Florida International University
FIU Digital Commons

FIU Electronic Theses and Dissertations | University Graduate School

7-11-2011

Critical Evaluation and Life Course Change: The Development of the Critical Problem-Solving Skills Scale – Qualitative Extension

Brent M. Maximin
Florida International University, bmaxi001@fiu.edu

Follow this and additional works at: http://digitalcommons.fiu.edu/etd

Recommended Citation
Maximin, Brent M., "Critical Evaluation and Life Course Change: The Development of the Critical Problem-Solving Skills Scale – Qualitative Extension" (2011). *FIU Electronic Theses and Dissertations.* Paper 418.
http://digitalcommons.fiu.edu/etd/418

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

CRITICAL EVALUATION AND LIFE COURSE CHANGE: THE DEVELOPMENT OF THE CRITICAL PROBLEM-SOLVING SKILLS SCALE – QUALITATIVE EXTENSION

A thesis submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

in

PSYCHOLOGY

by

Brent M. Maximin

2011

To: Dean Kenneth Furton
College of Arts and Sciences

This thesis, written by Brent M. Maximin, and entitled Critical Evaluation and Life Course Change: The Development of the Critical Problem-Solving Skills Scale – Qualitative Extension, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

Mary Levitt

Dionne Stephens

William M. Kurtines, Major Professor

Date of Defense: July 11, 2011

The thesis of Brent M. Maximin is approved.

Dean Kenneth Furton
College of Arts and Sciences

Interim Dean Kevin O'Shea
University Graduate School

Florida International University, 2011

DEDICATION

For Amanda.

ACKNOWLEDGMENTS

I would first like to acknowledge my colleagues in the Miami YDP for their support and camaraderie. I also owe a great deal of thanks to the members of my committee - Dr. Stephens for her enthusiasm and valued feedback, and Dr. Levitt for her extraordinary patience and understanding. Most of all, I remain in the debt of my mentor Dr. Kurtines, without whose guidance, instruction, and unfailing encouragement, this project would not be possible.

ABSTRACT OF THE THESIS

CRITICAL EVALUATION AND LIFE COURSE CHANGE:

THE DEVELOPMENT OF THE CRITICAL PROBLEM-SOLVING

SKILLS SCALE – QUALITATIVE EXTENSION

by

Brent M. Maximin

Florida International University, 2011

Miami, Florida

Professor William M. Kurtines, Major Professor

The thesis serves as an evaluation of the psychometric properties of a measure of cognitive competence used with a multi-ethnic, adolescent sample. The primary goal of this study was the development of the Critical Problem Solving Skills Scale – Qualitative Extension, using Relational Data Analysis (RDA). This study builds on previous work that has been conducted to provide evidence for the reliability and validity of the RDA framework in evaluating youth development programs (Kurtines et al., 2008).

Inter-coder percent agreement among the TOC and TCC coders for each of the category levels was moderate to high, with a range of .76 to .94. The Fleiss' kappa across all category levels was from substantial agreement to almost perfect agreement, with a range of .72 to .91. The correlation between the TOC and the TCC demonstrated medium to high correlation, with a range of $r(40)=.68$, $p<.001$ to $r(40)=.79$, $p<.001$.

TABLE OF CONTENTS

CHAPTER PAGE

I. INTRODUCTION 1

II. LITERATURE REVIEW 1
Identity Exploration Style 1
Identity Exploration Competence (Problem Solving) 2
Critical Thinking and Critical Discourse 3
The Development of a Free-Response Performance-Based Indicator of Critical Evaluation 5
Relational Data Analysis: The Qualitative Extension (QE) Method 6
Changing Lives Program 6

III. METHODS 9
Participants 9
Measures 10
Procedure 11

IV. RELATIONAL DATA ANALYSIS 12
Conceptual Open Coding (COC) 12
Theoretical Open Coding (TOC) 13
Theoretical Classification Coding (TCC) 15

V. RESULTS 25
Conceptual Open Coding (COC) 25
Theoretical Open Coding (TOC) 26
Theoretical Classification Coding (TCC) 31

VI. DISCUSSION 37
Limitations 41
Conclusion 42

REFERENCES 43

APPENDIX 47

I. INTRODUCTION

Cognitively-focused identity exploration has been conceptualized as an important mediator of the development of a positive sense of identity and as an intervention target for positive youth development (PYD) interventions (Albrecht, 2007). Recent findings provide evidence that PYD interventions may be effective in promoting differential intervention change in self-reported self-construction among troubled adolescents in the short-term (Eichas et al., 2010).

II. LITERATURE REVIEW

Identity Exploration Style

Extensive literature has established a link between identity exploration style and identity status. Literature has found Informational Identity Style to be positively related with Moratorium and Achievement (Berzonsky, 1989), Normative Style with Foreclosure (Berzonsky and Neimeyer, 1994), and Diffuse/Avoidant Style with Diffusion (Berzonsky, 1989; Streitmatter, 1993). A study by Schwartz et al. (2000) yielded similar findings such that Identity Achievers and those in Moratorium status scored significantly higher in Informational Identity Style, while Foreclosures scored significantly higher in Normative Identity Processing Style and participants in Diffused status scored highest for the Diffuse/Avoidant Style (Schwartz et al., 2000).

However, this extensive literature is built upon cognitively-focused identity exploration data collected through the use of fixed-response self-report quantitative measures of identity *style*. No studies have evaluated the effectiveness of PYD interventions in promoting cognitively-focused identity exploration collected through the

use of free-response open-ended qualitative measures of critical thinking as a cognitively-focused identity exploration *competence* as part of the identity formation process. One reason for this may be the relative lack of measures that tap the use of cognitive competence and psycholinguistic meaning making in addressing life challenges.

Identity Exploration Competence (Problem Solving)

The availability of a free-response cognitive performance measure of the use of critical thinking *competence* as a cognitive self-construction identity formation process, and a free-response narrative expressive performance measure of the use of meaning making as a psycholinguistic self-construction identity formation process, each postulating differing but complementary exploration processes (and with distinct component-appropriate measurement techniques), opens up considerable potential for knowledge development concerning the exploration process. Berman et al. (2001), for example, have noted that cognitive competence constructs have been historically measured using performance-based "tests of limits" (e.g., Flavell and Markman, 1983) rather than self-report methods, and these performance-based measures often lacked content relevance for identity issues and revealed only moderate success establishing a link between cognitive competence and identity status. The current identity formation literature, as a consequence, has tended to focus on non-competence indicators of cognitively focused identity exploration. However, recent conceptual advances in the understanding of critical thinking and critical discourse (Habermas, 1979) and advances in the application of Relational Data Analysis (RDA) (Kurtines, et al., 2008) have opened up the specific possibility of the measurement of critical thinking. RDA is a ready at hand

method for identifying free-response qualitative categorical linguistic properties of developmental research concepts and constructs that extend and complement the dimensional properties evaluated by existing fixed-response quantitative measures.

Critical Thinking and Critical Discourse

In developing the theoretical framework for a measure of critical thinking, this paper draws on, and extends, work on critical thinking and critical discourse (Habermas, 1979). In doing so, the focus has been on the process by which individuals acquire a complex set of cognitive and communicative competencies, including the capacity for critical thinking and critical communication (Habermas, 1979). In the current study, the focus has been primarily on critical thinking.

The psychosocial developmental life course approach defines change as developmentally driven only up through childhood, emphasizing instead the self-directed nature of the developmental process in adolescence and adulthood (Brandstadter and Lerner, 1999; Lerner and Busch-Rossnagel, 1981). Thus, after adolescence, the emergence of critical cognitive competencies is hypothesized to provide the basis for the emergence of a self-directed agentic self (Sheldon et al., 2003), conceptualized for the purposes of the current study as a construction of the human mind that serves a wide variety of important functions, such as, (a) feeling and thus hopefully determining what is "right" for the organism, (b) generating and selecting long-term goals and plans for the organism, and (c) motivating behavior over long periods of time, particularly adaptive behavior which is not necessarily enjoyable for its own sake.

Drawing on this work, Kurtines (2001) has identified three problem-solving processes hypothesized to facilitate the exploration process and the emergence of a self-directed agentic self. *Creativity* is the degree to which the individual is innovative or inventive in generating alternatives for life choices encountered during the exploration process. *Suspension of judgment* represents the degree to which the individual is capable of adopting multiple perspectives with respect to life choices, that is, considering positive and negative aspects (i.e., "pros" and "cons") for each alternative. *Critical evaluation* represents the degree to which the individual is capable of questioning or challenging alternatives and willing to change one's original choice in the context of "a more viable alternative."

Accordingly, a cognitive competence measure that specifically uses both quantitative (dimensional) performance-based and free-response qualitative (linguistic content) data to tap domains relevant to identity issues has been developed, the Critical Problem-Solving Skills Scale (CPSS) (Ferrer-Wreder, 2002). In administering this performance-based measure, participants are encouraged to 1) generate as many potential alternatives as possible for solving life-choice dilemmas (thinking creatively), 2) to suspend judgment by providing justifications for alternatives they disagree with (thinking hypothetically), and 3) to indicate a willingness to question or challenge one's own choices (thinking critically). Thus, the four subscales of the CPSS tap the three previously discussed problem-solving processes: creativity, suspension of judgment, and critical evaluation (Berman, Schwartz, Kurtines& Berman, 2001). Much of the previous work on cognitive competence has tended to focus on creative and hypothetical thinking. Given this, the work on cognitive competence described here has been specifically on the

development of a free-response performance-based measure of the willingness to question or challenge one's own choices (critical thinking), rather than on creative or hypothetical thinking.

To do so, this study used a Relational Data Analysis (RDA) approach in the development and evaluation of a qualitative measure of a core positive developmental construct, critical thinking. Relational Data Analysis is a multidimensional, multiphasic framework for unifying data analytic strategies across domains (quantitative, qualitative) and phases of analyses (conceptual, theoretical, and research analyses) (Kurtines et al., 2008). The study data were drawn from a gender inclusive, multi-problem, multi-ethnic sample of adolescents in urban alternative high schools in a positive youth development program.

The Development of a Free-Response Performance-Based Indicator of Critical Evaluation

As Rinaldi et al. (2011) have noted, the use of qualitative free-response measures in developmental research broadens the scope of the investigation beyond the examination of properties identified as theoretically meaningful prior to conducting the research, i.e., as is usually done under cross-sectional and longitudinal quantitative research designs using fixed-response measures (Eichas et al., 2010). Although there are many advantages to the use of such methods of data collection, an important limitation of their use is that fixed-response measures rule out in advance the possibility of detecting response properties that are uniquely meaningful (ordinary language meaning, theoretical meaning, or both) within a specific population not previously studied, qualitatively

transformed, or involving temporal change (e.g., developmental, historical, longitudinal, or intervention change) resulting in newly emergent properties in a previously studied population (Rinaldi et al., 2011).

Relational Data Analysis: The Qualitative Extension (QE) Method

In response to the need for the inclusion of measures in positive identity research capable of detecting response properties that are uniquely meaningful within a population of interest, RDA uses the QE method in the construction and evaluation of qualitative extensions for available fixed-response measures. In doing so, RDA uses a qualitative extension module (QEM) to facilitate the creation of qualitative extensions and for maximizing the likelihood of collecting the fullest range of participant generated open-ended content properties. Specifically, the QEM was designed to provide a set of standardized meaning and significance questions and probes intended to provide a ready-at-hand method for adding a qualitative data collection component to each of the core battery of quantitative measures used in our program of research. The use of the QE method with these measures has allowed us to elicit participant free-response data that provide access to the subjective meaning and significance of the content of the theoretical constructs targeted by the diversity of questionnaires, scales, dimensions, etc. in our quantitative core battery.

Changing Lives Program

Recent reviews of positive youth development programs reveal an accumulation of evidence that (a) these programs can have an impact on young people and (b) that the

field has made considerable strides, including increased methodological rigor and sophistication. (Catalano et al., 1999; Ferrer-Wreder et al., 2002). Although a relationship has been shown to exist between participating in positive youth development programs and positive intervention outcomes using quantitative measures, the use of these measures have been limited to self-report responses (Ritchie, 2007). Despite the recent availability of new methodology and the extended growth of qualitative research, the use of qualitative measures and its analysis in evaluating the impact of youth developmental programs continues to be limited (Albrecht, 2007). In this context, a call has been made to move in new research directions - particularly, outreach research - that expand the capacity to generate useful knowledge (Jensen, Hoagwood, & Trickett, 1999; Lerner, Fisher, & Weinberg, 2000). The measures and analytic strategies that have been developed as part of the Miami Youth Developmental Project (YDP) represent a response to this call (Kurtines et al., 2008).

The Miami Youth Development Project (YDP) is a community-supported positive youth development program that draws on community-university collaboration and principles consistent with the outreach research model (Kurtines et al., 2008) to foster positive youth development among the culturally diverse and multi-problem adolescents attending alternative high schools in the Miami-Dade County Public Schools (M-DCPS), the fourth largest school system in the United States. The YDP applies a developmental intervention approach that fuses the intervention change goals of the three main approaches to intervention (treatment science, prevention science, and developmental science) to create, refine, and implement in real-world settings effective, feasible,

affordable, and sustainable intervention programs that meet the community needs (Kurtines et al., 2008).

The intervention currently being implemented by the YDP within the Miami-Dade school system is a selective/indicated PYD program known as The Changing Lives Program (CLP). The CLP provides on-site counseling services in all four of the M-DCPS alternative high schools. As described in Montgomery et al. (2008), the primary intervention goal is to create contexts that empower troubled adolescents to transform basic features of their sense of self and identity (e.g., life goals, direction and purpose, etc.) and take control and responsibility over their lives in ways that also result in positive change in problem domains, thereby changing their "negative" life pathways into positive ones.

III. METHODS

Participants

Participants for this study were drawn from the Miami Youth Development Project archival data. The Miami Youth Development Project (YDP) is an outreach research program that promotes positive youth development via a partnership between Florida International University in Miami, Florida and local public schools and community resources. Participation in the Changing Lives Program (CLP) is limited to self-referrals, or students referred by the school counselor or teachers. As part of each school's counseling program, the CLP counseling groups are organized and implemented through each school's administration. Alongside the group of students identified to partake in the CLP, a separate group was also selected for the comparison control condition. All of the students that participated in the comparison control condition were selected randomly from a pool of students not referred or self-referred by school counselors or administrators. The participants in the comparison control condition were further selected for not having participated in any of the counseling and guidance programs prior to or during their involvement with the YDP. All study participants and a parent or legal guardian completed an Internal Review Board (IRB) approved parent consent and student assent form before participating in any research.

My study utilized an archival sample of approximately 250 students who have participated in the Changing Lives Program of the YDP since 2002. The multiethnic sample of intervention participants were compromised of approximately 46% African-American, 35% Hispanic, 8% White/Non-Hispanic, and 6% Other - with 56% females

and 44% males – a constitution that is typical of the schools' demographic make-up. The age of the participants ranged from 14 to 18 years of age.

Measures

The Critical Problem-Solving Skills Scale (CPSS)

The Critical Problem-Solving Skills Scale (CPSS) was used to assess participants' problem-solving competence. The CPSS is a group administered performance-based measure that contains four subscales that tap three problem-solving processes: creativity, suspension of judgment, and critical evaluation (Berman, Schwartz, Kurtines & Berman, 2001). The concept of creativity includes the degree to which one is innovative in generating alternatives for approaching life change goals. Creativity is tapped by the Generation of Alternatives (GA) score, which is the total number of different and distinct alternatives generated for both dilemmas. The participants were asked to generate a number of approaches to achieving a previously indicated life change goal, and their performance appraisal was based on (i) the number of alternatives that were created (quantitative scores), as well as (ii) how many of those alternatives were judged to be qualitatively different from each other, i.e., possessing some unique content properties. Suspension of judgment is the extent to which one is capable of adopting multiple perspectives with respect to these alternatives. Suspension of judgment is tapped by the Decentering, Positive Alternatives (DPA) and the Decentering, Negative Alternatives (DNA) scores (the total number of "cons" and "pros" each participant generated for her/his own "best" and "worst" alternatives). The participants were asked to select one of their alternatives as a "best choice," and another as their "worst choice." The participants

were then asked to suspend their judgment with respect to their evaluation of these two alternatives and list as many “good things” about their “worst choice” as they are able to, as well as list as many “bad things” about their “best choice” as they can. Critical evaluation is the degree to which one is capable of questioning or challenging the alternatives and willing to change one’s original choice in the context of a more viable alternative. Critical evaluation is tapped by the Modification (MO) score (the number times a participant changes (modifies) his/her original choice and provides a ‘‘justification’’ for the change based on his/her previous “pros” and “cons” for that alternative). In other words, this final item on the CPSS asks participants to reconsider their “best choice” in light of the critical thinking exercise.

Procedure

Every intervention group is structured and implemented by an intervention team that consists of one group facilitator, one co-facilitator, and one or two group assistants. All group facilitators and co-facilitators are graduate level students enrolled in either a doctoral developmental science program, or a master’s level mental health counseling program. Undergraduate psychology students who have been trained in the administration of the measures and in participant tracking procedures serve as group assistants. The group facilitators and co-facilitators serve in a counseling capacity that utilizes the CLP’s participatory transformative approach (Montgomery et al., 2008). The intervention groups meet for approximately 45 minutes to 1 hour every week for approximately 8 to 12 weeks during the fall and spring semesters.

Participants' assessments are carried out by undergraduate psychology students serving as research trainees at school grounds and during school hours in the week preceding the commencement of the semester session and the week after its end. The training administered to those undergraduates includes instruction concerning confidentiality issues, assessment administration, dress code, high school regulations, interviewing strategies, and role-playing of interviews.

IV. RELATIONAL DATA ANALYSIS

The Data Analyses Plan for this study was comprised of the complete RDA three-step psychometric analyses (reliability and validity estimates) of the RDA Coding Template constructed during the RDA Theoretical Analysis Task. A condensed description of the RDA process is provided below.

Conceptual Open Coding (COC)

During the Conceptual Analysis phase of RDA a set of theoretically neutral conceptual coders (coders systematically selected to represent no particular theoretical perspective) will be assembled to use the grounded theory concept of "open" coding and the method of constant comparison (the process of the comparison of content properties of participant response data) for similarities and differences (i.e., the process of the comparison of the properties of participant response data for creating and eliminating categories) to identify the unique properties that define the basic pool of conceptual categories in the raw response data, to use those properties to identify all qualitatively different (non-overlapping) conceptual categories in the data set, and to formulate and document an explicit description of the unique property that the response data of each

conceptual category share in common (similarity) and that they do not share with any other categories (difference) (Lewis Arango, et al., 2008; Kurtines, et al., 2008). This open coding process draws on an ordinary language perspective (Wittgenstein, 1953) and is operationalized as the comparison of the properties of participant response data for the purpose of creating and eliminating "ordinary language" content groups of responses with each group of responses defined by a single unique ordinary language content property. In other words, this first set of coders identify all of the unique content properties in a particular data set and thus break the data down into the largest possible set of basic elements, with each element representing the most basic conceptually meaningful units of content from an ordinary language perspective (Eichas, 2010).

Theoretical Open Coding (TOC)

During the Theoretical Analysis phase of RDA, a set of theoretically committed coders (coders systematically selected to be representative of a particular theoretical perspective) will be assembled to work collaboratively on five tasks, each of which generates a particular type of outcome. Specifically, during this phase, the theoretical coders will be asked to use the method of constant comparison to review and discuss the content categories identified in the previous phase from the perspective of the guiding theory.

Theoretical Analysis Task 1: Identifying Theoretical Categories

Using the developmental life course perspective that is described above, the theoretically committed coders are tasked with identifying the smallest number of theoretically meaningful categories and sub-categories from the initial pool of conceptual

categories identified in the raw data during the conceptual analysis, and generating a set of property descriptions of the unique properties that define each category (and associated sub-categories). As part of generating this outcome, theoretical coders are asked to formulate and document an explicit description of the unique property that the response data of each theoretical category share in common (similarity) and that they do not share in common with any other categories (difference) (Lewis Arango et al., 2008; Kurtines et al., 2008).

Theoretical Analysis Task 2: Identifying Relations Between Categories and Identifying Structural Organizational Properties Between and Within Theoretical Categories

During Task 2, the theoretically committed coders are asked to identify a theoretically hypothesized structural organization between the identified theoretical categories and sub-categories (flat, nested, hierarchical, etc.) and to construct a Structural Tree Chart (STC). An STC is a means for visually representing the structural organization among the categories, sub-categories, and properties identified during the TOC. (Lewis Arango et al., 2008; Kurtines et al., 2008).

Theoretical Analysis Task 3: Identifying Mechanisms of Change

In Task 3, the theoretically committed coders are tasked with identifying and specifying hypothesized plausible mechanisms that provide a theoretically meaningful (and plausible) explanation of change over time (causal/functional, structural/transformational, etc.) in the theoretical categories. This third outcome is only generated when the theoretical analysis involves a temporal analysis of change (e.g., a developmental analysis, an historical analysis, etc.) (Lewis Arango et al., 2008; Kurtines et al., 2008).

Theoretical Analysis Task 4: Constructing a Decision Tree Chart (DTC)

During Task 4, the unique properties that define each category (and associated sub-categories) are translated into decision rule formats and the formulated rules are used in the construction of a Decision Tree Chart (DTC). For this task, the theoretical coders construct a DTC to be used for classification coding of un-coded free-response data into the categories and sub-categories that emerged out of the open coding process during the theoretical analysis phase (Kurtines, et al., 2008).

Theoretical Classification Coding (TCC)

If the psychometric analysis during the Theoretical Analysis Phase yields moderate to high estimates for the reliability of the RDA coding template and preliminary evidence for its discriminant and concurrent (external) validity, the Theoretical Analysis Phase ends and the RDA will begin the transition to the third and final phase, the Research Analysis Phase. During this phase, the theoretical coders assembled now function as an advisory panel of theoretical "researchers" whose theoretical expertise falls within the domain of study's guiding theory. The theoretical researchers will work collaboratively to generate initial research hypotheses that are refined, elaborated, extended, and subjected to appropriate research analysis within the context of the project's research design, population, measures, etc., evaluating both quantitative and qualitative research hypotheses. These tentative research hypotheses are subsequently further refined, elaborated, and extended by the ongoing flow of findings generated during the research analysis phase using analytic methods drawn from both the qualitative and quantitative research traditions (Kurtines et al., 2008).

The TCC is conducted as the last step in the RDA Theoretical Analysis Task 4 by theoretically neutral coders for use as part of the Theoretical Analysis Task 5 psychometric analyses. A new set of theoretical neutral coders is assembled to perform the TCC task of the RDA process. The task of training the TCC coders is essentially the same as the task of training the COC coders, with appropriate modifications. For this task, the copies of the same Sample Coding Deck used during the COC and TOC, the Coding Glossary, DTC, and the category descriptions developed during the TOC are provided. The panel then conducts the second coding procedure of the initial sample set using all the information developed during the TOC phase of the RDA process.

Theoretical Analysis Task 5: Psychometric Analyses: Reliability and Validity Analyses of an RDA Coding Template (RDA-CT)

The last task in the RDA Theoretical Analysis Phase is to conduct a three-step preliminary psychometric analysis (reliability and validity estimates) of the DTC constructed in the previous phase, a process referred to as a Psychometric Analysis of the Theoretical Classification Coding Decision Tree Chart (PA-DTC).

Theoretical Analysis Task 5, Step 1: Estimating inter-coder reliability. In the quantitative/experimental research tradition, reliability is the extent to which a measurement method is repeatable and yields consistent results. Given that the most important outcome of the conceptual and theoretical analysis phases of RDA is the identification of conceptual and theoretical categories, an important psychometric property of identified coding categories is the consistency with which they can be used to classify accurately participant response data into the theoretical categories (i.e., the reliability of the coding categories). For purposes of estimating inter-coder consistency of

an RDA-CT, RDA standard psychometric analysis of coding templates uses the inter-coder agreement percent among the coders for each category level to provide an estimate of inter-coder reliability for the TOC and TCC (Ritchie, 2007, Kurtines, et al., 2008).

Inter-coder agreement is also estimated using Fleiss' kappa. Fleiss' kappa incorporates a correction for chance agreement. Furthermore, the use of Fleiss' kappa allows for greater generalizability as it can be used to measure the significance of the agreement among multiple raters with the significance level being adjusted relatively to the number of comparisons made, as opposed to Cohen's kappa which only tests for the significance of two raters at a time (i.e., in pairs of ratings) (Fleiss, 1973). If the agreement is found to be high, it is assumed that the ratings accurately reflect the dimensions they are presumed to. If kappa is found to be low or 0, this would indicate a high degree of measurement error (Fleiss, 1973). Agreement from 0 to .20 assumes slight agreement, .21 to .40 assumes fair agreement, .41 to .60 assumes moderate agreement, .61 to .80 assumes substantial agreement, and .81 to 1.00 assumes almost perfect agreement.

Theoretical Analysis Task 5, Step 2: Estimating construct validity. In the quantitative/experimental research tradition, validity is said to be the extent to which a measurement method measures what it is supposed to measure. The concept of construct validity was first introduced by Cronbach and Meehl in 1955 to extend the two major types of psychometric validity recognized up to that time, "content" and "predictive" validity. From the time it was introduced, "construct" validity was recognized as the most important type of psychometric validity as well as the most difficult to establish

(Anastasia & Urbina, 1997). Construct validity requires the accumulation of evidence that a measurement method is linked to the theoretical construct it is hypothesized to represent. As a result, estimating construct validity is a process is that is often long and complex, involving at least three basic types of evidence: criterion-related (includes predictive and concurrent), convergent, and discriminant validity (Anastasia & Urbina, 1997; Campbell & Fiske, 1959; Nunnally, 1978).

In the case of construct validity, the evidence that a measurement method is a "valid" measure of a particular construct tends to be indirect, though not necessarily so. In the case of the two IQ measures, for example, if one of the measures is a "newly" developed measure and the other has been more extensively validated, the same correlation coefficient that might be interpreted as providing convergent validity might also be interpreted as providing concurrent validity for the newly developed measure, i.e., that the newer measure (with previously unknown properties) has now been found to share a significant proportion of variance with an already established measure of the same construct. Moreover, if the same high and positive inter-correlations were to be found in multiple IQ measures using multiple methods of measurement, these correlation coefficients might be interpreted as providing construct validational evidence for a particular "theory" of intelligence (e.g., a single factor theory) in addition to whatever evidence it provided for a particular measurement method as an indicator of IQ.

The process of construct validation is not only often long and complex; it is never really finished, i.e., there never comes a time when a measurement method can be finally declared valid. On the contrary, it is generally recognized that the construct validation is a process in which the degree of validity of a particular measurement method for a

particular theoretical construct is most accurately represented by the accumulated evidence regarding the multiple psychometric properties of a measurement method within the supporting nomological network of related concepts and constructs (i.e., the theory) that defines the theoretical construct (Campbell & Fiske, 1959). It is for this reason that at this level of analysis the boundaries between the validation of a method for measuring a theoretical construct and the validation of the theory that defines the theoretical construct (and potential measurement models and methods) is blurred, with obtained results often providing (or failing to provide) evidence in support of neither or both (Nunnally, 1978).

In this context, within RDA we consider the level of inter-coder agreement across all of the categories to move beyond the methodological issue of estimating the reliability of each of the categories identified by the emerging grounded theory to begin to address the theoretical issue of the construct validity of the theoretical framework of related concepts and constructs of the emerging grounded theory, i.e., relations among the categories. In grounded theory, the content properties identified by the conceptual analysis are viewed as defining the basic elements of the emerging theory and the theoretical properties of the theoretical categories identified by the theoretical analysis as defining the basic the concepts or constructs of the framework of the emerging theory. The structural organization of the theoretically meaningful categories identified by theoretical coders, in turn, is viewed as defining the theoretical relation among the concepts and constructs that make up the emerging theory. As noted, in RDA the second step in theory construction is the identification of the smallest number of qualitatively different (non-overlapping) theoretically meaningful categories and an explicit

description of the properties that define those categories in a specific data set and distinguish them from each other.

Further, because the aim of grounded theory is to construct qualitative theories with respect to the organization of and relations among qualitatively different variables, the theoretical expectation is that in the process of constructing a theoretical structural organization of the categories, each of the identified categories will have a unique (non-overlapping) property. Evidence that each construct in the theory is unique and non-overlapping with every other category (construct) that makes up the theoretical framework provides confirmatory construct evidence for the validity of the emerging theory while evidence that each construct in the theory is not unique and not non-overlapping with every other category provides non-confirmatory evidence for the validity of the emerging theory. The inter-coder agreement for each category thus provides an estimate of the inter-coder reliability for that specific category, and the average inter-coder agreement across all the categories provides a global index of the "goodness of fit" between a theoretically hypothesized set of relations among the identified theoretical categories and the actual relation among them (Kurtines et al., 2008).

At a concrete and specific level, for example, if the theoretical structure identified by theoretical coders is comprised of four categories (A, B, C, D), a high average inter-coder agreement across all of the categories for provides evidence for the construct validity of the hypothesized structural organization of the categories. In this case, the structural organization includes all the unique (non-overlapping) categories that make up the theoretical framework of the emerging theory. This theoretical framework constitutes

the theory within which each identified category is embedded and which the classification coders, explicitly or implicitly, make use of in generating their comparative judgments when coding and classifying participant responses into each category in relation to all the other categories. A high average inter-coder agreement across all of the categories means that the theoretical classification coders were able to classify accurately each of the responses into its appropriate categories. In doing so, each theoretical classification coder's decision to classify a specific response as a belonging to specific category (e.g., Category A) involves a simultaneous judgment that the property that defines that particular response is the same property (described in the property descriptions) that the theoretical coders identified as defining Category A and that that property is different from the property (described in the property descriptions) that defines Category B, different from the property that defines Category C, and different from the property that defines Category D. A decision to classify a response as Category B similarly involves simultaneous judgment of the same form, i.e., same as B, different from C, different from D, and different from A. Classifying a response as Category C similarly involves the simultaneous judgment of the same form, i.e., same as C, different from D, different from A, different from B, and Classifying a response as Category D involves the simultaneous judgment: same as D, different from A, different from B, different from C (Kurtines et al., 2008).

A moderate average inter-coder agreement across all of the categories, in contrast, means that the theoretical classification coders were only partially able to classify accurately each of the responses into its appropriate categories and provides moderate evidence for the construct validity of the hypothesized structural organization (Kurtines

et al., 2008). A low average inter-coder agreement across all of the categories, in turn, means that the theoretical classification coders were not able to classify accurately each of the responses into its appropriate categories and provides evidence for the absence or lack of construct validity for of the hypothesized structural organization (Kurtines et al., 2008). Finally, a variable inter-coder agreement for each of the categories (e.g., some categories displaying high and some displaying low inter-coder agreement) means that the theoretical classification coders were only able to classify some of the responses into its appropriate categories and provides evidence for differential construct validity for a subset of the hypothesized structural organization (with some displaying more construct validity than others) (Kurtines et al., 2008).

In the context of variable inter-coder agreement, high inter-coder agreement for a particular category is interpreted as providing evidence that the particular category (and the subset of the theoretical framework to which that particular category is linked) is theoretically meaningful. Low inter-coder agreement for a particular category, in contrast, is interpreted as providing evidence that the particular category (and the subset of the theoretical framework to which that particular category is linked) is theoretically meaningless (Kurtines et al., 2008). Thus, the second step of provides preliminary evidence (or lack of evidence) for the construct validity of a particular way of organizing the phenomena explained by the emerging grounded theory.

It should be further noted that in those cases where Task 1, Step 2 yields either consistently high or consistently low average inter-coder agreement across all categories, an examination of the results of the inter-coder agreement for each specific identified category in Task 1, Step 1 will yield information redundant with the results of Task 1,

Step 2, i.e., if the average inter-coder agreement across all categories are high (or low), inter-coder agreement across all the specific categories will also be respectively either high (or low). However, if the average inter-coder agreement yields mid to moderate inter-coder agreement, the results provide a more useful index of goodness of fit in that an examination of the results of the inter-coder agreement for each specific identified category may yield significant theoretically and empirically useful information with respect to the relative construct validity of each specific identified category, which categories merit continuing on to the research analysis phase of the RDA, and which categories need to be returned to an earlier phase of the relational data analysis cycle to undergo further conceptual or theoretical analysis and refinement before moving to the research analysis phase (Kurtines et al., 2008).

Theoretical Analysis Task 5, Step 3: Estimating criterion-related (concurrent) validity. Criterion-related validity consists of concurrent and predictive validity. Concurrent validity estimates the degree to which a measure relates to other manifestations of the same theoretical construct (e.g., the correlation between a measure of patients' self-reported depression symptoms and clinician ratings of depression symptoms in a treatment study). Predictive validity evaluates whether a measure relates to manifestations of other constructs the measurement method is theoretically hypothesized to predict (e.g., the correlation between accounting staff scores on a measure of math aptitude and the accounting supervisor's job performance ratings of the accounting staff). Within RDA, the third step is to use the correlation between the categories identified by the theoretical coders and the theoretical category classifications generated by a second set of theory neutral conceptual coders to estimate the concurrent

(external) validity of the coded categories. In Step 3, the resulting correlation coefficient is interpreted as a coefficient of concurrent validity (as a type of criterion-related validity) obtained using a variant of the widely recognized multitrait-multimethod matrix method first introduced by Campbell and Fiske (1959) (Ritchie, 2007; Kurtines, et al., 2008).

V. RESULTS

The data analysis for the current study was comprised of the two RDA phases of analyses and the same three-step psychometric analysis (reliability and validity estimates) as described above. The following section reports the psychometric properties generated by the RDA for the Critical Problem-Solving Skills Scale – Qualitative Extension (CPSS-QE).

Conceptual Open Coding (COC)

Conceptual open coding was conducted on each of the study's participant's "raw" interview response data, called a Macro Interview Response (MIR). Macro Interview Responses consisted of the transcription of all the words, phrases, and sentences a specific participant used to describe the meaning and significance of the experience (topic, issue, question, etc.) under investigation. For the first task in the Conceptual Analysis Phase, the material for the COC, which consists of creating MIR coding cards for all the response data to be used for the COC was assembled. A set of 40 "Sample Coding Cards," a theoretically representative set, was then selected from the larger sample under study. Finally, a set of five theoretically neutral conceptual coders was assembled to use the grounded theory concept of "open" coding and the method of "constant comparison" for similarities and differences in the comparison of the properties of participant response data to identify the unique content properties that define the basic pool of conceptual categories in the raw response data, use those properties to identify all qualitatively different (non-overlapping) conceptual categories in the data set, and formulate and document an explicit description of the unique content property that the

response data of each conceptual category share in common (similarity) and that they do not share with any other categories (difference) (Kurtines et al., 2008) .

After a brief orientation to RDA (and the current study) that included a general explanation of the goals of coding (i.e., to identify and classify answers to interview questions into conceptually meaningful groups or categories) and a brief overview of the process that would be used (i.e., by sorting cards containing a transcription of various segments of the interviewee's answers to the questions), the coders reviewed the sample set and identified a preliminary initial set of seven non-overlapping conceptual categories (Academics, Self-Preservation, Relationships, Conflict Avoidance, Self-Improvement, Approval of Others, Stress Reduction) contained in this particular data set, working blind to Condition.

Theoretical Open Coding (TOC)

The initial set of conceptual categories identified in the COC provided the data for the theoretical analysis conducted in this phase. A panel of five theoretical coders was assembled to review the entire CPSS-QE sample set (40 MIRs) for the seven conceptual categories, to gain an understanding of what the conceptual coders identified as the properties of the conceptual categories. The theoretical coders then discussed to consensus the theoretically meaningful categories, the organizational structure (STC) of the categories, translated the unique properties of the categories into the DTC, and later performed an open coding exercise using the sample set.

Theoretical Analysis Task 1: Identifying Theoretical Categories

During the first task in the TOC stage, a theory saturated panel of coders utilized Elder and Paul's Stage Theory of Critical Thinking (described in more detail in previous sections) as the primary theoretical perspective for this theoretical analysis. Theoretical coding differs from content category coding performed in the COC, in that during the conceptual analysis phase, open coding is done by coders with no particular theoretical perspective (or perhaps more accurately, coders with multiple implicit theoretical perspectives). During the Theoretical Analysis phase, in contrast, open coding is performed by coders who explicitly share the same identified and extensively articulated theoretical perspectives.

This first phase of the theoretical analysis was conducted by a panel of five theoretical coders who are given the task of identifying theoretically relevant and meaningful categories contained in the pool of content categories identified by the neutral coders in the COC. The theoretical coders were provided with the sample MIR coding cards grouped into the categories formulated in the COC, as well as descriptions for the categories. They were then asked to review the cards and the content property descriptions and to refine the basic content categories and the structure of the groupings generated by the content coders during the COC, identify theoretically meaningful properties, and organize them into groups (categories) and respective sub-categories in ways that are meaningful with respect to the identified theoretical framework. The theoretical analysis of the CPSS-QE yielded two (Level One) categories - Critical Thinking, and Uncritical Thinking. Uncritical Thinking yielded two (Level Two) sub-

categories - Unreflective Thinking, and Challenged Thinking. Critical Thinking yielded three (Level Two) sub-categories - Rudimentary Thinking, Advanced Thinking, and Accomplished Thinking. Appendix A presents the structural organization of the theoretical categories and sub-categories as well as the property descriptions of the unique categories.

Constructing an RDA Structural Tree Chart (STC). The last step in finalizing the identification of the theoretical categories was the construction of a visual representation of the structural organization of the categories. In RDA, the Structural Tree Chart (STC) that emerges out of the TOC process not only provides a conceptual framework (a visual tool) for reporting structural organizations, but also for facilitating the identification of theoretical categories during the open coding process (see Appendix A for STC constructed during this study).

Theoretical Analysis Task 2: Identifying Relations Between Categories and Identifying Structural Organizational Properties Between and Within Theoretical Categories

The second phase of the theoretical analysis had two aims: 1) structural organization of the relations between all of the identified categories, and 2) structural organization of the narrative response data within each of the identified categories (Kurtines et al., 2008). The theoretical coders first examined the pattern of relations between the theoretical categories (i.e., the structural organization) and refined and re-defined them in terms of a structural organization (hierarchical) meaningful with respect to the guiding theoretical framework. Then the theoretical coders identified the structural organization of the narrative response data within each of the theoretical categories

within each MIR. Drawing on the basic directionality of change over time in critical competence – as conceptualized in much of the current critical thinking literature (Elder & Paul, 1996; Sun, 2006) – the theoretical coders identified a theoretically meaningful developmental progression. Each of the five (Level Two) sub-categories in the following succession - Unreflective Thinking, Challenged Thinking, Rudimentary Thinking, Advanced Thinking, and Accomplished Thinking – represents an advancement in critical evaluation *performance* (if not necessarily capacity). It is important to note here that although the terms "performance" and "capacity" are not interchangeable, Norris maintains that, in order to be considered a "critical thinker," one must possess both the ability and the disposition to put that particular competence to use (1989). Another one of the primary structural organizational properties identified in RDA of the CPSS-QE was Structural Coherence – the degree to which the overall structure and organization of an MIR may be considered a cohesive and coherent. The term "consolidated" was used to describe a MIR that was characterized as cohesive and coherent. Consolidation is a term that describes a merging of the multiple elements or components (of the self) into of an integrated whole. The panel further identified three sub-categories of consolidation for the MIR data for the RDA CPSS-QE: unconsolidated, partially consolidated, and fully consolidated (see Appendix A for a more detailed explanation of consolidation). Structural Coherence was placed in the DTC at Level Three.

Theoretical Analysis Task 3: Identifying Mechanisms of Change

The aim of the third phase of the theoretical analysis is to identify and specify plausible mechanisms of change (e.g., causal/variational, structural/transformational, etc.)

in categories/variables that provide theoretically meaningful explanations from the perspective of a particular theoretical framework. Because of the lack of temporal investigation in the current study, the third phase of theoretical analysis was not performed.

Theoretical Analysis Task 4: Constructing a Decision Tree Chart (DTC) RDA Coding Template (RDA-CT)

Constructing an RDA-CT is normally undertaken as penultimate step in an RDA Theoretical Analysis and uses the STC that is constructed during the completion of steps 1 and 2 of the theoretical analysis. The first step in constructing an RDA-CT is to construct a Decision Tree Chart (DTC) from the relevant RDA-STC and accompanying set of property descriptions. A DTC is typically derived during the construction of an RDA-CT for a specific measure for a specific population, problem, program, etc. RDA-CTs are intended to provide all the basic information needed for conducting an RDA Theoretical Classification Coding (TCC). The components of this study's RDA-CT are as follows: (1) CPSS-QE-STC (2) Theoretical Property Descriptions, (3) Theoretical Category Structural Organization, (4)RDA CPSS-QE DTC, and (5) Coding Glossary.

The basic goal of Theoretical Analysis Task 4 is thus to create a structure for the RDA Decision Tree Chart that sequentially and systematically guides the TOC and later, the TCC coders' decision making with respect to the appropriate category properties to be used at each choice point in the decision making process, so that each MIR is assigned to an appropriate category or sub-category (Kurtines et al., 2008). The logic of decision

tree flow charts provide a useful tool for identifying and mapping the most efficient and effective structure for guiding the decision making process.

Level One includes the categories of Critical Thinking and Uncritical Thinking, Level Two includes the five respective sub-categories, and level three includes the consolidation (see Appendix A). A TOC was then conducted by the panel to classify the MIRs of the sample set using the coding template to place the free responses of the sample set into the corresponding categories and sub-categories identified in the previous tasks of the theoretical analysis phase of RDA. Levels were included in CPSS-QE RDA-CT to evaluate inter-coder reliability at each level as well as overall category reliability. Relational Data Analysis makes it possible to easily and readily switch between poles of the splits (qualitative → quantitative, structural → causal → structural) based on findings or results obtained at any of the phases of analysis and at any level of analysis (theory and data). In this context, the current study implemented the use of category levels to distinguish and further investigate the theoretical meanings of the theoretical categories identified in the TOC versus the empirical meaning of the categories identified in the TOC.

Theoretical Classification Coding (TCC)

A TCC was conducted as the last step in the RDA Theoretical Analysis Phase 4 by a panel of theoretically neutral coders. A new set of theoretically neutral coders was assembled to perform the TCC phase of the RDA process. The task of training the TCC coders is essentially the same as the task of training the COC coders, with appropriate modifications. For this task, the copies of the same Sample Coding Deck used during the

COC and TOC, the Coding Glossary, DTC, and the category descriptions developed during the TOC were provided. The panel then conducted the second coding procedure of the initial sample set using all the information developed during the TOC phase of the RDA process.

Theoretical Analysis Task 5: Reliability and Validity Analyses of an RDA Coding Template

The last task in the RDA Theoretical Analysis Phase is to conduct a three-step preliminary psychometric analysis (reliability and validity estimates) of the Decision Tree Chart (DTC).

Theoretical Analysis Task 5, Step 1: Estimating inter-coder reliability. The most important outcome of the conceptual and theoretical analysis phases of RDA is the identification of conceptual and theoretical categories; an important psychometric property of identified coding categories is the consistency with which they can be used to accurately classify participant response data into the theoretical categories (i.e., the reliability of the coding categories). For purposes of estimating inter-coder consistency of an RDA CT, psychometric analysis of the Coding Template uses the inter-coder agreement among the TCC coders during the TCC coding, and the inter-coder agreement among the TCC coders of the second set of theory neutral coders, thereby providing an independent estimate of inter-coder reliability for both the theory saturated coders and a second set of theory neutral coders blind to the theoretical meaning and significance of the categories identified by the theoretical coders.

For this study, coder agreement was evaluated by category, sub-category, and consolidation, as identified by the TOC coders. Inter-coder agreement among the five theory-saturated coders of the TOC for each of the specific categories was found to be substantial to high, with a range of .81 to .95. Inter-coder percentage agreement for the TCC level one was .95, for level two, .81, and for level three, .88, with a total percent agreement across all levels of .88. Inter-coder agreement among the five theory neutral coders of the TCC for each category was also found to be substantial to high, with a range of .75 to .95. Inter-coder percent agreement for the TCC level one was .95, level two was .75, and level three was .94, with a total percent agreement across all levels of .88. Inter-coder agreement across the TOC and the TCC was also moderate to high, with a range of .76 to .94. Inter-coder percent agreement for level one across the TOC and the TCC was .94, level two was .76, and level three was .91.

In addition to agreement percentage, Fleiss' kappa was used to estimate the inter-coder reliability of each of the levels correcting for chance. Fleiss' kappa for the TOC level one was estimated at .91 showing almost perfect agreement, level two was estimated at .74 showing substantial agreement, and level three was estimated at .73, again showing substantial agreement. Fleiss' kappa for the TCC level one was estimated at .90 showing almost perfect agreement, level two was estimated at .78 showing substantial agreement, and level three was estimated at .72, again showing substantial agreement.

Theoretical Analysis Task 5, Step 2: Estimating construct validity. As noted above, the average inter-coder agreement across all the categories is interpreted as

providing an indirect estimate of the degree to which the conceptual properties for each of the identified categories are unique and qualitatively different from all of the other identified categories of the emerging grounded theory.

The overall average percent agreement across all the category levels was moderate to high, with a range of .76 to .94, providing evidence for a relatively high degree of construct validity for the identified categories. Given the particular type of theory used as a framework for the RDA CT reported here (i.e., a type of grounded theory in which all identified theoretical categories are not only hypothesized to be uniquely different from each other but also to represent a specific theoretically hypothesized structural organization), which yields consistently high inter-coder agreement for all the categories - i.e., a high average inter-coder agreement – it is interpreted as not only providing evidence for high reliability for each identified category, but also as evidence for a high level of construct validity for all the categories within the theoretical structural organization identified by the theoretical coders (Kurtines et al., 2008). In the case of grounded theory, this includes all the other categories that make up the theoretical framework within which each identified category is embedded and which the classification coders, explicitly or implicitly, make use of in generating their comparative judgments with respect to each category. Thus, the overall accuracy of the coders' classifications of the participants' open-ended response data based on the coders' evaluation of the specific property of each individual category, made in comparison to the properties of all the other categories, provides evidence in support of the theoretically hypothesized structural organization of the categories. Specifically, a high average inter-coder agreement provides evidence that each category has the

theoretical meaning it is claimed to have within the context of a theoretically generated structural organization that is defined by all the other categories within which it is embedded. The psychometric properties of the CT thus provide a method for evaluating the hypothesized structural organization of the identified categories generated by the theoretical coders as well as preliminary and indirect evidence for the construct validity of the specific properties of specific categories (Kurtines et al., 2008).

Theoretical Analysis Task 5, Step 3: Estimating criterion-related (concurrent) validity. As noted above, the third step of Task 5 is to use the correlation between the categories identified by the theoretical coders and the category classifications generated by a second set of theory neutral conceptual coders to estimate the concurrent (external) validity of the coded categories. During this stage, the resulting correlation coefficient is interpreted as a coefficient of concurrent validity (as a type of criterion-related validity) (Kurtines et al., 2008). The focus of the analysis is on the concurrent validation of multi-manifestations of the same theoretical construct generated by multiple methods (open coding by theory-laden coders versus classification coding by an independent sample of theory neutral coders) rather than predictive validity (Kurtines et al., 2008). The correlation between the modal Coding Category, by level, that the theory-laden coders (TOC) assigned to each participant's MIR and the modal Coding Category, by level, that the theory neutral coders (TCC) assigned to each participant's MIR was, $r(40) = .79$, $p < .001$, for level one, $r(40) = .75$, $p < .001$ for level two, and $r(40) = .68$, $p < .001$ for level three, providing evidence for medium to high concurrent (external) validity for the identified theoretical categories. As previously noted, theoretical categories consensually identified in the theoretical analysis are not only theoretically "meaningful" but also

rooted in content properties that actually exist in the raw data, (i.e., they have conceptual meaning independent of the theoretical meaning ascribed by the theoretical coders). Thus, in contrast to conceptual and theoretical analysis, the concurrent validational analysis does not use open coding to *identify* concepts (categories); instead, in concurrent validational analysis of the coding categories, the TCC is used to *classify* the response data into the categories (concepts) identified during the theoretical analysis (Kurtines et al., 2008). The correlation between the theoretical categories generated by the theoretical coders' open coding of the original raw data responses, and the classification of the same original raw data responses into the same theoretical categories by the second set of theory neutral coders provide an estimate of the concurrent (external) validity of the categories.

VI. DISCUSSION

Prior to the current study, there has been little research done to evaluate the effectiveness of PYD interventions in promoting cognitively-focused identity exploration collected through the use of free-response qualitative measures of critical thinking as a cognitively-focused identity exploration "competence" as part of the identity formation process. The construct of cognitive competence has clear implications in positive youth development, particularly in disempowered populations, but it has still managed to be severely understudied within the context of adolescent development. Even within the limited literature on problem-solving skills as a developmental asset, much of the previous work has tended to focus on creative and hypothetical thinking. Given this, the results of my study go some way towards addressing a gap in the literature by focusing instead on the willingness to question or challenge one own choices (critical thinking), rather than on creative or hypothetical thinking. The CPSS-QE also taps into the exploration of psycholinguistic meaning making in addressing the frequently overwhelming life challenges multi-problem youth face, whereas the majority of problem-solving measures in use in adolescent populations lack similarly relevant "real world" content.

The development of the CPSS-QE advances cognitive competence research on troubled youth by extending a measure that, while useful, had previously only produced quantitative estimates of creativity and hypothetical thinking. The CPSS-QE not only introduces a much needed qualitative component to this body of work, but it also

attempts to capture a more developmentally advanced subset of problem-solving skills than is usually studied – critical evaluation.

Theoretical and Methodological Contribution

At the broadest level, the research on the development of free-response qualitative measures of critical thinking as a cognitively-focused identity exploration "competence" reported here is part of a broader effort to advance the unification of psychological/developmental science and sociological/anthropological science. The employment of a grounded theory approach in the development of the CPSS-QE emphasizes theory *construction*, as it uses data collected on human participants in field or community settings. This is in contrast to the experimental tradition that evolved in psychological/developmental research emphasizing the *evaluation* of theory (vis-à-vis hypothesis testing) using data collected under rigorously controlled clinic/lab settings and frequently intended to generalize to human and nonhuman species (i.e., species that are not language using, culture bearing, meaning making). Consequently, qualitative research methods, such as theoretical sampling and saturation (as developed for use in grounded theory), have tended to be used to develop and refine theory regarding the linkages between inner psychological states (e.g., subjective meaning and significance of life course experiences) and observed human behavior in real time and in real world human ecologies, as interpreted by the researcher and, more important, by the participants themselves - a context in which theory development is justifiably considered still in the discovery stage (Kurtines et al., 2008). In this way, the work undertaken as part of the current study has the potential to enhance the evolution of new theory through the

development and refinement of a unified data analytic strategy (RDA), and through the development of a measure that captures in "real time" the temporal dimension of intervention, developmental, and historical change.

The main objective of RDA research analysis is to test and evaluate evolving theoretical perspectives against the evolving developmental and socio-cultural historical reality revealed by the changing content and structural organizational properties of the meaning and significance of the content of participants' free-response data as captured and rendered explicit and intelligible by the COC and TOC coding processes of RDA (Kurtines et al., 2008). In other words, RDA allows for the construction and evaluation of an array of developmentally appropriate unstructured response measures that are capable of capturing change in the meaning and significance of the life course experiences of participants in the CLP intervention, their positive development, and historical change.

As noted by Kurtines et al. (2008), and in the context of the distinctly different historical evolution of contemporary psychological and developmental science and sociological and anthropological science, our broadest aim is to investigate the analytic utility of uniting (within an RDA framework) qualitative and quantitative methods representative of these two distinct research traditions, the quantitative/experimental research tradition associated with clinic/lab based psychological/developmental sciences, and the qualitative/field research tradition associated ethnographic/community based sociological/anthropological sciences. Most important in terms of contributing to the broader process of knowledge development, when RDA is used within a repeated measures design to analyze unstructured free-response data, it helps to ensure that in such research designs the COC and TOC phases of the RDA analysis will uncover new

categories of the meaning and significance of life course experience as they emerge at the leading edge of developmental and historical change (Kurtines et al., 2008). Thus, RDA is designed to be used in ways that maximize the likelihood that new content categories identified during conceptual analysis will not be obscured or overlooked. The challenge for the investigator and the theoretical research team is to transform these conceptual categories into theoretically meaningful constructs and concepts and, ultimately, into data driven theories to be evaluated by both quantitative and qualitative data analytic strategies.

Limitations

As tends to be the case in many studies on similar populations, limitations on the generalizability of the findings arise when one takes into consideration the nature of the sample utilized, i.e., multi-ethnic, multi-problem background adolescents in an urban setting. Although minority groups are largely understudied, a necessary caveat of studies such as this one is that the results may not be replicable in non-minority populations, adult populations, or adolescent populations from less maladaptive backgrounds.

Conclusion

The study reported in this paper, undertaken as part of a psychometric evaluation of measures in the evaluation of the Changing Lives Program (CLP), provides evidence that the measure under development (the CPSS-QE) has acceptably high psychometric properties (i.e., reliability and validity), particularly at the first order level of theoretical categories. These findings provide evidence in support of undertaking a full-scale short-term controlled outcome trial, and evidence for a framework in which to utilize the CPSS-QE in evaluation of positive development programs for troubled multi-ethnic adolescents.

REFERENCES

Albrecht, R. E. (2007). *A model of self-transformative identity development in troubled adolescent youth*. Unpublished Dissertation. Florida International University, Miami.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*, 469-480.

Benson, P. L., Scales, P. C., Hamilton, S. F., Sesma, A. Jr. (2006). Positive youth development: theory, research, and application. In W. W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology*, Volume 1, Theoretical models of human development. New York: John Wiley.

Berman, A. M., Schwartz, S. J., Kurtines, W. M., & Berman, S. L. (2001). The process of exploration in identity formation. The role of style and competence. *Journal of Adolescence, 24*, 513-528.

Berzonsky, M. D. (1989). Identity style: Conceptualization and measurement. *Journal of Adolescent Research, 4*, 267-281.

Brandtstadter, J. & Lerner, R. M. (1999). *Action and self-development: theory and research through the life-span*. Thousand Oaks, CA: Sage.

Byrne, B. M. (2001). Structural equation modeling with AMOS: Basic concepts, applications, and programming. Mahwah, NJ: Laurence Erlbaum Associates, Inc.

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (1999). *Positive youth development in the United States: Research findings on evaluations of positive youth development programs*. Washington, D.C.: U.S. Department of Health and Human Services.

Csikszentmihalyi, M. (1998). The development of the person: An experiential perspective on the ontogenesis of psychological complexity. In Damon, W. & Lerner, R.M. (Eds.)*, Handbook of child psychology, Vol 1: Theoretical models of human development*. New York, NY: Wiley.

Damon, W. (2004). What is positive youth development? *Annals of the American Academy of Political & Social Science.Special Positive Development: Realizing the Potential of Youth, 591*, 13-24.

Damon, W., Menon, J., &Bronk, K. C. (2003). The development of purpose during adolescence.*Applied Developmental Science, 7*(3), 119-128.

de Rosenroll, D. A. (1988). Goal Attainment Scaling: *A Vehicle for Group Support in Career and Life Issues Exploration*. 19.

Eichas, K., Albrecht, R.E., Garcia, A., Ritchie, R.A., Varela, A., Garcia, A., Rinaldi, R., Wang, R., Montgomery, M.J., Silverman, W.K., Jaccard, J. &Kurtines, W. K. (2010). A developmental intervention approach to promoting positive development: Pathways of intervention change. *Child & Youth Care Forum, 39(4),*211-237.

Emmerson, G. J. and Neely, M. A. (1988). Two Adaptable, Valid, and Reliable Data-Collection Measures: Goal Attainment Scaling and the Semantic Differential. *Counseling Psychologist, 16(2)*, 261-271.

Ferrer-Wreder, L. A., Cass-Lorente, C., Kurtines, W. M., Briones, E., Bussell, J. R., Berman,S. L. (2002).Promoting identity development in marginalized youth. Journal of Adolescent Research, 17, 168-187.

Ferrer-Wreder, L., Palchuk, A., Poyrazli, S., Small, M. L., & Domitrovich, C. E. (2008). Identity and adolescent adjustment. *Identity: An International Journal of Theory and Research, 8*(2), 95-105.

Flavell, J. H. and Markman, E. M. (1983).*Handbook of child psychology: cognitive development (Vol. 3)*. New York: Wiley.

Habermas, J. (1979). *Communication and the evolution of society*. Boston: Beacon.

Jaccard, J. and Guilamo-Ramos (2002). Analysis of variance frameworks in clinical child and adolescent psychology: Basic issues and recommendations. *Journal of Clinical Child and Adolescent Psychology, 31,* 130-146.

Jaccard, J. & Turrisi, R. (2003). *Interaction effects in multiple regression.* Thousand Oaks, CA: Sage.

Jensen, P., Hoagwood, K., & Trickett, E. (1999). Ivory tow¬ers or earthen trenches? Community collaborations to foster "real world" research. *Applied Developmental Science, 3(4*), 206-212.

Kiresuk, T.J., Smith, A. & Cardillo, J.E. (1994), (Eds.). *Goal attainment scaling: Applications, Theory, and Measurement*. Hillsdale, NJ: Lawrence Eribaum Associates.

Kurtines, W. M., Ferrer-Wreder, L., Berman, S. L., Lorente, C. C., Silverman, W. K., & Montgomery, M. J. (2008). Promoting positive youth development: New directions in developmental theory, methods, and research. *Journal of Adolescent Research, 23*(3), 233-244.

Kurtines, W. M., Ferrer-Wreder, L., Berman, S. L., Lorente, C. C., Briones, E., Montgomery, M. J., Albrecht, R., Garcia, A. J., &Arrufat, O. (2008). Promoting positive youth development: The Miami Youth Development Project (YDP). *Journal of Adolescent Research, 23*(3), 256-267.

Kurtines, W. M., Montgomery, M. J., Eichas, K, Ritchie, R., Garcia, A., Albrecht, R., Berman, S., Ferrer-Wreder, L., &Lorente, C. C. (2008). Promoting positive development in troubled youth: A Developmental Intervention Science outreach research approach. *Identity: An International Journal of Theory and Research, 8*, 125-138.

Lerner, R. M. & Busch-Rossnagel, M. A. (1981). Individuals as producers of their development: conceptual and empirical bases. In *Individuals as producers of their development: a life span perspective*, R. M. Lerner and M. A. Busch-Rossnagel (Eds). New York: Academic Press, 1–36.

Lerner, R. M., Fisher, C. B., & Weinberg, R. A. (2000). Toward a science for and of the people: Promoting civil society through the application of Developmental Science. *Child Development, 71(1)*, 11-20.

Lewis Arango, L, Kurtines, W. M., Montgomery, M. J., & Ritchie, R. (2008). Promoting Positive Youth Development: The Reliability, Validity, and Utility of the RDA Life Course Interview (RDA-LCI) in the Evaluation of the Changing Lives Program (CLP). *Journal of Adolescent Researc, 23(3),* 310-342.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*(1), 83-104.

Maher, C. A. and Barbrack, C. R. (1984). Evaluating the individual counseling of conduct problem adolescents: The Goal Attainment Scaling method. *Journal of School Psychology, 22(3),* 285-297.

Montgomery, M. J. (2005). Psychosocial intimacy and identity: From early adolescence to emerging adulthood. *Journal of Adolescent Research, 20*(3), 346-374.

Montgomery, M. J., Kurtines, W. M., Ferrer-Wreder, L., Berman, S. L., Lorente, C. C., Briones, E., Silverman, W., Ritchie, R., &Eichas, K. (2008). A Developmental Intervention Science (DIS) outreach research approach to promoting youth development: Theoretical, methodological, and meta-theoretical challenges. *Journal of Adolescent Research, 23*(3), 268-290.

Moyer, L. S.and de Rosenroll, D. A. (1984). Goal Attainment Scaling: Its use with pregnant and single-parent teenagers in an alternative education setting. *Canadian Counsellor, 18(3)*, 111-116.

Muthen, L. K., & Muthen, B. O., (2004). Mplus: The comprehensive modeling program for applied researchers. User's guide (3rd ed.). Los Angeles: Muthen & Muthen.

Overton, W. (1998). Developmental psychology: Philoso¬phy, concepts, and methodology. In, R. M. Lerner (Ed.), W. Damon (Series Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (5th ed., pp. 107-187). New York: Wiley.

Paritzky, R. S. and Magoon, T. M. (1982). Goal Attainment Scaling Models for Assessing Group Counseling. *Personnel and Guidance Journal, 60(6),* 381-384.

Rausch, J. R., Maxwell, S. E., & Kelly, K. (2003) Analytical methods for questions pertaining to randomized pretest, posttest, follow-up design. *Journal of Clinical Child and Adolescent Psychology, 32*(2), 467-86.

Rinaldi, R. (2011). A Developmental Intervention Science Outreach Research Approach to Promoting Positive Youth Development. Unpublished Dissertation. Florida International University, Miami.

Ritchie, R. (2007). The Development and Refinement of a Generic Coding Template for Enhancing the Utility, Reliability and Validity of Relational Data Analysis

Theoretical Classification Coding (RDA-TCC).Unpublished Thesis. Florida International University, Miami.

Roth, J., Brooks-Gunn, J., Murray, L., & Foster, W. (1998). Promoting healthy adolescence: Synthesis of youth development program evaluations. *Journal of Research on Adolescence*, 8(4), 423-459.

Seligman, M., Steen, T., Park, N., & Peterson, C. (2005) Positive psychology progress: Empirical validation inventions. *American Psychologist. 60*(5), 410-421.

Schwartz, S. J. (2001). The evolution of Eriksonian and Neo-Eriksonian identity theory and research: A review and integration. *Identity: An International Journal of Theory and Research, 1*(1), 7-58.

Schwartz, S. J. (2002). In search of mechanisms of change in identity development: Integrating the constructivist and discovery perspectives in identity. *Identity: An International Journal of Theory and Research, 2*(4), 317-339.

Schwartz , S. J., , Montgomery, M. J., &Kurtines, W. K. (2005). A comparison of two approaches for facilitating identity exploration processes in emerging adults: an exploratory study. *Journal of Adolescent Research, 20*(3), 309-345.

Schwartz, S. J., Mullis, R. L., Waterman, A. S., & Dunham, R. M. (2000). Ego identity status, identity style, and personal expressiveness: An empirical investigation of three convergent constructs. *Journal of Adolescent Research, 15(4),* 504-521.

Sheldon, K. M., Joiner, T. E., Jr., Pettit, J. W., & Williams, G. (2003). Reconciling humanistic ideals and scientific clinical practice. *Clinical Psychology: Science and Practice, 10*, 302-315.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.

Waterman, A. S. (1993). Finding something to do or someone to be: A eudaimonist perspective on identity formation. In L. Kroger (ed.), *Discussions on Ego Identity*. (p. 147-167). Hillsdale, NJ:Lawrence Erlbaum Associates, Inc.

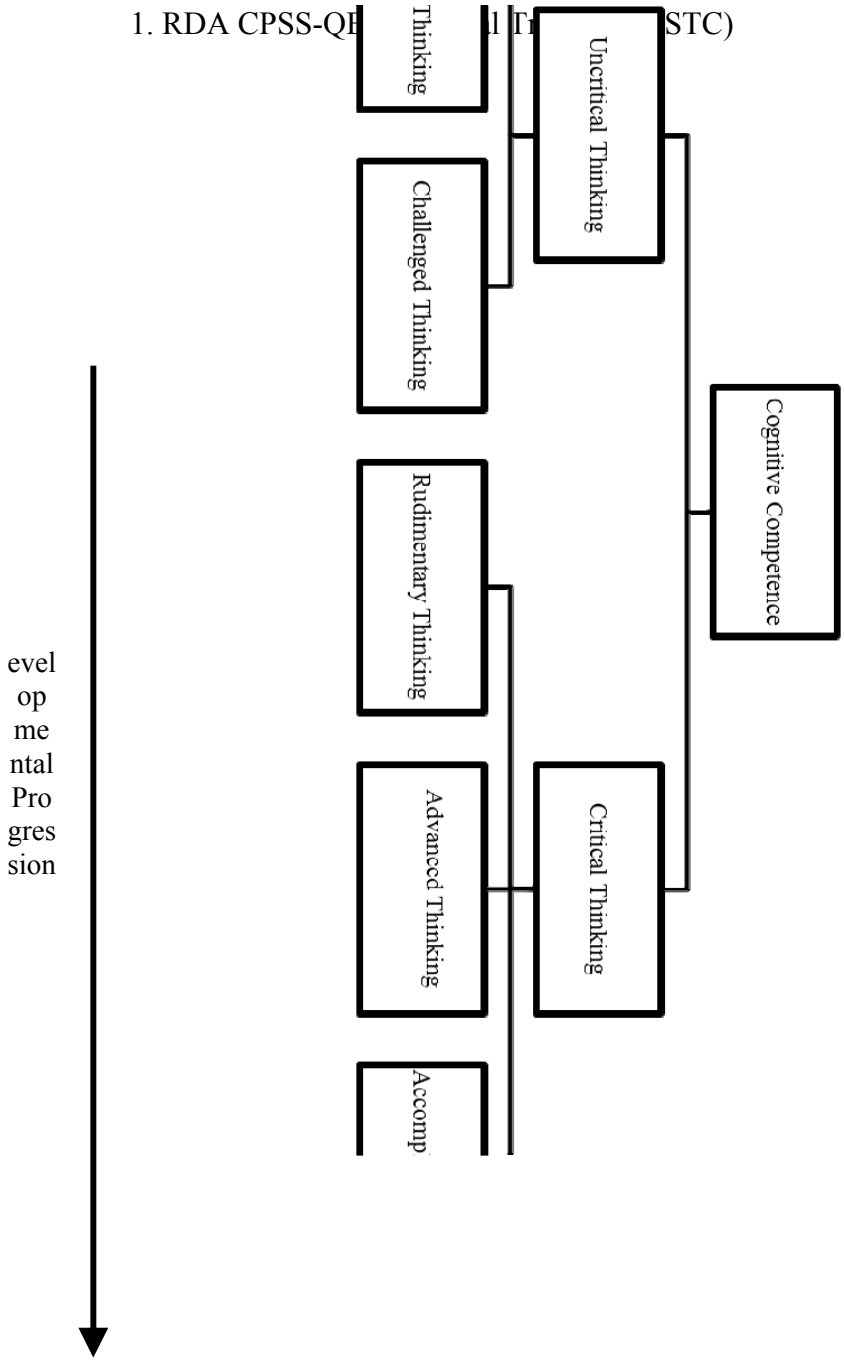
Waterman, A. (2004). Finding someone to be: Studies on the role of intrinsic motivation in identity formation. *Identity: An International Journal of Theory and Research, 4*(3), 209-228.

APPENDIX

Appendix A: CPSS-QE-RDA-CT

1. The CPSS-QE STC
2. Theoretical Category Property Descriptions
3. Theoretical Category Structural Organization
4. Coding Glossary

1. RDA CPSS-QE l Tr STC)



2. Theoretical Category Property Descriptions

- **Uncritical Thinking**: The key property of the Uncritical Thinking category is that the responses demonstrate an unwillingness or inability on behalf of the participant to question his/her original choice. The coders agreed that the properties of this category that made it different from the Critical Thinking responses in this data set is that the Uncritical Thinking responses exclusively included any of the above criteria. The Uncritical Thinking category included two sub-categories, *Unreflective Thinking*, and *Challenged Thinking*.

    - *Unreflective Thinking*: The unique properties of this sub-category were a lack of ability on the participant's behalf to assess his/her thinking, and that the quality of the response was undermined by the expression of prejudices and misconceptions. Responses in this sub-category indicated a lack of awareness of the thinking process.

    - *Challenged Thinking*: Responses in this category demonstrate an awareness of the thinking process (distinct from Unreflective Thinking), and some awareness of the role that critical thinking plays in life course change. The unique property of the responses in this category is the difficulty in recognizing the problems inherent in poor thinking.

- **Critical Thinking**: The key property of the Critical Thinking category is that the responses demonstrate a willingness and ability on behalf of the participant to question his/her original choice. The coders agreed that the properties of this category that made it different from the Uncritical Thinking responses in this data set is that the Critical Thinking responses exclusively included any of the above criteria. The Critical Thinking category included three sub-categories, *Rudimentary Thinking*, *Advanced Thinking* and *Accomplished Thinking*.

    - *Rudimentary Thinking:* The responses in this category show initial signs of attempting to understand, and improve, problematic thinking. Responses in this category have the unique property of possessing some insight into the critical thinking process, but without a plan for adjusting it.

    - *Advanced Thinking:* Responses in this category show evidence of a deeper insight into the consequences of problematic thinking, and also articulate the benefits of critiquing the participant's original choices.

    - *Accomplished Thinking:* Responses in this category demonstrate a "taking charge" of the problem-solving process by the participant. They have the unique property of an effective articulation of the strengths and weaknesses of not only the original choice, but also of the logic underlying it.

3. Critical Problem-Solving Skills Scale – Qualitative Extension

Theoretical Category Structural Organizational Property Description Overview

- **Macro Interview Response (MIR)**: a narrative transcription that includes all the words, phrases, and sentences the participant used in describing the meaning and significance of the experience (topic, issue, question, etc.) under investigation.
- **Structural Coherence**: The primary structural organizational property we have identified for the LCI is Structural Coherence – the degree to which the overall structure and organization of an MIR may be considered a cohesive and coherent expression of a sense of "Who you are." – One's sense of identity
- **Consolidation**: a merging of the multiple elements or components of the self into of an integrated whole

We have identified three sub-categories of **Consolidation**: Unconsolidated, Partially Consolidated, and Fully Consolidated. Each sub-category is defined by variation along the following structural organizational properties: Diversity/Variety, Elaboration/-Articulation, and Differentiation/Integration

Diversity/Variety (MIRs are classified according to the number or frequency of distinct and different RCUs an MIR contains and the degree of variability within each different type of RCU)

- **Diversity** – refers to the number or frequency of distinct and *different* RCUs. That is, number of explicit references to distinct or different personality variables, affective variables, cognitive variables, etc. that the MIR contains.
- **Variety** – refers to the degree of variability within each different type of RCU an MIR contains. For instance, variability in different types of personality variables (e.g., traits, goals, accomplishments, etc.), types of affective variables (e.g., feelings or mood states such as anger, confusion, helplessness, etc.), types of cognitive variables (e.g., thoughts, ideas, attitudes, belief, etc). An MIR may contain a *single* RCU describing only *one* personality variable, but with extensive primary and secondary elaboration; two*, three, or more* different types of personality variables with primary and secondary elaboration *and no* affective or cognitive variables; or *two, three, or more* different types of personality variables *and two, three, or more* different types of affective variables with primary and secondary elaboration; etc.

Elaboration/Articulation (MIRs are classified according to the number or frequency of elaboration (primary and secondary) and the degree to which the elaborations are articulated)

- **Primary Elaboration (PE):** refers to an initial reason, justification, rationale for a particular narrative response.
- **Secondary Elaboration (SE):** refers to additional reasons, justifications rationales given to provide more thoughtful attention to parts or details of reason already given.
- **Articulation:** refers to the degree to which the joining or bringing together words and utterances in elaborations (primary or secondary) is expressed in a creative or refined way *and/or* is intricate and rich in detail *and/or* with a great deal of thoughtful specificity and explicitness.

Differentiation/Integration

- **Differentiation (DIF):** refers to the degree to which an MIR contains primary or secondary elaboration making distinctions and/or discriminating among RCUs along either a horizontal or vertical temporal dimension. A horizontal temporal dimension refers to making distinctions among two or more RCUs (e.g., personality variables) at the *same* point in time (either past, present, or future). For instance, "My [current] goals in life are to accomplish A and B, and B is more [or less] important [significant, meaningful, etc.] to me." A vertical temporal dimension refers to making distinctions among two or more RCUs (e.g., personality variables) *across* (over) time (e.g., from the past to the present or the future). For instance, "When I was younger my goals in life were to accomplish A and B; however, B is no longer [or less] important [significant, meaningful, etc.] to me now than A." Primary or secondary elaboration making distinctions and/or discriminations involves making comparative (e.g., more than, less than, etc.) or evaluative (e.g., better than, worse than, etc.) judgments between RCUs or across RCUs over time.
- **Integration (INT):** refers to the degree to which an MIR contains primary or secondary elaboration that describes the relations between RCUs in an effort to identify links or connections between them; unify; or join or unite them with each other. For instance, "When I was younger I used to think that I had to choose between a career and a family; however, now I think I can do both."

4. Critical Problem-Solving Skills Scale – Qualitative Extension Decision

Coding Glossary

Articulation: joining or bringing together words and utterances in a creative or refined way.

Category: a group of RCUs that share one non-overlapping (unique) property.

Consolidation: merging multiple elements or entities.

Decision Tree Chart (DTC): a "diagram" for facilitating classification coding decision-making processes and its "associated descriptive matrix of classification criteria."

Differentiation: to make distinctions; discriminate.

Diversity/Variety: involving multiple elements or entities.

Elaboration, Primary: the term used to code the initial or (if stated) primary reason, justification, rationale for a particular narrative response.

Elaboration, Secondary*:* the term used to code additional reasons, justifications, rationales for a particular narrative response for which a Primary Elaboration has already been coded.

Macro Interview Response (MIR): a narrative transcription that includes all the words, phrases, and sentences the participant used in describing the meaning and significance of the experience (topic, issue, question, etc.) under investigation.

Organization: something made up of elements with varied functions that contribute to the whole and to collective functions

Positivity: confident, optimistic, and focusing on good things rather than bad

Response Content Unit (RCU): units of response content (i.e. the ideas, incidents, examples, persons, events, things, actions/interactions, etc.) that are made up of the specific words, phrases, and sentences used in response to the interview questions and probes**.**

Structure: the interrelation or arrangement of parts in a complex entity.

Structural Tree Chart (STC): a "diagram" for visually representing the structural organization of an identified set of theoretically meaningful categories.