to promote the writing of accurate descriptions of the interactive effects of two variables on behavior.

As noted above, the writing of complete and accurate descriptions of graphs that depict one type of interaction necessitates control by all of the features of the graph, all 16 of which of are listed in the leftmost column of Table 3, for each of the four types of interaction. Control by particular features of graphs and their accompanying descriptions will be referred to as targeted conditional control by the graphs and particular elements of the descriptions: the words in the Interact:d (Id) phrases, the Interact:p (Ip) phrases, and the Intersect (Is) phrases. In addition, once established, these graph-phrase relations must also exert joint stimulus control in combination with all of the non-relational terms in the printed descriptions. Because there are four types of interaction, there are 64 forms of conditional and joint stimulus control that must be established to link all of the graphs and their corresponding textual representations.

For this reason, the participants in a different training group received training with some of the components listed in Table 3 to enhance the writing of accurate description of graphs that depicted the interactive effects of two variables on behavior. The participants in this group were first given a writing pre-test and then received training that was designed to establish conditional control by some of the elements in the graphs and their corresponding textual descriptions; these forms of control are listed in column A, and this training will be referred to as A training. Specifically, conditional control was established by the graphs themselves and some undefined
features of the printed descriptions, through traditional training. Targeted conditional control was then established between the graphs and the terms within the Id phrases themselves, and, once established, joint control training was used to establish control by the terms in the Id phrases in combination with IV1, IV2, DIR:HI and DIR:LO. Having acquired these conditional relations, participants were given a writing post-test that was identical to the writing pre-test, and wrote descriptions that were 70% accurate on average with little inter-subject variability. These scores were significantly higher than the pre-test scores (~20%) and the post-test scores of the traditional training group (~20%). Clearly, then, the establishment of targeted conditional control by particular elements and joint stimulus control by combinations of relational and non-relational elements in the graph-text conditional relations enhanced the writing of accurate descriptions of the interactive effects of two variables on behavior. The fact that writing performances did not reach maximum accuracy suggests that some or all of the other forms of targeted conditional and joint stimulus control that had not been established by Spear and Fields (submitted) were not available to influence the emergence of written descriptions of the interactive effects of two variables on behavior; these forms of stimulus control are represented by the empty cells in column A.

The present experiment, then, sought to improve the writing of accurate descriptions of interaction-based graphs by the establishment of some of the forms of targeted conditional and joint stimulus control that were not established by Spear and Fields (submitted). These forms are symbolized by some of the empty cells in the A column in Table 3. In addition to using the training conducted by Spear and Fields (traditional, Id, and Id + non-relational), further training was conducted to establish targeted conditional control by the Ip phrases and the Is phrases, as indicated in the B column of Table 3. This training will be referred to as B training. Finally, two
yoked control groups were studied to isolate the effects on the emergence of accurate written
descriptions of the many forms of targeted and joint stimulus control established by training.
Method

Participants

All participants were drawn from the Psychology 101 subject pool at Queens College/CUNY. Participants received course credit for participating in the first training session, and were paid $50 for participating in a follow-up training session. A total of 52 participants were recruited, and were assigned to each of four groups in a block-randomized basis without replacement, resulting in 13 participants in each group. Participants who learned the conditional discrimination trained in Session 1 were invited to participate in session 2, which was held within seven days of session 1. A participant was not invited to participate in session 2 if she/he did not acquire all of the trained relations in session 1.

Materials

Pretest and posttest booklets. The pretest and posttests were identical, and consisted of four page booklets, printed on 8.5 x 11 inch pages. The top of each page contained a graph of one type of interaction. Each of the four graphs is shown on the left side of Figure 1. On the top of each page was the instruction, “Please write a complete and accurate description of the following graph.”

Setting, hardware and software. Each participant worked independently in a cubicle measuring 2.5m x 2.5m. During training, all stimuli were presented on a PC based desktop computer. Custom software developed in our laboratory was used to conduct this trial-based experiment. The software collected information about all stimuli presented in each trial and all responses made to the stimuli.

Stimuli. As noted in the last paragraphs of the Introduction, two training procedures were used. The forms of conditional and joint control established by each of these two procedures are
indicated by the A and B columns of Table 3. During training, each trial featured one of the four graphs in Figure 1 as a sample stimulus. Each graph sample was presented with two, three, or four printed descriptions as comparisons. One of these was always a complete and accurate description of the graph sample, and is referred to as the [Co+ +]. This comparison contained accurate references to all of the relational elements of the graph (Id, Ip, and Is, indicated by the first “+”) and all of the non-relational elements of the graph (IV1, IV2, DIR:LO, and DIR:HI, indicated by the second “+”). The [Co+ +] for each graph sample is shown on the right side of Figure 1.

**Traditional training stimuli.** As shown in the first row of Table 3, A training included traditional training, which established conditional control by the graph sample stimuli. During traditional training, a graph sample was presented with two printed descriptions as comparisons: one was the [Co+ +], and the other was a complete and accurate description of one of the three other graphs, referred to as the [Co- -]. This description contained inaccurate references to the relational elements (Id, Ip, and Is, indicated by the first “-“) and inaccurate references to the non-relational elements (IV1, IV2, DIR:LO, and DIR:HI, indicated by the second “-“). During traditional training, for instance, a graph sample, such as graph #1 in Figure 1, would be presented with its [Co+ +] (comparison #5 in Figure 1), and its [Co- -] (comparison #6 in Figure 1). Because these two comparisons differed with respect to a number of relational and non-relational elements, the selection of the [Co+ +] would indicate conditional control by at least one of the relational or non-relational elements.

**Id training stimuli.** As shown in the second row of Table 3, A training also included training to establish targeted conditional control by the graphs and the terms constant or growing difference of the interaction statement (Id). During Id training, a graph sample was presented
with two printed descriptions as comparisons. One was the [Co+ +], which contained accurate references to all the relational and non-relational elements. The other was an incorrect description, and is referred to as the [Co- +]. This description contained an inaccurate reference to one of the relational elements (Id), indicated by the “-,” and accurate references to all of the non-relational elements, indicated by the “+.” Because the only inaccurate relational element in the [Co- +] was Id, this comparison is referred to as [Id-]. The [Id-] stimuli for all four graphs are shown in Table 4. During training to establish targeted conditional control by Id, a graph sample, such as graph #1 in Figure 1, would be presented with its [Co+ +] (comparison #5 in Figure 1) and its [Id-] (comparison # 9 in Table 4). Because the [Co+ +] and [Id-] differed only with respect to Id, consistent selection of the [Co+ +] would indicate conditional control by Id.

**Ip training stimuli.** As shown in the third row of Table 3, both A and B training included the establishment of targeted conditional control by the graphs and the terms produced or reversed of the interaction statement (Ip). During Ip training, the graph sample was again presented with two comparisons. One was the [Co+ +]. The other comparison was the [Co- +], as it contained an inaccurate reference to one of the relational elements, and accurate references to all of the non-relational elements. Because the only inaccurate relational element in the [Co- +] was Ip, this comparison is referred to as [Ip-]. The [Ip-] stimuli for all four graphs are shown in Table 5. During training to establish control by Ip, a graph sample, such as graph #1 in Figure 1, would be presented with its [Co+ +] (comparison # 5 in Figure 1) and its [Ip-] (comparison #13 in Table 5). Because the [Co+ +] and [Ip-] stimuli differed only with respect to Ip, consistent selection of the [Co+ +] would indicate conditional control by Ip.

**Is training stimuli.** As shown in the fourth row of Table 3, B training included the establishment of targeted conditional control by graphs and the intersection of the two functions
The structure of Is training trials was similar to that of Id and Ip trials, described above, in that they featured the presentation of the [Co+ +] and [Co- +]. However, the [Co- +] contained an inaccurate reference only to Is, and so is referred to as [Is-]. The [Is-] stimuli for all four graphs are shown in Table 6. During training to establish targeted conditional control by Is, then, a graph sample (such as graph #1 in Figure 1) would be presented with its [Co+ +] (comparison #5 in Figure 1) and its [Is-] (comparison #17 in Table 6). Because the [Co+ +] and [Is-] differed only with respect to Is, consistent selection of the [Co+ +] would indicate conditional control by Is.

*Non-relational training stimuli.* As shown in rows 5-8 of Table 3, A training included the establishment of joint stimulus control by Id in combination with each of the four non-relational elements. Id + non-relational training trials were structured similarly to Id training trials, described earlier, but also featured an additional incorrect comparison, referred to as the [Co+ -]. The [Co+ -] accurately referenced all of the relational elements of the graph sample (indicated by the “+”) but inaccurately referenced one of the four non-relational elements (indicated by the “-”). To establish joint control by Id in combination with each of the four non-relational elements, then, four different [Co+ -] stimuli were used for each graph. Each contained an inaccurate reference to only one of the four non-relational elements. Thus, one contained an inaccurate reference to IV1, another to IV2, another to DIR:LO, and another to DIR:HI. The [Co+ -] stimuli for each of the four graphs are shown in Tables 7-10. Table 7, for instance, shows the four [Co+ -] stimuli presented with a graph depicting no interaction.

To illustrate the establishment of joint control by Id in combination with IV1, as shown in row 5 of Table 3, a graph sample, such as graph #1 in Figure 1, would be presented with three comparison stimuli. One would be its [Co+ +] (comparison #5 in Figure 1), and the other two
would be incorrect comparisons. One of these would be the [Id-] (comparison #9 in Table 4) and the other would be the [Co+ -] that contained an inaccurate reference to IV1 (comparison #21 in Table 7). In a trial such as this, the only comparison that accurately referenced all of the elements of the graph sample was the [Co+ +]. The [Id-] contained an inaccurate reference to Id, and the [Co+ -] contained an inaccurate reference to IV1. Thus, consistent correct selection of the [Co+ +] would indicate joint stimulus control by Id in combination with IV1. A lack of control by either Id or IV1 would result in selection of one of the incorrect comparisons. For instance, conditional control by IV1 but not Id would result in equally frequent selection of the [Co+ +] and [Id-], as these both contained accurate references to IV1. Conversely, conditional control by Id but not IV1 would result in equally frequent selection of the [Co+ +] and [Co+ -], as these both contained accurate references to Id. Thus, consistent selection of the [Co+ +] would require joint control by the combination of Id and IV1. Trials that established joint control by Id in combination with the three other non-relational elements (IV2, DIR:LO, and DIR:HI) were informed by the same logic, and differed only with respect to the content of the [Co+ -], which contained an inaccurate reference to either IV2, DIR:LO, or DIR:HI.

**Procedure**

**Experimental design.** The experiment determined how the emergence of written descriptions of graphs was influenced by the two training procedures designed to establish the forms of targeted conditional and joint stimulus control shown in Table 3. The experiment was conducted in two sessions, and used four groups. Each of these groups experienced three stages, and is named after those stages: the x-x-x group, the x-A-B group, the x-A-A group, and the x-x-B group.

The first stage in each group included no intervention (denoted by the first “x” in the
string used to denote a group) followed by a writing pre-test. These test performances provided a baseline measure of the accuracy of writing descriptions of graphs that depicted the interactive effects of two variables on behavior in the absence of any experimental intervention. The pretest was composed of 4 graphs, and each graph depicted one type of interaction. Each graph was presented on the top of a separate page, and participants were instructed to write a complete and accurate description of each graph on the bottom of each page. The four graphs featured on the pretest were identical to those presented during training. Each description was evaluated using the rubric shown in Figure 2. One point was earned for each term if it was included in the description of a given graph. Inter-observer agreement (IOA) data were collected from three independent raters for all pretest and posttest performances using a trial-by-trial IOA (Cooper, Heron, & Heward, 2007).

The second stage varied across groups and involved no intervention for two groups (x-x-x and x-x-B) or A training for two groups (x-A-B and x-A-A). In all groups, these conditions were followed by a second administration of the writing test. Differences in the writing tests after the completion of stage 2 would indicate how the emergence of accurate written descriptions of the data portrayed in the interaction graphs was influenced by the effects of test repetition alone (in the x-x-x and x-x-B groups) and by A training (in the x-A-A and x-A-B groups).

The third stage involved another no-intervention condition in the in the x-x-x group, re-administration of A training (in the x-A-A group), or B training (in the x-x-B and x-A-B groups). In all cases, these conditions were followed by a third and last administration of the writing test. Differences in the writing tests at the end of stage 3 would indicate how the emergence of accurate written descriptions of the data portrayed in the interaction graphs was influenced by test repetition alone (in the x-x-x group), by the overtraining of the A conditional relations (in the
x-A-A group), by B training alone (in the x-x-B group), and by the combination of A and B training (in the x-A-B group). A comparison of the final performances of the three training groups, then, would indicate the relative effects of A and B training alone, as well as their combined effects (A + B).

The experiment was conducted in two sessions. During the first session, all four groups received the writing pre-test and the first writing post-test. Between the two writing tests, the x-A-A and x-A-B groups received A training. During the second session, the x-A-B and x-x-B groups received B training and the x-A-A group received A training. All four groups then received the third, and final, writing test. The design and content of the A and B training procedures are elaborated below.

**Trial structure, contingencies, and feedback.** All training was conducted using trials that contained one graph as a sample stimulus along with two, three, or four written descriptions of graphs as the comparison stimuli. The sample stimulus (a graph) was located in the top middle of the screen with the comparison stimuli below the sample stimulus. If there were two comparison stimuli, they were located below and to the left and right of the sample; if there were three, they were located below and to the left, right, and directly below the sample; if there were four, they were located equally spaced below the sample. The comparison stimuli were presented with their positions randomized across trials. In addition, the letters ‘A,’ ‘B,’ ‘C,’ and ‘D’ were presented, one each, below the comparisons and their positions were also randomized with respect to each comparison. Thus, there was no correlation between a comparison paragraph and the ‘A,’ ‘B,’ ‘C,’ or ‘D’ letters used as comparison labels. Prior to training, the following instructions were presented to participants:

  In this experiment, you will be presented with graphs and descriptions.
A graph will appear on the top of the screen, and descriptions will appear on the bottom. Press the key on the keyboard that corresponds to the description that best matches the graph: either A, B, C, or D.

Call the experimenter if you have any questions.

If not, press "Enter" to continue.

During trials that were followed by informative feedback, selection of the correct comparison was immediately followed by presentation of the word “Right!” and selection of an incorrect comparison was immediately followed by presentation of the word “Wrong!” Participants were then required to press either the ‘R’ or ‘W’ key, respectively, to progress to the next trial. When informative feedback was not provided, selection of any comparison was immediately followed by presentation of the words “No Feedback,” and participants were required to press the ‘E’ key to progress to the next trial.

**Training to establish the ‘A’ stimulus control relations (‘A’ training procedure).** ‘A’ training used the logic described above to establish the forms of targeted conditional and joint stimulus control listed in the A column of Table 3 and consisted of a 14 phase training sequence that is detailed in Table 11.

**Traditional training.** ‘A’ training began with the establishment of conditional control by each of the four graphs in combination with the printed descriptions, which was accomplished by traditional training. This training was conducted in Phases 1 and 2. Trials in these Phases featured the presentation of the [Co+ +] and [Co- -], and ensured discrimination among the graphs used as sample stimuli. Trials in Phase 1 contained graphs and printed descriptions of no- and crossover interaction; these trials are represented symbolically by rows 1 and 2 of Table 9. The no-interaction graph sample is shown in Figure 1 (graph #1), as is the [Co+ +] (comparison
and the [Co-] (comparison #6). Trials in Phase 2 contained graphs and printed descriptions of divergent- and synergistic interactions; these trials are represented symbolically by rows 3 and 4 of Table 11.

**Establishment of control by Id.** Traditional training was followed by Id training during Phases 3 and 4, which are represented symbolically by rows 5-8 of Table 11. These phases constitute targeted conditional discrimination training that involved the presentation of the [Co+] and [Id-], and established conditional control by Id of no- and crossover interaction and synergistic and divergent interaction, respectively. The two trials presented in each Phase were administered four times each, for a total of eight trials per Phase. Integration of conditional control by Id of all four types of interaction was ensured by Phase 5, which is represented by row 9 of Table 11, and contained the four trials presented during Phases 3 and 4, presented twice each for a total of eight trials. All the aforementioned Phases were repeated up to five times or until a participant achieved 100% accuracy in a Phase, and all trials were followed by informative feedback.

**Establishment of control by non-relational elements.** Phases 6 and 7 were structured similarly to Phases 3 and 4, but Stage 6 established targeted conditional control by IV1 and DIR-HI of graphs and descriptions of all four types of interaction, and Phase 7 established conditional control by IV2 and DIR-LO of graphs and descriptions of all four types of interaction. These Phases featured the presentation of the [Co++] and [Co+-], and are represented symbolically by rows 10-25 of Table 11. Conditional control by IV1, IV2, DIR:LO, and DIR:HI was ensured by Phase 8, which featured the presentation of the four trials presented during Phases 6 and 7. All trials were followed by informative feedback, and were repeated up to five times or until a participant achieved 100% accuracy.
Integration and maintenance of previously established discriminations: joint stimulus control. Phase 9 featured the presentation of all trials featured in Phases 3-8. Each of the four graphs was presented in four trials: on each trial, the [Co+ +], [Id-], and one of the four [Co+ -] stimuli was presented. Because four different trials were presented for each graph, a total of 16 trials and their comparisons were presented once each. All trials were followed by informative feedback, thus ensuring integrity of joint control by Id in combination with all of the non-relational elements. Phase 10 was similar to Phase 9, but no informative feedback was presented, thus ensuring maintenance of joint control by all elements in the absence of feedback. If participants responded with at least 94% accuracy on Phase 10, they progressed to the first post-test. Otherwise, they progressed through a series of feedback reduction Phases (11-14) in which feedback was gradually eliminated. In these stages, the trials were the same as those presented in Phase 10.

Training to establish the ‘B’ stimulus control relations (‘B’ training procedure). ‘B’ training used the logic described earlier to establish the forms of targeted conditional and joint stimulus control listed in the B column of Table 3, and consisted of a 16 phase training sequence that is detailed in Table 12.

Establishment of control by Id. B Training began with three phases designed to establish targeted conditional control by Id for all four types of interaction, and these Phases are represented by the first five rows of Table 12. Each graph sample was presented with two printed descriptions as comparisons: the [Co+ +] and [Id-]. Phases 1 and 2 each featured two different trials presented four times each, for a total of eight trials. Phase 1 featured graphs and descriptions of no and crossover interactions, and Phase 2 featured graphs and descriptions of divergent and synergistic interactions. These were followed by Phase 3, which featured all four
trials from Phases 1 and 2 presented twice each, to insure maintenance of all conditional discriminations established during Phases 1 and 2.

**Establishment of control by Ip.** Phases 4 and 5 were structured similarly to Phases 1 and 2, but were designed to establish targeted conditional control by Ip for all four types of interaction, and are represented by rows 6-10 of Table 12. Each graph sample was presented with two printed descriptions as comparisons: the [Co+ +] and [Ip-]. Phase 6 was structured identically to Phase 3, and featured the presentation of all trials from Phases 4 and 5, to insure maintenance of all conditional discriminations established during Phases 4 and 5.

**Integration and maintenance of Id and Ip discriminations.** Phase 7 featured four different trials presented twice each, for a total of eight trials. Each trial featured one of the four graph samples, along with three comparisons: the [Co+ +], [Id-], and [Ip-]. Thus, Phase 7 insured maintenance of conditional control by Id and Ip for all four types of interaction.

**Establishment of control by Is.** Phases 8 and 9 were structured identically to Phases 4 and 5, but were designed to establish targeted conditional control by Is, and so each graph sample was presented with the [Co+ +] and [Is-]. Phase 10 then featured the presentation of all trials from Phases 8 and 9, twice each, to insure maintenance of all conditional discriminations established during Phases 8 and 9.

**Integration and maintenance of Id, Ip, and Is relations: joint stimulus control.** Phase 11 was designed to insure maintenance of all previously established conditional discriminations, and so featured four different trials, presented twice each. Each trial featured one of the four graph samples presented with four comparisons: the [Co+ +], [Id-], [Ip-], and [Is-], and all trials were followed by informative feedback. Thus, Phase 11 insured maintenance of joint control by Id, Ip, and Is. Participants were required to respond with 100% accuracy during the
aforementioned Phases, and could repeat each Phase up to five times if they failed to do so. Failure to respond with 100% accuracy within five repetitions of a Phase resulted in dismissal from the experiment.

Phase 12 was identical to Phase 11, but no informative feedback was presented following any trial. If participants responded with 100% accuracy in Phase 12, they proceeded to the third, and final, writing test. If they failed to achieve 100% accuracy in Phase 12, they progressed through a series of Phases (13-16) during which informative feedback was reinstated and thinned. Each of these Phases could be repeated up to five times, or until a participant responded with 100% accuracy.
Results and Discussion

Survival of Participants

Of the 52 students who began the experiment, 43, or 83%, completed it. Of the nine who did not complete the experiment, four did not acquire the conditional discriminations in session 1 (two each assigned to the x-A-A and x-A-B groups), and five acquired all of the conditional discriminations in session 1 but did not return for session 2 (one each in the xxx, x-A-A, and x-A-B groups, and two in the x-x-B group). As a result, the data reported below were collected from 12 participants in the x-x-x group, 10 in the x-AA group, 10 in the x-A-B group, and 11 in the x-x-B group.

Acquisition of Conditional Discriminations

The average number of trials required to acquire one conditional discrimination in each group is shown in Figure 3, along with the scheduled minima for each condition. Acquisition data from participants who progressed through the feedback reduction Phases are included in the analyses. All groups acquired the conditional discriminations rather rapidly and in somewhat more trials than the scheduled minima. As noted in the Procedure, the A condition involved the establishment of conditional control by the graphs, targeted conditional control by Id and the non-relational elements, and joint control by Id in combination with the non-relational elements. Since there were no apparent differences in the number of trials to acquisition for each of these types of conditional relations, their data were aggregated and treated together. The B condition involved the establishment of targeted conditional control by Id, Ip, and Is, and joint control by Id, Ip, and Is in combination. Once again, since there were no apparent differences in the number of trials to acquisition, their data were aggregated and treated together. In contrast, however, there were some small but significant differences found within and across the three experimental...
The x-A-A group required significantly more trials than the scheduled minimum to acquire the discriminations during initial A training, as confirmed by a one-sample t test, $t(9) = 2.3$, $p < .05$, and during repetition of A training, the x-A-A group required a similar number of trials to re-acquire the trained relations, as confirmed by a paired samples t test, $t(9) = 1.76$, $p > .05$. Thus, the initial acquisition of the A-based conditional relations did not shorten the re-acquisition of the same relations. Participants in the x-x-B group also required significantly more trials than the scheduled minimum to learn the B conditional discriminations, as confirmed by a one-sample t test, $t(10) = 3.5$, $p < .01$. Last, the participants in the x-A-B group also required significantly more trials than the scheduled minimum to learn the conditional discriminations during initial A training, as confirmed by a one-sample t test, $t(9) = 2.7$, $p < .05$. In addition, significantly fewer trials were needed to acquire the conditional discriminations during B training than during the previously administered A training, as confirmed by a paired-samples t test, $t(9) = 2.4$, $p < .05$. Finally, significantly fewer trials were needed to acquire the conditional discriminations during B training in the x-A-B condition than the equivalent training in the x-x-B group, as confirmed by an unpaired t test, $t(19) = 2.8$, $p < .01$.

**Inter-Observable Agreement**

All of the written descriptions in all three tests in the experiment were assessed by three independent raters using a trial-by-trial IOA (Cooper, Heron, & Heward, 2007). Agreement was defined as the awarding of the same number of points for a given element (either 1 or 0), using the rubric shown in Figure 2. IOA was calculated by dividing the number of agreements by the number of agreements + disagreements, and multiplying that result by 100. The IOA for each type of sentence is shown in Table 13. IOAs varied from 92% to 99% for all types of sentences.
across all three writing tests.

**Writing Performances**

The main goal of the present research was to determine how four intervention packages that contained combinations of two different training protocols influenced the emergence of accurate written descriptions of the interactive effects of two variables on behavior. One was a control procedure (x-x-x) that involved the administration of three writing tests in which participants wrote descriptions of the interactions depicted in graphs and received no feedback. The three different interventions designed to induce the emergence of accurate written descriptions study (x-A-B, x-A-A, and x-x-B) used conditional discrimination training procedures to establish stimulus control by a variety of elements of graphs and printed descriptions.

Regardless of the intervention, a complete and accurate written paragraph that described a graph had to contain three types of statements. Each described one of the three primary features of the graph: the interaction of the two independent variables, the intersection of the functions, and the directionality of the two functions. Further, each of these statements had to include accurate specifications of both the relational elements (those describing the slopes of the functions, the effect of the interaction, or the intersection of the functions) and the non-relational elements (the independent and dependent variables represented in the graphs). Thus, each statement had to include both relational and non-relational elements.

The results of the experiment were analyzed on four levels; in each, the accuracy of written descriptions was calculated as the percentage of possible points earned according to the rubric shown in Figure 2. First, Figure 4 shows the percentage of participants in each condition whose final descriptions reached one of four constant performance criteria: at least 60, 70, 80,
and 90% accuracy. Second, Figure 5 analyzes the global accuracy of written descriptions, with writing accuracy measured at the end of each of the four intervention packages, and averaged across participants in a condition. Third, Figures 6 and 7 present a process analysis of the effects of retesting on the accuracy of each statement (Figure 6), and of the effects of each component of training (x, A, A repetition, and B) on the accuracy of each statement (Figure 7). These analyses show accuracy aggregated across relational and non-relational elements in a written statement and averaged across the participants in a condition. Fourth, Figure 8 presents a process analysis of the separate effects of each training procedure on the emergence of the relational and non-relational elements in each of the three statements mentioned above, and will also be averaged across the participants in a condition.

**Yields by training intervention.** Figure 4 displays the percentage of participants in each group who achieved four different accuracy levels (at least 60, 70, 80, or 90%) on the final writing test. These values correspond to letter grades of D, C, B, and A, respectively.

In the x-x-x group, all participants wrote descriptions that were less than 60% accurate, a traditional academic grade of D. In the three training groups, on the other hand, more than half the participants wrote descriptions that were at least 60% accurate. In the x-A-A group, however, all of the participants wrote descriptions that were less than 70% accurate. Therefore, when testing was repeated, or when A training was administered, no students wrote descriptions of the graphs that rose to the level of a traditional academic grade of C.

In contrast, the combination of A and B training in the x-A-B group resulted in the writing of descriptions that were at least 80% accurate by 90% of the participants, and were at least 90% accurate by 80% of the participants. Most of the participants in this condition, then, earned letter grade equivalents of B or A. A Fishers Exact Test showed that the improvement
relative to the test repetition only or A training alone was likely to occur by chance with a probability of less than .0001. The outcomes after B training alone were more complex. 63% of participants wrote descriptions that were at least 70% accurate, thus earning a letter grade of C, while 23% wrote descriptions that were at least 80% accurate, or equivalent to a letter grade of B. No participants wrote descriptions that were at least 90% accurate, which is equivalent to a letter grade of A.

**Overall effects of interventions on writing accuracy.** The overall accuracy of participants’ descriptions at the end of all training would reflect the accuracy of descriptions of all of the elements, relational and non-relational, for all of the three statements: interaction, intersection, and directionality. Overall accuracy scores and those for each of the three types of statements are presented in Figure 5. Significant differences between groups’ performances, as found by a post-hoc test conducted after an independent groups ANOVA using an alpha level of .05, are reported. The top panel of Figure 5 shows the overall accuracy of all written descriptions in each of four interventions.

**Test repetition only.** In the x-x-x- condition, the writing test was administered three times with no interventions between tests. In the last writing test, the participants wrote relatively inaccurate and incomplete descriptions of the information conveyed in the interaction-based graphs.

**A-B training.** When A training was followed by B training, the written descriptions were 92% accurate on average. When compared to the effects of test repetition alone, this increment in writing accuracy was significant, and could have occurred by chance with a probability less than .0001. Thus, targeted conditional and joint stimulus control training administered in an automated matching-to-sample training format induced the writing of paragraph-length
descriptions of the interactive effects of two variables on behavior with four forms of interaction. This outcome was achieved with no direct training of writing behavior.

**A-A training.** The high level of accuracy of the written descriptions in the x-A-B condition could have been due to additional A training alone, B training alone, or the combination of A and B training together. To determine whether additional A training alone produced the effect shown by the x-A-B condition, an x-A-A condition was conducted in which additional A training was administered after the writing test that followed initial A training.

When A training was repeated, the accuracy of the written descriptions, while significantly greater than that of the x-x-x condition, increased only to an intermediate value. Therefore, extended A training alone did not account for the very high levels of writing accuracy produced by the combination of A and B training.

**B training.** It is also possible that B training alone could have produced virtually complete and accurate written descriptions, even in the absence of prior A training. To assess this possibility, the participants in the x-x-B condition received B training only. Writing accuracy after B training alone was significantly greater than that seen in the x-x-x condition, but was lower than that obtained after x-A-B training. Therefore, B training alone did not account for the very high levels of writing accuracy produced by the combination of A and B training.

**Summary.** The results of all four conditions indicate that the virtually complete level of accurate writing was driven by the combined effects of A and B training. The establishment of graph-printed text conditional relations that engendered joint control by the relational and non-relational elements of the graphs and printed descriptions induced the writing of accurate and complete descriptions of the interactive effects of two variables on behavior, with no direct training of the writing of such descriptions.
Writing accuracy by sentence type. The written descriptions of each graph should contain sentences that characterize the interaction of the two independent variables (Interact), the intersection of the two functions (Intersect), and the functions and their slopes (Direction). Thus, the effects shown in the top panel of Figure 5 could reflect the differential effects of the A and B training procedures on the production of any or all of the three types of sentences. For this reason, the effects of each training condition on the production of each of the three types of sentence were analyzed and presented in the lower three panels of Figure 5.

Accuracy of the interaction statements. With respect to the accuracy of the interaction statements, test repetition resulted in low levels of accuracy. The A training alone and B training alone each produced the same increments in accuracy and were significantly different from that measured in test repetition. Finally, training with the combination of A and B produced written descriptions of interactions that were completely accurate. The accuracy under the x-A-B condition was equal to the sum of accuracies engendered by A training alone and B training alone. Thus, the emergence of accurate written descriptions of the interactive effects of two variables on behavior was produced by prior establishment of conditional and joint stimulus control during A training in combination with B training.

Accuracy of the intersection statements. With respect to the accuracy of the intersection statements, test repetition resulted in a minimal level of accuracy, as did A training alone. In contrast, B training alone produced a large and significant increase in writing accuracy, that was equal to that produced by B training in combination with A training. Thus, the increment in the writing of accurate descriptions of intersection were based on B training alone. Because B training established targeted conditional control by the intersection statements, B training was responsible for the emergence of very accurate written descriptions of the intersection
Accuracy of directionality statements. With respect to the directionality statements, test repetition resulted in intermediate levels of accuracy, which indicated the presence of stimulus control by some elements of the directionality statements prior to any experimenter based training. B training alone did not produce an improvement of the accuracy of descriptions of the directionalities of the function in the interaction-based graphs. On the other hand, the A training alone, or in combination with B training, produced nearly maximal levels of written accuracy. Therefore, the emergence of accurate written descriptions of the directionalities of the functions in the graphs was likely produced by the A conditional discrimination training alone. Indeed, this reflects the fact that the contingencies in the A training explicitly established conditional control by the directionalities of the functions in the graphs and the printed descriptions of the directionalities of these functions.

Summary. The effects of A and B training alone and together had differential effects on the accuracies of written descriptions of the interactive effects of two variables on behavior, on the intersection of functions in interaction based graphs, and on the directionalities of the functions in interaction based graphs. In total, these data suggest that neither A nor B training alone produced nearly complete accuracy of written descriptions; rather, the combination of A and B training was responsible for this improvement.

Process analysis: accuracy of entire written sentences. The data in Figure 5 showed the overall effects of all training on the accuracy of written descriptions. Each intervention (x-x, x-A-A, x-x-B, and x-A-B), however, was conducted in three stages, with a writing test administered after the completion of each stage. Thus, the results shown in Figure 5 did not document the changes in writing accuracy that emerged after each stage of training.
As mentioned above, a complete written description of the interactive effects of two variables on behavior should include sentences that characterize the interaction of the two independent variables (Interact), the intersection of the two functions (Intersect), and the functions and their slopes (Direction). Thus, the effects of each stage of training in each intervention could have differential effects on the accuracy of writing each type of sentence.

The following figures show their serial effects on writing accuracy for each type of sentence for each intervention. Figure 6 illustrates the results of the x-x-x- condition and shows how test repetition alone influenced writing accuracy for each type of sentence. Figure 7 shows the effects of each stage of training in the x-A-A, x-x-B, and x-A-B condition on writing accuracy for each type of sentence.

**Effects of repeated testing.** Figure 6 presents the results produced by the repetition of the writing test alone. Because this procedure was the same across experimental conditions, accuracies were aggregated for all participants and averaged separately for the first, second, and third test, respectively. Across all three writing tests, the paragraphs were on average 28% accurate, and accuracy did not change with test repetition. The second through fourth panels show accuracy for each sentence type in each test. Again, accuracy did not change with test repetition, but level of accuracy differed across sentence types.

**Effects of the training procedures by type of sentence.** Figure 7 displays the effects of each training condition on the production of each of the three types of sentence. Each row displays the accuracies on written descriptions for one of the three types of sentences, and each column displays the scores for participants in one group on each of the three writing tests. For example, the top row of Figure 7 shows the accuracy of participants’ descriptions of Interactions, and the left-most column shows the accuracy of the three types of sentences produced by
participants in the x-A-B group. Each panel in Figure 7 also contains horizontal lines that span various pairs of bars in a condition. The lines indicate a significant difference in the values obtained in the two conditions shown underneath the endpoints of each line. Each is based on a post-hoc test conducted after a repeated-measures ANOVA, using an alpha level of .05.

Descriptions of interactions (Interact). The top row of Figure 7 shows the effects of the three intervention conditions on the accuracy of the written descriptions of the interactions depicted in the graphs. The left panel of in this row shows the accuracy of written descriptions after A and subsequent B training. In the first writing test, the participants wrote inaccurate and incomplete descriptions of the interactive effects of the two variables on behavior, with minimal variability in participants’ scores. After the initial A training, which explicitly established joint control by the Id in combination with the two independent variables and the two directionality elements, the written descriptions of Interact more than doubled in accuracy. The subsequently administered B training, which established conditional control by all the remaining elements of the interaction and intersection, resulted in an a further increase to nearly 100% accuracy for most of the participants, with little inter-subject variability in written accuracy.

The effects of A-B training could have been achieved by more A training as opposed to the addition of B training. If so, re-exposure to A training would also result in an improvement in the accuracy of the descriptions of the interaction equal to that resulting from B training. This possibility was assessed by the performance of the x-A-A group, as represented in the middle graph on the upper row of Figure 7. During pre-testing, participants in the x-A-A group produced largely inaccurate and incomplete descriptions of the interaction that were equivalent to those observed during pre-testing in the x-A-B group. The initial A training alone produced results similar to those produced by the A training in the x-A-B group. Because the repetition of A
training did not produce any additional improvements in participants’ descriptions of the interaction, the terminal performances of the x-A-B group could not be accounted for by additional A training alone.

It is also possible that the increase in the accuracy of the written descriptions of the interaction statements after B training could have resulted from B training alone, rather than the combined effects of B and A training. This possibility was evaluated in the x-x-B condition, which included only B training, the results of which are presented in the right-most panel of the top row of Figure 7. In the first two writing tests, participants wrote descriptions of the interactions that were generally inaccurate and incomplete, and were quite similar to the initial test performances of the x-A-B and x-A-A groups. After the completion of B training, which established conditional control by elements of the interaction and intersection, the participants’ descriptions increased to approximately 50% accuracy. This effect was very similar to the effect of A training alone. Because the improvements in description of the interaction produced by B training alone, then, did not approximate those produced by the combination of A and B training, the terminal performances of the x-A-B training could not be accounted for by B training alone.

When the effects of all three training conditions are considered, the effect of A-B training resulted from the combination of the individual effects of A training and of B training. Indeed, the writing accuracy produced after A and B training together was approximated by the simple addition of the effects produced by each type of training alone.

*Descriptions of the intersection (Intersect).* The middle row of Figure 7 shows the effects of the three intervention conditions on the accuracy of the written descriptions of the intersections depicted in the graphs. The left panel shows the accuracy of written descriptions after A and B training. Before initial A training, participants generally did not produce
descriptions of the intersection, and scores across participants varied minimally. Following initial A training, which did not establish conditional control by elements of the intersection, participants’ descriptions improved somewhat to an average of 20% accuracy, and variability also increased, though not significantly. Because A training did not establish conditional control by elements of the intersection, this change in performance indicates generalization of the effects of training. Subsequent B training, which established conditional control by the elements of the intersection, resulted in a significant increase to an average of 90% accuracy and a substantial reduction in variability in scores across individuals.

As was the case for descriptions of the interaction, it is possible that the improvement in participants’ descriptions of the intersection observed following A-B training could have resulted from either additional A training alone or from B training alone. These possibilities were again assessed by the performances of the x-A-A and x-x-B groups. As shown in the middle panel, the participants in the x-A-A group generally failed to produce descriptions of the intersection during pretesting. Following initial A training, the accuracy of participants’ descriptions improved slightly, to approximately 10% on average, and the variability of scores also increased, but remained minimal. This improvement was not as great as that shown by participants in the x-A-B group following A training, and performance remained generally unchanged following additional exposure to A training, which implied that significant improvement in descriptions of the intersection was due to B training.

This interpretation was confirmed by the performance of participants in the x-x-B group, who also generally did not produce complete and accurate descriptions of the intersection during the first two writing tests. After B training, however, which established conditional control by the elements of the intersection, the accuracy of participants’ descriptions improved significantly
to approximately 90%, and variability of scores also increased. This performance was similar to that of the x-A-B group following A and B training. Taken together, then, these performances indicate that significant improvements in descriptions of the intersection could be attributed to B training alone.

*Descriptions of the directionalities.* The bottom row of Figure 7 shows the effects of the three intervention conditions on the accuracy of writing descriptions of the directionalities of the functions depicted in the graphs. The left panel shows the accuracy of descriptions after A and B training. During pretesting, participants in the x-A-B group produced descriptions that were approximately 65% accurate with little variability in scores. Following initial A training, which established conditional control by the directionalities, accuracy increased to nearly 100%, with little variability. Following subsequent B training, which did not establish conditional control by the directionalities, the accuracy of participants’ descriptions decreased slightly, and variability increased.

The performance of the x-A-B group suggested that improvements in descriptions of the directionalities resulted from A training alone, a conclusion that was supported, first, by the performance of the x-A-A group, shown in the middle panel. Initial A training resulted in a significant improvement, and a decrease in variability, in the accuracy of participants’ descriptions. A second round of A training produced nearly complete accuracy in participants’ descriptions. This was the case since A training involved the direct establishment of conditional control by DIR:LO and DIR:HI.

The performance of the x-x-B group, shown in the right panel, provided further support for the conclusion that A training alone was responsible for improvements in participants’ descriptions of the directionalities. The first two writing tests produced the same low accuracies
of the written descriptions. B training, which did not establish conditional control by the
directionalities, also produced virtually no change in the accuracy of the written descriptions,
though variability increased slightly. Significant improvement in the descriptions of the
directionalities, then, could be attributed to extensive A training.

**Summary.** The data presented in Figure 6 indicate that repetition of the writing test had no
effect on the accuracy of participants’ descriptions of the interaction, intersection, or
directionalities. Further, the data presented in Figure 7 indicate that the overall accuracy scores
depicted in Figure 5 were produced by differential effects of the different training procedures on
the accuracy of the different sentences: while the greatest overall improvements in all three types
of sentences were produced by the combination of A and B training, some of these
improvements could be attributed to training with A alone or B alone.

**Effects of each training procedure across types of sentence.** The data in Figure 7 can
also be analyzed by row, to determine the accuracy of the three types of sentences produced by
participants in the each group. Analyzing the data in this way reveals the effects of each of the
training procedures across the three types of sentence.

**Effects A-B training.** The left-most column of Figure 7 shows the effects of A-B training
on the accuracy of the three types of sentences. During pre-testing, participants’ accuracy scores
were at or near zero for both the interaction and intersection sentences, as shown in the top and
middle panels. Accuracy scores for the directionality sentences, however, were at an
intermediate level, as shown in the bottom panel. A training improved accuracy scores for the
interaction significantly, to an intermediate level, produced a slight and non-significant
improvement in accuracy of the intersection sentence, and produced a significant improvement,
to near 100%, in accuracy of the directionality sentences. Subsequent B training produced
significant improvements in accuracy, to near 100%, in the interaction and intersection sentences, and did not significantly change accuracy of the directionality sentences, as these descriptions were already nearly 100% accurate.

Effects of A-A training. The middle column of Figure 7 shows the effects of A-A training on the accuracy of the three types of sentences. Participants’ pre-test scores were similar to those of participants in the x-A-B group, and A training produced similar and significant increments in the accuracy of the interaction and directionality sentences, and a slight and non-significant increment in the accuracy of the intersection sentence. Repetition of A training did not significantly change the accuracy of any of the three types of sentence; rather, performances observed following initial A training remained largely intact.

Effects of B training. The right-most column of Figure 7 shows the effects of B training alone on the accuracy of the three types of sentences. Participants’ pre-test scores, for all three types of sentences, were similar to those of participants in the x-A-B and x-A-A groups, and there was no change in the accuracy of any of the three sentences with re-administration of the writing test. Subsequent B training, which established conditional control by the relational elements of the interaction and intersection, produced a significant improvement in accuracy of the interaction sentence, as shown in the top panel, to an intermediate level, and in accuracy of the intersection sentence, as shown in the middle panel, to near 100%. Accuracy of the directionality sentences, as shown in the bottom panel, did not change significantly following B training, and remained at an intermediate level.

Process analysis: written descriptions of IV/DV and relational elements in all sentences. Figure 7 considered the effects of each training procedure on the accuracy of the written descriptions of the interactive effects of two variables on behavior, on the intersection of
the functions in the interaction based graphs, and on the directionality of the functions in the interaction based graphs. Each of these three statements consists of the names of the independent and dependent variables (non-relational elements) and elements that indicate effects of one variable on another (relational elements). In the directionality statements, the relational elements are the terms describing the functions (either inverse or direct); in the interaction statements, the relational elements are those previously described as Interact (Ip and Id); in the intersection statements, the relational elements are those previously described as Intersect. All three statements also contain references to the non-relational elements (IV and DV).

Figure 8 shows the effects of each intervention on IV/DV and relational elements of each of the three types of statements. In the interaction statements, accuracy of descriptions of both relational elements (Ip and Id) was consistently similar, and so is represented by a single bar. Each of the nine graphs contains three pairs of bars; the left bar of each pair indicates accuracy of descriptions of IV and DV, and the right bar indicates accuracy of descriptions of the relational elements—either the interaction terms, the intersection terms, or the directional terms. The first pair of bars in each graph indicates performance before any intervention; the other two indicate accuracy after the interventions indicated on the abscissae.

A total of 15 pairs of bars in Figure 8 indicate performances after one of the interventions. Of these 15 pairs, five indicate approximately equal effects of the interventions on the accuracy of relational elements and IV/DV. The other 10 pairs, in shaded boxes, indicate different post-intervention accuracies of the relational elements and IV/DV, and will be discussed below.

*Elements of the interaction statements.* The top row of Figure 8 shows the effects of each intervention on the accuracy of the written descriptions of IV/DV and the relational elements...
elements in the interaction statements. After A training (see the x-A-B and x-A-A groups in the left and middle panels), participants’ descriptions of IV/DV were more accurate than those of the relational elements. This difference in accuracy between the relational elements and IV/DV, however, was proportionate to that seen before intervention for both groups (bars above the first “x” on the abscissa). Thus, the effect of A training was to increase accuracy of both relational elements and IV/DV in equal degree but from different baseline levels.

The pre-intervention descriptions of the B only group (shown in the right panel) also contained more accurate descriptions of IV/DV than the relational elements. B training improved accuracy of descriptions of both types of elements to approximately 60% accuracy. Since relational accuracy was lower than that seen for IV/DV, B training induced a greater gain for relational than for the non-relational elements. Because B training did not establish conditional control by IV/DV, the change in accuracy of IV/DV reflects generalization of the effects of training.

**Elements of the intersection statements.** The middle row of Figure 8 shows the effects of each intervention on the accuracy of descriptions of the elements of the intersection statements. As shown in the left and middle panels, A training produced a greater improvement in accuracy of descriptions of the relational elements than IV/DV. This was the opposite effect of A training on the interaction statements, where it produced a greater improvement in the accuracy of IV/DV. Additionally, the improvement in accuracy of the intersection statements reflects generalization, as A training did not establish conditional control by elements of the intersection statements.

**Elements of the directionality statements.** The bottom row of Figure 8 shows the effects of each intervention on the accuracy of descriptions of elements of the directionality statements.
As shown in the left and middle panels, initial A training produced significant improvement in descriptions of the relational elements of the directionality statements. In contrast, because the accuracy of these descriptions was nearly 100% prior to intervention, A training produced much smaller improvements in the descriptions of IV/DV. B training alone, as shown in the right panel, produced a larger improvement in accuracy of the relational elements. Because B training did not establish conditional control by elements of the directionality statements, this outcome reflected generalization of the effects of training.

**Contextual control of IV/DV descriptions.** The data in Figure 8 can also be viewed down a column to illuminate the contextual control exerted by particular sentence types over writing of descriptions of particular types of elements. Specifically, the leftmost bar in the bottom panel of the left-most column showed the presence of very accurate descriptions of IV/DV, prior to training, in the directionality statements. In contrast, the same descriptive repertoire was not observed in the sentences that described the interactions between functions or the intersection of functions, as seen in the upper and middle panels, respectively. The same can be said regarding the panels in the other columns of Figure 8. All of these data imply that IV and DV exerted control over description writing prior to training; however, this control was contextually bound by the content of the sentence in the descriptive paragraph. As the accuracy of descriptions of IV and DV improved following training, then, training reduced the control of IV and DV’s control by sentence context. In contrast, for the relational elements, the effect of training was to induce accurate writing regardless of sentence context.

**Summary.** The results as depicted in Figure 8 showed that each component of each intervention had varying effects on the accuracy of the written descriptions of the relational elements and IV/DV within and across sentences. These results, then, document the complexity
of the effects of joint stimulus control training on the writing of accurate and complete
descriptions of the interactive effects of two variables on behavior, and the extensive network of
stimulus control that influences the emergence of these complex written repertoires. These
differential effects were masked in the overall measures of accuracy shown in Figures 5 and 7.
These results, however, do not nullify the outcomes of the results as depicted in those figures:
rather, they highlight the complexity of effects produced by the various intervention components.
Regardless of this complexity, the A-B training package resulted in the emergence of complete
and accurate written descriptions of the interactive effects of two variables on behavior.
**General Discussion**

**Synopsis**

The results of the present experiment and those of Spear and Fields (submitted) indicate that the establishment of visual-visual conditional relations resulted in the emergence of accurate and complete paragraph-length written descriptions of interaction-based graphs. The relations established were targeted conditional discriminations between graphs and key elements of their printed descriptions and joint stimulus control by the graphs and key elements of their printed descriptions in combination with a number of non-relational elements of the printed descriptions. Further, description-writing was never directly trained. The training involved the establishment of selection-based repertoires, indicated by acquisition of graph-text conditional discriminations. Writing, on the other hand, is an expressive or production-based repertoire. Therefore, the establishment of a selection-based repertoire resulted in the immediate emergence of a rather extensive production-based repertoire.

**Effects of Components of A and B Training Alone and Together**

**Effects of B alone.** As noted in the Introduction, the production of accurate descriptions of the graphs required attention to many features of the graphs and their printed descriptions. More specifically, the production of complete and accurate descriptions of the interactions and intersections had to reflect at least three stimulus control topographies: the phrases describing the interaction (Ip and Id) and the phrase describing the intersection of the functions in the graphs (Is). These had to be present for each of the four graphs, for a total of four “sets” of stimulus control topographies. When these topographies were established in the B training protocol, the participants wrote descriptions that were about 50% accurate. Therefore, establishing targeted conditional control by graph-Id, graph-Ip, and graph-Is relations together was not sufficient to
induce the writing of accurate descriptions of the interaction-based graphs.

**Effects of A alone.** A training involved the establishment of conditional relations between the graph and any element of printed text (traditional training), targeted conditional relations between graphs and one set of phrases that described the interactive effects of the variables on behavior, (the Id relations), and joint stimulus control by the Id relations in combination with four non-relational elements. To what extent did each of these components induce accurate writing of descriptions of the interactions?

Part of the A training protocol involved the use of traditional conditional discrimination training to establish a discrimination among the various graphs used as sample stimuli and conditional relations between each graph and one unspecified element of the printed paragraphs used as comparison stimuli. While it was not studied in isolation in the present experiment, it was used in a virtual replication of the present experiment by Spear and Fields (submitted), and produced very inaccurate written descriptions. In the present experiment it is reasonable to assume that the traditional conditional discrimination training used in the A-training component did not influence the emergence of accurate written descriptions of interactions. This finding is also consistent with the failures to describe medical diagnoses and features of research designs by the participants in the experiments reported by Walker, et al. (2010) and Walker and Rehfeldt (2012).

The B training protocol established control by each of the four sets of relational elements in the context of only one set of non-relational elements. It did not establish joint control by the combination of relational elements, such as Id, Ip and Is, and the non-relational elements of graphs and descriptions that would also be featured in accurate descriptions of the interactions. Thus, it is possible that the establishment of control by the relational elements in the context of
more than one set of non-relational elements would have produced accurate written descriptions of the interactions. This issue was addressed in part with the A training protocol that established control by the Id relational elements in the context of four sets of non-relational elements. This training also resulted in the production of descriptions that were about 50% accurate. Thus, the establishment of joint control by one of the elements of the interaction (Id) on combination with the non-relational elements was partially responsible for the induction of accurate written descriptions.

**Effects of A and B training together.** The very high level of writing accuracy that emerged only after the establishment of the conditional relations established during A and B training consisted of improvements in the production of descriptions of all of the features of the graphs: the directionalities of the functions, the interaction of the independent variables, the intersection of the functions, and the specification of the non-relational variables. These results demonstrate that only the combination of A and B training established all sources of stimulus control necessary to produce nearly complete and accurate descriptions of the interactions.

**Sources of stimulus control after A and B training.** The exact sources of stimulus control established by the A and B protocols are shown in Table 3. There are a total of 16 combinations, or forms, of control that can be established for each of the four types of interaction, for a total of 64 forms of conditional stimulus control. Theoretically, all of these forms of conditional stimulus control would have to be intact for an individual to produce a complete and accurate description of a graph. Traditional training, as reported by earlier studies (Walker, et al.; Walker & Rehfeldt; Spear & Fields), establishes one of these sources of control for each of the four interaction types (graph and any correlated element of the printed description of the graph) for a total of four of the 64 forms of control. This procedure resulted in the writing
of descriptions that were approximately 20% accurate.

The A training protocol, on the other hand, establishes six conditional stimulus control topographies for each of the four types of interaction, for a total of 24 of the 64 forms of control. These 24 forms of control included targeted conditional stimulus control for one graph-phrase relation (Id), and four joint stimulus control topographies, consisting of Id and the four non-relational elements. This training resulted in the writing of descriptions that were approximately 70% accurate. The addition of B training, which established two additional forms of control for each of the four types of interaction, or a total of eight forms of control, results in the writing of near completely accurate descriptions. Thus, establishing control by 32 of the 64 possible forms of control, as accomplished by A and B training together, resulted in nearly complete writing accuracy. These performances occurred only when the training protocols established three targeted graph-phrase conditional relations (Id, Ip, and Is) and joint control by one of the three already established graph-phrase relations (Id) in combination with the four non-relational elements that characterized the interactive effects of two variables on behavior.

Notably, it is possible that the 32 forms of conditional and joint control established by the current procedures are not necessary to induce the production of accurate descriptions, and that the establishment of a different set of 32 forms of control—including, for instance, Is in combination with each of the four non-relational elements—might suffice. It is also possible that the establishment of an as-yet-unknown smaller number of forms of control would also result in the production of accurate descriptions. Both of these issues require further research.

**Mechanisms of Effect: Emergent Repertoires and Function Transfer**

**Emergent Repertoires.** As stated previously, the production of complete and accurate descriptions of necessity implies the presence of all 64 forms of targeted conditional stimulus
control and joint stimulus control in some combination. The combination of A and B training established only half of these, and yet resulted in the writing of near completely accurate descriptions. The fact that this outcome was achieved indicates generalization of writing performance from trained to untrained forms of control. For instance, control by IV1 was established by A training, and control by Is was established by B training. Thus, control by each of these elements was established in isolation. After the combination of A and B training, however, these two elements exerted control in combination, facilitating production of an accurate description of the intersection that included accurate references to both Is and IV1. The emergence of an accurate description of these two elements in combination is an example of adduction (Andronis, Layng, & Goldiamond, 1997), or recombinative generalization (Goldstein, 1983; Goldstein, Angelo, & Mousetis, 1987), as these two elements had not been directly linked during prior conditional discrimination training but, nonetheless, together exerted control over writing behavior.

**Function Transfer.** Recombinative generalization, however, does not explain the main finding of the current research: a generalized improvement in writing, a production-based response, following selection-based training. A possible explanation for this outcome is based on transfer of function: participants entered the experiment with a text-copying repertoire (i.e., the ability to produce a complete and accurate copy of text presented to them), and the training procedures used in the current study likely facilitated the transfer of this text-copying function from the printed paragraphs to the graphs. In combination with recombinative generalization, the current selection-based procedures could produce a generalized improvement in writing behavior. If this account is accurate, failure of the current training procedures to produce an improvement in writing could result from a lack of either function transfer or generalization, and
could be treated by conducting additional training to establish the some or all of the additional 32 forms of control not established by the current procedures.

Most studies of function transfer have been conducted in the context of equivalence classes: some function is acquired by one class member and then generalizes with a constant high probability to the other class members. In the present experiment, equivalence class formation was not used to study function transfer. Rather, a set of graph-text conditional relations was established, after which the behaviors presumed to occur in the presence of one of the stimuli generalized to the other member of the conditional relation. In addition, the stimuli in the visual-visual conditional relations were quite complex and the response that generalized involved the production of paragraph-length written descriptions. Thus, function transfer did not depend on the prior formation of equivalence classes, but rather of conditional discriminations.

It would be of interest, however, to expand the behavioral preparation by establishing a second type of conditional discrimination: for instance, between the name of the type of interaction and the graph of that type of interaction. Thereafter, two important questions arise: first, would the presentation of any one of the stimuli representing one of the types of interaction evoke the selection of the other stimuli representing that interaction, i.e., would equivalence classes form? Second, would the presentation of any one of the stimuli, such as the name of the interaction, evoke the writing of a complete and accurate description of the graph depicting that type of interaction, i.e., would the writing function transfer to all members of the same class?

**Summary and Conclusions**

In the current experiment, joint stimulus control can be described as attention to all the features of complex stimuli. Conditional discrimination training procedures that established this sort of control by many elements of complex stimuli were found to produce an improvement in
writing behavior. The method of measuring joint control in the current study is fundamentally similar to those of Reynolds (1961), Gibson, Wasserman, Gosselin, and Schyns (2005), and Lazareva, Freiburger, and Wasserman (2006). Further, the results obtained by Reynolds and others (Stromer, McIlvane, Dube, & Mackay, 1993; Critchfield & Perone, 1993) indicate that simple discrimination training with complex discriminative stimuli does not establish control by all elements of those stimuli, and so are similar to those of traditional conditional discrimination training procedures as reported by Walker, Rehfeldt, and Ninness (2010), Walker and Rehfeldt (2012), and Spear and Fields.

The current research extends previous research that has addressed the control of behavior by complex stimuli over complex responses by the use of more complex forms of each (e.g., Danforth, Chase, Dolan, & Joyce, 1990). While successful, these results also raise several questions that can only be answered by additional empirical research. First, will the reported procedures produce similar improvements in descriptions of novel academic content? Second, will training such as this result in improvements in other repertoires such as verbally describing the interactions displayed in the graphs used in the present experiment? Third, can improvements in performances described in the present experiment be accomplished by use of simpler methods?

It is not known whether the current training would result in improved accuracy of descriptions of graphs that were not presented during training. Specifically, how would the accuracy of descriptions vary with graphs that differed from those used in training with respect to content, slopes of the functions, or both? Similarly, would the current training protocol result in improvement in other repertoires, such as the production of spoken descriptions? The documentation of positive generalization of the effects of training would be a prerequisite for the
application of the current training procedures to a practical venue such as the teaching of relevant college level courses.

It is also possible that similar outcomes could be produced with a more efficient form of training: it is not known whether all of the forms of joint control established in the current study are necessary to facilitate the production of written descriptions. It is thus possible that a significantly shorter form of joint control training might produce the same outcomes. Further, it is possible that other forms of training, such as practicing text-copying, listening to a lecture, or reading a textbook chapter, might also produce notable improvements in the accuracy of written descriptions. A comparison of outcomes of these interventions would provide a critical measure of the efficacy of the present and refined training protocols using more traditional measures as the evaluative benchmarks.
Table 1

A glossary of terms and abbreviations used to denote the elements of graphs and their accompanying printed descriptions. Examples refer to the features of a graph depicting no interaction, as shown in the top row of Figure 1. The relevant terms of each sentence are italicized.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Name</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV1</td>
<td>Independent variable 1</td>
<td>Independent variable listed on the abscissa</td>
<td>Antidepressant dose</td>
</tr>
<tr>
<td>IV2</td>
<td>Independent variable 2</td>
<td>Independent variable listed in the legend</td>
<td>Rising time</td>
</tr>
<tr>
<td>DIR:LO</td>
<td>Directionality of the lower function</td>
<td>Slope of the function with the lower y-intercept</td>
<td>Depression score was a direct function....</td>
</tr>
<tr>
<td>DIR:HI</td>
<td>Directionality of the higher function</td>
<td>Slope of the function with the higher y-intercept</td>
<td>Depression score was a direct function....</td>
</tr>
<tr>
<td>Ip</td>
<td>Interaction: produced or reversed</td>
<td>Produced or reversed term of the interaction statement</td>
<td>Rising time produced a constant difference...</td>
</tr>
<tr>
<td>Id</td>
<td>Interaction: difference</td>
<td>Difference term of the interaction statement</td>
<td>Rising time produced a constant difference....</td>
</tr>
<tr>
<td>Is</td>
<td>Intersection of the functions</td>
<td>Intersection term of the intersection statement</td>
<td>Depression score functions did not intersect....</td>
</tr>
<tr>
<td>DV</td>
<td>Dependent variable</td>
<td>Variable listed on the ordinate</td>
<td>Depression score</td>
</tr>
</tbody>
</table>
Table 2

The values of each element of each graph presented during the traditional training protocol.

<table>
<thead>
<tr>
<th>Non-Relational</th>
<th>IV1</th>
<th>Antidepressant</th>
<th>Exercise</th>
<th>Yoga</th>
<th>Salt Eaten</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relational</td>
<td>IV2</td>
<td>Rising Time</td>
<td>Education</td>
<td>Income</td>
<td>School</td>
</tr>
<tr>
<td>DV</td>
<td>Depression Score</td>
<td>Hours Slept</td>
<td>Anxiety</td>
<td>Pulse Rate</td>
<td></td>
</tr>
<tr>
<td>DIR:LO</td>
<td>Direct</td>
<td>Inverse</td>
<td>Inverse</td>
<td>Direct</td>
<td></td>
</tr>
<tr>
<td>DIR:HI</td>
<td>Direct</td>
<td>Direct</td>
<td>Direct</td>
<td>Direct</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relational</th>
<th>Interact:d</th>
<th>Constant</th>
<th>--</th>
<th>--</th>
<th>Growing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interact:p</td>
<td>produced</td>
<td>Reversed</td>
<td>Reversed</td>
<td>produced</td>
<td></td>
</tr>
<tr>
<td>Intersect</td>
<td>Did not</td>
<td>Intersected</td>
<td>Did not</td>
<td>Did not</td>
<td></td>
</tr>
<tr>
<td>Intersect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intersect</td>
</tr>
</tbody>
</table>

[Graphs showing values for different elements]
Table 3

All possible conditional and joint stimulus control relations among the relational and non-relational elements, and those established during the intervention components. “y” indicates the training of the form of stimulus control listed in the leftmost column, while “n” indicates an untrained stimulus control relation.

<table>
<thead>
<tr>
<th>Train Type</th>
<th>Element</th>
<th>A+B</th>
<th>A</th>
<th>B</th>
<th>x-x-x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional CD</td>
<td></td>
<td>y</td>
<td>y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Targeted CD</td>
<td>Interact:d</td>
<td>y</td>
<td>y</td>
<td>y</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Interact:p</td>
<td>y</td>
<td>n</td>
<td>y</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Intersect</td>
<td>y</td>
<td>n</td>
<td>y</td>
<td>n</td>
</tr>
<tr>
<td>Joint Ctrl</td>
<td>Id &amp; IV1</td>
<td>y</td>
<td>y</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Id &amp; IV2</td>
<td>y</td>
<td>y</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Id &amp; DIR:LO</td>
<td>y</td>
<td>y</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Id &amp; DIR:HI</td>
<td>y</td>
<td>y</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>Joint Ctrl</td>
<td>Ip &amp; IV1</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Ip &amp; IV2</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Ip &amp; DIR:LO</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Ip &amp; DIR:HI</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>Joint Ctrl</td>
<td>Is &amp; IV1</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Is &amp; IV2</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Is &amp; DIR:LO</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Is &amp; DIR:HI</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>DV</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
</tbody>
</table>
**WRITING WITHOUT WRITING**

Table 4

The stimuli used as incorrect comparisons to establish targeted conditional control by Interact:d (Id) for each of the four types of interaction. The incorrect words in each description are bolded. This bolding did not appear during training.

<table>
<thead>
<tr>
<th>No</th>
<th>XOVR</th>
<th>DVR</th>
<th>SYN</th>
</tr>
</thead>
<tbody>
<tr>
<td>#9. Depression score was a direct function of antidepressant dosage for late risers. In addition, depression score was a direct function of antidepressant dosage for early risers. For each level of antidepressant dosage, the depression score was an inverse function of rising time. Increasing the antidepressant dosage produced a <strong>growing</strong> difference in depression scores for late risers relative to early risers. Depression score functions for late- and early-risers did not intersect at any level of antidepressant dosage.</td>
<td>#10. For high school graduates, the number of hours slept was an inverse function of the number of hours exercised per day. For college graduates, the number of hours slept was a direct function of the number of hours exercised per day. For different numbers of hours spent exercising, then, education produced a <strong>constant</strong> difference in the effect of time spent exercising on the number of hours slept. The hours slept functions for individuals of different education levels intersected at an intermediate level of hours spent exercising.</td>
<td>#11. The anxiety level was a direct function of days of yoga per week for low-income individuals, and an inverse function of days of yoga per week for high-income individuals. For each number of days of yoga per week, the anxiety level was greater for low- than high-income individuals. Income level produced a <strong>constant</strong> difference in the directional effect of days of yoga on anxiety level. Finally, the anxiety level functions for high- and low-income individuals did not intersect at any level of days of yoga.</td>
<td>#12. Pulse rate was a direct function of salt consumption for graduate students. In addition, pulse rate was a direct function of salt consumption for college students. For each level of salt consumption, type of school attended produced higher pulse rates for graduate students relative to college students. Increasing the amount of salt eaten produced a <strong>constant</strong> difference in pulse rates for graduate students relative to college students. Pulse rate functions for graduate and college students did not intersect at any level of salt consumption.</td>
</tr>
</tbody>
</table>
Table 5

The stimuli used as incorrect comparisons to establish targeted conditional control by Interact:p (Ip) for each of the four types of interaction. The incorrect words in each description are bolded. This bolding did not appear during training.

<table>
<thead>
<tr>
<th>No</th>
<th>XOVR</th>
<th>DVR</th>
<th>SYN</th>
</tr>
</thead>
<tbody>
<tr>
<td>#13. Depression score was a direct function of antidepressant dosage for late risers. In addition, depression score was a direct function of antidepressant dosage for early risers. For each level of antidepressant dosage, the depression score was an inverse function of rising time. Increasing the antidepressant dosage did not produce a constant difference in depression scores for late risers relative to early risers. Depression score functions for late- and early-risers did not intersect at any level of antidepressant dosage.</td>
<td>#14. For high school graduates, the number of hours slept was an inverse function of the number of hours exercised per day. For college graduates, the number of hours slept was a direct function of the number of hours exercised per day. For different numbers of hours spent exercising, then, education did not reverse the effect of time spent exercising on the number of hours slept. The hours slept functions for individuals of different education levels intersected at an intermediate level of hours spent exercising.</td>
<td>#15. The anxiety level was a direct function of days of yoga per week for low-income individuals, and an inverse function of days of yoga per week for high-income individuals. For each number of days of yoga per week, the anxiety level was greater for low- than high-income individuals. Income level did not reverse the directional effect of days of yoga on anxiety level. Finally, the anxiety level functions for high- and low-income individuals did not intersect at any level of days of yoga.</td>
<td>#16. Pulse rate was a direct function of salt consumption for graduate students. In addition, pulse rate was a direct function of salt consumption for college students. For each level of salt consumption, type of school attended produced higher pulse rates for graduate students relative to college students. Increasing the amount of salt eaten did not produce a growing difference in pulse rates for graduate students relative to college students. Pulse rate functions for graduate and college students did not intersect at any level of salt consumption.</td>
</tr>
</tbody>
</table>
The stimuli used as incorrect comparisons to establish targeted conditional control by Intersect (Is) for each of the four types of interaction. The incorrect words in each description are bolded. This bolding did not appear during training.

<table>
<thead>
<tr>
<th>No</th>
<th>XOVR</th>
<th>DVR</th>
<th>SYN</th>
</tr>
</thead>
<tbody>
<tr>
<td>#17. Depression score was a direct function of antidepressant dosage for late risers. In addition, depression score was a direct function of antidepressant dosage for early risers. For each level of antidepressant dosage, the depression score was an inverse function of rising time. Increasing the antidepressant dosage produced a constant difference in depression scores for late risers relative to early risers. Depression score functions for late- and early-risers intersected at an intermediate level of antidepressant dosage.</td>
<td>#18. For high school graduates, the number of hours slept was an inverse function of the number of hours exercised per day. For college graduates, the number of hours slept was a direct function of the number of hours exercised per day. For different numbers of hours spent exercising, then, education reversed the effect of time spent exercising on the number of hours slept. The hours slept functions for individuals of different education levels did not intersect at any level of hours spent exercising.</td>
<td>#19. The anxiety level was a direct function of days of yoga per week for low-income individuals, and an inverse function of days of yoga per week for high-income individuals. For each number of days of yoga per week, the anxiety level was greater for low- than high-income individuals. Income level reversed the directional effect of days of yoga on anxiety level. Finally, the anxiety level functions for high- and low-income individuals intersected at an intermediate level of days of yoga.</td>
<td>#20. Pulse rate was a direct function of salt consumption for graduate students. In addition, pulse rate was a direct function of salt consumption for college students. For each level of salt consumption, type of school attended produced higher pulse rates for graduate students relative to college students. Increasing the amount of salt eaten produced a growing difference in pulse rate for graduate students relative to college students. Pulse rate functions for graduate and college students intersected at an intermediate level of salt consumption.</td>
</tr>
</tbody>
</table>
Table 7

Printed descriptions used to establish targeted conditional control by the non-relational elements of graphs and descriptions depicting no interaction (Interact: NO). The incorrect words in each description are bolded. This bolding did not appear during training.

Depression score was a direct function of **hours exercised** for late risers. In addition, depression score was a direct function of **hours exercised** for early risers. For each level of **hours exercised**, the depression score was an inverse function of rising time. Increasing the number of **hours exercised** produced a constant difference in depression scores for late risers relative to early risers. Depression score functions for late- and early-risers did not intersect at any level of **hours exercised**.

<table>
<thead>
<tr>
<th>#21</th>
<th>Id, Ip, Is</th>
<th>-- IV1</th>
</tr>
</thead>
</table>

Depression score was a direct function of antidepressant dosage for **high school graduates**. In addition, depression score was a direct function of antidepressant dosage for **college graduates**. For each level of antidepressant dosage, the depression score was an inverse function of education. Increasing the antidepressant dosage produced a constant difference in depression scores for **high school**- relative to **college graduates**. Depression score functions for **high school** and **college graduates** did not intersect at any level of antidepressant dosage.

<table>
<thead>
<tr>
<th>#22</th>
<th>Id, Ip, Is</th>
<th>-- IV2</th>
</tr>
</thead>
</table>

Depression score was a direct function of antidepressant dosage for late risers. In addition, depression score was an **inverse** function of antidepressant dosage for early risers. For each level of antidepressant dosage, the depression score was an inverse function of rising time. Increasing the antidepressant dosage produced a constant difference in depression scores for late risers relative to early risers. Depression score functions for late- and early-risers did not intersect at any level of antidepressant dosage.

<table>
<thead>
<tr>
<th>#23</th>
<th>Id, Ip, Is</th>
<th>-- DIR:LO</th>
</tr>
</thead>
</table>

Depression score was an **inverse** function of antidepressant dosage for late risers. In addition, depression score was a direct function of antidepressant dosage for early risers. For each level of antidepressant dosage, the depression score was an inverse function of rising time. Increasing the antidepressant dosage produced a constant difference in depression scores for late risers relative to early risers. Depression score functions for late- and early-risers did not intersect at any level of antidepressant dosage.

<table>
<thead>
<tr>
<th>#24</th>
<th>Id, Ip, Is</th>
<th>-- DIR:HI</th>
</tr>
</thead>
</table>
Table 8

*Printed descriptions used to establish targeted conditional control by the non-relational elements of graphs and descriptions depicting crossover interaction (Interact: XOVR). The incorrect words in each description are bolded. This bolding did not appear during training.*

For high school graduates, the number of hours slept was an inverse function of antidepressant dosage. For college graduates, the number of hours slept was a direct function of antidepressant dosage. For different levels of antidepressant dosage, then, education reversed the effect of antidepressant dosage on the number of hours slept. The hours slept functions for individuals of different education levels intersected at an intermediate level of antidepressant dosage.

For high late risers, the number of hours slept was an inverse function of the number of hours exercised per day. For early risers, the number of hours slept was a direct function of the number of hours exercised per day. For different numbers of hours spent exercising, then, rising time reversed the effect of time spent exercising on the number of hours slept. The hours slept functions for individuals of different rising times intersected at an intermediate level of hours spent exercising.

For high school graduates, the number of hours slept was a direct function of the number of hours exercised per day. For college graduates, the number of hours slept was a direct function of the number of hours exercised per day. For different numbers of hours spent exercising, then, education reversed the effect of time spent exercising on the number of hours slept. The hours slept functions for individuals of different education levels intersected at an intermediate level of hours spent exercising.

#25
+ Id, Ip, Is
-- IV1

#26
+ Id, Ip, Is
-- IV2

#27
+ Id, Ip, Is
-- DIR:LO

#28
+ Id, Ip, Is
-- DIR:HI
Table 9

*Printed descriptions used to establish targeted conditional control by the non-relational elements of graphs and descriptions depicting divergent interaction (Interact: DVR). The incorrect words in each description are bolded. This bolding did not appear during training.*

| The anxiety level was a direct function of **salt consumption** for low-income individuals, and an inverse function of **salt consumption** for high-income individuals. For each level of **salt consumption**, the anxiety level was greater for low- than high-income individuals. Income level reversed the directional effect of **salt consumption** on anxiety level. Finally, the anxiety level functions for high- and low-income individuals did not intersect at any level of **salt consumption**. | #29  
+ Id, Ip, Is  
-- IV1 |
|---|---|
| The anxiety level was a direct function of days of yoga per week for **graduate students**, and an inverse function of days of yoga per week for **college students**. For each number of days of yoga per week, the anxiety level was greater for **graduate students** than for **college students**. **Type of school attended** reversed the directional effect of days of yoga on anxiety level. Finally, the anxiety level functions for **graduate and college students** did not intersect at any level of days of yoga. | #30  
+ Id, Ip, Is  
-- IV2 |
| The anxiety level was a direct function of days of yoga per week for low-income individuals, and a **direct** function of days of yoga per week for high-income individuals. For each number of days of yoga per week, the anxiety level was greater for low- than high-income individuals. Income level reversed the directional effect of days of yoga on anxiety level. Finally, the anxiety level functions for high- and low-income individuals did not intersect at any level of days of yoga. | #31  
+ Id, Ip, Is  
-- DIR:LO |
| The anxiety level was an **inverse** function of days of yoga per week for low-income individuals, and an inverse function of days of yoga per week for high-income individuals. For each number of days of yoga per week, the anxiety level was greater for low- than high-income individuals. Income level reversed the directional effect of days of yoga on anxiety level. Finally, the anxiety level functions for high- and low-income individuals did not intersect at any level of days of yoga. | #32  
+ Id, Ip, Is  
-- DIR:HI |
Table 10

*Printed descriptions used to establish targeted conditional control by the non-relational elements of graphs and descriptions depicting synergistic interaction (Interact: SYN). The incorrect words in each description are bolded. This bolding did not appear during training.*

<table>
<thead>
<tr>
<th>Pulse rate was a direct function of <strong>days of yoga</strong> per week for graduate students. In addition, pulse rate was a direct function of <strong>days of yoga</strong> per week for college students. For each number of <strong>days of yoga</strong>, type of school attended produced higher pulse rates for graduate students relative to college students. Increasing the number of <strong>days of yoga</strong> produced a growing difference in pulse rate for graduate students relative to college students. Pulse rate functions for graduate and college students did not intersect at any number of <strong>days of yoga.</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>#33</td>
</tr>
<tr>
<td>+ Id, Ip, Is</td>
</tr>
<tr>
<td>-- IV1</td>
</tr>
<tr>
<td>Pulse rate was a direct function of salt consumption for <strong>low-income individuals</strong>. In addition, pulse rate was a direct function of salt consumption for <strong>high-income individuals</strong>. For each level of salt consumption, <strong>income level</strong> produced higher pulse rates for <strong>low-income</strong> relative to <strong>high-income individuals</strong>. Increasing the amount of salt eaten produced a growing difference in pulse rate for <strong>low-</strong> relative to <strong>high-income individuals</strong>. Pulse rate functions for <strong>low-</strong> and <strong>high-income individuals</strong> did not intersect at any level of salt consumption.</td>
</tr>
<tr>
<td>#34</td>
</tr>
<tr>
<td>+ Id, Ip, Is</td>
</tr>
<tr>
<td>-- IV2</td>
</tr>
<tr>
<td>Pulse rate was a direct function of salt consumption for graduate students. In addition, pulse rate was an <strong>inverse</strong> function of salt consumption for college students. For each level of salt consumption, type of school attended produced higher pulse rates for graduate students relative to college students. Increasing the amount of salt eaten produced a growing difference in pulse rate for graduate students relative to college students. Pulse rate functions for graduate and college students did not intersect at any level of salt consumption.</td>
</tr>
<tr>
<td>#35</td>
</tr>
<tr>
<td>+ Id, Ip, Is</td>
</tr>
<tr>
<td>-- DIR:LO</td>
</tr>
<tr>
<td>Pulse rate was an <strong>inverse</strong> function of salt consumption for graduate students. In addition, pulse rate was a direct function of salt consumption for college students. For each level of salt consumption, type of school attended produced higher pulse rates for graduate students relative to college students. Increasing the amount of salt eaten produced a growing difference in pulse rate for graduate students relative to college students. Pulse rate functions for graduate and college students did not intersect at any level of salt consumption.</td>
</tr>
<tr>
<td>#36</td>
</tr>
<tr>
<td>+ Id, Ip, Is</td>
</tr>
<tr>
<td>-- DIR:HI</td>
</tr>
</tbody>
</table>
Table 11

A symbolic representation of the ‘A’ training protocol. Numbers in the “Interaction type” column refer to the graphs shown in Figure 1; numbers in the “Comparisons” columns refer to the descriptions shown in Figure 1 and Tables 3-8. “FB” refers to the percentage of trials followed by informative feedback, and “FBR” refers to feedback reduction.

<table>
<thead>
<tr>
<th>Type of Training</th>
<th>FB</th>
<th>Phase</th>
<th>Sample Interaction type</th>
<th>Comparisons</th>
<th>+ - element</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Row #</td>
<td>NO</td>
<td>XOV R</td>
</tr>
<tr>
<td>Trad</td>
<td>100</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>(Id)</td>
<td>100</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>Phases 3 and 4 combined</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>6</td>
<td>10</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>11</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>12</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>13</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>14</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
<td>15</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>16</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17</td>
<td>17</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Serial</td>
<td>100</td>
<td>7</td>
<td>18</td>
<td>18</td>
<td>5</td>
</tr>
<tr>
<td>Id + Nonrel</td>
<td></td>
<td>19</td>
<td>19</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>20</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>21</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>22</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>23</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>24</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>25</td>
<td>5</td>
<td>All Phases 6 and 7 combined</td>
</tr>
<tr>
<td>All</td>
<td>100</td>
<td>9</td>
<td>26</td>
<td>26</td>
<td>All Phases 6 and 7 combined</td>
</tr>
<tr>
<td></td>
<td></td>
<td>27</td>
<td>27</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28</td>
<td>28</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29</td>
<td>29</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>30</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Probe</td>
<td>0</td>
<td>10</td>
<td>31</td>
<td>31</td>
<td>Same as Phase 9</td>
</tr>
<tr>
<td>FBR</td>
<td>100</td>
<td>11</td>
<td>32</td>
<td>32</td>
<td>Same as Phase 9</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>12</td>
<td>33</td>
<td>33</td>
<td>Same as Phase 9</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>13</td>
<td>34</td>
<td>34</td>
<td>Same as Phase 9</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>14</td>
<td>35</td>
<td>35</td>
<td>Same as Phase 9</td>
</tr>
</tbody>
</table>
A symbolic representation of the ‘B’ training protocol. Numbers in the “Interaction type” column refer to the graphs shown in Figure 1; numbers in the “Comparisons” columns refer to the descriptions shown in Figure 1 and Tables 3 and 4. “FB” refers to the percentage of trials followed by informative feedback, and “FBR” refers to feedback reduction.

<table>
<thead>
<tr>
<th>Type of Training</th>
<th>FB</th>
<th>Phase</th>
<th>Sample Interaction type</th>
<th>Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Row #</td>
<td>NO</td>
</tr>
<tr>
<td>(Id)</td>
<td>100</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>Phases 1 and 2 combined</td>
<td></td>
</tr>
<tr>
<td>(Ip)</td>
<td>100</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>Phases 3 and 4 combined</td>
<td></td>
</tr>
<tr>
<td>Id, Ip</td>
<td>100</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
<td>Phases 7 and 8 combined</td>
<td></td>
</tr>
<tr>
<td>(Is)</td>
<td>100</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>Phases 9 and 10 combined</td>
<td></td>
</tr>
<tr>
<td>Id, Ip, Is</td>
<td>100</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Id, Ip, Is</td>
<td>100</td>
<td>25</td>
<td>Same as Phase 11</td>
<td></td>
</tr>
<tr>
<td>Probe</td>
<td>0</td>
<td>12</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>FBR</td>
<td>100</td>
<td>13</td>
<td>Same as Phase 11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>14</td>
<td>Same as Phase 11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>15</td>
<td>Same as Phase 11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>16</td>
<td>Same as Phase 11</td>
<td></td>
</tr>
</tbody>
</table>
Table 13

*Inter-observer agreements for each sentence type and all sentences on the pre- and post-tests.*

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Pre-test</th>
<th>Post-test 1</th>
<th>Post-test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directionalities</td>
<td>98</td>
<td>99</td>
<td>93</td>
</tr>
<tr>
<td>Interaction</td>
<td>96</td>
<td>96</td>
<td>92</td>
</tr>
<tr>
<td>Intersection</td>
<td>99</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>All sentences</td>
<td><strong>98</strong></td>
<td><strong>97</strong></td>
<td><strong>93</strong></td>
</tr>
</tbody>
</table>
Figure 1. Graphs depicting interactions and their corresponding accurate printed descriptions.
**NO Interaction**

DIR:LO (4 pts): Depression score (1) was a direct function (1) of antidepressant dosage (1) for late risers (1).

DIR:HI (4 pts): In addition, depression score (1) was a direct function (1) of antidepressant dosage (1) for early risers (1).

Interact (6): Increasing the antidepressant dosage (1) produced (1) a constant difference (1) in depression scores (1) for late risers (1) relative to early risers (1).

Intersect (5): Depression score (1) functions for late (1) and early-risers (1) did not intersect (1) at any level of antidepressant dosage (1).

**CROSSOVER**

DIR:LO (4 pts): For college graduates (1), the number of hours slept (1) was a direct function (1) of the number of hours exercised (1) per day.

DIR:HI (4 pts): For high school graduates (1), the number of hours slept (1) was an inverse function (1) of the number of hours exercised (1) per day.

Interact (4 pts): For different numbers of hours spent exercising, then, education (1) reversed (1) the effect of time spent exercising (1) on the number of hours slept (1).

Intersect (5 pts): The hours slept (1) functions for individuals of different education levels (1) intersected (1) at an intermediate level of hours spent exercising (1).

**DIVERGENT**

DIR:LO (4 pts): The anxiety level (1) was a direct function (1) of days of yoga (1) per week for low-income individuals (1).

DIR:HI (4 pts): The anxiety level (1) was an inverse function (1) of days of yoga (1) per week for high-income individuals (1).

Interact (4 pts): Income level (1) reversed (1) the directional effect of days of yoga (1) on anxiety level (1).

Intersect (5 pts): Finally, the anxiety level (1) functions for high- (1) and low-income individuals (1) did not intersect (1) at any number of days of yoga (1).

**SYNERGISTIC**

DIR:LO (4 pts): In addition, pulse rate (1) was a direct function (1) of salt consumption (1) for college students (1).

DIR:HI (4 pts): Pulse rate (1) was a direct function (1) of salt consumption (1) for graduate students (1).

Interact (6 pts): Increasing the amount of salt eaten (1) produced (1) a growing difference (1) in pulse rate (1) for graduate students (1) relative to college students (1).

Intersect (5 pts): Pulse rate (1) functions for graduate and college students (1) did not intersect (1) at any level of salt consumption (1).

**Figure 2.** The grading rubric used to evaluate participants’ written descriptions. Points are earned for the presence of bold font terms. The numbers of points possible for an entirely correct written production of each sentence are shown in parentheses. Writing accuracy scores were calculated by totaling the number of points earned for entire descriptions and for each type of sentence.
Figure 3. The average number of trials required to reach mastery criterion by participants in the three groups: x-A-A (left-most set of bars), x-x-B (middle bar) and x-A-B (right-most set of bars). The dashed line indicates the minimum number of trials needed to reach mastery.
Figure 4. The percentage of participants in each group that achieved each of four different constant performance criteria on the final writing test. 60, 70, 80, and 90% accuracy correspond to letter grades of D, C, B, and A, respectively.
Figure 5. The average accuracy of all elements of all statements following each intervention package (top panel), and of each type of statement (lower three panels). Error bars represent the standard errors of the mean.
Figure 6. The average accuracy of all statements (top panel) and each type of statement (lower three panels) of descriptions written by the x-x-x group. Error bars represent the standard errors of the mean.
Figure 7. The average accuracy of each of the three statements following each intervention. The three bars on each graph represent scores on the first, second, and third writing tests, respectively, and significant differences between test scores are indicated by the endpoints of the horizontal bars. Striped bars represent performances reflecting generalization, and error bars represent the standard error of the mean.
Figure 8. Accuracy following each intervention of IV and DV (white bar in each pair) and relational elements (gray bar in each pair) of each of the three statements. Striped bars represent performances reflecting generalization, and boxes surround bars that represent instances of differential effects of training on the two types of elements, as discussed in the text. Error bars represent the standard error of the mean.
Bibliography


