Graduate Theses and Dissertations   Graduate School

2004

The importance of program-delivered differential reinforcement in the development of classical music auditory discrimination

Gudmundur Torfi Heimisson
University of South Florida

Follow this and additional works at: http://scholarcommons.usf.edu/etd

Part of the American Studies Commons

The Importance of Program-Delivered Differential Reinforcement in the Development of

Classical Music Auditory Discrimination

by

Gudmundur Torfi Heimisson

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Arts
Department of Applied Behavior Analysis
College of Graduate Studies
University of South Florida

Major Professor: Darrel E. Bostow, Ph.D
Kevin Murdock, Ph.D.
Ken Richter, Ph.D.

Date of approval:
July 12, 2004

Keywords: Programmed instruction, education, computer-assisted learning, web-based
instruction, experimental analysis of teaching methods.

Table of contents

The importance of program-delivered differential reinforcement in the development of classical music auditory discrimination.

Gudmundur T. Heimisson

ABSTRACT

Posttest performances after two forms of Web-based tutorial instruction were compared. Both forms were designed to teach students to identify musical compositions that typify renaissance, baroque, classical, romantic, and 20th century music. The first treatment condition was a series of Web pages with text and accompanying hyperlinks to musical selections matched to the text. In this condition, students read and listened at their own discretion — without Web software program restrictions. The second treatment contained exactly the same text and musical selections, but students in this condition read the text in small portions while being required to fill in missing words in the text presented. No time constraints were placed on participants. The essential difference between the conditions was 1) movement with the instruction content without restriction, and 2) advancement through the program being dependent upon correct responses to the text material (which included discriminative responding to accompanying musical examples). A statistically significant difference between pretest and posttest was found in both experimental conditions, but a difference in posttest scores between the two conditions was not found. Implications of the study and suggestions for future research were discussed.

List of Tables

List of Figures

Chapter One

Introduction

Education is probably one of the most hotly debated topics in our time. Most people have an opinion about education; how it should or should not be, and how it should be improved. Sadly, it could be argued that modern education systems do not seem to advance and improve in proportion to the energy spent discussing them. B. F. Skinner argued that modern school systems were inherently punitive and restrictive, frustrating fast and slow learners alike by shaping all students into the same mold of mediocrity (Skinner, 1984; 1968). In Skinner's view, educational development and reform is constantly misguided; money for educational institutions is too often allocated to things of trivial academic consequence (such as buildings and gadgetry), instead of to the heart of instruction – the scientific analysis of the contingencies that shape academic behavior (Skinner, 1984). Later authors have strongly echoed Skinner's sentiment; for example, fingers have been pointed the at educational establishments' purported rigidity against adopting scientific, empirically proven, and effective teaching methods (Engelmann, 1991; Lindsley, 1992). Still, despite the apparent lack of interest from educational establishments, a handful of effective scientific and empirical teaching practices have survived, one of them programmed instruction.

The origins of programmed instruction date back to Sidney Pressey's work in the 1920s (e. g. Pressey, 1927), but B. F. Skinner's work in the 1950s (e. g. Skinner, 1954) is generally recognized as the breakthrough in the field (Holland, 1960). According to Skinner, programming means carefully arranging the relevant contingencies leading to the terminal performances, or outcomes, of education (Skinner, 1963). In programmed instruction, the learner goes through a series of pre-arranged materials presented in short units, and is required to respond overtly to each one to continue.

Holland (1960) elaborated on several important principles of programmed instruction: 1) Reinforcement is provided immediately after the student emits the correct response. 2) The student learns only when he or she emits a response that is reinforced. 3) Complex repertoires are built by starting to reinforce behavior already in the learner's repertoire, and continue reinforcing successive approximations to the terminal repertoire

in a gradual progression of developmentally ordered steps. At first, the program provides ample stimulus support, prompts or primes that give explicit or subtle hits about the answer. 4) Stimulus support is gradually withdrawn as the student becomes more proficient. This is a procedure known as fading. 5) Since the student needs to constantly interact with the program, his or her study behavior is constantly observed and controlled. Therefore, a minimal part of the learning process is left to chance. 6) Programmed instructions establishes discriminations, such as the difference between simple graphic forms, or complex abstract concepts. 7) The student writes the program. If the student responds incorrectly to a question in the program, or progresses too rapidly through it, the program has failed to accurately assess the student's initial behavior. Revisions to the program must be made contingent on the student's answers, and thus it is said that the student writes it.

Programmed instruction is based on the principles of learning, and is thus an effective teaching method by design. Unfortunately, programmed instruction never started the revolution of which B. F. Skinner dreamed, it never even became popular. Some authors even claim that programmed instruction outright failed and died (e.g. Benjamin, 1988). However, times may be changing, thanks in part to technological advances. Publications on programmed instruction got more prominent again in the late 1980s, when the personal computer – a most convenient delivery machine for programmed instruction (Thomas & Bostow, 1991) – started to emerge as a household item. Another spurt in publications came in the late 1990s, probably not least due to educators' apparent increased interest in using the Internet and computers for educational purposes. As this is written, most hardware problems have been solved, but a programmed instruction revival in educational circles still requires more wide-spread understanding of the approach's principles (Bostow, Kritch, & Tompkins, 1995). All things considered, a programmed instruction revival may be more imminent than previously thought. With more and more homes becoming connected to the Internet, interactivity has become a buzzword; the emphasis on interactivity, and public acceptance of the Internet may have opened the door for programmed instruction's second coming.

According to recent research, interactivity (i. e. responding overtly to instructional material, and constructing responses to questions) increases the effectiveness of instruction (Kritch & Bostow, 1998; Miller & Malott, 1997; Kritch, Bostow, & Dedrick, 1995; Tudor, 1995, Tudor & Bostow, 1991). In addition, Miller & Malott (1997) showed that programmed instruction materials are more effective than read-only materials, regardless of whether or not extraneous incentives such as extra credit points are offered. Early and recent research also demonstrates that programmed instruction increases in effectiveness as reinforcement for correct answers becomes more frequent (Terrace, 1966; Kritch & Bostow, 1998). The effectiveness of computer-programmed instruction has already been well-established for text-based stimuli (see e.g. Kritch & Bostow, 1993: Terrace, 1966; Tudor & Bostow, 1991; Kumar, Bostow, Schapira, & Kritch, 1993), but controversy still exists. Kritch & Bostow (1998) found that several prominent investigators do not agree on the definition of programmed instruction. Also, programs designed for commercial use tend to compromise educational properties for the sake of sales-inducing gimmicks (Kritch & Bostow, 1998). In addition, generality issues still exist. For example, Kritch & Bostow (1998) established proficiency in verbal discrimination, but noted that the same approach may not apply to sensory discrimination – an area that certainly merits research, but has seemingly been left untouched by researchers in programmed instruction so far.

This study was a modified replication of Kritch & Bostow's 1998 study: programmed instruction, using general principles of programmed instruction, was compared to non-programmed instruction. Instead of responding primarily to verbal stimuli, as in Kritch & Bostow (1998), participants were required to respond primarily to auditory stimuli. This was proposed to address generality issues with sensory discrimination (Kritch & Bostow, 1998). The subject matter was classical music history.

Music taste is probably acquired early on in life. As children grow up, they are exposed to music on the radio and television, music to which responses are likely to be reinforced differentially, either socially, intraverbally, or via respondent conditioning. However, in our times' consumer culture, not all music is equal. Popular music dominates the mainstream airwaves, and airplay of classical music is mostly left to

publicly funded radio stations with a small share of the market. Worldwide, classical music is only estimated to be 5% of all music sold, and a decline in sales has forced major record labels to downsize their classical divisions (Bambarger, 2002). The pop music industry is well known for using other means, such as sex appeal, to spur sales. To avoid losing their jobs, some classical musicians have resorted to similar tactics, changing their traditional formal appearance to skimpy outfits, and even adding pop music instruments and effects to classical pieces (Clennell, 2004).

It follows from the basic principles of behavior that people cannot acquire behavior if they have no opportunity to be reinforced for responding to it. Classical music is no exception. Public radio is almost the only kind of radio left to play classic works and has less marketing funds than commercial radio, as is evident by the fundraising drives characteristic of public radio. Adding to classical music's problems is the music itself, more complicated and difficult to the untrained ear than popular music which often is especially designed to be simple and immediately reinforcing. Another factor may be that classical music is not present where youngsters come together for social functions, and thus, classical music misses an opportunity to be reinforced via respondent conditioning. Long reinforcement histories with pop music are likely to change easily.

However, if people unfamiliar with classical music got easy access to it, were to be told about it in terms they know, and be reinforced differentially for responding to the music, people might eventually start valuing it. A web-based instructional program featuring short instructional text and clips of music might be a way to access those potential listeners, and two methods of doing so will be compared in the proposed study. To teach about classical music, it was judged helpful to use the common grouping of musical periods: Early music (the middle ages and the renaissance), the baroque era, the classical era, the Romantic period, and modern music (20th and 21st Century). The first condition featured a web page with short instructional paragraphs about prominent characteristics of each period. Each paragraph was accompanied with a relevant music clip, accessible by clicking on a link by the paragraph. No other reaction to the material was required, and participants could move freely around the web page at their own

discretion. The second condition was programmed instruction, in which participants were taken through the material in a linear, temporal order, from early to modern music.

Despite classical music's problems, there may be grounds for optimism if something is done. The current authors find it important that the science of behavior and the educational establishment join forces to teach about classical cultural treasures that would otherwise go unnoticed.

Chapter Two

Method

Participants

One hundred and sixteen undergraduate and graduate education students in various classes were randomized to experimental conditions with a computer program. Participants worked through the experimental music tutorials for academic course credit.

Setting

Previous research has revealed the importance of directly observing students as they work through computer delivered materials, this being necessary to ensure treatment integrity (R. Canton, personal communication, February, 2004). Therefore, the experiment was conducted in a computer lab at the university. Participants registered for appointments at a USF computer lab, during which the experimenter administered instructions, maintained appropriate lab conditions, and ensured completion of pretest, tutorials and posttest.

Institutional Review Board Approval

All procedures were approved by the Institutional Review Board prior to data collection. Participants were informed of the nature of the study at an orientation session for a program course. Taking part in the study was a course requirement, but all participants signed an informed consent form for the course, which included informed consent for taking part in the study (see Appendix C).

Materials

The computer lab contained internet-connected PC computers with headphones. Six computer-delivered programmed instruction tutorials of approximately 30-40 frames each were used to teach auditory discrimination between examples of music from five periods: 1) The Middle Ages and the Renaissance, 2) The Baroque era, 3) The Classical era, 4) The Romantic period, and 5) the $20^{th}/21^{st}$ Century. The tutorials were designed to progressively develop skills by constantly requiring overt responding, but did not explicitly follow the programming guidelines of the Ruleg system designed by Evans, Homme, and Glaser (1960). The tutorials were field tested and revised based on an earlier trial round with volunteer participants.

The tutorials and tests were delivered with software specially developed for this experiment, hosted on a USF server. The music samples for the tutorials were selected using properties widely-accepted and argued by musicologists to exemplify and distinguish these periods. A pretest and generalization test were composed from the same clip library as the tutorials. The test contained 30 examples of music that were not part of the tutorial program, yet were assumed to be composed of the defining musical properties taught within the tutorials. No time limits were imposed on participants when taking the tutorials and tests.

An eight-question survey was written to follow the experiment. The purpose of the survey was to give participants an outlet for feedback, and to assess participants' prior experience with classical music.

Treatments (Independent variables)

Programmed Instruction (PI): Five tutorials on music history were used as an independent variable. Music clips and textual instructional material was presented concurrently, frame by frame. In order to advance through the tutorials, participants typed in answers to questions in each frame. Textual feedback, "Correct!" or "Incorrect", was provided contingent on the submission of a typed answer. Two tries were allowed for open-ended questions, but only one try for multiple-choice or yes/no questions, as two tries for yes/no questions would have heightened the probability of haphazard guessing. No minimum score was necessary for advancement within or between tutorials, nor were there be any time contstraints to finish the tutorials.

Non-programmed Instruction: Instructional text identical to the one used in the programmed instruction tutorials was presented on a single web page, but all blanks and questions were modified to look like coherent instructional paragraphs. Participants were able to access the material in its entirety, in any order, by scrolling up and down the web page, non-contingent on discriminative responses. When musical clips accompanied a paragraph, they were available as hyperlinks by the respective paragraph. No time constraints were placed on participants.

Dependent variables

Time to complete all tutorials in the sequence, time to complete each tutorial, and to complete each individual frame, were all measured during the programmed instruction condition. During the PI treatment, the percentage of PI frames answered correctly was also measured. For the scrolling web page treatment, individual participant time going through the whole program was recorded. The number of clicks on each music link by each participant was also counted.

A 30-item multiple-choice pretest was constructed, which included five examples from each musical period. This test was in the form of a series of musical clips with a selection of appropriate choices on the screen. Clips were different selections from the tutorials, yet contained key properties of the various musical periods. Thus, a "generalization" posttest of musical clips was also employed, to assess the amount of learning during the tutorials in this study. The generalization test was the same as the pretest used, but practice effects from the test were considered minimal, since the testing items were not included in the tutorials. Demographic data concerning undergraduate/graduate status, age, gender, etc., were collected at the commencement of the course sections.

A survey was written to follow the experiment. On the survey, participants were asked to provide general feedback about the experiment, and provide information about potential hearing impairments that could have affected their learning. Participants were also asked about their prior experience with classical music, and about the probability of them seeking to purchase or listen to classical music in the wake of this experiment.

Experimental design and data analysis

As described previously, there were two experimental conditions: 1) Programmed Instruction (PI), and 2) Scrolling web page with short instructional paragraphs. A one-way analysis of variance (ANOVA) was used to check for a statistically significant difference between the conditions, and a dependent-sample t-test was used as a test of significance between pretesting and posttesting within conditions. A single-subject design was not deemed appropriate, nor even possible for treatment comparisons. This is due primarily to a changing substrate of discriminations required as the program advanced

8

and the potential confusion of effects produced by changing content rather than those produced by tutorial delivery differences.

Procedure

Requirements for participation were stated in course syllabi made available at the beginning of the course. Participants were required to sign up for their visit to a computer lab where the experiment was conducted, and a message with available times was posted on the computer-based bulletin board of the course. The experimenter supervised the computer lab at all times. Participants were led to computers as they arrived and were given a detailed instruction sheet (see Appendix A). In addition to the instructions on the sheet, the experimenter read the following message to participants: "When you have earned a score on the pretest, used the tutorial to which you were assigned, and earned a score on the posttest, your final course average will be boosted by 3 additional points in return for your participation. The score you earn on the posttest will NOT impact your course grade." Students who refused to attend or could not attend did not lose academic credit, and were not subjected to other tasks.

Participants accessed a web page per their instructions and logged on with the login names they used for their web-based course management system. A computer program automatically randomized participants to the experimental conditions. Once participants had logged on, a computer program automatically displayed a multiple-choice pretest in which they listed to music clips with headphones, and attempted to identify the period from which the music was dated. Upon completion of the pretest, grades were displayed to participants, and a computer program automatically took them to a screen with instructions on their assignment. Once participants pressed a key to continue, the program presented a screen with instructions on what to do next, and after pressing a button, the program opened the respective experimental condition. In the PI condition, participants first went through a brief introductory tutorial that taught how to go through the programmed instruction tutorials. Those who received the prose condition first saw an instruction screen and were taken directly to their experimental condition thereafter, as the only skills needed were scrolling up and down a web page and clicking hyperlinks.

In the PI condition, music clips and text were presented automatically, screen frame by screen frame, with a keyword missing within presented text. Each clip was around 30 seconds long, and the program constantly repeated it while the student was working on the respective frame. To advance from one frame to the next, participants were required to respond discriminately to the music by typing an answer in a text box on the web page. Answer modes were two: completing a statement about the music, or typing in letters corresponding to the correct answer in multiple-choice questions. Two tries were allowed for open-ended questions, and one try for double- or multiple-choice questions. Once participants had answered, they pressed the Enter key, and were taken to a screen where they received feedback. The feedback message for correct answers read: "Your answer [answer here] is correct. Press Enter to continue." The feedback for incorrect answers read: "Your answer [answer here] is incorrect. Press Enter to continue." The music continued playing through the feedback screen. After pressing Enter, participants were either taken to the same frame again for a second try, or the next question and music clip. If a participant would answer all allowed attempts incorrectly, the program would give the feedback for an incorrect answer and provide the correct answer, move on to the next frame, and lower the score.

Progression through the program was constantly visible to participants in the upper left corner of the frame window (e.g. "Frame 16 of 30"), and so was their current score in the program (e.g. "Score: 94%"). The number of tries for a given frame was also displayed (e.g. "Try: 1 of 2"). No time constraints were placed on participants. When the participants had finished all five tutorials, the program displayed an instruction screen to prepare them for the posttest to come. Pressing a button started the test.

In the prose web page condition, all the material was presented on one scrolling web page. The same instructional text and music as in the PI condition were used, but all blanks had been filled in, and multiple-choice items had been changed into normal paragraphs. Advancement was completely at participants' discretion and not contingent on any overt discriminative response. To hear the music sample relevant to a paragraph, participants clicked on a hyperlink by the paragraph. The clips were not on automatic repeat, but participants could click on each clip as often as they pleased. Clicking a

button reading "Done" would automatically open an instruction screen. After pressing another button, participants could start their posttest. The generalization test was the same across conditions and similar in form to the PI condition, but feedback was not presented after each frame. Upon completion of each condition, the delivery software presented a debriefing page, on which the experiment's details were displayed. After that, participants filled out an eight-question survey on paper to check for hearing impairments, former exposure to classical music, and opinions of the study. Participants had an option to answer the survey anonymously (see Appendix B).

Chapter Three

Results

Pretest data

The total sample was 116 participants, but 109 students showed up for the study. Data from four participants had to be removed due to unanticipated participant errors. Such errors included accidentally closing the web browser window and starting again, thereby distorting the data in the computerized database. The remaining sample was a total of 105 participants--with 46 receiving the prose treatment and 59 receiving programmed instruction.

On the pretest, the prose condition scores ranged from 6% - 55% with a standard deviation of 8.86. Learners who received programmed instruction scored from 16% - 58%, with a standard deviation of 8.68. The distributions of the scores are very similar in shape, but the main difference lies in the lower range of the grade scale (see figures 1 and 2).
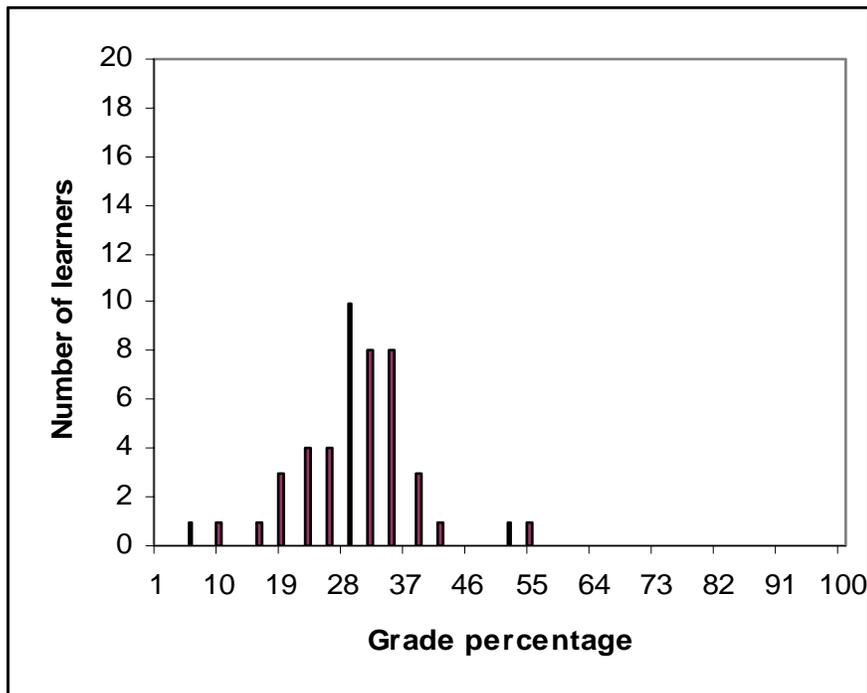

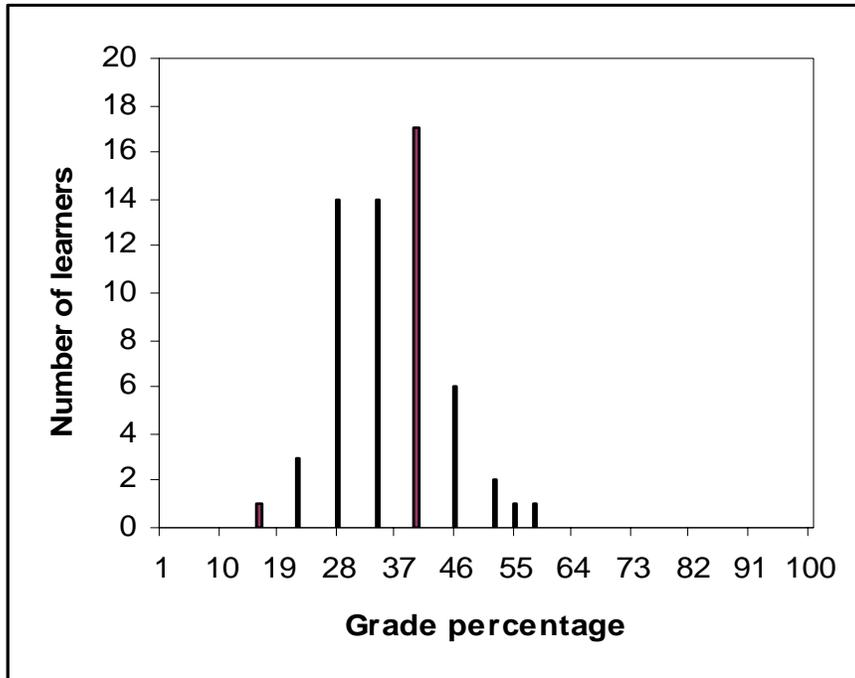**Figure 1.** Pretest score distribution, prose condition

Figure 2. Pretest score distribution, PI treatment

Analysis of tutorials

     Characteristics of programmed instruction tutorials were measured individually, but scoring and measurement of each particular segment of prose material was not possible on the prose page due to the nature of the prose page's free-form design. However, time taken browsing the prose web page was measured. The PI participants took more time, on the average, to finish their material (32 minutes, standard deviation 13.59) than did their counterparts with prose (29 minutes, standard deviation 11.55). This difference proved to be statistically significant, $t(58) = 18.03$, $p= .00$. See Table 1 for a listing of PI tutorial averages. Difference in median posttest times between conditions were not remarkable (Table 2).

Table 1. Averages for individual programmed instruction tutorials

	Renaissance	Baroque	Classical	Romantic	Modern	All Tutorials
Mean score (%)	88	89	90	82	89	88
Mean tutorial time (min)	5:12	6:73	6:32	8:17	5:37	6:33
Mean frame time (sec)	16:69	13:01	12:24	14:03	11:64	13:34

13

Number of frames	17	26	27	32	25	25.4

Table 2. Median times taking the pretest and posttests, across treatments

Treatment	Pretest (minutes)	Posttest (minutes)
Prose	6	7
PI	6	6

Pretest-posttest comparisons

The difference between pretest and posttest scores was analyzed with paired-sample t-tests. The difference between pretest and posttest scores proved statistically significant for the prose condition $t(45)= 6.11$, $p=.00$, as well as for the PI condition, $t(58) = 6.44$, $p=.00$. However, when analyzing posttest score distributions for a possible difference between the prose and programmed instruction methods, a one-way ANOVA did not indicate an effect, $F(1,104)= .108$, $p= .743$. A breakdown of pretest and posttest scores is listed in Table 3.

Table 3. Prose and PI pretest and posttest averages

Condition	Pretest score (%)	Posttest score (%)
Prose	29.9	39.7
PI	32.8	39.0
Total	31.6	39.3

The prose group's scores on the posttest had a wide distribution (see figure 3), ranging from 16% -77%, with a standard deviation of 12.46. However, students who received programmed instruction received a considerably narrower range of scores (see figure 4), from 19%-65%, standard deviation 9.59. The posttest's Kuder-Richardson reliability score was .66.

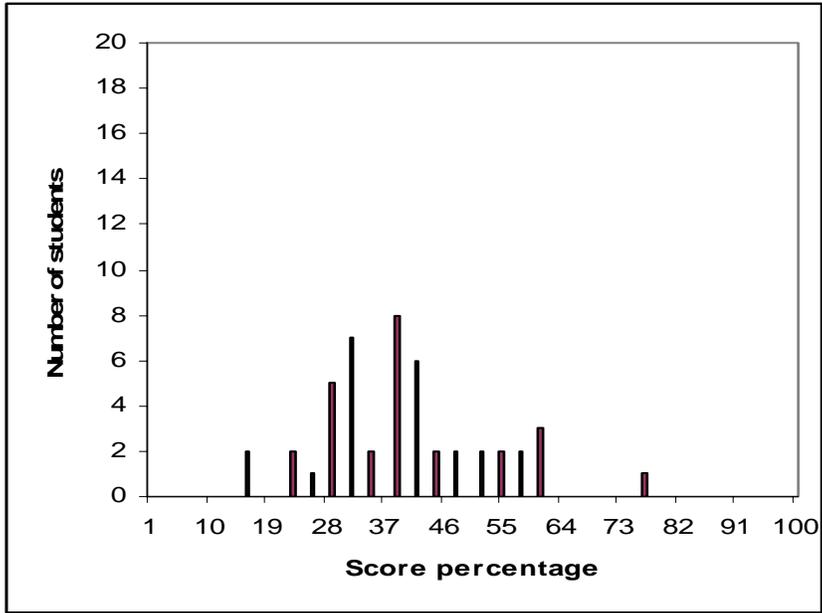**Figure 3.** Posttest score distribution, prose treatment
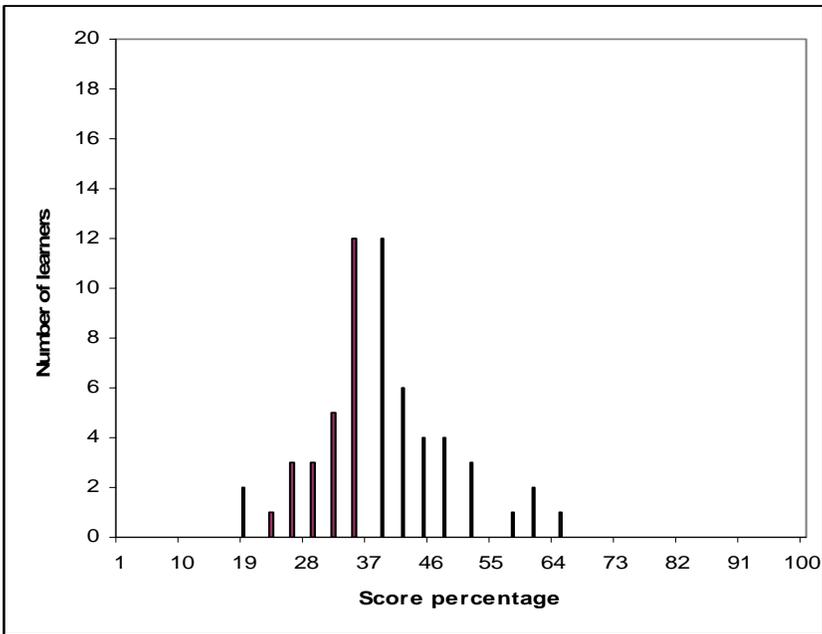

Figure 4. Posttest score distribution, PI treatment

Post-experiment survey

After the experiment, an eight-question survey was administered to assess participants' opinions of their respective learning experiences, their music tastes and classical music listening habits, and to gauge for potential hearing impairments that might

have compromised results. One question asked for outcome on the posttest, and since those results are described in more detail above, they will not be recounted in the present subsection. From 105 participants total, 99 surveys were retrievable. An analysis of the survey follows:

Q. 1: Have you got any major hearing impediments that you feel could have compromised your test results?

Hearing problems were reported in only three cases (see Table 4), but anonymity made filtering of the relevant participants' data impossible.

Table 4. Participants' reports of potential hearing problems

	Frequency	Percent
Yes	3	3.0
No	96	97
Total	99	100

Q. 2: What kind of instructional strategy did you receive?

Of the 99 participants who completed a questionnaire, 62 participants answered that they had received the programmed instruction treatment, and 36 marked that they had received the prose condition. The self-reporting is inaccurate and does not match the experiment sample, in which 59 students received programmed instruction, and 46 received the prose method. One participant answered as having both, so her answer was categorized as missing. Due to the discrepancy between these self-reported data and the computer-collected data, this item is deemed unusable for calculations of statistical significance.

Q. 3: In your opinion, the instructional strategy you got was...[effective vs. ineffective]

Regardless of experimental condition, 52 students judged their instructional method to be either very or rather effective, as opposed to 20 who thought it rather or very ineffective. The answers are broken down in Figure 5.
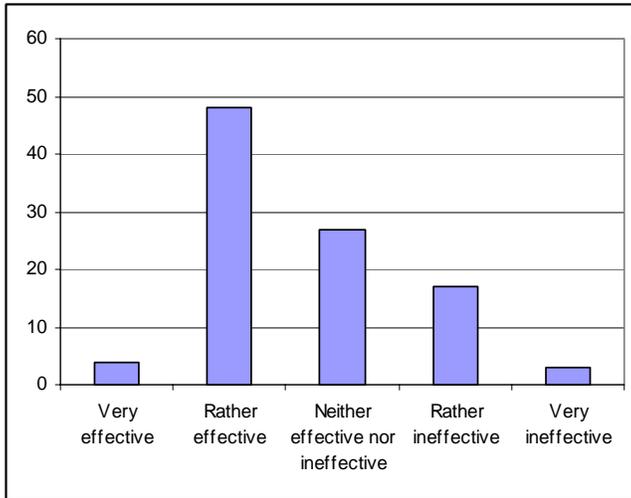
Figure 5. Participants' evaluations of their treatment's effectiveness

Q. 4: How interesting or boring was the lesson?

A majority, a total of 70 learners, found their treatment either interesting or very interesting. Twelve described it as boring or very boring (see Figure 6).
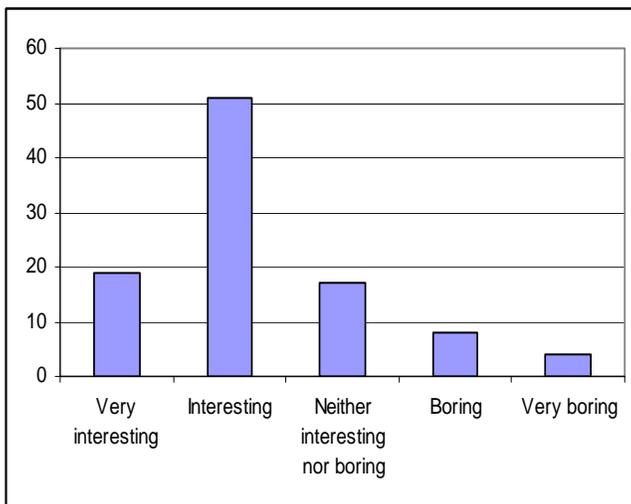

Figure 6. Participants' reports of how interesting or boring their lesson was

Q. 5: How often or seldom do you usually listen to classical music?

Most participants reported not listening often to classical music, only 10 claiming they listened to it often or every day. A third of participants reported listening seldom or never to classical music (Figure 7).
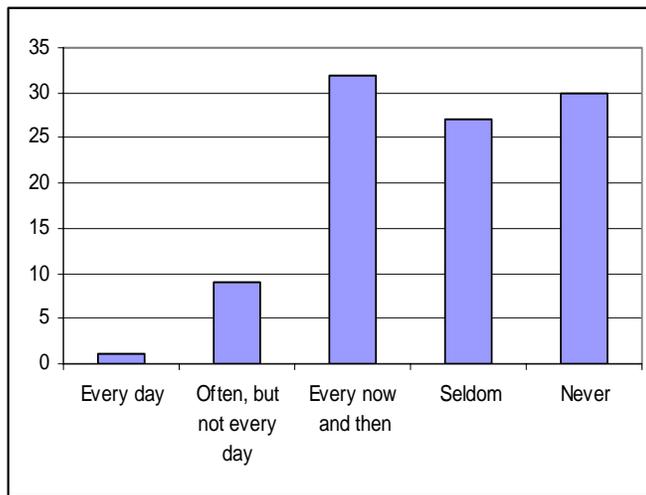
Figure 7. Classical music listening habits of participants

Q: 6. *How likely or unlikely is this lesson to increase your taste for classical music, if only slightly (e. g. tune in more often to classical stations, check out the classical selection at your local music store, look for it on the internet, etc.)?*

A total of 55 participants reported that they would be more likely than before to listen to classical music, as a function of this study. Twenty-six participants claimed that taking part in the study would not influence them to listen more to classical music (Figure 8.).
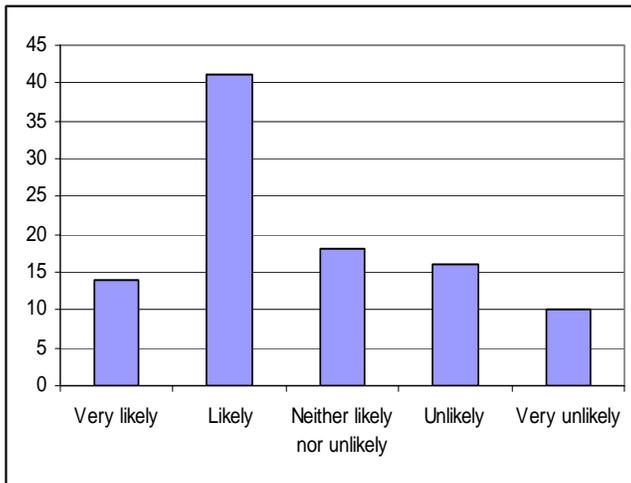
Figure 8. Participants' estimate of lesson to increase taste for classical music

Q. 7. Is there anything you would like to add?

This was an open-ended question to get feedback from students on the experiment. Fifty-six of the 99 students who turned in the survey answered the question. A content analysis revealed the following most prominent issues:

- Both the programmed instruction and prose tutorials were interesting and fun.

- There was too much material for one session.

- The posttest was difficult considering what students thought they had learned.

Chapter Four

Discussion

This study extended on the literature on programmed instruction by experimentally evaluating two approaches to auditory instructional materials. This was a departure from the line of research on textual and visual material that has dominated the field. Programmed instruction, in which participants advanced frame by frame, was compared with a "free form" prose web page that allowed the learners to browse all of the material at once. Along with grades on a posttest, posttest times were measured.

When compared to other instructional approaches, programmed instruction generally yields better results than the competing approach. However, previous research has mostly been on verbal and/or visual material and no research with the current study's particular focus is believed to exist. Therefore, it can only lay claim to be a pilot study. In the current study, a difference was not found between experimental conditions in posttest scores, nor in the time taken completing the posttest. Participants in the programmed instruction condition took three minutes longer than their prose condition counterparts to complete the tutorials, a difference that proved statistically significant, but bears trivial clinical significance, given that the groups performed evenly on the posttest.

The pretest/posttest had a modest Kuder-Richardson reliability score of .66, so inferences about instructional effects of the tutorials are presented with appropriate caution. Items on the test and items on tutorials were drawn from the same item pool, and the tutorials were pretested and edited after a trial run. However, pretesting and statistical analyses of individual pretest/posttest items were not conducted before the experiment, which may compromise the validity of the test. Such validation is recommended for future research.

PI learners averaged a score of 88% on the tutorials, a marked difference from their average posttest outcome of 39%. This discrepancy, along with the modest reliability of the test, presents a challenge to the tutorials' and/or the test's internal and external validity. There are several possible explanations. The first is that the tutorials may have been too easy or not adequate in scope. The creator of the PI tutorials was not an expert in the field, and further isolation of music characteristics may have been needed to pin

down the essential stimulus differences that define the periods of classical music. It is also possible that there were too many echoic frames that made guessing easy.

For replications of the current study, it is recommended that conceptual properties defining the musical periods be isolated in advance before revising the tutorials. Music samples on the posttest may also have been too subtle for discrimination—given the preparation resulting from the tutorials. Music styles tend to transcend time periods, and it is possible that the test was simply too difficult. The material, music history from medieval to modern times, may have been too extensive.

In addition, many aspects of possible participant behavior during the prose tutorial condition were uncontrolled variables--due to the nature of their instructional program. Although the prose condition software was recorded how much time prose participants took, it was impossible to monitor them adequately, or to estimate the differential effect of the verbal stimulation received by their counterparts in the PI group (I.e. the resulting feedback to each PI frame).

On the post-experiment surveys, participants frequently commented that there was too much material for one session, an arguably respectable claim, given the scope of the subject matter. For future study, a sensible approach may be to combine the more similar musical periods initially and then gradually create more subtle discriminations through careful programming. This strategy would arguably make initial discrimination easier. For example, instead of using the common five periods, 1) early music/renaissance, 2) baroque, 3) classical 4) romantic, and 5) modern, it would be possible to group the periods into three: 1) early music/renaissance/baroque, 2) classical/romantic, and 3) modern. For this purpose, it would be necessary to employ a professional musicologist's help to pick the quintessentially characteristic music for each period, and analyze its properties. Those samples would be used as prime examples of a given period, and gradual fading would be based on subtle deviances from their characteristics. The professional's help would also serve as a validity measure possibly lacking in the current study.

Another explanation for the discrepancy between tutorial scores and test scores may the issue of music characteristics vs. chronology. Classification into the five periods was

accomplished by using dates of the compositions rather than musical stimulus properties. In other words, the tutorials focused upon teaching characteristics of music by dates not stimulus dimensions. This emphasis on chronology may have made complicated the music discriminations. Oftentimes, discrimination between musical periods is difficult even for professionals. Even though laypeople may possibly tend to group musical periods by chronology, professionals may be more likely to group them by characteristics of the music.

A professional classical music director at a public radio station went through the posttest as a post-hoc validation measure and got all but two questions "correct" as defined by the creator of the posttest. However, if categorization by stimulus dimensions rather than date of composition had been used, the music director's score would have been 100%. (For example, a fugue is best known as a baroque musical form, but the composer Ludwig van Beethoven is best known as a classical period composer. When taking the test, the professional listened to a fugue by Beethoven and guessed it was baroque, but the experimenter had coded the piece as classical to be congruent with the composer's time period.) Similar issues evolve when comparing any periods with potential overlap: late baroque/early classical, late classical/early romantic, and so on. The third complication with the chronology approach is the definition of modern music. Modern, or 20th century, composers tend to write music that could be classified under any period, especially the romantic period. This makes it very difficult to tell modern from older music.

A related issue is the test design. Thirty-second music clips often do not adequately display characteristics of given musical periods, such as structure or lack of it, changes in dynamic range, orchestration, and so on. This can partly be solved by using several clips from different sections in a given piece, but the technique still calls for accurate working definitions and rules that describe characteristics of the music.

In general, participants liked the materials, and several reported that they would be more likely to seek out "classical" music in the future, as a function of taking part in the study. For follow-up in future research, collaboration with a music store might be established. This could be accomplished by giving participants discount coupons for

22

classical music CD purchases at the store. The number of coupons used would give an indirect measure of the strength of the new repertoire. Another possibility would be to give discounts to any type of music CD, and check how many classical vs. pop CDs are purchased. This would indirectly measure a new repertoire against a more established one.

The results of this pilot study clearly point to better ways of teaching music discrimination. It is quite possible that behavior analysis is a field uniquely prepared to face the challenge. It has the advantage of defining the dimensions of variables in the language of physics rather than metaphor (or chronology, in this case). In addition, behavior analysis focuses upon objective physical dimensions of stimuli in the effort to bring the verbal behavior of identification (or nomination) under the control of abstract qualities of stimulus configurations. The isolation of the stimulus properties that lead to the correct identification of musical periods (or for that matter ANY concept) should be greatly advanced by building upon the behavior analytical discrimination research literature and its conceptual analysis. The next step in research will require the isolation of stimulus dimensions that are the physical referents of the various periods of classical music.

References

Bambarger, B. (2002). Classical CD industry offers cautious optimism as sales plunge in 2001. *Andante magazine [web-based article]*:
http://www.andante.com/article/article.cfm?id=15643

Benjamin, L. T. Jr. (1988). A history of teaching machines. *American Psychologist, 43*, 703-712.

Bostow, D. E., Kritch, K. M., & Tompkins, B. F. (1995). Computers and pedagogy: Replacing telling with interactive computer-programmed instruction. *Behavior Research Methods, Instruments, & Computers, 27*, 297-300.

Clennell, A. (2004). Classical music sales resurgent thanks to 'easy listening brigade'. *The Independent [Web-based article]*:
http://enjoyment.independent.co.uk/music/news/story.jsp?story=499015

Engelmann, S. (1991). Change schools through revolution, not evolution. *Journal of Behavior Education, 1,* 295-304.

Evans, J. L., Homme, L. E., Glaser, R. (1960). *The ruleg system for the construction of programmed verbal learning sequences* (Report No. CRP-691-3). University of Pittsburgh: Cooperative Research Project, United States Office of Education.

Holland, J. G. (1960). Teaching machines: An application of principles from the laboratory. *Journal of the Experimental Analysis of Behavior, 3*, 275-287.

Kritch, K. M., & Bostow, D. E. (1993). Verbal responses to past events: Intraverbal relations, or tacts to private events? *The Analysis of Verbal Behavior, 11*, 1-7.

Kritch, K. M., & Bostow, D. E. (1998). Degree of constructed-response interaction in computer-based programmed instruction. *Journal of Applied Behavior Analysis, 31*, 387-398.

Kritch, K. M., Bostow, D. E., & Dedrick, R. F. (1995). Level of interactivity of videodisc instruction on college student recall of AIDS information. *Journal of Applied Behavior Analysis, 28*, 85-86.

Kumar, N. B., Bostow, D. E., Schapira, D. V., & Kritch, K. M. (1993). Efficacy of interactive, automated instruction in nutrition education for cancer prevention. *Journal of Cancer Education, 8*, 203-211.

Miller, M. L., & Malott, R. W. (1997). The importance of overt responding in programmed instruction even with added incentives for learning. *Journal of Behavioral Education, 7*, 497-503.

Lindsley, O. R. (1992). Why aren't effective teaching tools widely adopted? *Journal of Applied Behavior Analysis, 25,* 21-26.

Pressey, S. L. (1926). A simple apparatus which gives tests and scores-and teaches. *School and Society, 23*, 373-376.

Skinner, B. F. (1963). Reflections on a decade of teaching machines. *Teachers College Record, 65*, 947-954.

Skinner, B. F. (1968). *The Technology of Teaching*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1984). The shame of American education. *American Psychologist, 39*, 947-954.

Terrace, H. S. (1966). Stimulus control. In Honig, W. K. (Ed.), *Operant Behavior. Areas of Research and Application*. New York: Appleton-Century-Crofts.

Thomas, D. L., & Bostow, D. E. (1991). Evaluation of pre-therapy computer-interactive instruction. *Journal of Computer-Based Instruction, 18*, 66-70.

Tudor, R. M. (1995). Isolating the effects of active responding in computer-based instruction. *Journal of Applied Behavior Analysis, 28*, 343-344.

Tudor, R. M., & Bostow, D. E. (1991). Computer-programmed instruction: The relation of required interaction to practical application. *Journal of Applied Behavior Analysis, 24*, 361-368.

Vargas, E. A., & Vargas, J. S. (1991). Programmed instruction: What it is and how to do it. *Journal of Behavioral Education, 1*, 235-251.

Appendices

Appendix A: Experimenter instructions to participants

Experimenter instructions to participants:

Welcome. Read these instructions very carefully. If you have any questions after reading this, please save them until you've finished reading this.

Since we are teaching a course on learning and the application of learning principles, we need to practice what we preach and experimentally test our potential instructional programs experimentally. You are taking part in our evaluation of teaching methods, as a part of this course on learning. Your personal information and everything you submit will be held stricty confidential, as if it were your final exam grades. This evaluation has received IRB approval for protection of personal information.

You will be doing three activities today: A pre-test, an instructional program, and a posttest. When you have earned a score on the pretest, gone through the instructional program, and earned a score on the posttest, your final course average will be boosted by 3 additional points in return for your participation. The score you earn on the posttest today will NOT impact your course grade.

1) Plug your headphones into the green jack on the rear on the computer.

2) Open Internet Explorer and enter the URL address:

http://sirocco.coedu.usf.edu/mcohen/gummi/music

3) Write your WebCT login name in the text box.

4) Follow the on-screen instructions.

Appendix B: Student survey of music lesson

Student survey of music lesson

This is your opportunity to tell us what you thought of the lesson you just had. Your input is *extremely* important. For multiple choice questions, just check one. You may answer anonymously if you want.

Your name (optional) :
_____

Your course (optional) :
_____

1. What kind of instructional strategy did you receive?

[ ] A programmed tutorial that played music. I filled in blanks to answer.
[ ] A web page with many short paragraphs. I scrolled up and down the page and clicked on links to hear music examples.

2. What grade did you receive on the last test for this lesson? _____.

3. In your opinion, the instructional strategy you got was...

[ ] Very effective
[ ] Rather effective
[ ] Neither effective nor ineffective
[ ] Rather ineffective
[ ] Very ineffective

4. How interesting or boring was the lesson?

[ ] Very interesting
[ ] Interesting
[ ] Neither interesting nor boring
[ ] Boring
[ ] Very boring

5. How often or seldom do you usually listen to classical music?

[ ] Every day
[ ] Often, but not every day
[ ] Every now and then
[ ] Seldom
[ ] Never

6. How likely or unlikely is this lesson to increase your taste for classical music, if only slightly (e. g. tune more often in to classical stations, check out the classical selection at your local music store, look for it on the internet, etc.)?

[ ] Very likely
[ ] Likely
[ ] Neither likely nor unlikely
[ ] Unlikely
[ ] Very unlikely

7. Have you got any major hearing impediments that you feel could have compromised your test results?
[ ] Yes
[ ] No

8. Is there anything you would like to add (you can also write on the back of this sheet)?

_____
_____
_____
_____
_____
_____
_____

Appendix C: Demographic Form/Informed Consent Form

Demographic Form for Dr. Bostow's courses (Summer Term, 2004)

Circle which course: EDF 3214 (Human Development and Learning)
EDF 3228 (Human Behavior and Environmental Selection)
EDF 6211 (Psych. Foundations of Education)
EDF 6215 (Principles of Learning)
EDF 6217 (Behavior Theory and Classroom Learning)
EDF 7227 (Doctoral Course in Behavioral Analysis)

Your name: _(last)_____(first)_____Gender: (M/F) _____

Your Email address: [print clearly]: _____

Social Security No. _____ Age: _____

Your mailing address: (While at USF)   Home phone: _____
_____
_____ Work phone: _____
City: Zip

    Various instructional techniques may be tried in this course. The following information will be used to determine whether they correlate in any way with your success in this course. If you feel it is an unwarranted intrusion into your life, simply skip that question. However, we request that you answer all possible questions that could relate to your success in this course. The relationships, if any, will be used to adjust teaching techniques and counsel future students.

How many course hours are you taking this term?_____
What year are you (fresh., soph., jr., sr., other)? _____
What is you major or specialization? _____
What is your current GPA at USF (If transferred, list last GPA)? _____
If you have a computer at home, what type is it? (IBM, clone, Apple-Macintosh, other)? _____
Does your computer have a CD rom drive? _____
If you work, in addition to going to school, how many hours are you working per week? _____
Do you commute to campus classes? (Y/N)_____
If so, how many miles to you travel each way? _____
What type of internet connection do you have? I.e., broadband _____ or dial-up connection (speed)?
____

        I have listened carefully about the course EDF _____ in which I have enrolled. I understand that rules can change within the course, and that I will be notified ahead of time—this may include the deadline dates for quizzes throughout the course during various weeks. I understand that Dr. Bostow may run various sections with slightly different course requirements to compare how students perform. Further, I understand that I will participate in the evaluation of new instructional techniques as part of course requirements. I understand that any differences in the effectiveness of these trial techniques will not influence my final grade in the course. I have been given the opportunity to ask questions about the course operations and possible variations that may occur, and have received satisfactory answers. I understand the course requirements and agree to abide by them.

Student's signature: _____ Date: _____