USING STIMULUS EQUIVALENCE TECHNOLOGY TO TEACH RESEARCH DESIGN CONDITIONAL RELATIONS FOR UNDERGRADUATE STUDENTS

BY

ANA CAROLINA SELLA

Submitted to the graduate degree program in Applied Behavioral Science and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

_____

Chairperson (Glen W. White)

_____

Edward K. Morris

_____

Florence DiGennaro Reed

Date Defended: November 07, 2011

The Thesis Committee for ANA CAROLINA SELLA

certifies that this is the approved version of the following thesis:

USING STIMULUS EQUIVALENCE TECHNOLOGY TO TEACH RESEARCH DESIGN CONDITIONAL RELATIONS FOR UNDERGRADUATE STUDENTS

_____

Chairperson (Glen W. White)

Date approved: November 07, 2011

Abstract

The purpose of this study was to evaluate the effects of a stimulus equivalence instructional package on undergraduates' performance in conditional discrimination tasks that involved research design names, definitions, notations, and examples. Participants were four undergraduate students whose primary language was Portuguese. Participants remained in the study only if their percentage of correct responses in Probes 1, 2, and 3 was lower than 20%. Thirty-six experimental stimuli were used in the study. They were comprised of nine research design names, nine research design definitions, nine research design notations, and nine examples presented in a matching to sample format during teaching and emergent relations sessions. Probes consisted of nine open-ended questions on the taught conditional relations and new examples. All participants learned all conditional relations, showed emergence of symmetric and transitive relations, and generalized from the selection-based tasks (multiple-choice tasks) to the topography-based tasks (open-ended probes). Lessons learned from this study can help in programming effective instruction for higher education settings.

Keywords: stimulus equivalence, higher education, research designs.

Acknowledgements

To Glen W. White who gave me the opportunity to come back to KU and opened up doors so I could work on many different areas.

To Daniela M. Ribeiro who read all the open-ended questions, scored them according to my crazy rubrics, and helped make sense of the social validity questionnaire when considering the Brazilian reality.

To Dot Nary who read this manuscript many times and was sure to give kind feedbacks on clarity and readability and for all her help with the outlines and HSCL forms.

To Melissa Gard who helped me develop all the appendices that are here, and shared the anxiety of trying to finish a degree within a brief period.

To Jason Hirst and Dan Schober who helped me develop the review checklist in the form it is now.

To Shannon Tierney and Cara Smith who gave me feedback on the stimuli content for the Glossary.

To Chiaki Gonda for helping me with the first outlines and the HSCL forms.

To Jan Butain for answering all my endless e-mails with doubts about the HSCL process.

To Dick and Muriel Saunders for bringing me to KU and for supporting me since we first met at ABAI - Brazil in 2004.

To Florence DiGennaro Reed for being available to read documents and answer questions on different items described in this thesis.

To Derek Reed for making helpful suggestions to documents and providing great models on how describe a research design and on how to construct a cool figure.

To Edward K. Morris for giving great recommendations on people I could talk to and on references I could read, and for being available to listen to me for so many times.

To Daniel Bernstein for giving great advice on parts of this procedure and for giving me the chance to volunteer at the Center for Teaching Excellence which provided me enough information to be able to work with Adobe Captivate ® and distance learning.

To the RTC/IL team - especially Amanda Reichard - for being supportive and patient in moments of crises.

To ISSS and Legal Services for advising me on steps to be taken as an international student.

To the Applied Behavioral Science department and the Libraries for helping me navigate through KU's system.

To my family and friends (some of them are included above) who listened to me, helped me, and emotionally supported me through the endless adaptation process that includes trying to get things done through two different governmental, educational -and even health - systems.

To my participants who had a lot to do and took the time to participate in this study.

Table of Content

Table of Figures

Table of Appendices

"Since the programming of any skill or knowledge is a long and difficult process, a programmer will need a persistent commitment to changing student behavior and to gathering empirical evidence that he has done so" (Markle, 1969, p. vi).

Importance of Higher Education

Higher education is seen "as extremely important, and for most people, a college education has become the necessary admission ticket to good jobs and a middle-class lifestyle" (Immerwahr & Foleno, 2000, p. 1). Authors such as Heiman and Slomianko (1998) even argue that higher education provides more job choices and opportunities, allows for the development of different skills, and significantly increases income levels, demonstrating its importance to social equity and mobility (Brennan & Teichler, 2008). For example, according to the U.S. Census Bureau (2011a), in 2009, a female high school graduate aged 18 to 24, earned an average of $22,620 while someone with the same demographic characteristics, but with a bachelor's degree or more earned $32,103 on average. In addition, the unemployment rate for females from 2000-2010 was 9.0% for high school graduates and 4.7% for females with a bachelor's degree or more (U.S. Census Bureau, 2011b).

Some Criticisms of Higher Education

Despite the benefits higher education might provide, there are many critiques of this system. Specifically, Immerwahr, Johnson, Ott, and Rochkind (2010) reported that 60% of a national sample of the US population believed that colleges care more about money and business than the educational needs of their students. This same study found that 60% of the respondents agreed with the assertion that colleges could enroll many more students without lowering quality or raising prices (Immerwahr et al., 2010). Another criticism frequently made relates to graduation rates and number of degrees granted (Hacker & Dreifus, 2010). Low graduation rates and low numbers of degrees granted mean that students are dropping out of school and might encounter more difficulties than graduates to find employment (U.S. Census Bureau, 2011b).

How Effective Educational Methods Can Help Address Criticism to Higher Education

Pascarella, Salisbury, and Blaich (2011) and Laird, Chen, and Kuh (2008) suggest that effective instruction (e.g., course organization and preparation, instructional clarity, teacher expressiveness, and feedback to students) not only improves students' performance on standardized tests, but also increases the likelihood that students will re-enroll in future semesters (Braxton, Bray, & Berger, 2000).

Although behavior analytic studies tend to address isolated academic skills such as writing for concision (Dermer, Lopez, & Messling, 2009) or concept teaching (Critchfield & Fienup, 2010; Fienup, Covey, & Critchfield, 2010; Walker, Rehfeldt, & Ninness, 2010), the teaching procedures described in these studies could help to develop more effective college curricula, thus improving its quality.

Behavior Analysis and Effective Educational Methods

Behavior analysis has been involved in the development and evaluation of several educational methods (Cohen, Manion, & Morrison, 2007; Moran & Malott, 2004). Educational methods derived from behavior analysis that are considered evidence-based include, but are not limited to, programmed instruction (Davis, Bostow, & Heimisson, 2007; Fredrick & Hummel, 2004; Jaehnig & Miller, 2007; McDonald, Yanchar, & Osguthorpe, 2005; Moore, 1963); personalized systems of instruction (Eyre, 2007; Fox, 2004); and direct instruction (Adams & Engelmann, 1996).

Even though the literature does not explicitly assert that stimulus equivalence is evidence-based, its use has provided empirical evidence on its efficacy (Alves, Kato, Assis, & Maranhão, 2007; Araújo & Ferreira, 2008; de Rose, de Souza, & Hanna, 1996; de Rose, de Souza, Rossito, & de Rose, 1992; de Souza et al., 2009; Fienup et al., 2010; Fienup & Critchfield, 2010). Among the main criticisms remaining about this technology is the question of how generalizable its use is to everyday classroom instructional circumstances (Fienup & Critchfield, 2010; Walker et al., 2010), since most of the studies use computerized instructions. Even though stimulus equivalence is based on individual performance and is commonly computerized, it is important to highlight that the use of online instruction and hybrid courses in higher education has been increasing (Garrison & Kanuka, 2004; Seaman, 2011; Young & Duhaney, 2008) and that stimulus equivalence might provide important procedural steps for those studying effective instruction programming using mixed media.

Broadly, stimulus equivalence is seen as a way to understand human symbolic behaviors, including language development and maintenance (Almeida-Verdu et al., 2008; de Souza et al., 2009; Luciano, Gomez Becerra, & Rodriguez Valverde, 2007). Stimulus equivalence is an

attempt to explain how the myriad of arbitrary relations among signs and their referents, which characterize human symbolic functions, are formed (Green & Saunders, 1998; Sidman, 1971; Sidman & Cresson, 1973). Equivalence-based instructions are considered important because they aim at teaching generatively (Fienup et al., 2010; Fienup & Critchfield, 2010). This implies programming procedures in a way that involves directly teaching of a few conditional discriminations that will yield untaught performances (de Souza et al., 2009; Fienup et al., 2010; Green & Saunders, 1998; Marques & Galvão, 2010; Saunders, Saunders, Kirby, & Spradlin, 1988; Sidman, 1971, 1986, 1990; Sidman & Cresson, 1973).

Sidman and Tailby's (1982) seminal article defined the properties and parameters for testing whether a conditional-discrimination procedure generated equivalent relations. According to these authors, all classes of equivalent stimulus derive from well-established conditional relations; however, the equivalence classes supersede the conditional relations. Sidman and Tailby (1982) proposed that to determine if a given performance involves more than conditional relations among stimuli, three tests derived from modern elementary mathematics should be used. The first test assesses reflexivity. "To determine that the conditional relation, R, is reflexive, one must show that each stimulus bears the relation to itself" (Sidman & Tailby, 1982, p. 6). For example, when shown a picture of a flower (sample stimulus), an identical flower will be chosen as the correct comparison stimulus without direct training. The second test evaluates symmetry.

> To demonstrate that the relation, R, is symmetric, one must show that both aRb and bRa hold true. A subject who matches a sample to comparison b is then required, without further training, to match sample b to comparison a, reversing "if a, then b" to "if b, then a" (Sidman & Tailby, 1982, p. 6).

If a conditional relation between a lily (sample stimulus) and an orchid (“correct” comparison stimulus) is established, the response of choosing a lily when an orchid is the sample should be emitted without further training. Finally, the third test: transitivity. “To determine whether R is transitive requires a third stimulus, c. Once "if a, then b" and "if b, then c" have been established, transitivity requires "if a, then c" to emerge without differential reinforcement or other current instructions” (Sidman & Tailby, 1982, p. 6). An example of this relation would be after establishing the lily-orchid conditional relation, establish another conditional relation: orchid-daisies. The response of choosing daisies when a lily is the sample stimulus should emerge without further training. These three tests (reflexivity, symmetry, and transitivity) define equivalence classes (Sidman & Tailby, 1982).

A Brief Literature Review on Stimulus Equivalence

The literature on stimulus equivalence is vast and studies involve different populations, humans and other organisms, such as sea lions (Kastak & Schusterman, 2002), monkeys (Brino, Assumpção, Campos, Galvão, & McIlvane, 2010), and pigeons (Urcuioli, 2008). The studies in this area also involve different stimuli, such as visual (Fienup & Dixon, 2006; Luciano et al., 2007; Merwin & Wilson, 2005), auditory (da Silva et al., 2006; Toussaint & Tiger, 2010) , tactile (Toussaint & Tiger, 2010), olfactory (Fienup & Dixon, 2006; McAtamney & Annett, 2009), and gustatory stimuli (Hayes, Tilley, & Hayes, 1988; McAtamney & Annett, 2009). Stimulus equivalence also involve different settings, such as preschools (Pilgrim, Jackson, & Galizio, 2000), universities (Fienup et al., 2010; Fienup & Critchfield, 2010), and medical treatment facilities (Guercio, Podolska-Schroeder, & Rehfeldt, 2004). Additionally, procedures based in stimulus equivalence have been used to teach various skills such as math skills (Araújo & Ferreira, 2008; Lynch & Cuvo, 1995), basic reading skills (de Souza et al., 2009; Hubner,

Gomes, & McIlvane, 2009; Toussaint & Tiger, 2010), shopping skills (Taylor & O'Reilly, 2000), and musical skills (Arntzen, Halstadtro, Bjerke, & Halstadtro, 2010).

Due to the scope of the present thesis, the focus of this literature review was on studies involving college students and teaching complex verbal skills (i.e. concepts and/or their utilization). Searching for key studies involved several steps. First, a literature search was conducted using Google Scholar with the words "stimulus equivalence" within quotation marks. No date or location of the word in the article/book/report was specified in the advanced search mechanism. This search resulted in 3,710 entries. A new search was conducted within Google Scholar, now specifying dates by decades (see Appendix A for complete results). From 2000 to 2009, 1,180 entries were found and from 2010 onwards, 344 entries were found. To analyze the most recent studies, within the period of a thesis, only the papers published in 2010 or later were retrieved and analyzed.

In addition to Google Scholar, literature searches were conducted in the following databases: PubMed, ProQuest Research Library, Academic OneFile, and PsycINFO. First, the phrase "stimulus equivalence" was searched within quotation marks. Next, the search was narrowed to include studies within the date range from 2010 onward, including 2012 papers that are available online, but not in print. Finally, the words "college students" and the word "concept" were added to the database search engines, one at a time. If these steps yielded no results, the last word or phrase used (e.g., the word "concept") was deleted from the search, until at least one study was found. The articles described below represent those that were (a) found in at least two of the five databases, (b) peer reviewed, (c) empirical, (d) related to stimulus equivalence, and (e) within the date range specified above. These studies also involved college students as participants. One of the four articles found in two or more databases investigated

differences in acquisition of conditional discriminations in contexts using emotionally evocative *versus* arbitrary or neutral stimuli, and did not involve teaching complex verbal skills (i.e., concepts taught in higher education settings). This article (Adcock et al., 2010) was not included in the scope of the present thesis.

The next three paragraphs will describe the three studies that matched all criteria. Then, their main contributions to the equivalence literature will be briefly described. Finally, some of their limitations will be presented and ways to address these limitations will set up the stage for the present study.

Fienup et al. (2010) used equivalence-based instruction to establish relations among brain regions, their anatomical locations and psychological functions, and problems associated with them. Overall, the procedures included general instructions that described the computerized lessons, immediate feedback on performance, and training to mastery. Participants were four college undergraduates, ranging in age from 18 to 22 years, with self-reported grade point averages (GPAs) ranging from 2.4 to 3.7, and with American College Test (ACT®) examination scores ranging from 21 to 27. The stimuli were four sets of five stimuli encompassing brain regions, their anatomical locations, psychological functions, and psychological problems associated with the brain regions. Percentage of correct responses in multiple-choice tasks and time to complete the tasks were the variables of interest. The results showed that participants learned the directly taught relations and their performance showed emergence of new conditional relations among the stimuli. Additionally, the classes that shared a common member spontaneously merged, thereby increasing the number of emergent relations. Overall, students mastered more than twice as many relations as they were directly taught. The study demonstrated

the potential of equivalence-based instruction to reduce the amount of student investment required to master advanced academic topics.

Also using a match-to-sample procedure, Fienup and Critchfield (2010) investigated the effects of establishing contextual control to teach students the conditional application of concepts of statistical significance and hypothesis decision making. As in Fienup et al. (2010), the authors used computerized instructions to teach key content together with immediate feedback and training to mastery. Fienup and Critchfield (2010) also measured percentage of correct responses in multiple-choice tasks and time to complete the tasks. Participants were ten students, ranging in age 18 to 28 years, with GPAs ranging from 1.40 to 3.70. The authors used two sets of stimuli related to statistical inference. Overall, following the match-to-sample training, scores improved from below 70% of correct responses on the pretest to near 100% on the posttest. This study represents an additional illustration of the use of equivalence-based instruction to establish academic skills in higher education participants.

The third study identified in the literature search that used equivalence-based instruction was by Critchfield and Fienup (2010). These authors examined whether a group setting would adversely affect learning outcomes by using previously designed lessons on inferential statistics (Fienup & Critchfield, 2010) in a group setting. The authors also assessed whether learning gains would be similar to the ones obtained in previous studies when pre and posttests were administered in a paper-and-pencil multiple-choice format. The authors used a pretest-posttest design and match-to-sample procedures to teach conditional relations that contributed to the formation of equivalence classes involving statistical inference stimuli. As in both studies previously described, immediate feedback and training to mastery were included in the procedures. Participants were 27 undergraduates with an average age of 19.4 years. The

participants' average GPA was 3.0 and 23.5 for the SAT[1] college entrance exam score. The authors used two classes of six stimuli related to statistical inference, and measured the percentage of correct responses in multiple-choice tasks and time to complete the tasks. According to the authors, training proceeded quickly and with few errors: students achieved mastery (12 consecutive correct responses) in fewer than 20 trials. Overall, the authors argue that the lessons succeeded in building statistical inference skills, as measured on the paper-and-pencil tests. The authors suggest that future studies based on equivalence-based instruction should use repeated measures to evaluate student learning, should be used in more "natural settings", and should aim at evaluating the relative efficacy of equivalence-based instruction compared to other interventions.

These studies advanced the application of stimulus equivalence technology (i.e., teaching different complex verbal behaviors to a typically developing adult population) and they also advanced the analysis derived from their applications. Despite the advances, there are questions that still need to be addressed if equivalence-based instructions are to be used successfully in the context of higher education.

One point that has not been fully addressed in these studies is how the use of multiple-choice questions can affect performance in teaching and testing sessions. First, multiple-choice questions encompass forced-choice tasks that require selection-based behaviors (Chase, Johnson, & Sulzer-Azaroff, 1985; Michael, 1985; Polson & Parsons, 2000; Walker et al., 2010). Despite

[1] SAT used to stand for "Scholastic Aptitude Test". According to The Eduers.com (2009), "in 1993, the SAT was renamed as the SAT Reasoning Test (or known as SAT I). Meanwhile, the former Scholastic Achievement Test was renamed as the SAT Subject Tests (or known as SAT II)."

Michael's (1985) argument that selection-based behaviors require conditional discriminations, thus are harder to acquire, for people who already have this repertoire, pointing (selection-based behaviors) require only the response topography of pointing to be established. On the other hand, talking or writing (topography-based responses) require different response topographies to be reinforced. As highlighted by Walker et al. (2010) topography-based behaviors resemble more closely everyday behaviors and should be targeted in stimulus equivalence studies.

Another issue related to multiple-choice tasks is the fact that even though these tasks do not necessarily provide explicit feedback on correct responses, they provide exemplars and non-exemplars that could function as prompts for correct responses (Ribeiro, Pascualon, Sella, Bandini, & de Souza, 2009; Tiemann & Markle, 1990). Consequently, it is especially important to assure that performance will not improve through the exposure to the tasks, often referred to as a "practice effect". The three studies used either pre-posttest designs or one single probe for each conditional relation before the introduction of the independent variables. Having more than one baseline data point could help demonstrate that only when the intervention is presented, the participant's responses will change according to the experimental conditions (Bonfiglio, Daly, Martens, Lin, & Corsaut, 2004; Connell & Thompson, 1986; Horner & Baer, 1978; Murphy & Bryan, 1980).

One more limitation of these three studies includes the fact that participants had high scores in the pre-tests (they could emit up to 70%-75% of correct responses). When pretest scores are high, differentiation between pretest and posttest performances might not be so clear. Setting a pre-test criterion for a lower score (i.e., 20% or lower) could improve visual differentiation of pre and post-intervention performance.

The present study aimed at addressing the points described above by (a) using open-ended questions as probes, (b) using a multiple probe design, and (c) requiring that participant had less than 20% correct responses in all baseline probes for study eligibility. Three probes were presented before teaching sessions were delivered and the same probes were repeated after each of the three teaching parts. Overall, the purpose of the present study was to evaluate the effects of a stimulus equivalence instructional package on undergraduates' performance in conditional discrimination tasks and open-ended tasks that involved research design names, definitions, notations, and examples as discriminative and conditional stimuli.

Method

Participants

The experimenter contacted three professors from different state and federal universities in Brazil and asked them to refer students who could read articles in the English language. Ten undergraduate students were referred by Brazilian professors and were contacted by the experimenter through e-mail.

The e-mail contained a brief description of the study (see Appendix B for complete e-mail in Portuguese) and a consent form - as approved by the Human Subject Committee of Lawrence (HSCL # 19431, see Appendix C for approval letter). Participants were instructed to read, sign, scan, and return the consent form by e-mail if they wished to participate. Only participants who signed the consent form were contacted again to set up the initial interview (See Appendix D for the initial interview). Six participants returned the signed consent form by e-mail and completed the initial interview with the experimenter. Four participants remained in the study after the initial interview. At the end of this interview, a brief, online reading comprehension test (http://fcit.usf.edu/fcat10r/home/sample-tests/virtues-of-venom/index.html) was administered to assure that participants had the initial reading comprehension repertoire necessary to participate in the study (i.e., at least six out nine correct responses in the reading pretest).

In addition to scoring at least six out of nine questions correctly in the reading test, participants remained eligible for the study only if their percentage of correct responses in Probes 1, 2, and 3 was lower than 20% (probes are described below). The four participants had scores lower than this criterion, which was used to avoid ceiling effects that could preclude evidence

that the experimental procedures promoted learning. Participants ranged in age from 19 to 23 years. Their primary language was Portuguese and they were all enrolled in at least one class in the social or behavioral sciences. Additionally, all participants stated that they had not had any advanced classes on research methods and did not know much about research.

In exchange for participating, participants received a book (see Appendix E for references), a package with research articles on single-subject research, and a participation certificate.

Settings

Two virtual environments were used for data collection: Adobe Connect® and Skype ™. Adobe Connect® is a computer program that can be used to deliver information through presentations, online training materials, web conferencing, and to access functions. Adobe Connect® was used to host the session content, to present all trials, and to record the data. Because Adobe Connect® is based on Adobe Flash®, all participants were required to have Adobe Flash® installed in their computers to access the links to the sessions.

Skype ™ is a software application that allows users to make voice and video calls over the Internet, to exchange information, to share screens, and to access other functions. Skype ™ was used to hold the initial interview, including the reading test.

Dependent Variable

The dependent variables included (a) percent of correct responses in teaching sessions; (b) percent of correct responses in transitivity and symmetry of transitivity sessions; (c) percent of correct responses in open-ended questions - probes; and (d) percentage of correct responses in review checklists.

Experimental Stimuli

There were six sets of experimental stimuli: Learning stimuli, computerized instructions for teaching and matching to sample tasks, checklist for article review, article for review, open-ended questions, and the glossary. These sets of stimuli are described below.

Learning stimuli. A content analysis (Fox & Sullivan, 2007; Markle, 1975; Markle & Tiemann, 1970; Tiemann & Markle, 1990; Twyman, Layng, Stikeleather, & Hobbins, 2005) guided the learning stimulus choices. First, the experimenter searched for the most frequently cited books on behavioral and social research methods to find content related to the purpose of the study. Two books were identified as the most cited: Campbell and Stanley (1963) and Creswell (2009). Second, since both books did not use similar names or similar descriptions for the research designs they listed, the experimenter had to choose one of the two books. Both the Applied Behavioral Science department classes and a research primer developed at the Research and Training Center on Independent Living used the Campbell and Stanley book in classes and for development of a research primer, respectively. For these reasons, the Campbell and Stanley book was chosen. Third, items that were common for all research designs (i.e., they appeared in all book sections describing the research designs) became the key elements to comprise the stimulus classes. As highlighted by Tiemann and Markle (1990), finding the relevant properties of a stimulus allow for better exemplars and non-exemplars. Fourth, after identifying the key items, behavior and social science books and websites (ALLPsych Online, 2002; Connections, 2010; Fraenkel, Wallen, & Hyun, 2011; Kish, 2005; Wrench, 2009; Yu & Ohlund, 2010) were consulted to develop exemplars for the key elements: the research design names (set A), the research design definitions (set B), the research design notations (set C), and the research design

examples (set D). The most commonly used names, definitions, notations, and in the books and websites were used in the present study.

Thirty-six experimental stimuli (Appendix F) were developed through the content analysis described above. Stimuli were designated with the following alphanumeric symbols: the nine research design names were designated set A, the nine research design definitions, set B, the nine research design notations, set C, and the nine examples, set D. All stimuli related to the One-Shot Case Study received the number "1" after the letter (e.g., the design definition for the One-Shot Case Study received the denomination "A1"). Stimuli related to the One-Group Pretest-Posttest Design received the number "2" after the letter, etc (for the complete alphanumeric denominations, see Appendix F). All stimuli were comprised of printed sentences in English.

Computerized instructions for teaching and matching to sample tasks. All instructions were programmed through Adobe Captivate®. Instructions were divided into three "teaching parts" (i.e., Pre-experimental Designs – Part1, Quasi-experimental Designs – Part 2, and True experimental Designs – Part 3). Each of the teaching parts was comprised of three different research designs (for a graphic representation of each teaching part, see Appendix G). The first teaching part (Pre-experimental Designs) was comprised of One-Shot Case Study, the One-Group Pretest-Posttest Design, and the Static-Group Comparison. The second teaching part contained three Quasi-experimental Designs (the Nonequivalent Control-Group Design, the Counterbalanced Design, and the Multiple Time Series Design). Finally, the third teaching part was comprised of three True experimental Designs (the Pretest-Posttest Control Group Design, the Solomon Four-Group Design, and the Posttest-Only Control Group Design).

Each teaching part encompassed two sessions: one teaching session and one session in which symmetry, transitivity and symmetry of transitivity relations were tested (emergent relations session). The trials in these two types of sessions were presented in a matching-to-sample format made up of one sample stimulus at the top of the screen and three comparison stimuli at the bottom of the screen (Figure 1). All stimuli were simultaneously presented. The sample stimulus and the position of the correct answer were quasi-randomized: they were never presented in more than two consecutive trials.

The program instructed the participant to choose one comparison stimulus - among the three choices at the bottom - that best matched the information at the top of the screen (the sample stimulus). Mastery was achieved when the participant correctly answered 90% or more of the trials in both the teaching session and the emergent relations session.

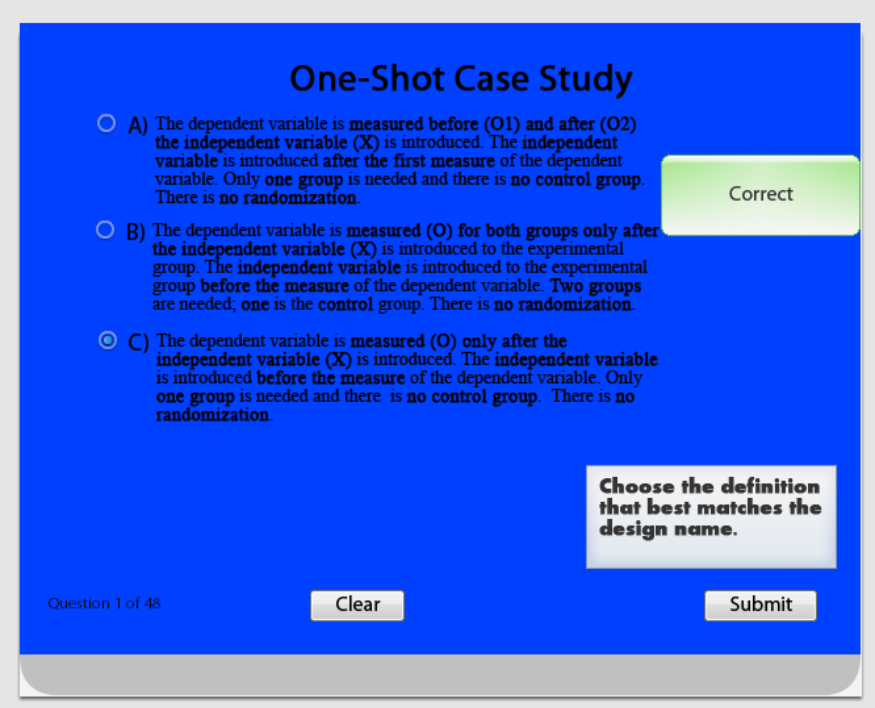


Figure 1. Screenshot from a teaching session trial reflecting format, font, and colors as they were presented to participants.

Checklist for article review. To avoid bias, the checklist used in the study was an adaptation from a peer reviewed article that suggested items to be reviewed in a manuscript (Roberts, Coverdale, Edenharder, & Louie, 2004). Since the study focus was on research designs, the items drawn from Roberts et al. (2004) represented design and methods. After the checklist was designed, three volunteers read the Whitehurst et al. (1988) article (see below) and answered the checklist questions. After reviewing the Whitehurst et al.'s article using the checklist, the volunteers suggested modifications to the checklist questions (e.g., readability). The final checklist was comprised of 21 multiple-choice questions, each with four possible answers. All questions were programmed in Adobe Captivate® and were presented through Adobe Connect®. The program instructed the participant to click on one of the possible answers. No feedback on accuracy was presented: both correct and incorrect responses resulted in the presentation of the next trial (for a complete list of the questions and instructions in the checklist, see Appendix H).

Article for review. An article by Whitehurst et al. (1988) was chosen for use in both article review sessions via a three step process. First, the experimenter wrote the nine research design names on nine separate slips and placed them into a container. Second, a volunteer drew one slip of paper from the container and read the design name aloud. Third, the experimenter typed the design name into Google Scholar, within quotation marks. Among the resulting entries in Google Scholar, the first article that (a) represented the behavioral sciences, more specifically, psychology, cognitive science, organization theory, psychobiology, social neuroscience, anthropology, organizational behavior, organization studies, or sociology; and (b)

was peer reviewed, was chosen. Whitehurst et al. (1988) was the first article that met both of these criteria.

Open-ended questions. The open-ended questions were comprised of nine questions, one on each research design. Each question had 12 sub-items (see Appendix I for complete set of questions and sub-items). Participants could answer the questions in their native language, in English, or in a combination of both languages. During the first three probes, there was no mastery criterion (baseline probes), but participants could not answer more than 20% of the questions correctly to remain eligible to participate. After a teaching session, mastery criterion was 80% correct responses on the questions related to the designs that had already been taught.

Failure to meet 80% of correct responses in open-ended questions related to A-B (name-definition relations) and A-C (name-notation relations), led to the presentation of new teaching sessions. These new teaching sessions contained only those relations in which the participant scored below 80%. For example, if the participant scored below 80% on items asking about the definition of a given design, but had 90% correct responses on items asking about notation and examples of this same design, only A-B relations were presented in the re-teaching session. If criterion was not met only A-D relations, there was no re-teaching, since it is argued that *new*, diversified examples are necessary when aiming at generalization of this type of response (Markle, 1975; Stokes & Baer, 1977; Tiemann & Markle, 1990).

Open-ended questions were scored based on three rubrics (one for each set of three designs) that contained several possibilities of answers for each one of the 12 sub-items contained in each one of the nine open-ended questions that comprised any given probe (see Appendices J, K, and L for the three rubrics). The rubrics were developed based on Freedman (1994) and Ebert-May (n. d.).

Glossary. This activity was developed to assure participants had access to basic research vocabulary (See Appendix M for complete list of terms and definitions covered in the Glossary). The activity was comprised of three slides containing two or three matching to sample trials per slide. The definition of the terms were presented in the left part of the screen, with the letters A, B, or C in front of the definition. The terms were present in the right part of the screen and the participant clicked on a drop down arrow to the right of the terms. When the participant clicked over the arrow, the three letters were displayed and the participant clicked on one of them. Accuracy feedback was provided on the screen: correct responses resulted in the presentation of a 3cm X 2cm green rectangle as the background for the word "correct" at the top right corner of the screen. Incorrect responses resulted in a 3cm X 2cm pink rectangle, with the words "incorrect- try again". This activity was optional and the correct answers were sent by e-mail to all participants, so they could refer to it at any moment throughout the study.

Experimental Design

A multiple probe design across the three teaching parts was implemented on an individual basis (Figure 2). Performance in all three parts (Pre-experimental Designs, Quasi-experimental Designs, and True Experimental Designs) was measured three times before the introduction of the intervention and again after each one of the teaching parts was presented to each participant. As highlighted before, this design can help to assess whether changes in the measures of the dependent variable are attributable to the introduction of the intervention (Barlow, Nock, & Hersen, 2009; Horner & Baer, 1978).

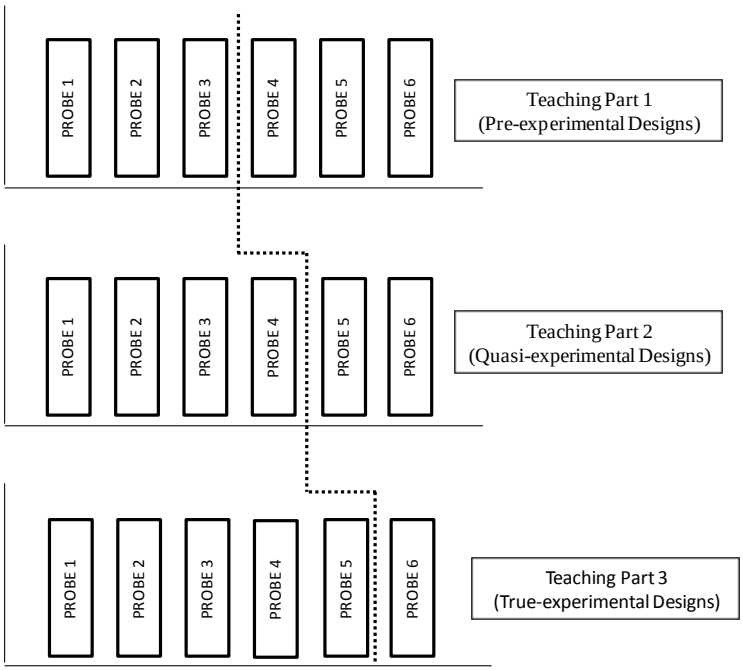


Figure 2. Schematic of the experimental design. The dashed lines represent the introduction of the independent variable in temporal relation to the probes.

General Procedures

Figure 3 depicts the experimental phases, performance criteria, and conditional relations involved in teaching sessions. Performance criteria were different for each phase and guided decisions regarding the presentation of a new phase or the re-teaching of a current phase. A detailed description of each phase is presented below Figure 3.

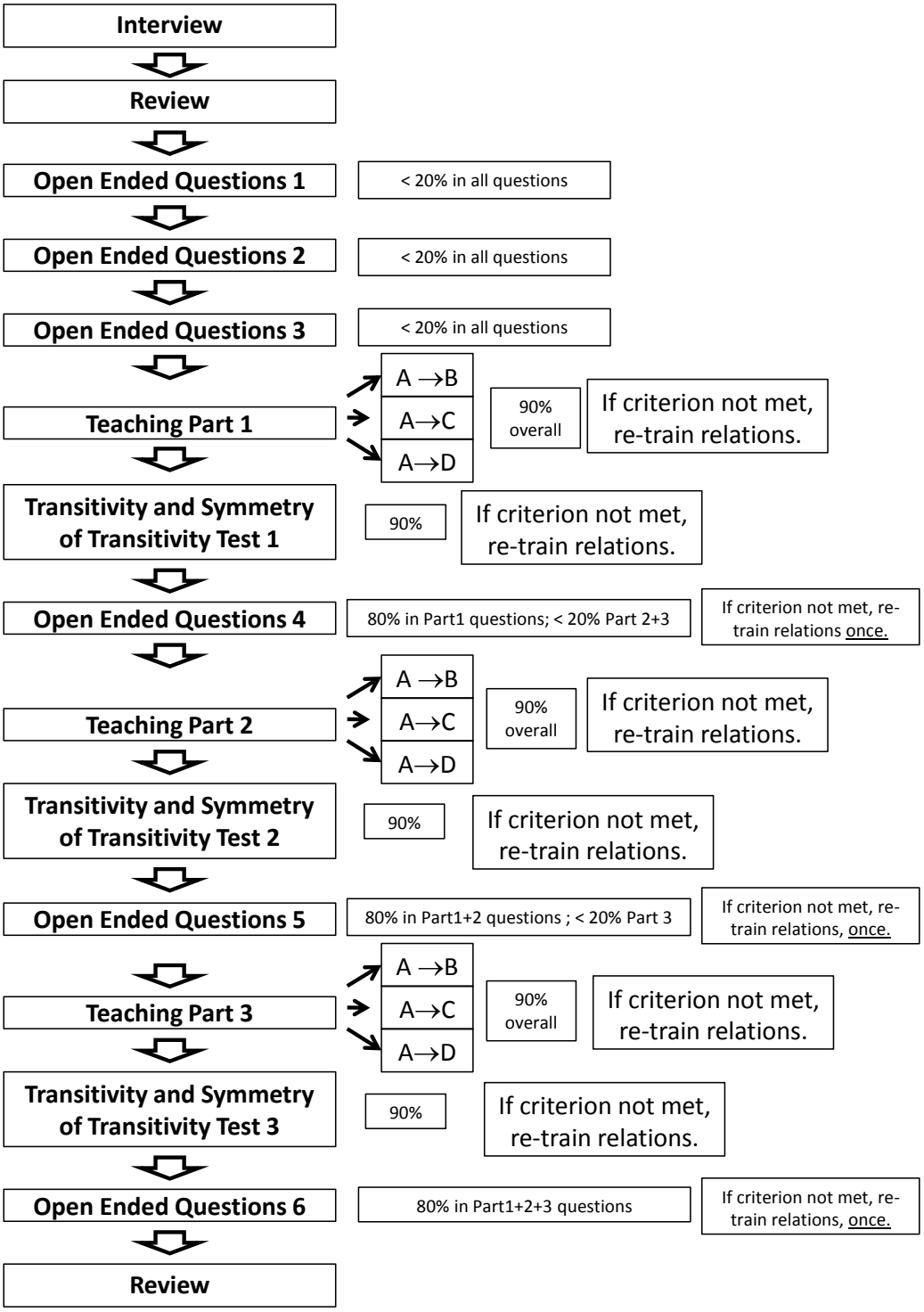


Figure 3. Experimental phases, performance criteria, and conditional relations involved in teaching sessions.

Skype interview sessions. The experimenter set up a Skype meeting with each participant by e-mail. The preferred Skype media for the meeting was video and voice; however, for participants who had limited internet connection speed, only voice or texting was used. The experimenter started by requesting demographic and academic information (see Appendix D for the questions) and by explaining the research procedures (see Appendix O, for explanation given on the research procedures). The experimenter then answered any questions the participant had regarding the study. Next, the experimenter gave the participant the choice of taking the reading comprehension test, then or at a later time. All participants took the reading comprehension test then. Finally, the experimenter explained that the link to the article review checklist and the article for review - in a .pdf format- would be sent through e-mail as soon as the meeting was over. About 5 minutes after the Skype meeting ended, the experimenter sent the link and the article for review.

Article review sessions. Article review sessions were conducted before baseline probes and after the last probe. Before starting these sessions, the participants received an e-mail with a link to the "Checklist for article review" and with the article to be reviewed (Whitehurst et al., 1988). The e-mail included instructions noting that participants should read the article before completing the checklist and that they could refer back to the article at any time while completing it. There was no mastery criterion and no feedback on accuracy was provided throughout this session. After participants finished this session, they e-mailed the experimenter to let her know they were ready for the next phase. If a given participant did not send an e-mail within five days, the experimenter contacted the participant again. In response to a participant's e-mail, the experimenter provided a general statement of appraisal: "Good job" or "you did well"

and provided the link to the next phase of the procedure: the first probe (see Appendix N for the information contained in the e-mail regarding the first probe).

Probe sessions. Probe sessions were comprised of two slides with general instructions about the open-ended questions and nine slides with the actual questions. The instructions in the first slide were: "Today you will answer open ended questions. You can: (a) Answer the questions in English; (b) Answer the questions in your own language; (c) Use a mix of languages if you need to borrow some expressions". The instructions in the second slide were: "Be sure to answer ALL items of the questions before you move onto the next question. If you do not know the answer, write: "I do not know the answer", before moving on to the next question. After the first two slides, the first open-ended question was presented. After answering the question, the participant clicked on "submit" and the next question was presented. After responding to the nine questions, the participant e-mailed the experimenter, to request the link to the next phase.

Teaching sessions. Each teaching session was comprised of at least 18 trials. These 18 trials were subdivided into three groups: the first six trials presented design name – design definition relations (A-B relations); the next six trials (trials 7-12) contained design name – design notation relations (A-C relations); and the last six trials encompassed design name-design example relations (A-D relations). During teaching sessions, accuracy feedback was provided on the screen: correct responses resulted in the presentation of a 3cm X 2cm green rectangle as the background for the word "correct" at the top right corner of the screen (Figure 1). This was followed by the presentation of the next trial. Incorrect responses resulted in a 3cm X 2cm pink rectangle, with the words "incorrect- try again" printed over it. Additionally, one correction trial was presented (the correction trial was not scored as correct or incorrect). Finally, the program automatically presented a new set of six trials containing the conditional relations in which

incorrect responding occurred; in these cases, the total number of trials was higher than 18, since a new set of six trials was presented every time an incorrect response occurred. Mastery criterion in teaching sessions was defined as six correct consecutive trials in all three subgroups of trials and at least 90% correct responses when considering all trials of one given session.

Transitivity and symmetry of transitivity sessions (emergent relations sessions). Each symmetry, transitivity and symmetry of transitivity relations session had 21 trials distributed among B-A, C-A, B-C, C-D, C-B, D-C, and D-A relations. Relations were randomly presented in each of these sessions. Correct or incorrect responses did not result in any type of accuracy feedback, just the presentation of the next trial. Mastery criterion was at least 90% correct responses.

Interobserver Agreement

As described by Boykin and Nelson (1981), interobserver agreement (IOA) is assessed when data collection relies on human observers. Most studies in stimulus equivalence are computerized and do not provide data on IOA (da Silva et al., 2006; Fienup et al., 2010; Fienup & Critchfield, 2010). Adobe Connect ® and Adobe Captivate® recorded the data and analyzed correct and incorrect responses for the review checklists, the teaching sessions, and the emergent relations sessions, thus IOA was not calculated for these sessions. However, since probes were comprised of open-ended questions that could not be automatically corrected, IOA was assessed for all six probes for at least 33% of all participants. Selection of a given participant's probe over another was random.

Interobserver agreement was calculated in two different forms. The first one consisted of assigning each item of the questions either "agreement" (value = "1") or "disagreement" (value = "0") and dividing the number of agreements by the agreements plus disagreements. Then, the

results were multiplied by 100. Using this calculation, IOA was 100% for Probes 1, 2, 3 and 6, 93.5% for Probe 4, and 91.7% for Probe 5.

In the second form of calculation - since items in the open-ended questions could be scored as “0”; “0.25”; “0.5”; “0.75”; or “1” - for items in which there was not exact agreement (exact agreement counted as “1”), the smaller score was divided by the higher score to find the partial agreement for a given item. The agreements and partial agreements were added and divided by 72 (total number of items in a probe), and multiplied by 100. Agreement was 100% for Probes 1, 2, 3, and 6, 95.1% in Probe 4, and 95.1% in Probe 5.

Social Validity

Participants were encouraged to give feedback on the experimental stimuli and the experimental phases at any moment: in all the emails sent to participants, the last sentences included requests to inform on any problems that might have occurred during sessions and/or suggestions on stimulus presentation. In addition, participants who finished the study were sent a social validity questionnaire through e-mail which contained nine affirmative propositions about the “tutorial” (i.e., instructional package). Six out of the nine propositions were Likert-type scaled. The scale ranged from 1 (strongly disagree) to 5 (strongly agree) and referred to (a) ease of use; (b) online sessions and time and space flexibility; (c) usefulness of information; (d) links sent on time; (e) importance of feedback on teaching sessions; (f) recommendation to other people. The open-ended questions asked about the most useful and the least useful features of the tutorial and requested additional suggestions for changes (see Appendix P for complete questionnaire).

The e-mail with the social validity questionnaire was individually sent (i.e., each participant received a personalized e-mail) and was comprised of (a) general instructions about

the questionnaire; (b) the questionnaire, and (c) instructions to send the questionnaire to the second observer (who scored the probes). The second observer was to receive the completed questionnaires and delete any information that could identify the participants and then send the unidentified questionnaires to the experimenter.

Results

All results are shown in terms of percent correct responses. Results of baseline and post-teaching probes, divided by each teaching part (1 – Pre-experimental Designs, 2 – Quasi-experimental Designs, and 3 – True experimental Designs), are summarized in Figure 4 and Figure 5. Results of the article review, baseline and post- teaching probes (not divided by each teaching part), the glossary, teaching sessions, and emergent relations sessions are summarized in Table 1. Because incorrect responses did not frequently occur in teaching and emergent relations sessions, the results on errors are embedded in the written description of these sessions. In addition, all participants answered the Glossary with 100% correct responses. Participants were given pseudo names: John, Mary, Barbara, and Sarah.

As depicted in Figure 4 and Figure 5, prior to teaching, all participants scored below mastery criterion in all three probes. Data on the left side of Figure 4 summarizes John probe results. Data on the right side of Figure 4 summarizes Mary probe results. Data on the left side of Figure 5 summarizes Barbara's probe results. Finally, data on the right side of Figure 5 summarizes Sarah's probe results. The only participants who scored above 0%, but still under 20%, were John (1.4% in Probe 1) and Sarah (5.9%, 18.1%, and 5.5% In Probe 1, 2, and 3, respectively). All participants increased their scores to over 50% correct responses after each specific teaching part was taught. Mary did not meet 80% correct responses in Probe 5, so another Part 2 teaching sessions was presented to her. Her errors were in notation and example items of the probe, so only these were presented in the new teaching session. It is important to note that this session did not contain new examples; it was just a repetition of a Part2 teaching session. Sarah did not meet 80% correct responses, but since her errors were in the examples, she did not go through re-training.

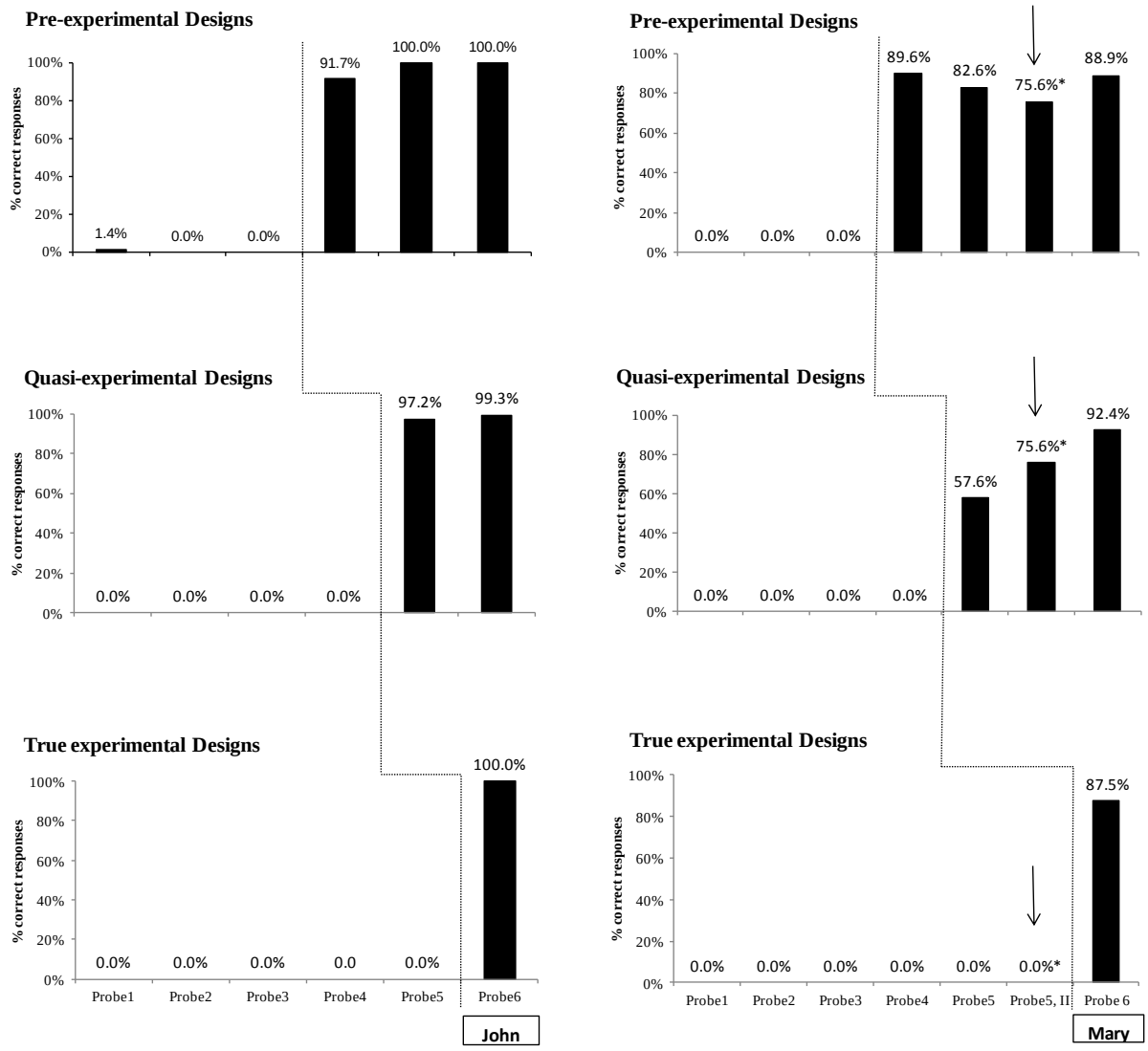


Figure 4. John's and Mary's percentage correct responses in probes, in each set of design: Pre-experimental Designs are at the top of the figure, Quasi-experimental Designs are in the center of the figure, and True experimental Designs are at the bottom of the figure. The arrow highlights Probe 5, II (probe after the Part 2 re-teaching session).

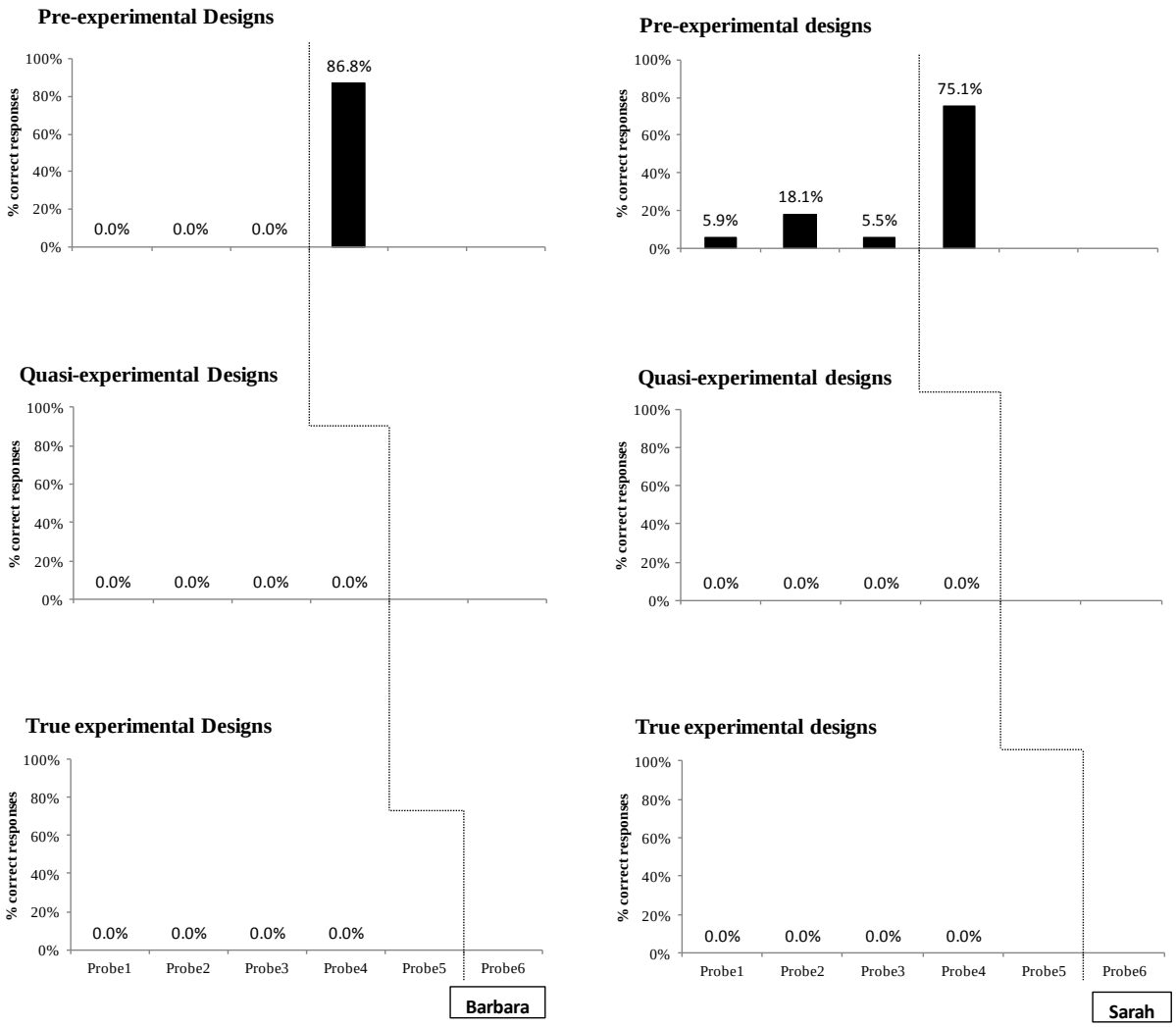


Figure 5. Barbara's and Sarah's percentage correct responses in probes, in each set of design: Pre-experimental Designs are at the top of the figure, Quasi-experimental Designs are in the center of the figure, and True experimental Designs are at the bottom of the figure.

Mastery was required in teaching and emergent relations sessions (at least 90% correct responses) and all participants met mastery criterion for both types of sessions (Table 1). Mary was exposed to two teaching sessions before she met criterion in Part 1 emergent relations session. Mary emitted nine incorrect responses during Part 1 emergent relation sessions. She was

also exposed to a modified teaching session in Part 2 due to her low scores in Probe 5. John and Barbara emitted two incorrect responses during Part 1 teaching sessions, both in A3-B3 relations (name-definition of the Static Group Comparison). The different total number of teaching trials was 56 and 80 for John and Mary – who finished the study - and 24 and 18 trials for Barbara and Sarah. These last two participants had serious events happen in their lives and had to leave the study.

Table 1

General Results For All Participants in All Experimental Phases

Phases	John	Mary	Barbara	Sarah
Article review (pre)	45.83%	58.3%	20.80%	50%
Probe1[a]	0.45%	0.0%	0%	2%
Probe2[a]	0%	0.0%	0%	6%
Probe3[a]	0%	0.0%	0%	1.8%
Glossary	100%	100%	100%	100%
Part 1 teach	95.83%	100.0%[b]	99.6%	100%
Part 1 transitivity	100%	66.7%	100%	95.2%
Part1 re-teach	N/A	100.0%	N/A	N/A
Part1re-transitivity	N/A	90.4%	N/A	N/A
Probe 4[a]	30.79%	29.8%	28.9%	25%
Part 2 teach	100%	100.0%	X	X
Part 2 transitivity	100%	100.0%	X	X
Probe 5[a]	66%	46.8%	X	X
Part2 re-teach	N/A	100.0%	X	X
Probe 5, II[a]	N/A	50.4%	X	X
Part 3 train	100%	100.0%	X	X
Part 3transitivity	100%	100.0%	X	X
Probe 6[a]	99.70%	89.6%	X	X
Article review (post)	41.67%	66.67%	X	X

Note. N/A indicates that a given phase was not applicable to a given participant and X indicates that the participant did not finish those phases of the study.

[a] Overall percentage correct response calculated by diving the sum of % correct responses in the nine questions comprising the probe by nine. Thus, results do not reflect data for each grouping of three designs (i.e., pre-experimental design, quasi-experimental designs, and true experimental designs).

[b] Participant's internet signal stopped after 2 trials of the 3rd sub section (name-example relations) and the program shut down.

Comparing results in both article reviews, for John, performance did not increase, even in the first five questions that related to research design. For Mary, her performance in the first five questions went from 60% to 80% correct responses. When analyzing the other responses changes, for both John and Mary, there was not a consistent pattern of change (i.e., the changes went from choosing alternative A to B, from C to A, among other possible combinations). Thus, teaching sessions did not consistently affect performance in the checklist.

All participants had spontaneous comments about the instructional package. The comments included suggestions for changes in the open-ended questions (Sarah) and suggestions for new topics to be included in the tutorial (John).

The two participants who answered the structured social validity questionnaire gave high ratings to the tutorial. John rated all six Likert-type scaled propositions as “5”, in a 1 (strong agree) to 5 (strongly agree) scale. Mary rated 4 of the six propositions as “5” and two as “4”. In the open-ended propositions, both participants provided information on what to keep in the multiple-choice questions formatting (John) and what to change in the probes – she suggested a decrease in the number of probes (Mary), because “the open-ended questions become tiring” (See Appendix Q for complete results of the questionnaire).

Discussion

The overall purpose of this study was to evaluate the effects of a stimulus equivalence instructional package on undergraduates' performance in conditional discrimination and open-ended tasks that involved research design names, definitions, notations, and examples. The discussion below is organized in the following order: (a) a summary of how the study contributes to the stimulus equivalence literature, (b) a broader discussion of the points under (a), (c) a description of the main limitations of the study, and d) a brief conclusion.

The present study adds to the literature on the applications of stimulus equivalence technology to higher education. First, like other recent studies using this technology (Critchfield & Fienup, 2010; Fienup et al., 2010; Fienup & Critchfield, 2010; Walker et al., 2010), these data show that all participants met performance criteria in directly taught and emergent conditional relations. Second, this study used a multiple probe design with three probes before the introduction of the intervention, thus better demonstrating that the percentage of correct responses did not increase with exposure to new probes alone. The differences between scores in baseline and post-intervention probes provide evidence that the participants emitted correct responses during probes only after teaching sessions were presented and mastered. Third, teaching of selection-based responses yielded predicted responses in the topography-based probes. Fourth, the content analysis involved the consideration of multiple exemplars - books, articles, and websites - to develop stimuli involved in the teaching procedures. Fifth, this study was conducted totally in online settings, which allowed space and time flexibility for participants and the experimenter. Sixth, teaching sessions were presented in English, but probes could be answered in Portuguese, English or using both languages. Finally, a checklist containing assessment questions for an article review was used to evaluate whether teaching conditional

relations on research designs could result in better overall performance in several items of the checklist.

In relation to the first topic presented above, as highlighted in many studies (Almeida-Verdu et al., 2008; de Souza et al., 2009; Fienup et al., 2010; Fienup & Critchfield, 2010), the use of stimulus equivalence technology promotes economy of teaching. In this study, participants were taught 27 conditional relations, nine relations in each one of the three teaching parts. Performances in emergent relations sessions show that for the two participants who finished the study, there was more than 90% correct responses among the 63 emergent relations that were explicitly tested[2]. For the other participants, results in emergent relations sessions also showed emergence of untaught conditional discriminations.

Although there was economy of teaching in relation to conditional relations, it should be highlighted that when aiming at teaching a concept, conditional discriminations in this matching-to-sample format may not be sufficient. As underscored by Tiemman and Markle (1990) and Markle (1975), to effectively teach a concept, several exemplars and non-exemplars of each concept must be presented: "Asking a student to repeat or recognize the definition is totally inadequate; asking him to generate an example or two of his own does not satisfy the requirements" (Markle, 1975, p. 3). In the present study, each research design was assigned one name, one definition, one notation, and one example; these were considered the exemplars of this given design. The stimuli that comprised two other research designs were the non-exemplars (in

[2] Eighteen conditional relations involving BD and BD were not tested since pilot data (see Appendix M for a brief summary) showed that not only B and D stimuli had very similar structures, but also the conditional discriminations involving these stimuli resulted in correct responses.

one given set of three designs). According to Tiemman and Markle (1990) and Markle (1975), more exemplars and non-exemplars must be presented to assure better inter-class discrimination and better intra-class generalization. The more exemplars and non-exemplars, the more likely it is that an abstraction will be yielded (Skinner, 1953, 1957/2002). One step that could have been taken to improve the inter-class discrimination: presenting stimuli from all nine designs together to allow each correct alternative to be compared to more non-exemplars. To improve intra-class generalization, additional examples, involving other dependent and independent variables, and participants, could have been used.

In relation to the second topic, this study used a multiple probe design with three probes before the introduction of the intervention. The more probes that are conducted before the introduction of an independent variable, the more information the experimenter has on the participants' initial repertoire to make a more informed decision on whether to expose the participant to a given intervention. When data show enough increase in performance during baseline probes, the intervention probably is not necessary. The three baseline probes showed that percentage of correct responses did not increase with exposure to new probes: all baseline scores were below 20% in all probes, for all participants. Probes provided enough data for a visual differentiation between baseline and post-intervention performance. This allowed the inference that the introduction of the independent variable resulted in an increase in the percentage of correct responses. General performance (Table 1) went from about 2% in baseline scores up to about 30% correct responses after Part 1 teaching sessions (Pre-experimental Designs), up to about 60% after Part 2 teaching sessions (Quasi-experimental Designs), and up to about 90% after Part 3 teaching sessions (True experimental Designs).

Third, in addition to symmetry and transitivity tests in a multiple-choice format, open-ended questions were used as probes. The decision to use open-ended questions as probes derived from two sources. The first source was the pilot data (see Appendix Q). These data showed that multiple-choice tests containing relations that would be taught and relations that were suppose to emerge, led to improved performance, without immediate feedback. The pilot participant reported that answers to the first questions were contained in the last questions of the probes, especially those questions that involved examples and definitions. The second source was Walker et al.'s (2010) study in which the authors used open-ended questions to evaluate emergence of topography-based responses when a selection-based teaching format was used. In the present study, all participants emitted correct responses in the open-ended probes after teaching sessions (i.e., they emitted written responses after being directly taught to emit selection-based responses in the multiple-choice tasks). In addition, all participants provided new examples that included all the items requested in the open-ended questions (see Appendix I for the open-ended questions). Probably, generalization from one type of responding to the other was due to the fact that the teaching session content was developed and organized based on the same prompts that were presented in the open-ended questions; there were "sufficient stimulus components occurring in common in both training" and probes (Stokes & Baer, 1977, p. 360). Future studies should investigate this question on common stimuli/prompts further by presenting open-ended questions based on similar prompts *versus* open-ended questions with different prompts.

Another aspect of data on selection-based versus topography-based behavior that should be emphasized, as stated in the introduction, is the fact that multiple-choice questions encompass forced-choice tasks. These tasks contain exemplars and non-exemplars that can provide prompts

for correct responses. As discussed by Walker et al., (2010), emitting correct responses in selection-based tasks does not imply correct responses in topography-based tasks. Participants in this study emitted correct responses in both types of tasks. This is important because, as argued by Walker et al., (2010), topography-based responses better reflect situations that are more commonly found "in everyday life" (Walker et al., 2010, p. 616), thus they are more likely to be socially relevant when compared to selection-based responses. Future studies should further examine not only the effects of multiple-choice tasks on topography-based behaviors, but also the effects of multiple-choice questions when no immediate feedback about correct or incorrect responses is provided. The delayed emergence literature has shown that after teaching some conditional discriminations, new relations emerge after repeated testing (Sidman, 1994; Sidman, Kirk, & Willson-Morris, 1985). It would be interesting to investigate the effects of repeated testing alone on the emergence of conditional relations (without any prior teaching with feedback).

In relation to the fourth topic, even though the steps taken to create the key features and key content that served as stimuli were not exhaustive, they can be used by researchers and teachers when they are making decisions on what is important to be taught. Instructors do not necessarily consult several books, articles, web pages, among other resources, before designing course materials. However, several exemplars and non-exemplars should be consulted before course materials are developed, since overlap in key features and in key content are good indicators of what is considered important regarding a given subject matter (Tiemann & Markle, 1990). Additionally, instructors might not attend to the fact that the key content to be taught should match objective, measurable instructional objectives. If content and instructional goals do not match, students will probably not meet the goals (Markle, 1969, 1990). Another issue that

should be underscored is the fact that several educators might provide readings for their students; however, as highlighted by Markle (1975), Mager (1997), and Vargas (2009), among others, readings are just another method of providing information to the student. Without instructional programming that allows (a) several opportunities to respond overtly, (b) immediate feedback; and (c) prompts in the form of exemplars and non –exemplars to avoid errors and facilitate abstraction, it should not be expected that student responses will match what was specified in the instructional objectives.

Fifth, this study was conducted totally in online settings. One of the most important features of an online course is time and space flexibility: participants, students, and instructors can access the links and results from anywhere, at any time. Having deadlines and contingencies in place is important to assure that students will master the content in a given time period, however - as long as deadlines and criteria are met – tasks can be completed in a flexible manner. In addition, as stated in the introduction, the use of online instruction and hybrid courses in higher education is increasing (Garrison & Kanuka, 2004; Seaman, 2011; Young & Duhaney, 2008). Several universities such as the University of Chicago, the University of Arizona, and Stanford University provide online instruction not only across cities and states, but also across countries (i.e., they are involved in the internationalization[3] process). Practices such as the creation of virtual campuses and the internationalization of higher education have been contributing to the wide spread use of online instruction. If the educational practices involved in

[3] Internationalization is defined as "the policies and practices undertaken by academic systems and institutions—and even individuals—to cope with the global academic environment" (Altbach & Knight, 2007, p. 290).

online education are to be effective, procedures that have been shown to be effective should be among the choices made by instructional designers.

In addition to the online feature of this study, teaching sessions were presented in English and probes could be answered in Portuguese, English, or using both languages. Online instruction that requests overt, topography-based responding from students might benefit from allowing students to answer questions in their native language - if foreign language skills are not at stake. Being able to answer in one's own language can prevent response errors that are derived from responding in a language that students have not mastered yet. In summary, the online sessions gave time, space, and language flexibility for participants and the experimenter. External validity and generality should be tested using similar sessions with other participants, settings, and languages.

In relation to the final topic, a checklist that contained assessment questions for the article review was used to evaluate whether teaching conditional discriminations on research designs could result in better overall performance in the checklist. As predicted, since there was no programmed teaching for the checklist items, there was no improvement in performance. Different from emergent relations and generalization to probes, which had stimuli in common with the teaching sessions, questions in the checklist did not contain similarities with the experimental phases. The checklist was used as a demonstration that, if certain stimuli are to control behavior, programming for stimulus control must occur. If people are to learn how to review a research study using items such as the ones presented in the checklist, teaching must be programmed to include all the items.

Limitations

One key study limitation refers to the type of contingencies that were employed. The experimenter was located in a different country (USA) from the students (Brazil) and was not a faculty member in any of the three Brazilian universities where the students attended school. There was no extra credit given contingent upon completion of the study in a timely manner. In addition, in Brazil, it is illegal to pay participants to take part in any study. The only incentives that could be used were items that could be sent either electronically or by international mail (see description of the items under "Participants"**,** p. 12). Researchers and educators trying to implement this type of online tutorial can probably gather more data in less time if they can have control (or influence) over environmental stimuli such as grades.

Another limitation that needs to be highlighted is that, even though teaching time for each design was very similar, when considering the six open-ended probes, participants spent more time writing about pre experimental designs than writing about the other two types of designs (pre-experimental designs teaching sessions were presented first). If the ultimate goal of the study was to have students apply the concepts by designing a well-controlled experiment, more emphasis would have to be given to the true experimental designs. Because the purpose of the present study was not to have student design a controlled experiment, Tiemann and Markle's (1990) approach was used: start with the simplest discrimination and gradually add on complexity until the most complex stimuli (true experimental designs) were presented.

One other limitation: when errors occurred, the program automatically presented a new set of six trials to the participant. A box that provided feedback on how many questions were left changed the display when participants were directed to the new block of trials. Two participants

reported that the box confused them when the re-direction happened. Researchers using Adobe Captivate® to collect data might consider not using the feedback box.

There were also problems specifically related to Adobe Connect® and Adobe Captivate®. The first problem was that, when using Adobe Connect®, if the end users (participants) have a problem with their internet connection signal, all data related to the session they are in will be lost. For John, for example, two open-ended questions were lost during Probe 6. To avoid requesting the participant to answer the 108 items all over again, the experimenter sent the two lost questions by e-mail. Also, while Mary was participating in the first teaching session, her internet signal stopped. The four last trials were not presented to her. She went on to the transitivity session (since her performance was 100% correct responses in the 14 trials she was exposed to) and most errors in this session were in the conditional relations that were not presented in full. She had to be exposed to a new teaching session before criterion was met in the emergent relations session for Part 1. When Mary was going through probes 5 and 6, her internet also faded in and out and the sessions just stopped. After re-starting the sessions more than 3 times, the participant wrote to the experimenter who provided the questions through e-mail. This change in the media (format of stimulus presentation) could also have influenced the participant's performance. The problems with the programs also represent a threat to the fidelity of implementation: despite the automated intervention delivery and data collection system, this problem in session delivery disrupted the implementation of the experimental phases. In addition, the programs (a) did not (and do not) record time to complete each trial and/or session, (b) demanded much time for mastery and (c) required the ability to program advanced variables when using Adobe Captivate® to set up performance criteria. In summary, unless all internet connections are reliable and full support for program use is provided to the researcher/educator,

it is recommended that additional instruments be used for data collection. Additional instruments can not only provide data, but also help to measure the fidelity of online program implementation. The use of paper-and-pencil formats, as suggested by Walker et al., (2010), might be an alternative, for instructors who are teaching in a traditional classroom. For educators looking to use hybrid courses as part of their classes, once the sessions are tested with an additional instrument to assure fidelity of implementation, the sessions can be used with different students, in different settings, and even for different courses, if they have common content to be delivered.

Another possible limitation is the fact that even though participants were instructed to send their questionnaires to the second observer, so they could not be identified, they sent it directly to the experimenter. There is always the possibility that participants did not provide negative feedback, because they were identified. However, since the participants addressed negative aspects of the tutorial (not only on the questionnaire, but also throughout the emails), their opened identification might not have caused bias in their feedback.

In addition, it is important to highlight that stimuli presented in one teaching part was never presented in other teaching parts. To assure maintenance of mastery, earlier material should be presented while new material is introduced (Markle, 1990).

One additional limitation refers to the number of studies that were retrieved when the literature review was conducted. As highlighted in the introduction, stimulus equivalence is broad in scope. Addressing all studies that are related to the topic of the present study would necessitate in a systematic review of the literature that would involve all studies that provided the foundations for the application of stimulus equivalence, such as de Rose et al., (1996), Sidman (1971), and Sidman and Cresson (1973), among many others. Besides the number of studies that

would have to be included, the decision on what should be included and excluded could also be questioned by different experts, since different people might value different studies. Additionally, conducting an unbiased systematic review would require objectively described keywords, inclusion and exclusion criteria, and human and other resources necessary to conduct this type of review (Clark & Castro, 2002; Pai et al., 2004). Considering all these factors, the scope of the present thesis would not allow a systematic review. Thus, it was decided to conduct a literature review that could be replicated if one was to use the same criteria as the ones laid out in the introduction. The criteria used narrowed the scope of studies to be described in the introduction. The three studies that were described do not represent the totality of studies that could have been included. However, the criteria described in the introduction can be used by other researchers in the field and should yield the similar results.

Conclusion

In recent years, the literature on the use of stimulus equivalence in higher education settings has been growing. There are still many questions to be answered on how broad this use can be. Examples of unanswered questions are related to: (a) making equivalence technology more user-friendly (so people do not need to have an extensive background in stimulus equivalence to use the technology); (b) investigating which formats might be appropriate for different higher education settings (in classroom *versus* online instructions); and most importantly, how can we change contingencies that control educators' behavior in the higher education setting so that they will dedicate more of their time to "changing student behavior and to gathering empirical evidence to support these attitudes" (Markle, 1969, p. vi).

Despite the limitations of the study, it advances the area of stimulus equivalence by: (a) using a multiple probe design; (b) requiring topography-based responses; (c) involving a content analysis, based on several examples; (d) conducting the study online; and (e) presenting the sessions in English, but allowing topography-based responses to be in the participants' language of choice.

In addition, there are only few studies using stimulus equivalence that have accessed social validity measures (e.g., Fienup & Critchfield, 2011). As described above, participants were encouraged to comment on stimuli and phases at any time. All participants made spontaneous comments and both participants who finished the study, answered the structured questionnaire. All suggestions, varying from comments on the open-ended questions to "confusion" generated by the feedback questions will be incorporated in future versions of the instructional package. Receiving high social validity scores from typically developing adults suggests that similar tutorials might be used in higher education settings.

In relation to the broader context, as highlighted in the introduction, higher education is under a lot of scrutiny. Protests, like the one at Purchase College (http://www.eduinreview.com/blog/2011/10/purchase-college-students-protest-overpriced-tuition/) have been discussing the fact that people make debts to pay for an education that does not guarantee jobs in the future. Even though effective online tutorials can take a long time to be programmed, once they are designed and tested, their use can decrease classroom time and students can work on important skills until they reach mastery without further costs (or even lower total costs).

References

Adams, G., & Engelmann, S. (1996). *Research in direct instruction: 25 years beyond DISTAR.* Seattle,WA: Educational Achievement Systems.

Adcock, A., Merwin, R., Wilson, K., Drake, C., Tucker, C., & Elliott, C. (2010). The problem is not learning: Facilitated acquisition of stimulus equivalence classes among low-achieving college students. *The Psychological Record, 60*(1), 43–56.

ALLPsych Online. (2002, March 21, 2004). Research methods, Chapter 5: experimental design Retrieved March, 2011, from http://allpsych.com/researchmethods/preexperimentaldesign.html

Almeida-Verdu, A., Huziwara, E., de Souza, D., De Rose, J., Bevilacqua, M., Lopes, J., Jr., . . . McIlvane, W. (2008). Relational learning in children with deafness and cochlear implants. *Journal of the Experimental Analysis of Behavior, 89*(3), 407-424.

Altbach, P., & Knight, J. (2007). The internationalization of higher education: Motivations and realities. *Journal of Studies in International Education, 11*(3-4), 290-305. doi: 10.1177/1028315307303542

Alves, K., Kato, O., Assis, G., & Maranhão, C. (2007). Leitura recombinativa em pessoas com necessidades educacionais especiais: Análise do controle parcial pelas sílabas. *Psicologia: Teoria e Pesquisa, 23,* 387-398.

Araújo, P., & Ferreira, P. (2008). Ensinando subtração para pessoas com deficiência mental com base em relações de equivalência de estímulos. *Psicologia: Teoria e Pesquisa, 24*, 313-322.

Arntzen, E., Halstadtro, L., Bjerke, E., & Halstadtro, M. (2010). Training and testing music skills in a boy with autism using a matching-to-sample format. *Behavioral Interventions, 25*(2), 129-143. doi: 10.1002/bin.301

Barlow, D., Nock, M., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior for change* (3rd ed.). Boston, MA: Pearson Education, Inc.

Bonfiglio, C. M., Daly, E. J., 3rd, Martens, B. K., Lin, L. H., & Corsaut, S. (2004). An experimental analysis of reading interventions: Generalization across instructional strategies, time, and passages. *Journal of Applied Behavior Analysis, 37*(1), 111-411. doi: 10.1901/jaba.2004.37-111

Boykin, R., & Nelson, R. (1981). The effects of instructions and calculation procedures on observers' accuracy, agreement, and calculation correctness. *Journal of Applied Behavior Analysis, 14*(4), 479-489

Braxton, J., Bray, N., & Berger, J. (2000). Faculty teaching skills and their influence on the college student departure process. *Journal of College Student Development, 41*(2), 215-227.

Brennan, J., & Teichler, U. (2008). The future of higher education and of higher education research. *Higher Education, 56*(3), 259-264. doi: 10.1007/s10734-008-9124-6

Brino, A., Assumpção, A., Campos, R., Galvão, O., & McIlvane, W. (2010). Cebus cf. apella exhibits rapid acquisition of complex stimulus relations and emergent performance by exclusion. *Psychology & Neuroscience, 3*, 209-215.

Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand-McNally.

Chase, P., Johnson, K., & Sulzer-Azaroff, B. (1985). Verbal relations within instruction: Are there subclasses of the intraverbal? *Journal of the Experimental Analysis of Behavior, 43*(3), 301-313.

Clark, O., & Castro, A. (2002). Searching the Literatura Latino Americana e do Caribe em Ciências da Saúde (LILACS) database improves systematic reviews. *International Journal of Epidemiology, 31*(1), 112-114. doi: 10.1093/ije/31.1.112

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed. ed.). Oxford, UK: Routledge Publishers.

Connections, R. (2010). Pre-experimental designs. *Child care & early education*, from http://www.researchconnections.org/childcare/datamethods/preexperimental.jsp

Connell, P., & Thompson, C. (1986). Flexibility of single-subject experimental designs. Part III: Using flexibility to design or modify experiments. *Journal of Speech and Hearing Disorders, 51*(3), 214-225.

Creswell, J. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Los Angles, CA: Sage.

Critchfield, T., & Fienup, D. (2010). Using stimulus equivalence technology to teach statistical inference in a group setting. *Journal of Applied Behavior Analysis, 43*(4), 763-768.

da Silva, W., de Souza, D., de Rose, J., Lopes, J., Jr , Bevilacqua, M., & McIlvane, W. (2006). Relational learning in deaf children with cochlear implants. *Experimental Analysis of Human Behavior Bulletin, 24*, 1-8.

Davis, D., Bostow, D., & Heimisson, G. (2007). Strengthening scientific verbal behavior: An experimental comparison of progressively prompted and unprompted programmed instruction and prose tutorials. *Journal of Applied Behavior Analysis, 40*(1), 179-184.

de Rose, J., de Souza, D., & Hanna, E. (1996). Teaching reading and spelling: Exclusion and stimulus equivalence. *Journal of Appled Behavior Analysis, 29*(4), 451–469.

de Rose, J., de Souza, D., Rossito, A., & de Rose, T. (1992). Stimulus equivalence and generalization in reading after matching to sample by exclusion *Understanding verbal relations* (pp. 69-82). Reno, NV: Context Press.

de Souza, D., de Rose, J., Faleiros, T., Bortoloti, R., Hanna, E., & McIlvane, W. (2009). Teaching generative reading via recombination of minimal textual units: A legacy of verbal behavior to children in Brazil. *International Journal of Psychology and Psychological Therapy*, *9*(1), 19-44.

Dermer, M., Lopez, S., & Messling, P., III. (2009). Fluency training a writing skill: Editing for concision. *The Psychological Record, 59*(1), 3-18.

Ebert-May, D. (n. d.). Classroom assessment techniques: Scoring rubrics Retrieved August, 2011, from http://www.flaguide.org/cat/rubrics/rubrics1.php

Eduers.com, T. (2009). What Does SAT Stand for? Retrieved October 12, 2011, from http://www.eduers.com/sat/whatissat.htm

Eyre, H. (2007). Keller's Personalized System of Instruction: Was it a fleeting fancy or is there a revival on the horizon? *Behavior Analyst Today, 8*(3), 317-324.

Fienup, D., Covey, D., & Critchfield, T. (2010). Teaching brain-behavior relations economically with stimulus equivalence technology. *Journal of Applied Behavior Analysis, 43*(1), 19-33. doi: 10.1901/jaba.2010.43-19

Fienup, D., & Critchfield, T. (2010). Efficiently establishing concepts of inferential statistics and hypothesis decision making through contextually controlled equivalence classes. *Journal of Applied Behavior Analysis, 43*(3), 437-462. doi: 10.1901/jaba.2010.43-437

Fienup, D., & Critchfield, T. (2011). Transportability of equivalence-based programmed instruction: Efficacy and efficiency in a college classroom. *Journal of Applied Behavior Analysis, 44*(3), 435-450. doi: 10.1901/jaba.2011.44-435

Fienup, D., & Dixon, M. (2006). Acquisition and maintenance of visual-visual and visual-olfactory equivalence classes. *European Journal of Behavior Analysis, 7*(1), 87 - 98.

Fox, E. (2004). The Personalized System of Instruction: A flexible and effective approach to mastery learning. In D. Moran & R. Malott (Eds.), *Evidence-based educational methods* (pp. 201-221). San Diego: Elsevier Academic Press.

Fox, E., & Sullivan, H. (2007). Comparing strategies for teaching abstract concepts in an online tutorial. *Journal of Educational Computing Research 37*(3), 307-330.

Fraenkel, J., Wallen, N., & Hyun, H. (2011). *How to design and evaluate research in education* (8 edition ed.). Columbus, OH: McGraw-Hill.

Fredrick, L., & Hummel, J. (2004). Reviewing the outcomes and principles of effective instruction. In D. Moran & R. Malott (Eds.), *Evidence-based educational methods* (pp. 9-22). San Diego: Elsevier Academic Press.

Freedman, R. (1994). *Open-ended questioning: A handbook for educators*. Menlo Park, CA: Innovative Learning TM, Addison-Wesley Alternative Publishing Group.

Garrison, D., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education, 7*(2), 95-105. doi: 10.1016/j.iheduc.2004.02.001

Green, G., & Saunders, R. (1998). Stimulus equivalence. In K. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229–262). New York: Plenum.

Guercio, J., Podolska-Schroeder, H., & Rehfeldt, R. (2004). Using stimulus equivalence technology to teach emotion recognition to adults with acquired brain injury. *Brain Injury, 18*(6), 593-601. doi: doi:10.1080/02699050310001646116

Hacker, A., & Dreifus, C. (2010). *Higher education: How colleges are wasting our money and failing our kids---and what we can do about it*. New York: St. Martin's Press.

Hayes, L., Tilley, K., & Hayes, S. (1988). Extending equivalence and membership to gustatory stimuli. *The Psychological Record, 38*, 473-482.

Heiman, M., & Slomianko, J. (1998). *Learning to Learn: Thinking skills for the 21st century*. Somerville: Learning to Learn, Inc.

Horner, R., & Baer, D. (1978). Multiple-probe technique: A variation on the multiple baseline. *Journal of Applied Behavior Analysis, 11*(1), 189-196.

Hubner, M., Gomes, R., & McIlvane, W. (2009). Recombinative generalization in minimal verbal unit-based reading instruction for pre-reading children. *Experimental Analysis of Human Behavior Bulletin, 27*, 11-17.

Immerwahr, J., & Foleno, T. (2000). Great expectations: How the public and parents—White, African American and Hispanic— view higher education. In Public Agenda (Ed.). New York: The National Center for Public Policy and Higher Education.

Immerwahr, J., Johnson, J., Ott, A., & Rochkind, J. (2010). Squeeze play 2010: Continued public anxiety on cost, harsher judgments on how colleges are run. In The National Center for Public Policy and Higher Education and Public Agenda (Ed.).

Jaehnig, W., & Miller, M. (2007). Feedback types in programmed instruction: A systematic review. *The Psychological Record, 57*(2), 219-232.

Kastak, C., & Schusterman, R. (2002). Sea lions and equivalence: Expanding classes by exclusion. *Journal of the Experimental Analysis of Behavior, 78*(3), 449-465. doi: 10.1901/jeab.2002.78-449

Kish, L. (2005). Designs for comparison. In B. Shube (Ed.), *Statistical Design for Research* (pp. 68-72). Hoboken, NJ: John Wiley & Sons, Inc.

Laird, T., Chen, D., & Kuh, G. (2008). Classroom practices at institutions with higher-than-expected persistence rates: What student engagement data tell us. *New Directions for Teaching and Learning, 2008*(115), 85-99. doi: 10.1002/tl.327

Luciano, C., Gomez Becerra, I., & Rodriguez Valverde, M. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *Journal of Experimental Analysis of Behavior, 87*(3), 349-365.

Lynch, D., & Cuvo, A. (1995). Stimulus equivalence instruction of fraction-decimal relations. *Journal of Applied Behavior Analysis, 28*(2), 115-126.

Mager, R. (1997). *How to turn learners on... without turning them off: Ways to ignite interest in learning* (3 ed.). Atlanta, GA: The Center for Effective Performance.

Markle, S. (1969). *Good frames and bad: Grammar of frame writing* (2nd ed.). New York: John Wiley & Sons Inc.

Markle, S. (1975). They teach concepts, don't they? *Educational Researcher, 4*(6), 3-9.

Markle, S. (1990). *Designs for Instructional Designers* (3 ed.). Chicago, IL: Stipes Publishing, LLC.

Markle, S., & Tiemann, P. (1970). "Behavioral" analysis of "cognitive" content. *Educational Technology, 10*(1), 41-45.

Marques, L., & Galvão, O. (2010). Development and evaluation of a software for programmed instruction [Desenvolvimento e avaliação de um sistema informatizado de instrução programada]. *Acta Comportamentalia, 18*(3), 347-360.

McAtamney, G., & Annett, J. (2009). Learning to associate compatible and incompatible pictures with food and non-food odours, within a stimulus equivalence paradigm. *Food Quality and Preference, 20*(3), 259-267. doi: 10.1016/j.foodqual.2008.11.001

McDonald, J., Yanchar, S., & Osguthorpe, R. (2005). Learning from programmed instruction: Examining implications for modern instructional technology. *Educational Technology Research and Development, 53*(2), 84-98. doi: 10.1007/bf02504867

Merwin, R., & Wilson, K. (2005). Preliminary findings on the effects of self-referring and evaluative stimuli on stimulus equivalence class formation. *The Psychological Record, 55*(4), 561-575.

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

Moore, R. (1963). Programmed instruction: The engineering of education. *IEEE Transactions on Education, 6*(2), 60-64.

Moran, D., & Malott, R. (2004). *Evidence-based educational methods: Advances from the behavioral sciences*. New York: Academic Press.

Murphy, R., & Bryan, A. (1980). Multiple-baseline and multiple-probe designs: Practical alternatives for special education assessment and avaluation. *The Journal of Special Education, 14*(3), 325-335. doi: 10.1177/002246698001400306

Pai, M., McCulloch, M., Gorman, J., Pai, N., Enanoria, W., Kennedy, G., . . . Colford Jr, J. (2004). Systematic reviews and meta-analyses: An illustrated, step-by-step guide. *The National Medical Journal of India, 17*(2), 86-95.

Pascarella, E., Salisbury, M., & Blaich, C. (2011). Exposure to effective instruction and college student persistence: A multi-institutional replication and extension. *Journal of College Student Development 52*(1), 4-19.

Pilgrim, C., Jackson, J., & Galizio, M. (2000). Acquisition of arbitrary conditional discriminations by young normally developing children. *Journal of the Experimental Analysis of Behavior, 73*(2), 177-193. doi: 10.1901/jeab.2000.73-177

Polson, D., & Parsons, J. (2000). Selection-based versus topography-based responding: An important distinction for stimulus equivalence? *The Analysis of Verbal Behavior, 17*, 105-128.

Ribeiro, D., Pascualon, J., Sella, A., Bandini, C., & de Souza, D. (2009). Avaliação de um procedimento de ensino de categorias estruturais de histórias [Evaluation of a procedure for structural categories of stories]. *Revista Semestral da Associação Brasileira de Psicologia Escolar e Educacional, 13*(2), 303-314.

Roberts, L., Coverdale, J., Edenharder, K., & Louie, A. (2004). How to review a manuscript: A "down-to-earth" approach. *Academic Psychiatry, 28*(2), 81-87. doi: 10.1176/appi.ap.28.2.81

Saunders, R., Saunders, K., Kirby, K., & Spradlin, J. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior, 50*(1), 145–162.

Seaman, J. (2011). Online learning trends in private-sector colleges and universities. In P. L. Solutions (Ed.), *Sloan Survey of Online Education*. Boston, MA.

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

Sidman, M. (1986). Functional analysis of emergent verbal classe. In T. Thompson & M. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213-245). Hillsdale, NJ: Erlbaum.

Sidman, M. (1990). Equivalence relations: Where do they come from? In D. Blackman & H. Lejeune (Eds.), *Behavior analysis in theory and practice: Contributions and controversies* (pp. 93-114). Hillsdale, NJ: Erlbaum.

Sidman, M. (1994). *Equivalence relations and behavior: A research history*. Boston, MA: Authors Cooperative.

Sidman, M., & Cresson, O., Jr (1973). Reading and crossmodal transfer of stimulus equivalences in severe retardation. *American Journal of Mental Deficiency, 77*(5), 515-523.

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional discrimination procedures. *Journal of the Experimental Analysis of Behavior, 43*(1), 21-42.

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

Skinner, B. F. (1953). The controlling environment *Science and human behavior* (pp. 129-140). New York: The Free Press.

Skinner, B. F. (1957/2002). The tact *Verbal behavior* (pp. 81-146). Cambridge, MA: B. F. Skinner Foundation.

Stokes, T., & Baer, D. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*(2), 349-367.

Taylor, I., & O'Reilly, M. (2000). Generalization of supermarket shopping skills for individuals with mild intellectual disabilities using stimulus equivalence training. *The Psychological Record, 50*, 49-62.

Tiemann, P., & Markle, S. (1990). *Analyzing instructional content* (4th ed.). Champaign, IL: Stipes Publishing Company.

Toussaint, K., & Tiger, J. (2010). Teaching early braille literacy skills within a stimulus equivalence paradigm to children with degenerative visual impairments. *Journal of Applied Behavior Analysis, 43*(2), 181-194. doi: 10.1901/jaba.2010.43-181

Twyman, J., Layng, T., Stikeleather, G., & Hobbins, K. (2005). A non-linear approach to curriculum design: The role of behavior analysis in building an effective reading program. In W. L. H. e. al (Ed.), *Focus on behavior analysis in education* (Vol. 3, pp. 55-68). Upper Saddle River, NJ: Merrill/Prentice Hall.

U.S. Census Bureau. (2011a). *Average earnings of year-round, full-time workers by educational attainment: 2009* Retrieved from http://www.census.gov/compendia/statab/cats/education.html.

U.S. Census Bureau. (2011b). *Unemployed and unemployment rates by educational attainment, sex, race, and hyspanic origin: 2000 to 2010*. U.S. Census Bureau.

Urcuioli, P. (2008). Associative symmetry, antisymmetry, and a theory of pigeons' equivalence-class formation. *Journal of Experimental Analysis of Behavior, 90*(3), 257-282.

Vargas, J. (2009). *Behavior analysis for effective teaching* (1 ed.). New York, NY: Routledge.

Walker, B., Rehfeldt, R., & Ninness, C. (2010). Using the stimulus equivalence paradigm to teach course material in an undergraduate rehabilitation course. *Journal of Applied Behavior Analysis, 43*(4), 615.

Whitehurst, G., Falco, F., Lonigan, C., Fischel, J., DeBaryshe, B., Valdez-Menchaca, M., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24*(4), 552-559.

Wrench, J. (2009). Basic research design: Understanding and using the designing a research project worksheet Retrieved February, 2011, from http://www.slideshare.net/JasonSWrench/basic-research-design-presentation

Young, A., & Duhaney, D. (2008). Hybrid learning and the principles of good practice in undergraduate education. *International Journal of Instructional Technology and Distance Learning 5*(11). Retrieved from International Journal of Instructional Technology and Distance Learning website: http://www.itdl.org/Journal/Nov_08/article04.htm

Yu, C., & Ohlund, B. (2010). Threats to validity of research design, from http://www.creative-wisdom.com/teaching/WBI/threat.shtml

Appendix A – Stimulus Equivalence Literature Review By Decades

Decades	Number of articles or book chapters located	Sample of journals and books pulled
1890-1899[a]	0	N/A
1900-1909 [a]	0	N/A
1910-1919	1	American Journal of Psychology
1920-1929	0	N/A
1930-1939	6	Psychological Review Journal of Experimental Psychology Pedagogical Seminary Journal of Genetic Psychology
1940-1949	57	Psychological Review Journal of Experimental Psychology American Journal of Psychology
1950-1959	63	Journal of Experimental Psychology Psychological Review American Journal of Psychology
1960-1969	177	Psychological Review Journal of Comparative and Physiological Psychology Journal of Verbal Learning and Verbal Behavior
1970-1979	253	Journal of Cross-Cultural Psychology Science Psychological Review
1980-1989	295	Journal of the Experimental Analysis of Behavior Analysis and Intervention in Developmental Disabilities Trends in Neurosciences
1990-1999	894	Dialogues on verbal behavior Behaviour analysis in theory and practice: Contributions and controversies Journal of Experimental Analysis of Behavior
2000-2009	1.180	Neuroreport Journal of the Experimental Analysis of Behavior Developmental Review
2010-2011	344	The Psychological Record Journal of Applied Behavior Analysis Journal of Cognitive Neuroscience

[a] The criterion to stop looking for earlier papers was: no match found in two consecutive decades.

Appendix B – Recruitment E-mail

Eu estou desenvolvendo um tutorial online para ensinar métodos de pesquisa para alunos de graduação.

Estou procurando bons alunos que leiam inglês e que tenham interesse em pesquisa. Abaixo segue uma descrição mais detalhada sobre o projeto.

O estudo terá início em cerca de 15 dias e durará um mínimo de 12 dias, pois o programa possui 12 sessões. Se você realizar uma sessão por dia, terminará seu projeto em 12 dias. O programa possui 12 sessões. Cada sessão dura em média 1 hora, às vezes menos, às vezes um pouco mais. Quando você tiver um horário disponível, é só me avisar e trabalharemos dentro destes horários.

Ao final do estudo você receberá uma cópia de um dos seguintes livros: 1) Cummulative Record: definite edition; Technology of Teaching; verbal behavior; Principles of Psychology; Schedules of Reinforcement. Você escolherá o livro que desejar.

Você também receberá um certificado no qual constará o conteúdo do tutorial.

Adicionalmente, enviarei um documento com as informações fornecidas no tutorial para que você possa retomar estas informações, quando precisar desenvolver suas próprias pesquisas.

Além disso, se terminar o estudo em 12 dias, receberá um e-mail com artigos relacionados à elementos de pesquisa em Análise do Comportamento.

Para poder participar da pesquisa, você deverá primeiramente, ler o termo de consentimento em anexo, assina-lo, escanea-lo e envia-lo por e-mail (carolsella@yahoo.com.br).

Como o treinamento será feito em inglês, mas você poderá responder as questões em português, você precisará passar por um teste de compreensão de leitura em inglês via Skype. Se

você for (ou não) elegível para participar da pesquisa, darei seu resultado por e-mail e marcaremos para iniciar a pesquisa.

Por favor, se for participar da pesquisa não comente com seus colegas e/ou amigos, pois isto pode trazer vieses para o resultado da pesquisa. Após a pesquisa terminar, marcaremos uma reunião para conversarmos sobre sua experiência e você poderá dar sua opinião acerca do que mudaria no tutorial.

Obrigada por considerar o projeto.

Abraços,

Carol

Appendix C – HSCL Approval Letter

KU RESEARCH & GRADUATE STUDIES
The University of Kansas

6/9/11
HSCL #19431

Ana Sella
ABS
4098 Dole Ctr.

The Human Subjects Committee Lawrence Campus (HSCL) has received your response to its expedited review of your research project

19431 Sella/White (ABS) Effects of a Research Primer on Undergraduates' Performance in Knowledge Tasks on Research Methods

and approved this project under the expedited procedure provided in 45 CFR 46.110 (f) (7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies. As described, the project complies with all the requirements and policies established by the University for protection of human subjects in research. Unless renewed, approval lapses one year after approval date.

The Office for Human Research Protections requires that your consent form must include the note of HSCL approval and expiration date, which has been entered on the consent form(s) sent back to you with this approval.

1. At designated intervals until the project is completed, a Project Status Report must be returned to the HSCL office.
2. Any significant change in the experimental procedure as described should be reviewed by this Committee prior to altering the project.
3. Notify HSCL about any new investigators not named in original application. Note that new investigators must take the online tutorial at http://www.rcr.ku.edu/hscl/hsp_tutorial/000.shtml
4. Any injury to a subject because of the research procedure must be reported to the Committeeimmediately.
5. When signed consent documents are required, the primary investigator must retain the signed consent documents for at least three years past completion of the research activity. If you use a signed consent form, provide a copy of the consent form to subjects at the time of consent.
6. If this is a funded project, keep a copy of this approval letter with your proposal/grant file.

Please inform HSCL when this project is terminated. You must also provide HSCL with an annual status report to maintain HSCL approval. Unless renewed, approval lapses one year after approval date. If your project receives funding which requests an annual update approval, you must request this from HSCL one month prior to the annual update. Thanks for your cooperation. If you have any questions, please contact me.

Sincerely,

Jan Butin
Associate Coordinator
Human Subjects Committee Lawrence

cc: Glen White

Human Subjects Committee Lawrence
Youngberg Hall | 2385 Irving Hill Road | Lawrence, KS 66045 | (785) 864-7429 | Fax (785) 864-5049 | www.rcr.ku.edu/hscl

Appendix D - Initial Interview

Address:

Name:

Date of interview:

Age:

Gender: ( ) male ( ) female

Primary language:

Year in higher education program:

Major:

Course from a social or behavioral science program:

Do you have advanced knowledge on research methods?

Do you have any disabilities that would require material accommodation?

How did you hear about the study?

What interested you about the study?

What do you expect to gain by participating in this study?

Appendix E – B. F. Skinner Books

Skinner, B. F. (1999). *Cumulative record: Definite edition*. Cambridge, MA: B. F. Skinner Foundation. (Original work published 1959)

Skinner, B. F. (2003). *The technology of teaching*. Cambridge, MA: B. F. Skinner Foundation. (Original work published 1968)

Skinner, B. F. (2002). *Verbal behavior*. Cambridge, MA: B. F. Skinner Foundation. (Original work published 1957)

Skinner, B. F. (1995). *Principles of psychology*. Cambridge, MA: B. F. Skinner Foundation. (Original work published 1950)

Skinner, B. F. (1997). *Schedules of reinforcement*. Cambridge, MA: B. F. Skinner Foundation. (Original work published 1957)

Appendix F – Experimental Stimuli

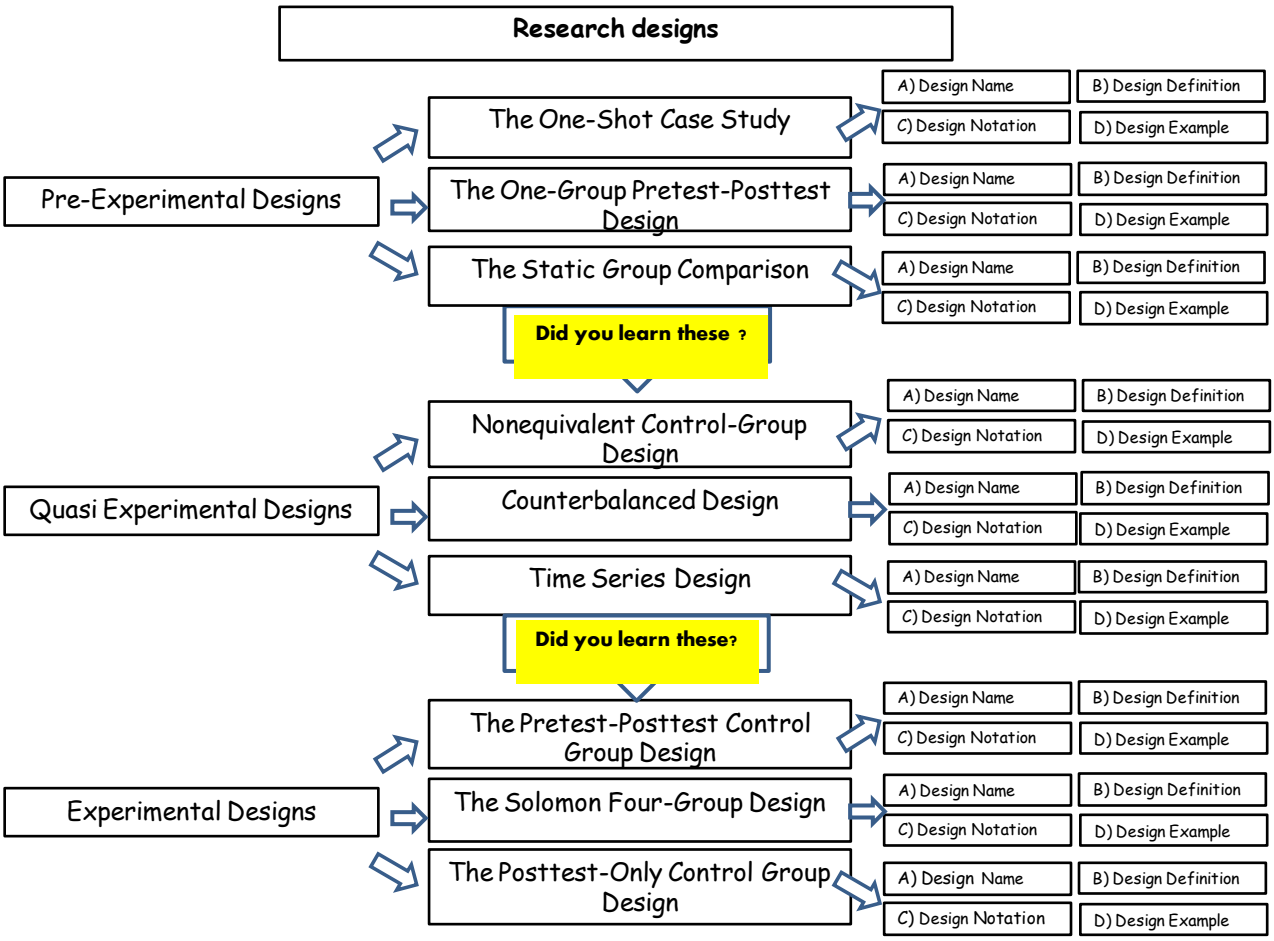
	Design name (set A)	Design explanation or definition (set B) What features define the design when the dependent variable is measured; when is the independent variable is introduced How many groups are needed and if there is a control group; if there is randomization."	Notation (set C) when the independent variable is introduced when the dependent variable is measured (observed); The group(s); The randomization, if applicable.	Example (set D) i. what is the dependent variable(s) and when it will be measured (observed). ii. what is the independent variable(s) and when it is introduced. iii. who are the participants, how many participants you will have, and if the participants will be divided into groups. iv. will there be randomization?
1	The One-Shot Case Study	The dependent variable is **measured (O) only after the independent variable (X)** is introduced. The **independent variable** is introduced **before the measure** of the dependent variable. Only **one group** is needed and there is **no control group**. There is no randomization.	X O1	The dependent variable is heart rate and it is measured after the independent variable is introduced. The independent variable is jogging and it is introduced before measuring the heart rates. The participants are 25 students who will not be divided into groups. There is no randomization.
2	The One-Group Pretest-Posttest Design	The dependent variable is measured before (O1) and after (O2) the independent variable (X) is introduced. The independent variable is introduced after the first measure of the dependent variable. Only one group is needed and there is no control group. There is no randomization.	O1 X O2	The dependent variable is heart rate and it is measured before and after the independent variable is introduced. The independent variable is jogging and it is introduced after the first measure of heart rates. The participants are 25 students who will not be divided into groups. There is no randomization.
3	The Static Group Comparison	The dependent variable is **measured (O) for both groups only after the independent variable (X)** is introduced to the experimental group. The **independent variable** is introduced to the experimental group **before the measure** of the dependent variable. **Two groups** are needed; **one** is the **control** group. There is **no randomization**.	X O1 — — — — O1	The dependent variable is heart rate and it is measured for both groups after the independent variable is introduced to the experimental group. The independent variable is jogging and it is introduced to the experimental group before heart rates are measured in both groups. The participants are 50 students who will be divided into two groups: experimental group and control group. There is no randomization.

	Design name (set A)	Design explanation or definition (set B)	Notation (set C)	Example (set D)
4	Nonequivalent Control-Group Design	The dependent variable is **measured for both groups before (O1) and after (O2) the independent variable (X)** is introduced to the experimental group. The **independent variable** is introduced to the experimental group **after the first measure** of the dependent variable. **Two groups** are needed; **one** is the **control group**. There is **no randomization**.	O1 __ X __ O2 O1 O2	The dependent variable is heart rate and it is measured for both groups before and after the independent variable is introduced. The independent variable is jogging and it is introduced to the experimental group after the first measure of heart rates. The participants are 50 students who will be divided into two groups: experimental group and control group. There is no randomization.
5	Counterbalanced Design	The dependent variable is measured for all groups (O), after each one of the four independent variables (X1, X2, X3, X4) is introduced for each experimental group. Each independent variable is introduced to all groups but in a different order for each group. Four groups are needed, but there is no "true control group", since the independent variable is introduced for all groups. There is no randomization.	Time 1 Time 2 Time 3 Time 4 Group A X1O X2O X3O X4O Group B X2O X4O X1O X3O Group C X3O X1O X4O X2O Group D X4O X3O X2O X1O	The dependent variable is heart rate and it is measured for all groups, after each independent variable is introduced for the groups. The independent variables can be jogging (X1), swimming (X2), dancing (X3), and walking (X4) and each one of them is introduced to all groups, but in a different order for each group. The participants are one hundred students who will be divided into four groups. There is no randomization.
6	The Multiple Time Series Design	The dependent variable is measured several times (O), for both groups, before and after the independent variable (X) is introduced to the experimental group. The independent variable is introduced to the experimental group after several measures of the dependent variable. Two groups are needed; one is the control group. There is no randomization.	O O O OXO O O O O O O O O O O O	The dependent variable is heart rate and it is measured several times, for both groups, before and after the independent variable is introduced. The independent variable is jogging and it is introduced to the experimental group after several measures of heart rates. The participants are 50 students who will be divided into two groups: experimental group and control group. There is no randomization.

	Design name (set A)	Design explanation or definition (set B)	Notation (set C)	Example (set D)
7	The Pretest-Posttest Control Group Design	The dependent variable is **measured for both groups before (O1) and after (O2) the independent variable (X)** is introduced to the experimental group. The **independent variable** is introduced to the experimental group **after the first measure** of the dependent variable. **Two groups** are needed; **one** is **the control group**. There is **randomization**.	R O1 X O2 R O1 O2	The dependent variable is heart rate and it is measured for both groups before and after the independent variable is introduced. The independent variable is jogging and it is introduced to the experimental group after the first measure of heart rates. The participants are 50 students who will be **randomly** assigned to either one of two groups: experimental group and control group. There is randomization.
8	The Solomon Four-Group Design	The dependent variable is measured for two groups before (O1 and O3) and after (O2 and O4) the independent variable (X) is introduced to the experimental groups. For the other two groups, the dependent variable is measured only after (O5 and O6) the independent variable is introduced to the experimental groups. The independent variable is introduced to the two experimental groups. For the first experimental group (Group A), the independent variable is introduced after the first measure of the dependent variable; for the other experimental group (Group C), it is introduced before. Four groups are needed; two are control groups. There is randomization.	Group A R O1 X O2 Group B R O1 O2 Group C R X O2 Group D R O2	The dependent variable is heart rate, it is measured before and after the independent variable is introduced for two of the four groups; it is measured only after in the other two groups. The independent variable is jogging and it is introduced to the experimental groups differently: for Group A it is introduced after the dependent variable is measured. For Group C, it is presented before the dependent variable is measured. The participants are 100 students who will be **randomly** assigned to either one of four groups: experimental group A or C, control group B or D. There is randomization.
9	The Posttest-Only Control Group Design	The dependent variable is **measured for both groups only after (O) the independent variable (X)** is introduced to the experimental group. The **independent variable** is introduced to the experimental group **before the measure** of the dependent variable. Two	R X O1 R O1	The dependent variable is heart rate and it is measured for both groups after the independent variable is introduced to the experimental group. The independent variable is jogging and it is introduced to the experimental group before heart rates (dependent variable) are measured in both groups. The participants are 50 students who will be **randomly assigned** to

	groups are needed; **one** is the **control group**. There is **randomization**.	either one of two groups: experimental group and control group. There is randomization.

Appendix G – Graphic Overview of Research Designs Taught in Each Teaching Part



Appendix H - Complete List Of The Questions And Instructions For Article Reviews

Instructions contained in the first slide:

"Please, read all instructions carefully. After you read the article (PDF in your e-mail), use the items below and check: yes, if the item is applicable and present in the article; no, if the item is applicable, but not present in the article; not applicable, if the item is not applicable to the article; I do not know, if you do not know.

Take your time to go through the items and the article."

	Item				
1	The research design is defined and clearly described, and is sufficiently detailed to permit the study to be replicated.	Yes	No	Not applicable.	I do not know.
2	The design is appropriate for the research question.	Yes	No	Not applicable.	I do not know.
3	The design has internal validity, potential confounding variables or biases are addressed.	Yes	No	Not applicable.	I do not know.
4	The design has external validity, including subjects, settings, and conditions.	Yes	No	Not applicable.	I do not know.
5	The design and conduct of the study are believable.	Yes	No	Not applicable.	I do not know.
6	The development and content of the independent variable are sufficiently described or referenced, and are sufficiently detailed to permit the study to be replicated.	Yes	No	Not applicable.	I do not know.
7	The dependent variables are clearly defined.	Yes	No	Not applicable.	I do not know.

8	The measures are appropriate given the study's variables; the scoring method is clearly defined.	Yes	No	Not applicable.	I do not know.
9	The psychometric properties and procedures are clearly presented and appropriate.	Yes	No	Not applicable.	I do not know.
10	The data set is sufficiently described or referenced.	Yes	No	Not applicable.	I do not know.
11	Observers or raters are sufficiently trained.	Yes	No	Not applicable.	I do not know.
12	Data quality control is described and adequate, i.e., monitoring and maintaining the quality of data during the conduct of the study.	Yes	No	Not applicable.	I do not know.
13	The population is clearly defined, sufficiently detailed to permit the study to be replicated.	Yes	No	Not applicable.	I do not know.
14	The experimental materials and stimuli are sufficiently detailed to permit the study to be replicated.	Yes	No	Not applicable.	I do not know.
15	The sampling procedures are sufficiently described.	Yes	No	Not applicable.	I do not know.
16	Subject samples are appropriate to the research question.	Yes	No	Not applicable.	I do not know.
17	Selection bias is addressed.	Yes	No	Not applicable.	I do not know.
18	Data analysis procedures are sufficiently described, and are sufficiently detailed to permit the study to be replicated.	Yes	No	Not applicable.	I do not know.
19	Data analysis procedures conform to the research design; hypotheses, models, or theory drives the data analyses.	Yes	No	Not applicable.	I do not know.

20	The assumptions underlying the use of statistics are fulfilled by the data, such as measurement properties of the data and normality of distributions.	Yes	No	Not applicable.	I do not know.
21	Statistical tests are appropriate (optimal).	Yes	No	Not applicable.	I do not know.
22	If statistical analysis involves multiple tests or comparisons, proper adjustment of significance level for chance outcomes was applied.	Yes	No	Not applicable.	I do not know.
23	Power issues are considered in statistical studies with small sample sizes.	Yes	No	Not applicable.	I do not know.
24	In qualitative research that relies on words instead of numbers, basic requirements of data reliability, validity, trustworthiness, and absence of bias were fulfilled.	Yes	No	Not applicable.	I do not know.

Appendix I – Complete Set of Questions

Question 1

Part A: "What features define THE ONE SHOT CASE STUDY design?(be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization."

Part B: "What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable."

Part C: "Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

Question 2

Part A: "What features define the ONE-GROUP PRETEST-POSTTEST DESIGN?(be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization."

Part B: "What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable."

Part C: "Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s)

and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

Question 3

Part A: "What features define the STATIC GROUP COMPARISON design?(be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization."

Part B: "What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable."

Part C: "Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

Question 4

Part A: "What features define THE NONEQUIVALENT CONTROL GROUP DESIGN?(be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization."

Part B: "What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable."

Part C: “Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

Question 5

Part A: “What features define the COUNTERBALANCED DESIGN design?(be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

Part B: “What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable.”

Part C: “Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

Question 6

Part A: “What features define the MULTIPLE TIME SERIES design?(be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

Part B: “What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The

group(s); iv. The randomization, if applicable."

Part C: "Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

Question 7

Part A: "What features define the PRETEST-POSTTEST CONTROL GROUP DESIGN?(be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization."

Part B: "What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable."

Part C: "Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

Question 8

Part A: "What features define the SOLOMON FOUR-GROUP design?(be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization."

Part B: “What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable.”

Part C: “Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

Question 9

Part A: “What features define the POSTTEST-ONLY CONTROL GROUP DESIGN?(be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

Part B: “What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable.”

Part C: “Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

Appendix J - Rubric for Pre-Experimental Designs

One-Shot Case Design

Part A: “What features define the One-Shot Case Design? (be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

The dependent variable is measured (O) only after the independent variable (X) is introduced.

score 1 point for this item if.

The answer is exactly the same as the one above OR

If the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) **AND** the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the observation/measure of the dependent variable will occur after the introduction of the independent variable.

score 0.5 points for this item if.

If the answer mentions the observation/measure of the dependent variable **and** the fact that it will occur as the last step in the study, BUT does not mention the independent variable.

score 0 points for this item if.

If the observation/measure of the dependent variable is mentioned without reference to the fact that it will occur last in the study.

If only the independent variable is mentioned.

The independent variable is introduced before the measure of the dependent variable.

score 1 point for this item if:

The answer is exactly the same as the one above OR

The answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) **AND** the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the introduction of the independent variable will occur before the observation/measure of the dependent variable.

score 0.5 points for this item if.

If the answer mentions that the independent variable will be introduced **and** that it will be the first step in the study procedures, BUT doesn't mention the dependent variable.

If the answer mentions the introduction of the independent variable, but only indirectly mentions that the independent variable will be introduced before the independent variable.

If both the dependent and the independent variables and their respective measure and introduction are mentioned only once in the definition.

score 0 points for this item if.

If the answer mentions the independent variable, but not the fact that it will occur first in the study OR

If only the dependent variable is mentioned.

Only one group is needed and there is no control group.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that only one group is necessary and that there is no control group. (This can be mentioned as a separate response item or together with any of the response items from Part A).

score 0.5 points for this item if only one of the two pieces of information below is mentioned:

Only one group is necessary OR

There is no control group.

score 0 points for this item if:

If none of the two pieces of information above is mentioned.

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is **no randomization** (This information can be mentioned as a separate response item or together with any of the response items from Part A).

score 0 points for this item if:

The lack of randomization is not mentioned.

Part B: **“**What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable.”

X O

score 4 points for this item if.

The observation/measure of the dependent variable is illustrated after the introduction of the independent variable.

The introduction of the independent variable is illustrated before the measure of the dependent variable.

There is one line, representing one group.

There are **no** symbols “R” representing randomization.

score 1 point for each item above (including the non-illustration of randomization)

score 0 points if there is no match to the items above.

Part C: **“**Please, provide an example of this design, different from the one given in the multiple choice questions, and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

My example: The dependent variable is heart rate and it is measured after the independent variable is introduced.

score 1 point for this item if.

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND

a dependent variable is explicitly specified/exemplified AND

the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that the observation/measure of the specified dependent variable will occur after the introduction of the independent variable.

score 0.5 points for this item if.

If the dependent variable is specified **AND** it is also mentioned that its observation/measure will occur after the introduction of the independent variable, BUT the words “dependent variable” or any of the synonyms are not used OR if the words “dependent variable” are not explicitly attached to the specification.

If the dependent variable is specified and the word “dependent variable” OR any of the synonyms are used, BUT there is no mention that the observation/measure will occur after the independent variable is introduced.

If the answer is embedded in the item above, gibe half of what the answer is.

score 0 points for this item if.

The dependent variable is specified, AND there is **no** mention that the observation/measure will occur after the independent variable is introduced AND the word “dependent variable” or any of the synonyms are **not** used.

If **no** dependent variable is specified AND there is **no** mention that the observation/measure will occur after the independent variable is introduced.

If there are information that contradict each other.

If there is an illogical dependent variable specified.

My example: The independent variable is jogging and it is introduced before measuring the heart rates.

score 1 point for this item if.

The answer contains the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

an independent variable is explicitly specified AND

the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) are mentioned AND

The answer mentions that the introduction of the specified independent variable will occur before the observation/measure of the dependent variable.

score 0.5 points for this item if.

If the independent variable is specified AND the answer mentions that the introduction of the specified independent variable will occur before the observation/measure of the dependent variable, BUT the words “independent variable” or any of the synonyms are not used.

If the independent variable is specified AND the word “independent variable” or any of the synonyms are used, but there is no mention that the introduction of the specified independent variable will occur before the observation/measure of the dependent variable.

If the answer is embedded in the item above, gibe half of what the answer is.

score 0 points for this item if.

The independent variable is specified, but "when it is introduced" is not mentioned AND the word "independent variable" or any of the synonyms are **not** used.

If **no** independent variable is specified AND "when it is introduced" is not mentioned.

My example: The participants are 25 students who will not be divided into groups.

score 1 point for this item if:

The answer mentions who are the participants, how many, AND that only one group is needed (or that the participants won't be divided into groups).

Score 0.5 for this item if

The answer mentions who are the participants and how many, BUT does not mention that only one group is needed (or that the participants won't be divided into groups).

The answer mentions who the participants are and that only one group is needed, BUT doesn't mention how many participants.

The answer mentions how many participants and that only one group is needed (or that the participants won't be divided into groups, BUT doesn't mention who are the participants.

score 0 points for this item if:

The answer mentions only how many participants.

The answer mentions only who are the participants.

The answer mentions only that only one group is needed (or that the participants won't be divided into groups).

My example: There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization in any of the other items above.

score 0 points for this item if:

The lack of randomization is not mentioned.

One-Group Pretest-Posttest Design

Part A: “What features define the One-Group Pretest-Posttest Design? (be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

The dependent variable is measured before (O1) and after (O2) the independent variable (X) is introduced.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the observation of the dependent variable will occur <u>before</u> and <u>after</u> the introduction of the independent variable.

score 0.5 points for this item if:

The answer mentions that observations of the dependent variable will occur twice.

score 0 points for this item if:

Observations of the dependent variable are mentioned, BUT when and how many times are not.

The independent variable is mentioned, BUT not the dependent variable.

The independent variable is introduced after the first measure of the dependent variable.

score 1 point for this item if:

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the introduction of the independent variable will occur after the first observation of the dependent variable.

score 0.5 points for this item if:

If the answer mentions that the independent variable will be introduced after one observation of the dependent variable, BUT the word dependent variable is not used.

If the answer is written together with the answer to the first item.

score 0 points for this item if:

If the introduction of the independent variable is mentioned, BUT not when it is introduced.

If only the dependent variable is mentioned.

Only one group is needed and there is no control group.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that only one group is necessary and that there is no control group. (This can be mentioned as a separate response item or together with any of the response items from Part A).

score 0.5 points for this item if only one of the two pieces of information below is mentioned:

Only one group is necessary OR

There is no control group.

score 0 points for this item if:

If none of the two pieces of information above is mentioned.

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization in any of the other items above.

score 0 points for this item if:

The lack of randomization is not mentioned.

Part B: **“**What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable.”

O1 X O2 (THE NUMBERS AFTER “O” ARE NOT MANDATORY)

score 4 points for this item if.

The observations/measures of the dependent variable are illustrated before and after the introduction of the independent variable.

The introduction of the independent variable is illustrated after the first measure of the dependent variable.

There is one line, representing one group.

There are **no** symbols "R" representing randomization.

score 1 point for each item above (including the non-illustration of randomization)

score 0 points if there is no match to the items above.

Part C: **"**Please, provide an example of this design, different from the one given in the multiple choice questions, and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

My example: The dependent variable is heart rate and it is measured before and after the independent variable is introduced.

score 1 point for this item if:

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND

a dependent variable is explicitly specified AND

the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that observations of the specified dependent variable will occur for both groups before and after the introduction of the independent variable.

score 0.5 points for this item if:

If the dependent variable is specified and "when it will occur is mentioned", BUT the words "dependent variable" or any of the synonyms are not used OR

If the dependent variable is specified and the word "dependent variable" or any of the synonyms are used, but there is no mention of when the dependent variable is observed.

score 0 points for this item if:

The dependent variable is specified, but its place in time is not mentioned AND the word "dependent variable" or any of the synonyms are not used.

If no dependent variable is specified and its place in time is not mentioned.

If the observations for both groups are not mentioned.

My example: The independent variable is jogging and it is introduced after the first measure of the heart rates.

score 1 point for this item if:

if the answer contains the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable)AND

an independent variable is explicitly specified AND

the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) are mentioned AND

The answer mentions that the introduction of the specified independent variable will occur after the first observation of the dependent variable.

score 0.5 points for this item if:

If the independent variable is specified and "when it is introduced" is mentioned, BUT the words "independent variable" or any of the synonyms are not used.

If the independent variable is specified and the word "independent variable" or any of the synonyms are used, but there is no mention of "when it is introduced".

If the answer is written together with the first one.

score 0 points for this item if:

The independent variable is specified, BUT "when it is introduced" is not mentioned NOR is the word "independent variable" or any of the synonyms are not used.

If no independent variable is specified NOR "when it is introduced" is not mentioned.

My example: The participants are 25 students who will not be divided into groups.

Score 1 point for this item if:

The answer mentions who are the participants, how many, AND that only one group is needed (or that the participants won't be divided into groups).

Score 0.5 for this item if

The answer mentions who are the participants and how many, BUT does not mention that only one group is needed (or that the participants will not be divided into groups).

The answer mentions who the participants are and that only one group is needed, BUT doesn't mention how many participants.

The answer mentions how many participants and that only one group is needed (or that the participants won't be divided into groups, BUT doesn't mention who are the participants.

score 0 points for this item if:

The answer mentions only how many participants.

The answer mentions only who are the participants.

The answer mentions only that only one group is needed (or that the participants won't be divided into groups).

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization in any of the other items above.

score 0 points for this item if:

The lack of randomization is not mentioned.

Static Group Comparison

Part A: "What features define the (name of the design) design (be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization."

The dependent variable is measured (O) for both groups only after the independent variable (X) is introduced to the experimental group.

score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that observations of the dependent variable will occur for both groups after the introduction of the independent variable to the experimental group.

score 0.5 points for this item if.

The answer mentions that observations the dependent variable will occur after the introduction of the independent variable to the experimental group, BUT it is not mentioned that the observation will occur for both groups.

The answer mentions that observations the dependent variable will occur for both groups, BUT "when the observation will occur" is not mentioned.

score 0 points for this item if:

Observations of the dependent variable are mentioned, BUT there is no mention to when the observation will occur" NOR to that it will occur for both groups.

The independent variable is introduced to the experimental group before the measures (observations) of the dependent variable.

score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the introduction of the independent variable will occur only to the experimental group before the observations of the dependent variable.

Score 0.75

If the independent variable is mentioned AND it is mentioned that it is introduced only to the experimental group, BUT when it will be introduced is mentioned only together with the dependent variable explanation.

score 0.5 points for this item if:

If the independent variable is mentioned AND when it will be introduced, BUT the experimental group exclusiveness is not mentioned.

If the independent variable is mentioned AND it is mentioned that it is introduced only to the experimental group, BUT when it will be introduced is not mentioned.

score 0 points for this item if:

If only the independent variable is mentioned and “when it will be introduced” is not mentioned and there is no mention to the experimental group.

Two groups are needed; one is the control group.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that two groups are necessary and that there is a control group (The response for this item can be either in a separate item or together with any of the items from Part A.

score 0.75

If there is a mention to two groups, AND to the control/experimental group, however it is not clear that both groups are necessary.

score 0.5 points for this item if:

Two groups are mentioned, but nothing is said about the control/experimental group.

The control group is mentioned, but nothing is said about the need for two groups.

score 0 points for this item if:

The information above is not mentioned.

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization in any of the other items above.

score 0 points for this item if:

Randomization (or the lack of it) is not mentioned.

Part B: **"**What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable."

X O1

O1

score 4 points for this item if.

The observations/measures of the dependent variable are illustrated after the introduction of the independent variable.

The introduction of the independent variable is illustrated before the measure of the dependent variable.

There are two rows (lines), representing the two groups.

There are **no** symbols "R" representing randomization.

score 1 point for each item above (including the non-illustration of randomization)

score 0 points if there is no match to the items above.

Part C: **"**Please, provide an example of this design, different from the one given in the multiple choice questions, and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

My example: The dependent variable is heart rate and it is measured for both groups after the independent variable is introduced to the experimental group.

score 1 point for this item if:

The answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND

a dependent variable is explicitly specified AND

the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that observations of the specified dependent variable will occur for both groups after the introduction of the independent variable to the experimental group.

score 0.5 points for this item if:

Observations/measures of the dependent variable are specified and when they will occur for both groups is mentioned, BUT the words “dependent variable” or any of the synonyms are not used AND

Observations of the dependent variable are specified for both groups and the word “dependent variable” or any of the synonyms are used, BUT there is no mention that they will occur after the introduction of the dependent variable.

score 0 points for this item if:

The dependent variable is specified, but its place in time is not mentioned NOR is the word “dependent variable” or any of the synonyms used NOR the observations for both groups are mentioned OR.

If **no** dependent variable is specified AND its place in time is not mentioned NOR is the word “dependent variable” or any of the synonyms used NOR the observations for both groups are mentioned.

My example: The independent variable is jogging and it is introduced to the experimental group before heart rates are measured in both groups.

score 1 point for this item if:

if the answer contains the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

an independent variable is explicitly specified AND

the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) are mentioned AND

The answer mentions that the introduction of the specified independent variable will occur only to the experimental group before the observation of the dependent variable in both groups.

score 0.5 points for this item if:

If the independent variable is specified, the experimental group is mentioned, and its place in time is mentioned, BUT the words “independent variable” or any of the synonyms are not used.

If the independent variable is specified, the experimental group is mentioned, and the word “independent variable” or any of the synonyms are used, but there is no mention of its place in time.

score 0 points for this item if:

The independent variable is specified, but its place in time is not mentioned AND the word “independent variable” or any of the synonyms are not used.

If no independent variable is specified and its place in time is not mentioned.

There is no mention of the experimental group.

My example: The participants are 50 students who will be divided into two groups: experimental group and control group.

score 1 point for this item if:

The answer mentions who are the participants, how many, AND that two groups are needed (or that the participants will be divided into groups).

Score 0.5 for this item if

The answer mentions who are the participants and how many, BUT does not mention that two groups are needed (or that the participants will be divided into groups).

The answer mentions who the participants are and that two groups are needed, BUT doesn't mention how many participants.

The answer mentions how many participants and that two groups are needed (or that the participants will be divided into groups), BUT doesn't mention who are the participants.

score 0 points for this item if:

The answer mentions only how many participants.

The answer mentions only who are the participants.

The answer mentions only that two groups are needed (or that the participants will be divided into groups).

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization in any of the other items above.

score 0 points for this item if:

The lack of randomization is not mentioned.

Appendix K – Rubric for Quasi-Experimental Designs

Nonequivalent Control-Group Design

Part A: “What features define the Nonequivalent Control-Group Design? (be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

The dependent variable is measured for both groups before (O1) and after (O2) the independent variable (X) is introduced to the experimental group.

Score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the observation of the dependent variable will occur before and after the introduction of the independent variable AND

The answer mentions that observations of the dependent variable will occur for both groups.

Score 0.5 point for this item if.

The answer mentions the observations of the dependent variable AND that it will be measured before and after the independent variable.

The answer mentions the observations of dependent variable AND that it will be observed in both groups.

Score 0 point for this item if.

The observations of the dependent variable are mentioned, BUT there is no mention to when it will occur or that it will occur for both groups.

The independent variable is introduced to the experimental group after the first measure of the dependent variable.

Score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the introduction of the independent variable will occur after the first observation of the dependent variable AND

The answer mentions that the introduction of the independent variable will occur only to the experimental group.

Score 0.5 point for this item if.

The answer mentions the introduction of the independent variable AND that it will be introduced after the first observation of the dependent variable, BUT it does not mentioned that if will be introduced only to the experimental group.

The answer mentions the introduction of the independent variable AND that it will introduced only to the experimental group, but it does not mentioned that it will be introduced after the first observation of the dependent variable.

Score 0 point for this item if.

The independent variable is mentioned, BUT there is no mention to when it is introduced NOR that it will occur only for the experimental group.

Two groups are needed; one is the control group.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that two groups are necessary and that there is a control group - either in a separate item or together with any of the items from Part A.

score 0.5 points for this item if:

The two groups are mentioned, but nothing is said about the control/experimental group.

The control/experimental group are mentioned, but nothing is said about the need of two groups.

score 0 points for this item if:

If the information described above is not mentioned.

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization together with any of the other items described above.

score 0 points for this item if:

The lack of randomization is not mentioned.

Part B: **“**What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable.”

O1 __ X __ O2

O1 O2

Score 4 points for this item if

The four observations/measures of the dependent variable are illustrated before and after the introduction of the independent variable for both groups.

The introduction of the independent variable is illustrated after the first measures of the dependent variable.

The groups are represented in two rows (lines).

There is nothing referring to randomization (letter R).

Score 1 point for each of the items above that are illustrated.

Score 0 points for this item if there is no match with the information above.

Part C: “Please, provide an example of this design, different from the one given in the multiple choice questions (during training), and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and

when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

My example: The dependent variable is heart rate and it is measured for both groups before and after the independent variable is introduced.

score 1 point for this item if:

The answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND

a dependent variable is explicitly specified AND

the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that observations of the specified dependent variable will occur for both groups, before and after the introduction of the independent variable to the experimental group.

score 0.5 points for this item if:

The dependent variable is specified AND it is mentioned that its observations/measures will occur before and after the introduction of the independent variable for both groups, BUT the words “dependent variable” or any of the synonyms are not used.

The dependent variable is specified AND the word “dependent variable” or any of the synonyms are used, AND the observations for both groups are mentioned, BUT it is not mentioned that the observations/measures will occur before and after the introduction of the independent variable .

The dependent variable is specified AND the word “dependent variable” or any of the synonyms are used, AND it is mentioned that the observations/measures will occur before and after the introduction of the independent variable, BUT observations for both groups are not mentioned.

score 0 points for this item if:

The dependent variable is specified, BUT it is not mentioned that its observations/measures will occur before and after the introduction of the independent variable for both groups NOR the word “dependent variable” or any of the synonyms are not used.

If no dependent variable is specified.

The independent variable is jogging and it is introduced to the experimental group after the first measures of heart rates.

score 1 point for this item if:

if the answer contains the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

an independent variable is explicitly specified AND

the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) are mentioned AND

The answer mentions that the introduction of the specified independent variable will occur only to the experimental group after the first observations/measures of the dependent variable in both groups.

score 0.5 points for this item if:

If the independent variable is specified, the experimental group is mentioned, and it is mentioned that the specified independent variable will be introduced to the experimental group after the first observation/measure of the dependent variable in both groups, BUT the words "independent variable" or any of the synonyms are not used.

If the independent variable is specified, the experimental group is mentioned, and the word "independent variable" or any of the synonyms are used, but it is not mentioned that the specified independent variable will be introduced to the experimental group after the first observation/measure of the dependent variable in both groups

score 0 points for this item if:

The independent variable is specified, but its place in time is not mentioned NOR the word "independent variable" or any of the synonyms is not used.

If no independent variable is specified.

My example: The participants are 50 students who will be divided into two groups: experimental group and control group.

score 1 point for this item if:

The answer mentions who are the participants, how many, AND that two groups are needed (or that the participants will be divided into groups).

Score 0.5 for this item if

The answer mentions who are the participants and how many, BUT does not mention that two groups are needed (or that the participants will be divided into groups).

The answer mentions who the participants are and that two groups are needed, BUT doesn't mention how many participants.

The answer mentions how many participants and that two groups are needed (or that the participants will be divided into groups), BUT doesn't mention who are the participants.

score 0 points for this item if:

The answer mentions only how many participants.

The answer mentions only who are the participants.

The answer mentions only that two groups are needed (or that the participants will be divided into groups).

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization in any of the other items above.

score 0 points for this item if:

The lack of randomization is not mentioned

Counterbalanced Design

Part A: “What features define the Counterbalanced Design? (be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

The dependent variable is measured for all groups (O), after each one of the four independent variables (X1, X2, X3, X4) is introduced for each experimental group.

score 1 point for this item if.

the answer is exactly the same as the one above OR

The answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the observations/measures of the dependent variable will occur after the introduction of each independent variable

The answer mentions that the observations/measures will occur for each one of all four groups.

Score 0.75 if (mentions 3 items out of the four above). Ex:

The answer mentions the observations/measures of the dependent variable AND the fact that they will occur after each independent variable, AND it mentions the word “dependent variable” or its synonyms, BUT it does not mention that they will occur for each group.

score 0.5 points for this item if.

The answer mentions the observations/measures of the dependent variable AND the fact that they will occur after each independent variable, BUT it does not mention that they will occur for each group NOR does it mention the word “dependent variable” or its synonyms.

The answer mentions the observations/measures of the dependent variable AND that they will occur for each group, BUT it doesn`t mention that they will occur after each independent variable NOR does it mention the word “dependent variable” or its synonyms.

The answer mentions the observations/measures of the dependent variable AND the word “dependent variable” or its synonyms, BUT it does not mention that they will occur for each group NOR that they will occur after each independent variable.

score 0 points for this item if.

If the observations/measures of the dependent variable are mentioned, BUT without explicitly using the word “dependent variable” OR reference to the fact that they will occur after the introduction of each independent variable, for each group.

If only the independent variables are mentioned.

Each independent variable is introduced to all groups but in a different order for each group.

score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the introduction of each independent variable will occur before each observation/measure of the dependent variable.

The answer mentions that all groups will be presented with all four independent variables.

Score 0.75 if three of the four items above are mentioned

score 0.5 points for this item if:

If the introduction of the independent variables are mentioned AND that each independent variable will be introduced for each group **before** the respective measures/observations of the dependent variable BUT the word "independent variable" or its synonyms i not used.

If the introduction of the independent variables are mentioned AND the word "independent variable" or its synonyms are used, BUT it is not mentioned that each independent variable will be introduced for each group.

score 0 points for this item if:

If the answer mentions the introduction of the independent variable, BUT not the fact that it will occur before each observation/measure of the dependent variable for each group AND the word "independent variable" or its synonyms are not used.

If only the dependent variable is mentioned.

Four groups are needed, but there is no "true control group", since the independent variable is introduced for all groups.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that four groups are necessary and that there isn`t a “real control group “ since the independent variable will be introduced to all groups (this item can be either a separate item or together with any of the items from Part A).

score 0.5 points for this item if:

The four groups are mentioned, but nothing is said about the “control/experimental groups”.

The “control group” is mentioned, but nothing is said about the need of four groups.

score 0 points for this item if:

None of the information above is mentioned.

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization in any of the other items above.

score 0 points for this item if:

Randomization (or the lack of it) is not mentioned.

Part B: **“**What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable.”

	Time 1	**Time 2**	**Time 3**	**Time 4**
Group A	**X1O**	**X2O**	**X3O**	**X4O**

— — — — — — — — — — — —

Group B X2O X4O X1O X3O

— — — — — — — — — — —

Group C X3O X1O X4O X2O

— — — — — — — — — — —

Group D X4O X3O X2O X1O

score 4 points for this item if:

The observations/measures of the dependent variable are illustrated after the introduction of the independent variables.

The introduction of the independent variables is illustrated before the observations of the dependent variable.

There are four rows (line), each representing one group.

There are **no** symbols (R) representing randomization.

score 1 point for each item above (including the non-illustration of randomization)

score 0 points if there is no match to the items above.

Part C: **“**Please, provide an example of this design, different from the one given in the multiple choice questions, and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

My example: The dependent variable is heart rate and it is measured for all groups, after each independent variable is introduced for the groups.

score 1 point for this item if:

The answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND

a dependent variable is explicitly specified AND

the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that observations of the specified dependent variable will occur after the introduction of each independent variable to the experimental groups.

score 0.5 points for this item if:

If the dependent variable is specified AND it is mentioned that its observations/measures will occur after the introduction of each independent variable to all experimental groups, BUT the words “dependent variable” or any of the synonyms are not used.

If the dependent variable is specified AND it is mentioned that its observations/measures will occur after the introduction of each independent variable AND the word “dependent variable” or any of the synonyms are used, BUT it is not mentioned that observations of the dependent variable will occur to all groups.

If the dependent variable is specified AND the word “dependent variable” or any of the synonyms are used, AND it is mentioned that the dependent variable will be observed/measure

in all groups BUT it is not mentioned that its observations/measures will occur after the introduction of each independent variable.

score 0 points for this item if:

The dependent variable is specified, BUT it is not mentioned that its observations/measures will occur after the introduction of the independent variable for all groups NOR the word “dependent variable” or any of the synonyms are not used.

If no dependent variable is specified.

My example: The independent variables can be jogging (X1), swimming (X2), dancing (X3), and walking (X4) and each one of them is introduced to all groups, but in a different order for each group.

score 1 point for this item if:

The answer contains the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable)AND

an independent variable is specified AND

the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) are mentioned AND

The answer mentions that the introduction of each specified independent variable to all experimental groups <u>before</u> the observation/measure of the dependent variable in all groups.

score 0.5 points for this item if:

The independent variables are specified, AND the four experimental groups are mentioned, AND it is mentioned that the specified independent variables will be introduced to all experimental groups before the observations/measures of the dependent variable in all groups, BUT the words "independent variable" or any of the synonyms are not used.

The independent variables are specified, AND the four experimental groups are mentioned, AND the word "independent variable" or any of the synonyms are used, BUT it is not mentioned that the specified independent variables will be introduced to the experimental groups before the first observation/measure of the dependent variable in all groups.

The independent variables are specified, AND it is mentioned that the specified independent variables will be introduced to all experimental groups before the observations/measures of the dependent variable in all groups AND the word "independent variable" or any of the synonyms are used, but the four experimental groups are not mentioned.

score 0 points for this item if:

The independent variables are specified together with only one of the following information: a) EITHER when they are introduced is not mentioned; 2) OR it is mentioned that the specified independent variables will be introduced to all experimental groups before the observations/measures of the dependent variable in all groups; 3) OR the word "independent variable" or any of its synonyms are used.

If no independent variable is specified.

My example: ***The participants are 100*** students who will be divided into four groups.

score 1 point for this item if:

The answer mentions who are the participants, how many, AND that four groups are needed (or that the participants will be divided into four groups).

Score 0.5 for this item if

The answer mentions who are the participants and how many, BUT does not mention that four groups are needed (or that the participants will be divided into four groups).

The answer mentions who the participants are and that four groups are needed, BUT doesn't mention how many participants.

The answer mentions how many participants and that four groups are needed (or that the participants will be divided into groups), BUT doesn't mention who are the participants.

score 0 points for this item if:

The answer mentions only how many participants.

The answer mentions only who are the participants.

The answer mentions only that four groups are needed (or that the participants will be divided into groups).

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization in any of the other items above.

score 0 points for this item if:

The lack of randomization is not mentioned.

The Multiple Time Series Design

Part A: “What features define the Counterbalanced Design? (be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

The dependent variable is measured several times (O), for both groups, before and after the independent variable (X) is introduced to the experimental group.

Score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that SEVERAL observations of the dependent variable will occur before and after the introduction of the independent variable AND

The answer mentions that observations of the dependent variable will occur for both groups.

Score 0.5 point for this item if.

The answer mentions the observations of the dependent variable AND that it will be measured SEVERAL TIMES before and after the introduction of the independent variable.

The answer mentions the observations of the dependent variable AND that it will be observed in both groups.

Score 0 point for this item if.

The observations of the dependent variable are mentioned, BUT there is no mention to how many times or when the observations/measures will occur nor that they will occur for both groups.

The independent variable is introduced to the experimental group after several measures of the dependent variable.

Score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the introduction of the independent variable will occur after SEVERAL observations/measures of the dependent variable AND

The answer mentions that the introduction of the independent variable will occur only to the experimental group.

Score 0.5 point for this item if.

The answer mentions the introduction of the independent variable AND that it will be introduced after the first observation of the dependent variable, BUT it does not mentioned that if will be introduced only to the experimental group.

The answer mentions the introduction of the independent variable AND that it will introduced only to the experimental group, BUT it does not mentioned that it will be introduced after the first observation of the dependent variable.

Score 0 point for this item if.

The independent variable is mentioned, BUT there is no mention to when it is introduced NOR that it will occur only for the experimental group.

The independent variable is mentioned, BUT there is no mention to when it will occur or that it will occur for both groups.

Two groups are needed; one is the control group.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that two groups are necessary and that there is a control group - either in a separate item or together with any of the items from Part A.

score 0.5 points for this item if:

The two groups are mentioned, but nothing is said about the control/experimental group.

The control group is mentioned, but nothing is said about the need of two groups.

score 0 points for this item if:

If the information described above is not mentioned.

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization together with any of the other items described above.

score 0 points for this item if:

The lack of randomization is not mentioned.

Part B: **“**What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable.”

O O O OXO O O O

— — — — — — — — —

O O O O O O O O

Score 4 points for this item if

The several observations/measures and when they happen (some before, some after the introduction of the independent variable) are represented.

The introduction of the independent variable and when it occurs is represented.

The two groups are represented in two rows (lines).

There is nothing referring to randomization.

Score 1 point for each of the items above that are illustrated.

Score 0 points for this item if none of the four items above are illustrated.

Part C: **“**Please, provide an example of this design, different from the one given in the multiple choice questions, and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who

are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

The dependent variable is heart rate and it is measured several times, for both groups, before and after the independent variable is introduced.

score 1 point for this item if:

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND

a dependent variable is specified AND

the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that SEVERAL observations of the specified dependent variable will occur for both groups before and after the introduction of the independent variable to the experimental group.

score 0.5 points for this item if:

If the dependent variable is specified and it is mentioned that its SEVERAL observations/measures will occur before and after the introduction of the independent variable AND the several observations/measures for both groups are mentioned, BUT the words “dependent variable” or any of the synonyms are not used.

If the dependent variable is specified and the word “dependent variable” or any of the synonyms are used AND the two groups are mentioned, BUT it is not mentioned that several observations/measures will occur before and after the introduction of the independent variable.

score 0 points for this item if:

The dependent variable is specified, but its place in time is not mentioned AND the word “dependent variable” or any of the synonyms are not used.

If no dependent variable is specified and the several observations before and after the introduction of the independent variable are not mentioned.

Observations of both groups are not mentioned.

The independent variable is jogging and it is introduced to the experimental group after several measures of heart rates.

score 1 point for this item if:

The answer contains the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable)AND

an independent variable is specified AND

the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) are mentioned AND

The answer mentions that the introduction of the specified independent variable will occur only to the experimental group <u>after</u> SEVERAL observations/measures of the dependent variable in both groups.

score 0.5 points for this item if:

The independent variable is specified, AND the experimental group is mentioned, AND it is mentioned that the specified independent variable will be introduced to the experimental group after SEVERAL observations/measures of the dependent variable in both groups, BUT the words "independent variable" or any of the synonyms are not used.

The independent variable is specified, the experimental group is mentioned, and the word "independent variable" or any of the synonyms are used, but it is not mentioned that the specified independent variable will be introduced to the experimental group after several observations/measures of the dependent variable in both groups

score 0 points for this item if:

The independent variable is specified, but its place in time is not mentioned NOR is the word "independent variable" or any of its synonyms used.

No independent variable is specified and its place in time is not mentioned.

There is no mention of the experimental group.

My example: The participants are 50 students who will be divided into two groups: experimental group and control group.

score 1 point for this item if:

The answer mentions who are the participants, how many, AND that two groups are needed (or that the participants will be divided into groups).

Score 0.5 for this item if

The answer mentions who are the participants and how many, BUT does not mention that two groups are needed (or that the participants will be divided into groups).

The answer mentions who the participants are and that two groups are needed, BUT doesn't mention how many participants.

The answer mentions how many participants and that two groups are needed (or that the participants will be divided into groups), BUT doesn't mention who are the participants.

score 0 points for this item if:

The answer mentions only how many participants.

The answer mentions only who are the participants.

The answer mentions only that two groups are needed (or that the participants will be divided into groups).

There is no randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is no randomization in any of the other items above.

score 0 points for this item if:

The lack of randomization is not mentioned.

Appendix L - Rubric for True Experimental Designs

The Pretest-Posttest Control Group Design

Part A: “What features define the The Pretest-Posttest Control Group Design ? (be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

The dependent variable is measured for both groups before (O1) and after (O2) the independent variable (X) is introduced to the experimental group.

Score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) and the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that observations of the dependent variable will occur before and after the introduction of the independent variable AND

The answer mentions that observations of the dependent variable will occur for both groups.

Score 0.5 point for this item if.

The answer mentions the observations of the dependent variable AND that it will be measured before and after the introduction of the independent variable BUT it does not mention that they will be observed in both groups.

The answer mentions the observations of the dependent variable AND that they will be observed in both groups BUT it doesn`t mention that it will be measured before and after the introduction of the independent variable.

Score 0 point for this item if.

The observations of the dependent variable are mentioned, BUT there is no mention to when the observations/measures will occur NOR will they occur for both groups.

The independent variable is introduced to the experimental group after the first measure of the dependent variable.

Score 1 point for this item if.

The answer is exactly the same as the one above OR

The answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the introduction of the independent variable will occur after the first observation of the dependent variable AND

The answer mentions that the introduction of the independent variable will occur only to the experimental group.

Score 0.5 point for this item if.

The answer mentions the introduction of the independent variable AND that it will be introduced after the first observation of the dependent variable, BUT it does not mentioned that if will be introduced only to the experimental group.

The answer mentions the introduction of the independent variable AND that it will introduced only to the experimental group, but it does not mentioned that it will be introduced after the first observation of the dependent variable.

Score 0 point for this item if.

The independent variable is mentioned, BUT there is no mention to when it is introduced NOR that it will occur only for the experimental group.

Two groups are needed; one is the control group.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that two groups are necessary and that there is a control group - either in a separate item or together with any of the items from Part A.

score 0.5 points for this item if:

The two groups are mentioned, but nothing is said about the control/experimental group.

The control/experimental group are mentioned, but nothing is said about the need of two groups.

score 0 points for this item if:

If the information described above is not mentioned.

There is randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is randomization together with any of the other items described above.

score 0 points for this item if:

The randomization is not mentioned.

Part B: **"**What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable."

R O1 X O2

R O1 O2

Score 4 points for this item if.

The observations/measures of the dependent variable are represented before and after the introduction of the independent variable.

The introduction of the independent variable is represented after the first measures of the dependent variable.

The two groups are represented in two rows (lines).

The letter R is present, illustrating the randomization.

Score 1 point for each of the items above that are illustrated.

Score 0 points for this item if none of the four items above are illustrated.

Part C: **“**Please, provide an example of this design, different from the one given in the multiple choice questions, and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

The dependent variable is heart rate and it is measured for both groups before and after the independent variable is introduced to the experimental group.

Score 1 for this item if.

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND

a dependent variable is specified AND

the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that observations of the specified dependent variable will occur for both groups before and after the introduction of the independent variable to the experimental group.

score 0.5 points for this item if:

If the dependent variable is specified and it is mentioned that its observations/measures will occur before and after the introduction of the independent variable, BUT the words “dependent

variable” or any of the synonyms are not used AND observations/measures for both groups are mentioned.

If the dependent variable is specified and the word “dependent variable” or any of the synonyms are used, but it is NOT mentioned that its observations/measures will occur before and after the introduction of the independent variable AND the observations for both groups are mentioned.

score 0 points for this item if:

The dependent variable is specified, but its place in time is not mentioned AND the word “dependent variable” or any of the synonyms are not used.

If no dependent variable is specified and its place in time is not mentioned.

Observations of both groups are not mentioned.

The independent variable is jogging and it is introduced to the experimental group after the first measure of heart rates.

score 1 point for this item if:

if the answer contains the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable)AND

an independent variable is specified AND

the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) are mentioned AND

The answer mentions that the introduction of the specified independent variable will be introduced only to the experimental group after the first observation/measure of the dependent variable in both groups.

score 0.5 points for this item if:

If the independent variable is specified, the experimental group is mentioned, and it is mentioned that the specified independent variable will be introduced to the experimental group after the first observation/measure of the dependent variable in both groups, BUT the words "independent variable" or any of the synonyms are not used.

If the independent variable is specified, the experimental group is mentioned, and the word "independent variable" or any of the synonyms are used, but it is not mentioned that the specified independent variable will be introduced to the experimental group after the first observation/measure of the dependent variable in both groups

score 0 points for this item if:

The independent variable is specified, but its place in time is not mentioned AND the word "independent variable" or any of the synonyms are not used.

If no independent variable is specified and its place in time is not mentioned.

There is no mention of the experimental group.

The participants are 50 students who will be randomly assigned to either one of two groups: experimental group and control group.

score 1 point for this item if:

The answer mentions who are the participants, how many, AND that two groups are needed (or that the participants will be divided into groups) AND that the groups will be randomly formed/ the participants will be randomly assigned.

Score 0.5 for this item if

The answer mentions who are the participants and how many, BUT does not mention that two groups are needed (or that the participants will be divided into groups) NOR that the groups will be randomly formed/ the participants will be randomly assigned.

The answer mentions who the participants are and that two groups are needed AND that the groups will be randomly formed/ the participants will be randomly assigned, BUT doesn't mention how many participants.

The answer mentions how many participants and that two groups are needed (or that the participants will be divided into groups) AND that the groups will be randomly formed/ the participants will be randomly assigned, BUT doesn't mention who are the participants.

score 0 points for this item if:

The answer mentions only how many participants.

The answer mentions only who are the participants.

The answer mentions only that two groups are needed (or that the participants will be divided into groups).

There is randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is randomization together with any of the other items described above.

score 0 points for this item if:

The randomization is not mentioned.

The Solomon Four-Group Design

Part A: “What features define the Solomon Four-Group Design? (be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

The dependent variable is measured for two groups before (O1 and O3) and after (O2 and O4) the independent variable (X) is introduced to the experimental groups. For the other two groups, the dependent variable is measured only after (O5 and O6) the independent variable is introduced to the experimental groups.

Score 1 for this item if.

the answer is exactly the same as the one above OR

The answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that observations of the specified dependent variable will occur for two groups before and after the introduction of the independent variable to the experimental group AND

That the observations/measures will occur for the other two groups only after the introduction of the independent variable to the experimental group.

score 0.5 points for this item if:

The answer mentions the observations of the dependent variable AND that its observations/measures will occur before and after the introduction of the independent variable

for two groups AND that the observations will occur only after the introduction of the independent variable for the other two groups, BUT the words “dependent variable” or any of the synonyms are not used.

The answer mentions the observations of the dependent variable AND that its observations/measures will occur before and after the introduction of the independent variable for two groups AND the words “dependent variable” or any of the synonyms are used, BUT it is not mentioned that the observations will occur only after the introduction of the independent variable for the other two groups.

The answer mentions the observations of the dependent variable AND the words “dependent variable” or any of the synonyms are used AND that the observations will occur only after the introduction of the independent variable for the other two groups, BUT it is not mentioned that observations/measures will occur before and after the introduction of the independent variable for two groups.

score 0 points for this item if:

The answer mentions the introduction of the independent variable, BUT not the fact that it will occur before each observation/measure of the dependent variable for each group AND the word “independent variable” or its synonyms are not used.

Only the dependent variable is mentioned.

The independent variable is introduced to the two experimental groups. For the first experimental group (Group A), the independent variable is introduced after the first measure of the dependent variable; for the other experimental group (Group C), it is introduced before.

Score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) and the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the introduction of the independent variable will occur after the first observation of the dependent variable for two and before the observation of the dependent variable for the other two groups AND

The answer mentions that the introduction of the independent variable will occur only to the experimental groups.

Score 0.5 point for this item if.

The answer mentions the independent variable AND that it will be introduced after the first observation of the dependent variable for two groups and before the observation of the dependent variable for the other two groups.

The answer mentions the independent variable AND that it will introduced only to the experimental groups.

Score 0 point for this item if.

The introduction of the dependent variable is mentioned, BUT there is no mention to when it will occur or that it will occur for both experimental groups.

Four groups are needed; two are control groups.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that four groups are necessary and that two are control groups - either in a separate item or together with any of the items from Part A.

score 0.5 points for this item if:

The four groups are mentioned, but nothing is said about the control groups.

The control groups are mentioned, but nothing is said about the need of four groups.

score 0 points for this item if:

None of the information above is mentioned.

There is randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is randomization together with any of the other items described above.

score 0 points for this item if:

The randomization is not mentioned.

Part B: **"**What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable."

Group A R O1 X O2

Group B R O1 O2

Group C R X O2

Group D R O2

Score 4 points for Part B if

The observations/measures of the dependent variable are represented before and after the introduction of the independent variable for Groups A and B. It is represented only after the introduction of the independent variable for groups C and D.

The introduction of the independent variable ONLY to the experimental groups, after the first observation of the DV for group A and before the observation of the DV for Group C.

The groups are represented in four rows (lines).

The letter R is present, illustrating the randomization.

Score 1 point for each of the items above that are illustrated.

Score 0 points for this item if none of the four items above are illustrated.

Part C: **“**Please, provide an example of this design, different from the one given in the multiple choice questions, and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

The dependent variable is heart rate, it is measured before and after the independent variable is introduced for two of the four groups; it is measured only after in the other two groups.

Score 1 for this item if.

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND

a dependent variable is specified AND

the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that observations of the specified dependent variable will occur for two groups before and after the introduction of the independent variable to the experimental group and only after the introduction of the independent variable to the experimental group for the other two groups.

score 0.5 points for this item if:

If the dependent variable is specified and it is mentioned that its observations/measures will occur before and after the introduction of the independent variable for two groups and only after the introduction of the independent variable to the experimental group for the other two groups , BUT the words “dependent variable” or any of the synonyms are not used.

If the dependent variable is specified and the word “dependent variable” or any of the synonyms are used AND observations for the four groups are mentioned, but it is NOT mentioned that its observations/measures will occur before and after the introduction of the independent variable.

score 0 points for this item if:

The dependent variable is specified, but its place in time is not mentioned AND the word “dependent variable” or any of the synonyms are not used.

If no dependent variable is specified and its place in time is not mentioned.

Observations of four groups are not mentioned.

The independent variable is jogging and it is introduced to the experimental groups differently: for Group A it is introduced after the dependent variable is measured. For Group C, it is presented before the dependent variable is measured.

score 1 point for this item if:

if the answer contains the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable)AND

an independent variable is specified AND

the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) are mentioned AND

The answer mentions that the introduction of the specified independent variable will be introduced only to the experimental groups after the first observation/measure of the dependent variable is conducted for two groups, but not to the other two.

score 0.5 points for this item if:

If the independent variable is specified, the experimental groups are mentioned, and it is mentioned that the specified independent variable will be introduced to the experimental groups after the first observation/measure of the dependent variable in two groups, BUT the words “independent variable” or any of the synonyms are not used.

If the independent variable is specified, the experimental groups are mentioned, and the word “independent variable” or any of the synonyms are used, but it is not mentioned that the

specified independent variable will be introduced to one of the experimental groups only after the first observation/measure of the dependent variable in two groups

score 0 points for this item if:

The independent variable is specified, but its place in time is not mentioned AND the word "independent variable" or any of the synonyms are not used.

If no independent variable is specified and its place in time is not mentioned.

There is no mention of the experimental groups.

The participants are 100 students who will be randomly assigned to either one of four groups: experimental group A or C, control group B or D.

score 1 point for this item if:

The answer mentions that four groups are needed AND that the groups will be randomly formed/ the participants will be randomly assigned AND that two groups will be experimental groups and two will be control groups.

Score 0.5 point for this item if

The answer mentions that four groups are needed AND that the groups will be randomly formed/ the participants will be randomly assigned OR

The answer mentions that four groups are needed AND that two groups will be experimental groups and two will be control groups.

score 0 points for this item if:

There is no information about groups.

The word "group" is not mentioned.

There is randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is randomization together with any of the other items described above.

score 0 points for this item if:

The randomization is not mentioned.

The Posttest-Only Control Group Design

Part A: “What features define the Posttest-Only Control Group Design ? (be sure to write: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. How many groups are needed and if there is a control group; iv. if there is randomization.”

The dependent variable is measured for both groups only after (O) the independent variable (X) is introduced to the experimental group.

Score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that observations of the dependent variable will occur after the introduction of the independent variable AND

The answer mentions that observations of the dependent variable will occur for both groups.

Score 0.5 point for this item if.

The answer mentions the observations/measures of the dependent variable AND that they will occur after the introduction of the independent variable BUT it does not mention that they will be observed in both groups.

The answer mentions the observations/measure of the dependent variable AND that they will be observed in both groups BUT it does not mention that the measures will occur after the introduction of the independent variable.

Score 0 point for this item if.

The observations/measures of the dependent variable are mentioned, BUT there is no mention to when they will occur or that it will occur for both groups.

The independent variable is introduced to the experimental group before the measure of the dependent variable.

Score 1 point for this item if.

the answer is exactly the same as the one above OR

if the answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) and the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) AND

The answer mentions that the introduction of the independent variable will occur before the observations of the dependent variable AND

The answer mentions that the introduction of the independent variable will occur only to the experimental group.

Score 0.5 point for this item if.

The answer mentions the independent variable AND that it will be introduced before the observations of the dependent variable BUT it does not mention that it will introduced only to the experimental group.

The answer mentions the independent variable AND that it will introduced only to the experimental group BUT it does not mention that it will be introduced before the observations of the dependent variable.

Score 0 point for this item if.

The dependent variable is mention, BUT there is no mention to when it will occur or that it will occur for both groups.

Two groups are needed; one is the control group.

score 1 point for this item if:

The answer is exactly the same as the one above OR

The answer mentions that two groups are necessary and that there is a control group - either in a separate item or together with any of the items from Part A.

score 0.5 points for this item if:

The two groups are mentioned, but nothing is said about the control/experimental group.

If the control/experimental group is mentioned, but nothing is said about the need of two groups.

score 0 points for this item if:

If the information above is not mentioned.

There is randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is randomization together with any of the other items described above.

score 0 points for this item if:

The randomization is not mentioned.

Part B: **“**What is the notation that represents this design? (be sure to illustrate: i. when the dependent variable is measured; ii. when the independent variable is introduced; iii. The group(s); iv. The randomization, if applicable.”

R X O1

R O1

Score 4 points for Part B if

The observations/measures of the dependent variable are represented after the introduction of the independent variable to the experimental group.

The introduction of the independent variable to the experimental group is represented before the observations of the dependent variable.

The two groups are represented in two rows (lines).

The letter R is present, illustrating the randomization.

Score 1 point for each of the items above that are illustrated.

Score 0 points for this item if none of the four items above are illustrated.

Part C: **“**Please, provide an example of this design, different from the one given in the multiple choice questions, and be sure to include: i. what is the dependent variable(s) and when it will be measured (observed); ii. what is the independent variable(s) and when it is introduced; iii. who are the participants, how many participants you will have, and if the participants will be divided into groups; iv. will there be randomization?

The dependent variable is heart rate and it is measured for both groups after the independent variable is introduced to the experimental group.

Score 1 for this item if.

The answer contains the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) AND

a dependent variable is explicitly specified AND

the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable) are mentioned AND

The answer mentions that observations of the specified dependent variable will occur for the two groups <u>after</u> the introduction of the independent variable to the experimental group.

score 0.5 points for this item if:

If the dependent variable is specified AND it is mentioned that its observations/measures will occur after the introduction of the independent variable to the experimental group BUT the words “dependent variable” or any of the synonyms are not used NOR it is mentioned that the observations will occur for both groups.

If the dependent variable is specified AND it is mentioned that observations will occur for both groups AND the word “dependent variable” or any of the synonyms are used, BUT it is not mentioned that its observations/measures will occur after the introduction of the independent variable.

score 0 points for this item if:

The dependent variable is specified, BUT when its observations/measures will occur is not mentioned, NOR is the fact that these observations will occur for both groups NOR is the word “dependent variable” or any of the synonyms used.

No dependent variable is specified NOR are its observations.

No dependent variable is specified NOR is the fact that its observations will occur for both groups.

The independent variable is jogging and it is introduced to the experimental group before heart rates (dependent variable) are measured in both groups.

score 1 point for this item if:

if the answer contains the words (independent variable, intervention, "predictor variable," "regressor," "controlled variable," "manipulated variable," "explanatory variable," "exposure variable)AND

an independent variable is explicitly specified AND

the words (behavior, dependent variable, outcome variable, response variable," "regressand," "measured variable," "observed variable," "responding variable," "explained variable," " and/or "output variable) are mentioned AND

The answer mentions that the introduction of the specified independent variable will be introduced only to the experimental group before observations/measures of the dependent variable in both groups.

score 0.5 points for this item if:

The independent variable is specified, AND the experimental group is mentioned, AND it is mentioned that the specified independent variable will be introduced to the experimental group before the observations/measures of the dependent variable in both groups, BUT the words "independent variable" or any of the synonyms are not used.

The independent variable is specified, AND the experimental group is mentioned, AND the word "independent variable" or any of the synonyms are used, BUT it is not mentioned that the specified independent variable will be introduced to the experimental group before the observation/measure of the dependent variable in both groups.

The independent variable is specified, AND it is mentioned that the specified independent variable will be introduced to the experimental group before the observations/measures of the dependent variable in both groups, AND the word "independent variable" or any of the synonyms are used, BUT the experimental group is not mentioned.

score 0 points for this item if:

The independent variable is specified, BUT when it will be introduced is not mentioned AND the word "independent variable" or any of the synonyms are not used.

If no independent variable is specified AND when it will be introduced is not mentioned.

If no independent variable is specified AND there is no mention of the experimental group.

The participants are 50 students who will be randomly assigned to either one of two groups: experimental group and control group.

score 1 point for this item if:

The answer mentions who are the participants, how many, AND that two groups are needed (or that the participants will be divided into groups) AND that the groups will be randomly formed/ the participants will be randomly assigned.

Score 0.5 for this item if

The answer mentions who are the participants and how many, BUT does not mention that two groups are needed (or that the participants will be divided into groups) NOR that the groups will be randomly formed/ the participants will be randomly assigned.

The answer mentions who the participants are and that two groups are needed AND that the groups will be randomly formed/ the participants will be randomly assigned, BUT doesn't mention how many participants.

The answer mentions how many participants and that two groups are needed (or that the participants will be divided into groups) AND that the groups will be randomly formed/ the participants will be randomly assigned, BUT doesn't mention who are the participants.

score 0 points for this item if:

The answer mentions only how many participants.

The answer mentions only who are the participants.

The answer mentions only that two groups are needed (or that the participants will be divided into groups).

There is randomization.

score 1 point for this item if:

the answer is exactly the same as the one above OR

the answer mentions that there is randomization together with any of the other items described above.

score 0 points for this item if:

The randomization is not mentioned.

Appendix M – Glossary

Activity 1 Stimuli

Match the definitions with their "names":

1) There is a **causal** relationship between a dependent variable and an independent variable, in which one variable (independent variable) sets up the occurrence or non-occurrence of another variable (dependent variable). **CAUSE AND EFFECT**

2) There is a relationship between the dependent variable and the independent variable. It is not necessarily a **"cause-and-effect"** relationship, but the variables occur together and are related. **CORRELATION**

3) What is being measured in a study, the outcome of interest. When you want to define what this variable is, you need to answer the question "What will I observe/measure?" In the behavioral sciences, this variable is the BEHAVIOR of interest. **DEPENDENT VARIABLE**

4) It is a way to graphically represent the design. Observations or Measures are symbolized by an '**O**'; Treatments or Programs are symbolized with an '**X**'; and each **group** is given its own line (e.g., if there are three lines, there are three groups in the design). **NOTATION**

5) Variable whose value determines that of the dependent variable. It is the variable which an experimenter deliberately manipulates in order to observe its relationship to the observed changes in the dependent variable. We say there is a relationship between these variables and the dependent variables when changes in the first one are correlated with or produce changes in the second one. This variable can be a treatment or an intervention (manipulated variables) or it can be variables of the participants of study (age, race, gender, etc). **INDEPENDENT VARIABLE**

6) This is the consistency of your measurement or the degree to which an instrument or observer measures the same phenomenon each time it is used under the same condition with the

same participants. In short, it is the repeatability or consistency of your measurement.

RELIABILITY

7) This refers to the degree to which the results are accurate and the degree that they measured what was proposed to be measured in the study. This can be assessed by answering the following two questions: "Did the measurement system accurately measure what it claimed to measure?" and "Did the results provide an answer to the research question?" **VALIDITY**

Activity 2 Stimuli

Match the definitions with their "names":

1) Is the demonstration that the independent variable accounts for the changes in what is being measured (dependent variable). It means the researcher has identified what is either **correlated with** or **causes** changes in the dependent variable. **EXPERIMENTAL CONTROL**

2) A group of participants that closely resembles the group that receives the intervention or treatment in demographic variables such as age, gender, years of education, etc. This group does not receive the intervention or the factor under study, thus serving as a comparison group when treatment results are evaluated. **CONTROL GROUP**

3) A method based on chance alone by which study participants are assigned to either a control or a treatment group. This process minimizes the differences among groups by equally distributing people with particular characteristics among all the groups. By doing this, researchers reduce the chance that one group would inherently get better (or worse) results than another. **RANDOMIZATION**

4) A method of generating a random sample. In this method, a table of numbers is generated in an unpredictable sequence. Use you this table by randomly picking a page of the table and

dropping your finger on the page with your eyes closed. Then, you choose a direction in which to read (up to down, left to right, or right to left) and you select the numbers you need. **RANDOM NUMBERS TABLE**

5) Subset of an entire target- population of the study. These are used because to research an entire population is very costly, might take a long time and the population is dynamic in that the individuals making up the population may change over time. **SAMPLE**

6) When a researcher creates (unintentionally or intentionally) a preference by the way that participants are selected for inclusion in the study, and even in the way participants are selected for each one of the groups of a study. **SELECTION BIAS**

Appendix N – E-mail Sent After the Article Review (pretest)

Oi (participant's name). Boa noite!!

Obrigada por responder ao checklist.

Analisei seus dados e ainda não posso passar seus dados exatos (só ao final do estudo vou mandar os gráficos e resultados com seus desempenhos), mas adianto você foi bem. Lembre-se que vou passar todos os seus dados em gráficos ao final do estudo.

O próximo passo será responder às nove questões abertas e seus subitens pela primeira vez. Aqui está o link:

Por favor, tente responder às questões o mais rápido possível. Se você não souber as respostas, responda "eu não sei a resposta" ou dê qualquer indicação de que você não sabe, mas você terá de responder algo para passar para as perguntas seguintes.

Leia bem as instruções antes de responder às questões. Por favor, não procure informações acerca das perguntas, nem durante, nem após responder às questões, pois a pesquisa se refere, em sua maior parte, ao ensino destas informações.

Qualquer problema que encontrar, é só me avisar.

Grande abraço e bom trabalho!

Carol

Appendix O – Research Procedures Explanation on Skype

Agora eu vou explicar os passos pelos quais vamos passar com o tutorial.

Revisão de artigo: Eu enviarei um PDF de um artigo e um link por e-mail. Você vai ler o artigo e fazer a avaliação dele através do link. Instruções mais detalhadas virão no e-mail.

Questões abertas: depois da revisão do artigo, você receberá 3 links diferentes com questões abertas sobre métodos de pesquisa. Por que 3 vezes? Porque eu tenho que ter certeza que você não conhece o conteúdo que vou ensinar. Só um pré-teste não me dá certeza do seu desempenho.

Depois das 3 vezes em que você responderá às perguntas, você receberá um link para atividades de glossário, ou seja, atividades que não contarão para a pesquisa, mas te ajudarão com o vocabulário científico. Depois que você tiver feito a atividade, a re-mandarei por e-mail para que você possa ter as palavras e os significados delas ao seu lado durante o tutorial.

No mesmo dia do glossário, você iniciará o tutorial em si. Mandarei dois links. O primeiro é do tutorial em si: são perguntas de múltipla escolha que vão te fornecer feedback imediato e que estão programadas para voltar para o começo do tutorial cada vez que você escolher uma alternativa incorreta. Então, procure prestar atenção e ler as alternativas com calma, mesmo que pareca repetitivo. O segundo link se refere a uma generalização do que você aprendeu no tutorial. Nesta "generalização" você responderá a perguntas de múltipla escolha similares às que você viu no tutorial, mas com pequenas diferenças na ordem de apresentação da informação.

Depois dos dois links, você responderá a perguntas de múltipla escolha novamente. Elas servirão para me mostrar se o tutorial foi efetivo em te fornecer as informações necessárias para

os 3 primeiros delineamentos de pesquisa os quais você acabou de aprender no tutorial. Se você não acertar as questões relativas aos 3 primeiros tutoriais, você terá de retornar ao ensino. Se acertar, passará para a segunda parte do tutorial de ensino.

O mesmo se repetirá mais duas vezes.

A ultima coisa que você fará é uma nova revisão do artigo e, se quiser, poderá fornecer sugestões para melhorar o tutorial.

Qualquer pergunta durante a pesquisa, você pode me escrever ou pedir uma reunião no Skype que esclarecerei suas dúvidas.

Appendix P – Social Validity Questionnaire[4]

Answer the following questions using a 1 to 5 scale. 1 means that you strongly disagree, 2 that you disagree, 3 that you do not agree, 4 that you agree, and 5 that you strongly agree.

1) The tutorial was easy to use.	**1**	**2**	**3**	**4**	**5**
2) The fact that the tutorial was online allowed me/ gave me flexibility to go through the sessions from where I wanted and when I wanted.	**1**	**2**	**3**	**4**	**5**
3) The information provided in the tutorial was useful to me.	**1**	**2**	**3**	**4**	**5**
4) The experimenter sent the links to the sessions within the agreed period of time.	**1**	**2**	**3**	**4**	**5**
5) The feedback provided during the teaching sessions were important.	**1**	**2**	**3**	**4**	**5**
6) I would recommend the tutorial to other people.	**1**	**2**	**3**	**4**	**5**

7) What was the most useful feature of this tutorial? _____

8) What was the least useful feature of this tutorial? _____

9) If you have any additional comments or suggestions, please, describe them below.

_____

_____

_____

_____

_____

[4] The actual questionnaire was presented in Portuguese. A copy of the original version can be obtained from the author.

Appendix Q - Ratings in the Social Validity Questionnaire

Proposition	Rating John	Rating Mary
1	5	4
2	5	5
3	5	5
4	5	5
5	5	5
6	5	4
7	I believe that the method used in the tutorial, in which the multiple-choice questions are presented since the beginning and feedback is provided letting me know if my responses are right or wrong, was the featured that contributed the most for my learning of the content. After establishing the pattern between the "correct associations" – "correct response" – which took about two or three questions to be established, the long series of multiple-choice questions was important to "fixate" the terms that I would have to transcribe in the open-ended questions.	The teaching of new research designs.
8	I cannot point to any feature as "less useful", since this tutorial was a unique experience for me; I cannot compare it to other tutorials.	None.
9	One feature that I found interesting throughout the tutorial, specifically in the multiple-choice questions with feedback, was the fact that the paragraphs that constituted the answer alternatives, kept their comprising sentences the same, however, their "shapes" were modified, so I could not choose an alternative based on its "format" or "shape".	The tutorial is easy to be used and it makes learning easy, however, the open-ended questions become tiring because they are repeated at the end of every part of the procedure, since the beginning. Thus, it becomes a little boring to do it all over again at the end of each part.

Appendix R – Pilot Study Summary

Method

Participant

The participant was a 20 years-old male whose primary language was Chinese. He was taking Level 5 classes at the Applied English Center and passed the CAT web-based reading comprehension pre-test (http://fcit.usf.edu/fcat10r/home/sample-tests/virtues-of-venom/index.html). He was going to start taking Psychology classes as soon as he took the last test from the Applied English Center. Additionally, the participant stated that he had not had any research methods classes and did not know much about research. The participant was paid U$5.00 per each session he attended and an additional U$5.00 for each article review he performed using the checklist. The participant read and filled out the consent form before participating in any research activities.

Setting, Materials, and Stimuli

The research was conducted in an office at a state university in Kansas. The materials used included a desktop computer, a desk, two chairs (one for the participant and one for the experimenter), pencils and erasers, two copies of the review paper and of the review checklist; two knowledge quizzes; and four sets of open-ended questions.

Experimental Design, Independent and Dependent Variable

A multiple probe design across teaching parts was used. The independent variables were equivalence-based instructions, training to mastery in teaching sessions, and immediate feedback for correct and incorrect responses in teaching sessions. The dependent variable was the percent of correct responses, measured during teaching and testing sessions.

General procedures

Experimental phases were presented in the order shown in Figure O1. First, the participant reviewed the Whitehurst et al. (1988) article. The checklist and the article were presented in a paper format. Second, the participant was asked to answer two 54-question multiple-choice knowledge quizzes. These had been programmed to be the probes, however, not only participant's performance increased from the first to the second knowledge quiz, but the participant reported that all answers were embedded in the questions. It is important to highlight that the experimenter did not provide feedback on specific responses to the quiz questions; only social praise was provided after the participant finished responding all questions (e.g., "Good job", "Well done", "Great"). Third, nine open-ended questions were design to serve as a new probe. These open-ended questions were similar to the ones presented in the main study; however, the prompts presented in the questions were not divided into sub-items. After the one set of open-ended questions was presented, teaching sessions were presented. As with the main study, teaching was divided into three parts. Each teaching part was comprised of a minimum of three and a maximum of six teaching sessions (these three sessions were compiled into one in the main study). As in the main study, after each teaching part was mastered, a probe was presented.

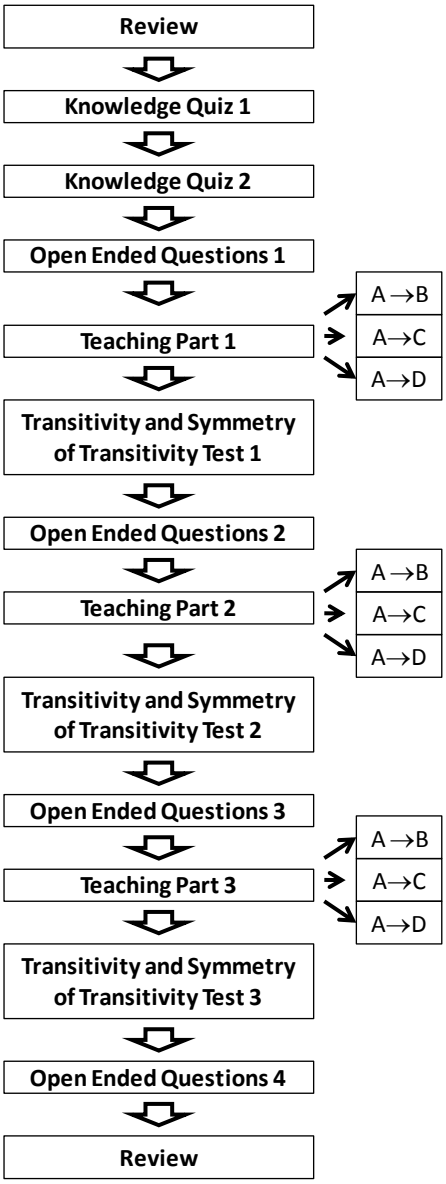


Figure O1. Flowchart illustrating the pilot study experimental phases.

Results and Discussion

Performance in the knowledge quizzes were 55% in the first one and 77% in the second one. The participant reported that answers were contained in the test.

The participant met criterion for Teaching Part 1 after four teaching sessions (one name-definition session, two name-notation sessions, and one name-example session). The participant met criterion after five teaching sessions in Part 2 and in Part 3.

Performance in transitivity and symmetry of transitivity tests were 100%; 87.5%; and 100% after each one of the teaching sessions (from Part 1, Part 2, and Part 3, respectively). The second test was conducted with a week delay from the teaching session, while the other two tests were presented right after a teaching session.

Performance in open-ended questions: 6.9%; 33.3%; 58.3%; 58.3%. Until the third probe, as it can be seen on Figure 2, performance was according to predicted. After the participant spaced out his sessions, his performance on designs that he was already trained in decreased. As with other complex "informational content", there is no maintenance. It is important to highlight that this participant had been exposed to all stimuli during the knowledge quizzes, which might explain the difference between this study and the main study.

Performance in the article reviews went from 43.5% to 60.8%. A change was not expected since peer review was not directly tackled.

The participant reported that the Pre-experimental Designs and True experimental Designs names allowed the inference on parts of the definition, notation, and examples.

This pilot study served the purpose of validating and adjusting the content of testing and teaching tasks. A table of changes from the pilot to the main study can be found on page 154.

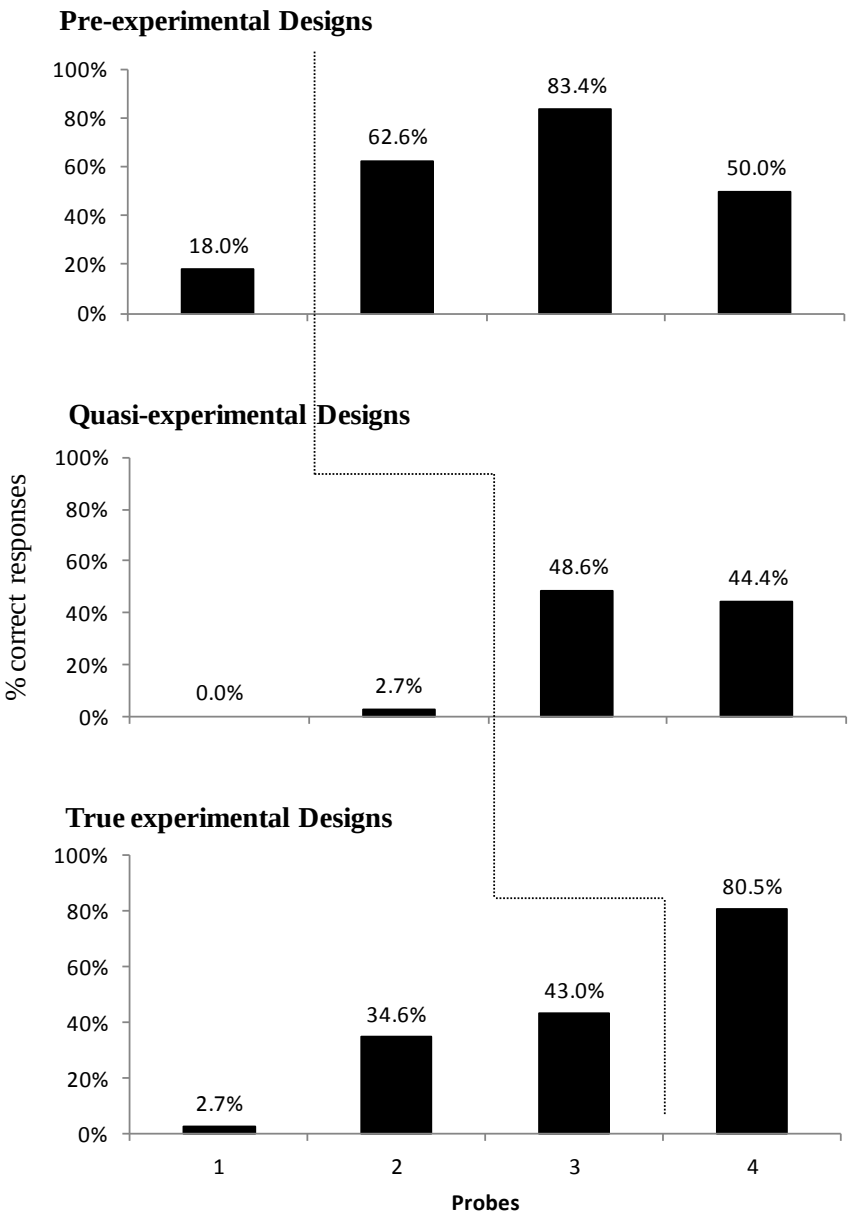


Figure O2. ZYI performance in the open ended questions, separated by teaching parts. The first graph represents performance in the Pre-experimental Designs during probes number 1, 2, 3, and 4. The second and third represent performance in Quasi-experimental and True experimental Designs, respectively.

Table O1

Changes from the pilot study to the main study

Item	Pilot	Future study	Rational for the change
Participant recruitment	We recruited freshmen and sophomores from The University of Kansas, Lawrence Campus.	Participants will be recruited from KU and from four Brazilian universities (UFGD, UNCISAL, UEL, UFSCar). They can be at any year; however they cannot have advanced knowledge on social and behavioral science research designs.	Only international students responded to the fliers. Only one international student finished the study. Former Brazilian students always write the experimenter for information on research and possible online classes on behavior analysis and research.
Participant incentives	The participant was paid U$5.00 for every session and additional U$5.00 for each review.	Brazilian participants will receive a certificate naming the research designs they were in contact with and the number of hours spent in the research study. Also, they will receive a copy of one of Skinner's books (they will be able to choose among five titles). Most Portuguese translations are not well done and access to the originals is difficult in Brazil. Participants from KU will be paid U$10.00 per session and U$20.00 per review.	It is illegal to pay participants with money in Brazil. At the same time, it is very difficult for students in Brazil to get access to organized and clear information about research designs (this is not the focus of Psychology programs). Thus, information will be used as an incentive instead of money. For American participants: the incentives predicted in the pilot were not enough to call any American student's attention, so, we will increase the incentives.
Language	The participant had to write in English.	Psychology students in Brazil tend to read most of their papers and books in English, however, the discussions and written products are all in Portuguese; writing in English might require a skill that is not the target of this intervention. Thus, the participants might be allowed to write in Portuguese and reliability will be done by a second observer from Brazil.	The participant asserted that he did not know how to write some things about the design in English. Thus, Brazilian students will be allowed to write the open-ended questions in Portuguese. Colleagues from the experimenter will do the reliability on those questions.
Setting	An office at the Research and	Virtual environments: Adobe Captivate® sessions (review	The participant stated that it was hard to keep coming for in

	Training Center on Independent Living.	and training) will be published through Adobe Connect®. Open-ended questions will be presented through Skype and/or Connect and will be recorded through tools in the software.	person sessions, since he had to take the bus. Online sessions will allow participants to go through training sessions whenever they want. The probes will have a scheduled time, but they will not need to be at a certain place to go through the probes.
Knowledge quizzes	Two knowledge quizzes were presented	No knowledge quizzes.	Forced choice tasks, like quizzes, give the participant the opportunity not only to get a correct response by chance, but also provide information about the content. Thus, these tasks are not appropriate as probes.
Open ended Questions (probe)	Were presented on paper and did not contain specific prompts for each part of the written response.	**Part A of the questions:** What features define the (name of the design) design (be sure to write: i. when is the independent variable is introduced and ii. when the dependent variable is measured; iii. How many groups are needed and if there is a control group; iv. if there is randomization. (4 items to be evaluated) **Part B of the questions:** What is the notation that represents this design? (be sure to illustrate: i. when the independent variable is introduced and ii. when the dependent variable is measured; iii. The groups; iv. The randomization. (4 items to be evaluated) **Part C of the questions:** Please, provide an example of this design, different from the one given in the multiple choice questions, and be sure to include: ii. who are the participants and how many participants you will have, iii.	The participant did not explicitly write all the items that the experimenter was expecting. To make sure that all items are in the answer, there will be prompts, as pointed out in the middle column.

		if the participants will be divided into groups and how (will there be randomization?), iv. what is the independent variable(s), v. what is the dependent variable(s) and when it will be measured.	
Open ended questions	No criterion was set.	For content that will have been taught: 80%. If criterion is not met for taught content, training will be re-presented. For content that will not have been taught: must remain under 20%**.**	In the pilot it was assessed of participants would increase performance during probes just by having the criterion within training sessions. Since the pilot showed that participants might not maintain performance, a criterion will be in place for the next study.
Reliability	Was not measured, but during knowledge quizzes and open ended questions.	The review and training will be delivered by Adobe Captivate. The program will be calibrated before data collection to assure that the answers are being reliably recorded, the content is being delivered accordingly and all instruction are understandable. For the open-ended questions, all participants' answers will be typed out on Skype and/or Connect and will be checked by a second observer.	The participant stated that he did not feel comfortable being recorded. Thus, the tasks that had a permanent product could have reliability for them, but training could not. With captivate and Connect, all training will be delivered and recorded by the computer programs. For open-ended questions, if participants do not want to get their voices recorded either, the permanent product will still be available for reliability.