Testing the Correlation Between Response Latency, Derivation, and Complexity

Jeffrey Ryan Oliver
University of South Florida, jroliver@mail.usf.edu

Follow this and additional works at: https://scholarcommons.usf.edu/etd
Part of the Child Psychology Commons

Scholar Commons Citation
https://scholarcommons.usf.edu/etd/5378

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Testing the Correlation Between Response Latency, Derivation, and Complexity

by

Jeffrey Oliver

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts
Department of Child and Family Studies
College of Behavioral and Community Sciences
University of South Florida

Major Professor: Timothy Weil, Ph.D.
Kimberley Crosland, Ph.D.
Andrew Samaha, Ph.D.

Date of Approval:
June 25, 2013

Keywords: implicit relational assessment procedure, IRAP, derivation, complexity, applied behavior analysis
Copyright © 2014, Jeffrey Oliver
# Table of Contents

List of Figures  
Abstract  
Chapter One: Introduction  
  Implicit Measures  
  Applications  
  Variable Affecting the IRAP Effect  
  Relational Elaboration and Coherence (REC) Model  
  Coherence  
  Complexity  
  Derivation  
Chapter Two: Method  
  Participants  
  Setting and Equipment  
  Design  
  Procedure  
Chapter Three: Results  
Chapter Four: Discussion  
References  
Appendices  
  Appendix A: Target Stimuli
List of Figures

Figure 1: IAT 15
Figure 2: IRAP 15
Figure 3: Illustration of Complexity 16
Figure 4: A flowchart of progression of the study 25
Figure 5: Participant 1 3D Graph 28
Figure 6: Participant 2 3D Graph 29
Figure 7: Participant 3 3D Graph 30
Figure 8: Participant 4 3D Graph 31
Figure 9: Participant 5 3D Graph 32
Figure 10: Participant 6 3D Graph 33
Figure 11: Participant 7 3D Graph 34
Figure 12: Participant 9 3D Graph 35
Figure 13: Participant 10 3D Graph 36
Figure 14: Participant 11 3D Graph 37
Figure 15: Participant 12 3D Graph 38
Figure 16: Participant 13 3D Graph 39
Figure 17: Participant 14 3D Graph 40
Figure 18: Participant 15 3D Graph 41
Figure 19: Participant 16 3D Graph 42
Figure 20: Participant 18 3D Graph 43
Figure 21: Participant 19 3D Graph 44
Figure 22: Participant 20 3D Graph 45
Figure 23: Participant 21 3D Graph 46
Figure 24: Participant 22 3D Graph 47
Abstract

This study investigates the constructs of Derivation and Complexity and how they relate to latency. Derivation and Complexity are theoretical constructs that have been posited as two of the main factors in differences in latency to responding in implicit measures such as the Implicit Relational Assessment Procedure (IRAP) and the Implicit Association Task (IAT). This study trained participants to relate two groups of novel stimuli in a linear fashion and then tested their latency to responding to derived relations (relations based on previously trained relations, but not directly trained themselves). The study then analyzed participant’s latency to responding after dividing the responses based on derivation, complexity, and phase. The study found a significant relationship between phase and latency (p=.01), derivation and latency (p=.01), and complexity and latency (p=.04). This indicates that brief, immediate relational responses are influenced by both derivation and complexity as well as practice responding and these variables should be considered in future investigations into implicit attitudes.
Chapter One: Introduction

Relational Frame Theory (RFT) is an account of verbal behavior that asserts relational networks, with the necessary features of mutual entailment, combinatorial entailment, and transformation of stimulus function, are the basic unit of verbal behavior (Blackledge, 2003). Arbitrarily Applicable Derived Relational Responding (AADRR), an ability that is said to emerge through basic operant processes, is based on arbitrary characteristics assigned to stimuli by verbal communities and focuses on the relationship between these stimuli. Arbitrarily Applicable Derived Relational Responding is learned through multiple exemplar training starting with relational responding based on formal properties in a bidirectional way and is subsequently generalized to arbitrary variables that may come to control this behavior (Blackledge, 2003). Relational Frame Theory presents a pragmatic and functional approach to human language that has allowed behavior analysts to venture into domains of research typically dominated by cognitive psychologists, including treatment of psychopathology (Hayes, Strosahl, & Wilson, 1999), intelligence (Cassidy, Roche, & Hayes, 2011), executive function (Weil, Hayes, & Capurro, 2011; O’Neill & Weil, 2014), and implicit cognition (Barnes-Holmes, Hayden, Barnes-Holmes, & Stewart, 2008). Leading this surge in behavior analytic research related to implicit cognitions is a measurement tool called the Implicit Relational Assessment Procedure, or IRAP (Barnes-Holmes, Barnes-Holmes et al., 2006). The IRAP relies on response latencies to illuminate patterns of relational
responding. Of special note, it looks to record responding that may show questionable veracity in self-report measures.

**Implicit Measures**

The IRAP is a computer based assessment procedure used to assess the relationships that have been trained or derived between stimuli. The IRAP is based on a previous cognitive psychology tool called the Implicit Association Task (IAT). The IAT uses four separate groups of words or pictures associated with four response stimuli presented in a computer task. Two of these response stimuli are the test stimuli (in figure 1 either black patient or white patient) and the other two response stimuli are typically pleasant/unpleasant or bad/good (as is the case in figure 1) to allow for the interpretation of implicit attitudes. The four word groups are associated with one of the test stimuli or with the implicit attitudes stimuli (e.g. good or bad). During test phases, the participant responds to the target stimulus (a word or picture from one of the four groups; the picture of the black man in figure 1) by associating it with its correct response stimulus (i.e. “black patient” in figure 1). By alternating which response stimuli share a response key (e.g. “good” and “white patient” share a response key in figure 1, while “bad” and “black patient” share a response key; this would be swapped so that “good” and “black patient” shared a response key and “bad” and “white patient” shared a response key), researchers assert that they can then interpret which test response stimulus is associated most closely with good/pleasant or bad/unpleasant based on response latency (measured in milliseconds from the time the stimuli are presented on the screen to the time the individual responds by pressing either the d or k button on
the keyboard), thus uncovering the participants’ implicit attitude toward the test stimuli (Greenwald, McGhee, and Schwartz; 1998). The IRAP, as outlined in Barnes-Holmes, Barnes-Holmes, Stewart, and Boles (2010), is similar to the IAT in several ways: it requires participants to respond to stimuli using a computer-based presentation and it measures the latencies of these responses and the IRAP is presented on a computer screen using black letters on a white background. In the top middle of the screen a label stimulus is presented (in figure 2 a picture of a black male), in the center of the screen a target stimulus is presented (“good” in figure 2), and on the bottom left and right are two response options which are assigned a specific response key on the keyboard (“same” and “different” in Figure 2). All IRAP tests have at least two possible labels which are equivalent to the two response stimuli being tested in the IAT. The target stimuli in the IRAP function similarly to the target stimuli in the IAT that are used to test for positive and negative valence of attitudes and cognitions (these would be the target stimuli associated with the good/bad or positive/negative response stimuli in the IAT) and typically include anywhere form 6-12 possible target stimuli. However, the two response options presented on the bottom left and right of the screen in the IRAP give it an added level of analysis not available in the IAT.

Unlike the IAT which only permits responding to good/bad, the response options in the IRAP are typically specific relations such as more/less, same/different, before/after, etc. This allows IRAP researchers to assess the relationship between the label stimulus and the target stimulus presented on the screen through their evaluation of response latencies. The IRAP allows a more
detailed view of the subjects learning history with respect to the stimuli presented. Researchers are able to understand exactly how certain stimuli may be related as well as the strength of the response as indicated by latency measures. The IRAP allows for more varied and flexible responding by the participants and can illuminate patterns of responding to which the IAT is insensitive. Two separate conditions exist in the IRAP: one condition requires participants to respond in a way that would be consistent with a typical participant’s learning history and the other would be considered inconsistent with a typical participant’s learning history. As an example, Barnes-Holmes et al. (2008) used pleasant and unpleasant as the label stimuli, words such as love as positive target stimuli and items such as sickness as negative target stimuli. The two response options were similar and different. Using these particular stimuli all potential consistent IRAP trials are: Pleasant/love/similar, unpleasant/sickness/similar, pleasant/sickness/different, unpleasant/love/different. Conversely, all possible inconsistent IRAP trials are: pleasant/love/different, unpleasant/sickness/different, pleasant/sickness/similar, and unpleasant/love/similar. By combining the response latencies to all of the consistent trials and the response latencies for the inconsistent trials, researchers can uncover how quickly one type of relationship is emitted compared to another, thus indicating the response strength of a particular relationship between stimuli, similar to the IAT. This can allow the researcher to uncover positive or negative “attitudes” for each stimulus, unlike the IAT where a positive attitude toward one stimulus automatically indicates a negative attitude toward another. An important point to make concerning the IRAP compared to the IAT is that the IRAP requires
participants to respond directly to relationships between stimuli, whereas the IAT requires the participants to respond only to the stimuli themselves, without indicating a particular relationship between them (Barnes-Holmes, Waldron, Barnes-Holmes, & Stewart, 2009).

Applications

The IRAP may be useful in many clinical areas. It has shown to correlate with some clinically relevant behaviors, which are useful for predicting problematic overt behavior and as an assessment tool for therapists. Additionally, the IRAP can be useful for identifying potential problematic verbal repertoires and guide clinicians in constructing interventions, if not being used to affect change on these verbal repertoires itself. IRAP researchers have begun to investigate correlations between IRAP performance and clinical assessment tools for depression (Hussey & Barnes-Holmes, 2012), phobias (Nicholson and Barnes-Holmes, in press), adolescent smoking behavior (Vahey, Boles, & Barnes-Holmes, 2010), criminal behavior (Dawson, Barnes-Holmes, Gresswell, Hart, & Gore, 2009), cocaine addiction (Carpenter, Martinez, Vadhan, Barnes-Holmes, & Nunes, in press), intelligence testing (O'Toole & Barnes-Holmes, 2009), and have used the IRAP as a training tool (Hughes & Barnes-Holmes, 2011).

Variables Affecting the IRAP Effect

Barnes-Holmes, Murphy, Barnes-Holmes, and Stewart (2010) directly tested the effect of both accuracy and latency criterion on IRAP effects. They found that IRAP effects were more discernable when accuracy criterion was held at a higher percentage of correct responding and when latency was held to shorter time periods
between stimulus presentation and responding. It is interesting to note that latency is so important to IRAP effects that it caused what would appear to be opposite effects on responding to socially sensitive IRAP tasks. That is, as latency to responding criteria are lowered, responses become less affected by other seemingly important contextual variables. As an example, Barnes-Holmes, et al. (2010) directly tested participant responding in a private and public context for a black vs. white bias IRAP and found that participants in the public context responded as more pro-white and anti-black than those in the private condition. It is hypothesized that this was because the rule to respond quickly was more salient in the public condition when an experimenter was present.

**Relational Elaboration and Coherence (REC) Model**

The Relational Elaboration and Coherence Model (REC model) has been outlined by Barnes-Holmes, Barnes-Holmes, et al. (2010). The model gives a behaviorally consistent organization of what has been termed “implicit attitude” while eliminating any unnecessary appeal to pure associationism or metaphysical constructs. The REC model achieves this goal by pointing to arbitrarily applicable derived relational responding as the basic behavioral phenomenon being recorded in implicit measures (i.e., as implicit attitude) with implicit attitudes being those responses that occur at extremely short latencies.

The REC model makes use of the terms brief immediate relational responding (BIRR) and extended elaborated relational responding (EERR) to easily divide between relational responses of interest on implicit measures (measures that focus on verbal responses that occur at extremely short latencies) compared to
those of interest on explicit measures (measures that focus on verbal responses that occur at relatively long latencies). BIRRs are categorized based on responding that occurs at relatively short latencies, while EERRs are categorized based on responding that occurs at relatively longer latencies. There is no formal divide between BIRR’s and EERR’s—the IRAP could be used similarly to a pencil and paper measure if the response latency is not restricted—and relational responses are viewed on a continuum with BIRR’s occurring relatively faster compared to EERR’s. This lack of a formal divide leads BIRR’s and EERR’s to be defined based on the type of measure used to uncover them: whether an implicit measure, which focuses on responses of short latency, or an explicit measure, which focuses on responses of more considerable length, was used to uncover the demonstrated relationship. But Implicit measures and explicit measures are defined by their ability to measure BIRR’s and EERR’s, respectively. Instead, BIRR’s and EERR’s must be defined independently of the measure used to detect them, and our measures subsequently defined by their ability to measure BIRR’s and EERR’s.

Now let’s consider an example of potential BIRR’s and EERR’s. In this example we will use a hypothetical participant who is asked to respond to a picture of a black male as either good or bad (as in figure 2; Sabin, Rivara, & Greenwald, 2008). A typical BIRR response might be the individual responding as if the black male is bad (based on a hypothetical learning history that would be consistent with this responding). Given a picture of a black male, the most salient stimuli from that presentation—skin color—and the participant’s previous hypothetical learning history—black males are dangerous—responding that the black male is bad seems
the most likely immediate response. However, given more time, the participant might respond to other stimuli in the context—the man is smiling—and other parts of the learning history, particularly rules, may come into play: for example, people with glasses are intelligent and not bad; or, responding that the black male is scary appears racist and racists are bad. Implicit measures, according to the REC model, are designed to capture BIRR’s while explicit measures are designed to capture EERR’s.

The question requiring investigation is what causes this difference in latency between these hypothetical relational responses; essentially why is one response emitted as a BIRR and another response emitted as an EERR. The REC model asserts that arbitrarily applicable derived relational responses (AADRR) can be influenced by both reinforcement type effects (coherence) and discriminative stimuli type effects (contextual cues or elaboration in REC terms; Barnes-Holmes, Barnes-Holmes, et al., 2010; Hughes, Barnes-Holmes, & Vahey, 2012). The behavioral principles of reinforcement and discriminative stimuli provide explanations for the discrepancies in latency that we see when participants respond to implicit measurement tasks. Elaboration and Coherence are theoretical constructs, which are in-principle observable (Wilson, 2001), and are used to explain “IRAP effects” using basic behavioral principles combined with the relational frame theory outline of verbal behavior. These constructs are based on previously observed data (Hughes et al., 2012) and are meant to organize observed phenomenon. Derivation and complexity, the components of elaboration, as well as coherence are meant to make working with the IRAP, as well as understanding what
may or may not influence the latency of relational responding, easier to organize. Coherence, as will be discussed next in this paper, is in the early stages of investigation while derivation and complexity are yet to be directly investigated.

**Coherence.** Coherence refers to the history of reinforcement as it relates to rule following. Reinforcement can be delivered by another person for following a rule stated by that person (pliance) or by the natural environment, if that environment corresponds with the conditions specified by the rule (tracking; Torneke, 2010). Through this long history of reinforcement with respect to rule following, coherence—or a correspondence between a verbal rule and the perceived environment—becomes a conditioned reinforcer.

Recent work by Bordieri (unpublished manuscript) has demonstrated Coherence to be a valid construct because it can be used to predict participant responding. In Bordieri's study, participants were asked to relate different nonsense stimuli to categories of food based on nutritional value or shape. Once participants were trained to relate the nonsense stimuli with their specific food category, participants were given the option to respond non-coherently (not consistent with the participants’ learning during the study) or coherently (consistent with the participant’s learning during the study). Interestingly, participants responded coherently, even when that coherent responding was accompanied by a hypothesized aversive stimulus: an increasing inter-trial interval. It was believed to be aversive to the undergraduates completing the study because it directly affected their ability to escape the study and access other, potentially more appetitive, activities or items.
In accordance with Bordieri (unpublished manuscript), participants may be more likely to respond quickly in experimental phases where the individual trials are presented as coherent with their learning history. This is because coherent responses will have a more extensive history of reinforcement and will have been emitted much more frequently. If it is seen that coherent responses have the highest probability of being emitted, latency would then be a clear indicator of response strength. In fact, participants may create their own meaning on ambiguous tasks (tasks in which rules for responding are not clear, or there are no rules to dictate responding) on all types of measures, which may lead participants to create meaning on their own and result in unpredicted yet significant effects as participants respond coherently to their own derived rules (Bordieri, unpublished manuscript).

**Complexity.** Complexity is defined as the number of relational nodes between two stimuli (Hughes et al., 2012). Relational nodes refer to the number of intermediary stimuli that connect one stimulus to another. Relational nodes appear to take a structuralist approach to relational networks, but the network model is fully functional as it is based on learning history and pertains uniquely to an individual. Utilizing a visual representation of the stimuli and their relations is a useful way to organize the concept of complexity. For instance, if stimulus A is related to stimulus E through relations to- and between-stimuli B, C, and D, the nodal distance between stimuli E and A is three nodal points (see Figure 3). This way of conceptualizing complexity is not an issue when discussing the difference in performance between implicit measures such as the IAT and IRAP and explicit
measures such as self-assessments and interviews—individuals may indeed derive more novel relations as well as incorporate more complex relational networks as more stimuli become salient in the environment, as would be expected in an explicit response. For an example, consider the participant who was asked to respond to a picture of a black male as either good or bad (Hughes, Barnes-Holmes, & DeHouwer, 2011). To take the previous example of responding to a black male, the brief immediate response of “black males are dangerous” would be considered to be the least complex response. So too, Hughes et al. would indicate that responses with greater response latencies would necessarily involve greater nodal distance, and thus represent greater complexity. Another way of saying this would be that these responses involve more complex relations as well as a greater number of relations to be derived in order for these responses to be emitted. But it does not necessarily follow, in the authors view, that complexity should influence brief, immediate responses that are captured by implicit measures as participants are hypothesized to respond so quickly that only some socially relevant stimuli are salient for the individual (Hughes, Barnes-Holmes, & Vahey, 2012).

**Derivation.** Derivation refers to the number of times that a derived relationship is emitted (Hughes et al., 2012). Hughes et al. state that the more times a relationship is derived, the less time it takes for the behavior to be emitted. To return to our previous example, based on our hypothetical participants learning history the response “Black males are dangerous” would have been derived much earlier in his history and emitted much more frequently, thus being derived more times than other possible responses. This more often emitted (or derived)
relationship would then have a much shorter response latency and the latency would show an overall trend of shortening every time the relationship is emitted (or derived). It is unclear, however, what is the primary cause of this decrease in latency between the first emitted response and subsequent emitted responses. RFT defines derived responses as those responses that are emitted based on previously trained relationships, but have not come into contact with direct reinforcement—in this sense, they are novel. Coherence, as a generalized conditioned reinforcer, may serve to reduce the latency of responding as relations are derived. But the role this plays is unclear. Thus, all responses should have roughly equivalent response latencies as far as derivation alone is concerned.

Hughes and Barnes-Holmes (2011) observed what appeared to be reinforcement effects (see also Vaughn et al., 2011, for evidence from the IAT) which suggests that reinforcement may play a role in the speed of responding to derived relations. If reinforcement plays a part in the observed “practice” effect of the IRAP and IAT, then we must admit that these reinforced responses (this would hold true for any relational response, not just those targeted by implicit measures) are: no longer derived and instead directly trained, assert reinforcement does play a role in the responding but the specified relationship is not being reinforced _per se_ (perhaps a class of behavior is being reinforced, but not the specific derived relationship), or change our current definition to incorporate these responses that are originally derived without a history of reinforcement for doing so but then contact reinforcement after subsequent derivations.

Barnes-Holmes, Murphy, et al. (2010) found that a practice IRAP before
moving to identical test phases increased the speed and accuracy, as based on the
given rules and requested relationships on the task, on the test IRAP while still
demonstrating a relevant IRAP effect. Additionally, Cullen, Barnes-Holmes, Barnes-
Holmes, and Stewart (2009) found that responding was consistent on two identical
IRAP tasks administered 24 hours apart from one another. Latencies for each group
did not appreciably change from one day to the next and, overall, the same IRAP
effects were still observed for both groups in the study. Because an effect is still
demonstrated, even after allowing a practice condition in which certain latency and
accuracy criterion must be met and after completing two identical IRAP tasks
approximately 24 hours after one another, we cannot simply reduce derivation to
practice effects. Thus, derivation requires further explication if it is found to in fact
influence latency to responding.

Additionally, research conducted using the IRAP (as well as other implicit
measures) assumes a particular learning history on the part of participants—
specifically, the grouping of the stimuli into specific stimulus classes as well as the
possible relationships being tested between the stimuli. Implicit measures are
simply not flexible enough to account for varied learning histories of large groups of
participants. Finally, and most glaringly, there is little to no understanding of what
is responsible for the difference in response latency that is seen so consistently
across different implicit measures. Coherence, derivation, and complexity are an
attempt to create an RFT consistent account of implicit attitudes (covert verbal
behavior that occurs at extremely short latencies), but these constructs currently
lack a strong research foundation to support them. While these theoretical
constructs are based on observed phenomenon from countless studies, other than
one study on coherence, they have yet to be tested in their own right. This study
will seek to offer evidence for or against derivation and complexity by describing
the latency of derived relational responses to stimuli within a newly established
relational network and working to understand variations in latency as a result of
derivation and complexity.
Figure 1. An IAT trial using Black patient, White patient, good, and bad as response stimuli and a picture of a black male as a target stimulus (See Sabin, Rivara, & Greenwald, 2008, for an example of this type of IAT).

Figure 2. An IRAP trial with a picture of a black male as the target stimulus, Good as the comparison stimulus, and same and different as response stimuli.
Figure 3. Illustration of Complexity. This figure shows how derived relationships between stimuli (A-F), when organized spatially, appear to cover a greater distance based on the number of trained relationships necessary to arrive at a specific derived response.
Chapter Two: Method

Participants

This study included 23 participants recruited from a local university and the surrounding community. Three participants did not meet the accuracy criterion for the training phase and their data was not included in any analysis. This number of participants allowed for detection of significant results at $\alpha=.05$ with an effect size of .3, resulting in a power of .95. This was calculated using a total of 1 group and 6 measurements. Participants were recruited through personal contacts. They were between the ages of 18 and 36. Participants were not excluded based on any participant characteristics.

Setting and Equipment

The experiment was conducted in a private office at a local clinic or in a quiet area of the participant’s home. Participants were seated at a desk or table with a laptop placed on it. There were no distractions present during the study. The entire procedure was administered on a computer. The experimenter occupied a chair behind the participant or somewhere else in the environment to monitor participant performance and computer performance as well as answer any questions the participant may have as they arose but as not to a distraction to the participant.

The stimuli used were borrowed from Steele and Hayes (1991). The stimuli used were novel for all of the participants. For a complete list of stimuli see
Appendix B. Six target stimuli for each group were chosen randomly from a bank of 31 total target stimuli. The target stimuli were randomly assigned to target positions one to 12. The position of the targets affected what stimuli were directly trained to one another (i.e. 1 was trained to 2, 2 to 3, 3 to 4, and so on).

**Design**

This study was conducted using a within subjects group design. The dependent variable was latency to responding (as measured by a key press). The independent variables were derivation, complexity (both as described above), and phase number. Analysis of the data was conducted using a Repeated Measures ANOVA and the interaction between derivation and latency to respond, complexity and latency to respond, and phase and latency to respond were the focus of this analysis. Additionally, visual inspection was conducted utilizing a single-case, 3-dimensional scatterplot of participant latency in accordance with the three independent variable conditions.

**Procedure**

The procedure consisted of three phases. The first phase was a training phase in which two groups of six stimuli were trained. The training consisted of training equivalence relationships. The training only consisted of five relationships (for example, stimulus A will be trained to stimulus B, B to C, C to D, D to E, and E to F; see Figure 3) and only stimuli within each group were trained to each other (e.g. stimulus A1 were trained to B1 and so forth; stimulus A2 were trained to B2 and so on). The two groups were not related to one another in any way. The second phase was a fluency testing phase in which participants were tested for latency of
responding to the previously trained relationships. The third phase consisted of a derived relations testing phase in which six test trial blocks were presented (See Figure 4 for a visual separation of the phases and what they contained). After the completion of the final testing block, participants were dismissed. The entire experimental procedure had a time limit of 60 minutes. No participant reached the time limit of 60 minutes. All phases will be discussed in full detail later in this section.

The screen presentation consisted of a white background with black stimuli presented on the screen. The target stimulus was presented at the top center of the screen (for instance, A1) with the matching stimulus (B1) and a distractor stimulus, chosen randomly for stimuli from the other relational network, located at the bottom left or right of the screen (see figure 3). The location of the matching stimulus and the distractor stimulus were randomly selected with the criteria that a correct response did not appear in the same location—left or right—for more than 3 consecutive trials (Hughes & Barnes-Holmes, 2011). Participants were required to press the “D” or “K” key on the computer keyboard for either the left or right responses, respectively. This screen presentation was consistent across all phases of the study. During the first phase a correct response produced a green check in the center of the screen and the participant advanced to the next training trial after a 1000-ms inter-trial interval. An incorrect response produced a red x in the center of the screen and the participant was required to emit a correct response before moving to the next trial. All subsequent phases presented the green checkmark for a correct response and the program progressed to the next trial after the 400-ms
inter-trial interval while an incorrect response resulted in the red x as described above.

In the first phase the program presented the instructions on the screen as follows:

*During this study, you will participate in three phases: a training phase and two test phases. During the training phase, you will be trained to relate two separate groups of six objects to each other within their group.*

*The program will give you feedback on whether your answers are correct (you will see a green check mark for a correct answer and a red x for an incorrect answer). You will be required to give the correct answer before moving to the next training trial. Mastery criterion for this phase of the training is 20 consecutive correct answers (three correct answers for each trained relationship).*

*You will receive further instructions before moving on to the testing phases. If you no longer wish to participate in this study, please notify the experimenter immediately. If you wish to proceed with the training trials, please prepare to answer the questions and press the d or k key to continue.*

The participant then pressed “D” or “K” to continue to the first phase. The participant was exposed to all five relationship training trials from each group before the trials were restarted resulting in ten trial training blocks; however, the individual relationships were randomly presented within training trial blocks. Upon the participant emitting an incorrect response, the ten trial block were restarted.
Participants were required to correctly relate all six target stimuli from each group to their designated matching stimulus for a total of 10 correct responses in a row before moving to the next phase. If a participant failed to meet this criterion before reaching 100 trials they were dismissed and their data was excluded from the study.

After reaching mastery criterion during the training trials, the participants were prompted to take a 5-minute break if needed before continuing to the next phase. During this break participants were not allowed to leave the room except to get water or go to the bathroom if necessary. Participants were allowed to stand up and converse with the experimenter about topics other than the current research project if desired. A timer was kept by the experimenter to ensure a maximum of five minutes for a break. No participant reached the five minute limit. The instructions for the second phase were presented on the computer screen included the following:

*Congratulations!*  

*You have successfully achieved Mastery Criterion for the Training Phase!*  

*The next phase will begin with a testing phase. During this phase you will see the items from the previous phase on the screen again. You will relate these items based on the training you just completed. There is a correct answer for every trial.*  

*You will not receive feedback for correct answers, however incorrect answers will be followed by a red x and require you to respond*
correctly before continuing to the next trial. During the testing phases, you are asked to go as quickly as you can when responding.

After completing this phase you will have an opportunity to take a short break if you wish.

If you do not wish to continue with this study, please alert the Experimenter immediately.

If you wish to continue, please press the d or k key when you are ready to start the next phase.

After the participant has pressed the “D” or “K” button, the program automatically progressed to the next phase of the study. During this phase the participants were tested on the previously trained relationships only and received 30 trials (three presentations of each of the trained relationships randomly presented by the program without regard for presentation order) regardless of performance. During this phase, the program also recorded the participants’ latency of responding (recorded from the time the stimuli are presented on the screen to the time the participant engages in a correct response in the form of a button press). There were no criteria for the latency to responding for continuation in the study.

After all trials were presented the participant was prompted again to take a five minute break if needed. The procedure during the break was the same as described above. Once the participant was prepared to continue, the program presented the directions for the next phase of the study. The directions were as follows:

You have now completed the testing phases. If you wish, you may
take a short break.

The next phase will also be a testing phase. It will contain 6 separate trial blocks. You may see some items presented together that have no directly trained relationship, but there is a correct answer for every trial.

You will not receive feedback for correct answers, however incorrect answers will be followed by a red x and require you respond correctly before continuing to the next trial.

During this phase, you are asked to go as quickly as you can when responding.

After completing this trial block you will have another opportunity to take a short break if you wish.

If you do not wish to continue with this study, please alert the Experimenter immediately.

If you wish to continue, please press the d or k key when you are ready to start the next trial block.

The participants then pressed the “D” or “K” button and the experiment automatically started the next phase. During this phase participants were presented with six trial blocks in which 50 possible derived relations were presented randomly in every trial block. The trial blocks were divided so that 25 derived relations, half randomly chosen from the first group and half randomly chosen from the second group, were presented in the first three trial blocks and the remaining relationships were presented in the last three blocks. The instructions above were
be presented at the beginning of every trial during the third phase. Participants’ accuracy (of their first emitted response) and latency (from the presentation of the stimuli to the correct response in the form of a button push) to responding were recorded for every trial along with the target stimulus and the correct response stimulus. There was no accuracy or latency criterion for any trial blocks during this phase of the study and participants continued with this phase until it was completed or they reached the 60 min time limit for participation in the study. No participant reached the 60 minute time limit during the course of the study. Participants were allowed to take a 5-minute break as described above in between test trial blocks as needed. Once participants completed this phase, their participation in the study was complete and they were debriefed and thanked for their time. Upon a participant’s completion of or dismissal from the study they were immediately debriefed by the experimenter (participants could refuse this in person debriefing if they chose). They were then allowed to exit the room.
Phase 1
20-80 trials
Train Equivalence Relationship between novel Stimuli
Accuracy is Measure of interest

Phase 2
30 trials
Test previously trained relationships
Latency is measure of interest

Phase 3
Test Derived Relationships

Test trial 1
50 total trials
derived relationships

Test Trial 2
50 total trials
derived relationships

Test Trial 3
50 total trials
derived relationships

Test Trial 6
50 total trials
derived relationships

Test Trial 5
50 total trials
derived relationships

Test Trial 4
50 total trials
derived relationships

Latency is measure of interest

Figure 4. Flowchart of the study progression.
Chapter Three: Results

Statistical analysis was conducted using SPSS v. 21. The independent variables of interest were phase, derivation, and complexity. The dependent variable of interest was latency of responding. Raw latencies were averaged across phase, derivation level, and complexity level for each participant. These averages were then used for further analysis. No data was excluded in any part of this analysis. A repeated measures ANOVA was used for statistical analysis. Each individual relationship (phase and latency, derivation and latency, and complexity and latency) will be discussed in more detail.

For phase the Mauchly’s test indicated that the assumption of sphericity had been violated, $\chi^2(6)=57.170$, $p=.00$, therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon=.53$). The results show that there was a significant effect between phase and latency, $F(3.18, 60.372)=3.96$, $p=.01$. As phase increased the latency to responding decreased.

For derivation the Mauchly’s test indicated that the assumption of sphericity had been violated, $\chi^2(14)=34.01$, $p=.00$, therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon=.61$). The results show that there was a significant effect between phase and latency, $F(3.03, 54.47)=4.34$, $p=.01$. As derivation increased the latency to responding decreased.

For complexity the Mauchly’s test indicated that the assumption of sphericity had been violated, $\chi^2(9)=111.79$, $p=.00$, therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon=.57$). The results show that there was a significant effect between phase and latency, $F(3.03, 54.47)=4.34$, $p=.01$. As complexity increased the latency to responding decreased.
corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon=.288$). The results show that there was a significant effect between phase and latency, $F(1.15, 20.785)=4.74, p=.04$. As complexity decreased the latency to responding decreased.

Figures 5-24 show the three dimensional graphs for each participant. The graphs are shown at a thirty degree tilt down and rotated to permit the latency axis to align at the direct center of the graph area (a 45 degree rotation). All participants show a general trend of having longer latencies around the center of the graph, while shorter latencies tend to be towards the end. However, it should be noted that there are extremely short latencies toward the bottom center of the three dimensional graph. These latencies tend to be data points that reflect responses of lower complexity and greater derivation. Additionally, it should be noted that the responses on the lower left indicating lower levels of complexity and higher levels of derivation show the shorter latencies of almost all responses. Conversely, responses on the far left side of the graph indicating those responses of lesser derivation and higher complexity, have generally longer latencies.
Figure 5. Participant 1 3D graph.
Figure 2. Participant 2 3D Data
Figure 7. Participant 3 3d Data
Figure 8. Participant 4 3D Data
Figure 9. Participant 5 3D Data
Figure 10. Participant 6 3D Data
Figure 11. Participant 7 3D Data
Figure 12. Participant 9 3D Data
Figure 13. Participant 10 3D Data
Figure 14. Participant 11 3D Data
Figure 15. Participant 12 3D Data
Figure 16. Participant 13 3D Graph
Figure 17. Participant 14 3D Data
Figure 18. Participant 15 3D Graph
Figure 19. Participant 16 3D Graph
Figure 20. Participant 18 3D Data
Figure 21. Participant 19 3D Data
Figure 22. Participant 20 3D Data
Figure 23. Participant 21 3D Graph
Figure 24. Participant 22 3D Data
Chapter Four: Discussion

The results of the current study support the current definitions of derivation and complexity. Additionally, the study offers evidence that participants do respond faster on computer based assessments as they practice responding. The results indicate that there is a relationship between derivation and response latency and complexity and response latency. The higher the number of times the individual had engaged in a response indicative of a derived relationship, the faster the individual was able to engage in a subsequent response indicative of that relationship—and subsequently hypothesized to produce a quicker response under the contextual control of the relationship (Crel). Also, the less complex a relationship is the faster an individual is able to engage in a response indicative of the relationship. Unfortunately, this study offers no evidence as to why these effects occur, but some possible explanations will be discussed in more detail later.

It is also important to note that the decrease across levels of derivation occurred independently of phase. This was demonstrated in how the possible derived responses, fifty total, were split into two separate groups of 25. The first group of 25 was presented in the first three test phases and the second group of 25 was presented in the next three test phases. By dividing the responses this way, derivation and phase were separated from one another for analysis purposes. Because of this significant relationship between phase and latency and derivation and latency, it can be asserted that participants become faster at responding in
general—perhaps as a result of practicing the physical response of pressing buttons on a keyboard—as they progress in a computer based assessment, but that responding to specific relationships repeatedly—as described in derivation—also has an impact on the behavior of button pressing as it relates to the relationship in question.

Some possible explanations as to why participants become faster at emitting specific responses as described in derivation were listed in the introduction of this paper; the most probable cause of this is reinforcement. In the context of this study the reinforcer during the test trials was simply moving to the next trial—which may be verbally related to the distant consequences of finishing the study and even more distantly related to some other reinforcer that occurs outside of the context of the study but requires the completion of the study before it can be accessed—as well as avoiding the red X presented on the screen for incorrect responses. The red x also corresponds to an increase in the intertrial interval experienced by participants. If these derived relationships were being reinforced, but were emitted at least initially without any explicit prior training (as in the current study), then the definition of derived relational responding seems incomplete at best. Another possible explanation is that the participants are simply becoming faster at derived relational responding in general. That is, by practicing specific derived relational responses, the repertoire of derived relational responding becomes more fluent. This begs an entirely different set of questions such as how large and lasting of an effect does this have on an individual's repertoire; what does this mean for responses related to a single stimulus but with different outcomes? The race example described in the
introduction of this manuscript is an excellent example: does practicing the relational response that black males are bad also strengthen the response that black males are good (strengthen meaning decrease the response latency in this case)? A third possible explanation that takes into account complexity is that as the participant engages in the relation, the two stimuli involved become more closely related, and thus the nodal distance simplifies. This could explain the effects seen across levels of derivation, but leaves us to further explore complexity. Any combination of, or another extraneous factor not offered here could explain the decrease in latency across level of derivation observed in this study; however, it is beyond the scope of this study to determine what this cause may be. Future studies should seek to both confirm the current findings related to derivation and seek to offer a more complete explanation as to why derivation has an effect on response latency.

Complexity also showed a significant effect as it related to latency: participant responding decreased across increasing levels of complexity. This study can offer no explanation as to why this effect occurred. Future studies should, again, seek to confirm these findings and offer an explanation as to how complexity affects latency to responding.

One limitation of this study is that it only investigated equivalence relations and neglected other types of relations such as greater than and less than relationships, hierarchical relationships, or temporal relationships. It is unclear if the derivation and complexity phenomenon would be observed in these complex relations or if it may in fact be more pronounced. Future studies should investigate
the effects of derivation and complexity with relation to these more complex verbal relations. Future studies might also parse the effects of derivation and complexity from one another. Of specific interest might be describing the interaction between complexity and derivation, such as which variable is better able to predict participant responding and do the variables interact with one another in any way.

This study outlines the importance of understanding derivation and complexity. Their effects should be accounted for and leveraged when training, measuring, and building technologies around derived relations. Relational Frame Theory based teaching technologies would benefit from a better understanding of both derivation and complexity. Given the recent foray into intelligence training by RFT researchers, whose work is based on the correlation between latency of responding on IRAP tasks and performance on intelligence tests (Cassidy, Roche, & Hayes, 2011), an understanding of the variables as well as mediators and moderators that affect derived relational responding are crucial to making these technologies more effective. An understanding of derivation and complexity could fundamentally change how RFT is incorporated into training programs. Again, given the correlation between latency of responding and intelligence, it is logical that individuals desiring to train in an RFT consistent way should and would want to include techniques to reduce latency to responding during training.

Furthermore, computer based technologies springing from programs like the IRAP or the one used in this study may make Discrete Trial Training techniques for verbal behavior more effective either as a training tool or a maintenance tool that can be used to increase fluency of responding to certain types of verbal operants.
This could not only reduce the amount of time needed to train targets but also allow for programmed practice of previously taught targets outside of therapy sessions with little to no extra training and time required of parents. The Behavior Analysis community should investigate how to use technology to its advantage to both increase the effectiveness of its training tools and to provide more practice opportunities for learners outside of clinical and therapy settings.

Derivation and complexity are useful when describing how and why a decrease in latency is observed when individuals are asked to repeatedly engage in derived relational responding. By successfully describing this phenomenon, we are able to build training and measurement tools to both influence and track changes in this particular behavior. It is extremely important to understand derived relational responding more completely in order to build more successful behavioral techniques and become more effective as trainers and impact the lives of those we work with for the better.
References


Appendices
Appendix A: Stimuli Used
Appendix A cont’d