Spring 2012

Perceived responsibility for learning in college students: A construct validity study

Matthew S. Swain
James Madison University

Follow this and additional works at: https://commons.lib.jmu.edu/master201019
Part of the Psychology Commons

Recommended Citation

This Thesis is brought to you for free and open access by the The Graduate School at JMU Scholarly Commons. It has been accepted for inclusion in Masters Theses by an authorized administrator of JMU Scholarly Commons. For more information, please contact dc_admin@jmu.edu.
Perceived Responsibility for Learning in College Students:

A Construct Validity Study

Matthew S. Swain

A thesis submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Arts

Department of Graduate Psychology

May 2012
Acknowledgements

There are many people that have supported me in the development of this thesis as well as through my graduate school experience. First, I would like to acknowledge my advisor, Dr. Keston Fulcher, for encouraging me to pursue projects that were of personal interest, honing my writing skills, and for supporting me when I needed it. I would also like to thank the other members of my thesis committee, Dr. Kenn Barron and Dr. Sara Finney. Thank you, Kenn, for your thought-provoking questions and perspective. And a hearty thank you to Sara for teaching me more than I thought I could learn in two years.

I would also like to thank a few special peers who have made my graduate school experience extremely fulfilling. Special thanks to Jessica Flake, Bo Bashkov, Kathleen Casto, Kelli Samonte, Jerusha Gerstner, and Makayla Grays. Whether you answered my questions at 2 AM, forced me to go on a bike ride, presented with me, helped me through a tough decision, or laughed with me during our truly non-existent free time, each of you and many others have made the past two years memorable and enriching. I probably would not have achieved as much without your friendship.

I must acknowledge all that the Orientation Office has contributed to my growth having serendipitously “struck gold” with the best Graduate Assistantship at JMU. Thank you, Tisha McCoy-Ntiamoah, for being a superb supervisor and never forgetting to ask me how I am doing even when we have a lot of work to do. And thank you to Sarah Sunde for being my fellow conferee and trouble-maker.

And finally, for the most important people in my life: my parents. Thank you, Mom and Dad for talking me through the bumps in the road and for always being my biggest fans—I am truly proud of everything you are and fortunate for your continued
support these past two years. As the first two of many teachers in my life, you have taught me to be responsible, mindful, loving, and to work hard to achieve my potential. And for that, I love you both very much and am tremendously grateful.
# Table of Contents

Acknowledgements ........................................................................................................ ii
List of Tables ................................................................................................................ vi
List of Figures ............................................................................................................... vii
Abstract ...................................................................................................................... viii

## CHAPTER ONE: INTRODUCTION ............................................................................. 1
  Importance of Responsibility for Learning ............................................................. 1
  Measures of Responsibility for Learning .............................................................. 2
  Purpose of the Current Study................................................................................... 5
  Scale Validation ....................................................................................................... 6
  Validity Evidence for the PRL ................................................................................. 7
    Substantive evidence for the PRL ...................................................................... 7
    Structural evidence for the PRL ...................................................................... 8
    External evidence for the PRL ...................................................................... 11
  Research Questions ............................................................................................... 12
  Hypothesized Structural Models .......................................................................... 13
    One-factor model .......................................................................................... 13
    Three-factor model ......................................................................................... 13
    Bifactor model ............................................................................................... 13
  Hypothesized External Relationships .................................................................. 15

## CHAPTER TWO: LITERATURE REVIEW ..................................................................... 16
  Notions of Responsibility ..................................................................................... 16
  Ascription of Responsibility ................................................................................ 19
  Responsibility in Education ............................................................................... 21
  Student Affairs Assessment .................................................................................. 22
    Importance of responsibility programs ......................................................... 23
    Assessment of responsibility programs ......................................................... 23

## CHAPTER THREE: METHOD ................................................................................... 25
  Participants and Procedures ................................................................................. 25
  Measures .............................................................................................................. 25
    Perceived Responsibility for Learning ............................................................ 25
    Achievement Goal Questionnaire-Revised ..................................................... 26
  Data Screening .................................................................................................... 26
  Sample Split ......................................................................................................... 27
  Testing Hypothesized Structural Models ............................................................ 28
  Planned Data Analyses ....................................................................................... 29

## CHAPTER FOUR: RESULTS .................................................................................... 30
  Item-Facet Mapping ........................................................................................... 30
  Confirmatory Factor Analysis ............................................................................. 31
    Global fit indices ......................................................................................... 31
    Local model misfit ......................................................................................... 32
    Tests of hypothesized models .................................................................. 32

## CHAPTER FIVE: DISCUSSION .................................................................................... 35
  Future Directions for Measuring Responsibility for Learning ......................... 38
  Suggested revisions for the PRL ...................................................................... 38
Response scale concerns .................................................................39
Suggested analyses for the revised PRL ............................................39
Designing a New Scale ....................................................................40
Limitations of the Current Study ......................................................42
Summary and Concluding Remarks ..................................................43
Appendix A ....................................................................................45
Appendix B ....................................................................................46
Appendix C ....................................................................................47
Appendix D ....................................................................................54
References ....................................................................................56
List of Tables

Table 1. Correlation Matrix and Descriptive Statistics for Full Sample ..................47
Table 2. Correlation Matrix and Descriptive Statistics for Sample A .......................48
Table 3. Correlation Matrix and Descriptive Statistics for Sample B .......................49
Table 4. Correlation Matrix and Descriptive Statistics for Sample C .......................50
Table 5. Fit Indices of Three Hypothesized Models of Perceived Responsibility for Learning ..................................................................................................................................51
Table 6. One-Factor Model Correlation Residuals for Sample A ..........................52
Table 7. Bifactor Model Correlation Residuals for Sample A ..............................53
List of Figures

Figure 1. Three Hypothesized Models for PRL Scale ........................................54

Figure 2. The Triangle Model of Responsibility and the Accountability Pyramid ..........55
Abstract

Responsibility for learning is an important, foundational construct for students in higher education. Because of its importance, higher education officials often design programs to inform students of their academic responsibilities. In order to assess these programs, a valid measure of responsibility for learning must be selected. In order to assess program effectiveness, measurement specialists collect validity evidence to support score interpretations. The current study focuses on the validity evidence of the Perceived Responsibility for Learning (PRL) scale. Benson’s (1998) framework for construct validation was used to examine current validity evidence and direct the study. Competing factor structures of the PRL were examined. Results indicated that a factor structure is still unknown; however, this may be due to measurement issues with the current scale. Directions for future responsibility for learning research are provided.
Importance of Responsibility for Learning

When considering the important educational constructs of college graduates, one may think of scientific reasoning, communication and critical thinking. Indeed, these areas are often the focus of general education assessment. A foundational construct that may not initially come to mind is perceived responsibility for learning, which is defined as a volunteristic acceptance of student academic expectations (Neff, 1969). Davis and Murrell (2003) stated that academic responsibility is the key to all development and learning. Admittedly, however, anecdotal (Delvin, 2002; Weimer, 2011) and empirical evidence (Schmelzer, Schmelzer, Figler, & Brozo, 1987) suggest that students in higher education are less responsible for their learning and educational outcomes than desired by both teachers and researchers. Educators and administrators may find this undesirable because academically responsible students attain higher semester and cumulative GPAs (Schlenker, 1997). Realizing the importance of responsibility for learning, an American Psychological Association (APA) task force on Psychology and Education emphasized that personal and academic responsibility should be a chief educational goal for the 21st century (Sternberg, 2003). Higher education officials should, therefore, aim to increase college student responsibility for learning.

In response, many universities create interventions or programs to clarify these expectations. Following recommendations from the National Association of Student Personnel Administrators (NASPA; Berson et al., 1998) and the Council for the Advancement of Standards in Higher Education (CAS; 2008), such programs are often
implemented during first-year orientation. The idea is that students should realize, as soon as possible, their responsibility for learning in a college context.

In addition to creating programs to address the important construct of responsibility for learning, university officials may wish to evaluate the efficacy of these interventions. To do so, they may select an appropriate instrument—one that can accurately capture students’ perceived responsibility for learning. Or, stated from a measurement perspective, university officials should select an instrument with validity evidence for the purpose of measuring college student’s perceived responsibility for learning.

**Measures of Responsibility for Learning**

Few measurement scales target responsibility for learning. Criticisms abound for those that do. In 2009, Reiss developed the 8-item Reiss School Motivation Profile Honor Scale (RSMP) for secondary education students. The author proposed that the RSMP measures student responsibility; however, the scale focuses on the student’s valuation of ethics and morals and his or her propensity to cheat. The Intellectual Achievement Responsibility (IAR) Questionnaire (Crandall, Katkovsky, & Crandall, 1965), a 34-item scale, was developed to assess elementary students’ perceived responsibility for academic successes or failures. The scale, however, has low internal consistency.

Chowning and Campbell (2009) developed a 15-item measure of academic entitlement, which is the expectation of positive academic outcomes regardless of a demonstration of intellectual merit. This scale included a 10-item Externalized Responsibility (ER) subscale, which addresses the responsibilities of the student and
others in the learning process, and a 5-item Entitled Expectations (EE) subscale that concerns the grading policy expectations of the professor. The ER subscale had high internal consistency (α = .81) and correlated negatively with measures of personal control, need for cognition, and self-esteem, which the authors propose are adaptive qualities in an educational setting. However, the structure of the academic entitlement scale was assessed using questionable methods, and the items measure students’ externalizing responsibility to their peers and university as well as their professor. Although the RSMP, IAR, ER and EE scales attempt to measure various aspects of responsibility in academia, the focus of the current study is on yet another scale because of conceptual differences in the definition of responsibility for learning.

In 2005, Zimmerman and Kitsantas developed the Perceived Responsibility for Learning scale (PRL; Appendix A). It was used to study the effects of homework practices on academic achievement. Specifically, Zimmerman and Kitsantas examined the effects of quality and quantity of homework on GPA with self-efficacy for learning and perceived responsibility for learning as mediating variables. The PRL was developed using a sample of 180 female high school students and consists of 18 items asking students to indicate who is more responsible for certain academic outcomes. The items represent academic outcomes such as “motivation (e.g., going through the motions without trying), deportment (e.g., fooling around in class), and learning processes (e.g., not taking notes in class)” (p. 404).

Students rate their level of responsibility for these outcomes on an atypical 7-point Likert scale: 1 (mainly the teacher), 2 (definitely more the teacher), 3 (slightly more the teacher), 4 (both equally), 5 (slightly more the student), 6 (definitely more the
student), and 7 (mainly the student). Responses are forced on a continuum that leaves either the student or teacher more responsible as only one opportunity is provided for students to rate the level of responsibility between the student and the teacher. Midpoint responses on the scale are used to indicate that the academic outcome is a responsibility of both the teacher and the student. One limitation of this scale format is that students are unable to indicate that an academic outcome is neither the teacher’s nor the student’s responsibility. According to Zimmerman and Kitsantas (2005), higher scores indicate a higher sense of responsibility attributed to the student.

The PRL has been implemented rarely since its genesis. Zimmerman and Kitsantas (2007) used the PRL in a study of the predictive validity of the Self-Efficacy for Learning Form (SELF) using college students. They used multiple regression analyses to test the predictive power of SAT, SELF, and PRL scores on both homework quality and course grades. The inclusion of the PRL explained an additional 1% of the variance in homework quality above and beyond SAT and self-efficacy scores but this increase was not statistically significant \( p = .23 \). For course grades, the PRL predicted an additional 2% of the variance above and beyond SAT and self-efficacy scores, which was statistically significant \( p = .03 \).

Kitsantas and Zimmerman (2009) used the PRL in a replication of the Zimmerman and Kitsantas (2005) model of the effects of homework practices on academic achievement. In this model, academic achievement is mediated by self-efficacy for learning and perceived responsibility for learning. Unlike the original study, this model was tested on a sample of college students. Similar to the 2005 study, the path model fit the data well, \( \chi^2 (1, N = 223) = 1.34, p < 0.25, NFI = 0.99, CFI = 1.00, RFI = \)
0.96); however, perceived responsibility had a weaker relationship with academic achievement than in the high school sample. Kitsantas and Zimmerman (2009) suggest that this weaker relationship may be due to the level of self-directed work required in a collegiate setting. However, it is more plausible that the findings are due to different criterion variables. Although the models in both studies indicated that perceived responsibility explained a significant portion of variance in an academic outcome, a college course grade replaced GPA in the model in 2009. Comparisons between the studies are difficult as a student’s performance in one course may not fully capture the effect of responsibility for learning on academic outcomes in general (i.e., GPA). Further, from a methodological perspective, the high goodness of fit statistics related to the second study could be due to the model itself. Models with fewer degrees of freedom tend to fit better (Hoyle, 2005). In this case, the degrees of freedom is merely “1”.

No further work has been conducted on the 18-item PRL scale beyond the previous three studies (B. J. Zimmerman, personal communication, September 17, 2011). One might initially suppose that this is due to the scale being relatively new to the educational community. Additionally, the original studies intended to provide validity evidence for a measure of self-efficacy rather than perceived responsibility for learning. Assessment practitioners do not have a sufficient scale to measure this important construct. Thus, further validation of the PRL scale is needed to justify its use in colleges and universities.

**Purpose of the Current Study**

In response to this need, the purpose of the current study is to gather additional validity evidence for the PRL. The process of validation provides evidence for or against
the interpretation of scores on the scale. The PRL is one of the only scales available as an outcome measure of perceived responsibility for learning; therefore, the current study is needed to investigate the use of a scale with very few alternatives. Additionally, assessment practitioners at the author’s university have recently selected the scale to assess the effectiveness of an orientation program. However, the current interpretations of the scores should be evaluated prior to making programmatic decisions. In the current study, I primarily aim to investigate competing factor structures of the PRL as well as response scale issues.

Benson’s (1998) framework of construct validity will guide this study. She describes validation as “the process by which scores take on meaning” (p. 10). Benson suggested three stages for developing a strong program of construct validity: substantive, structural, and external. An overview of Benson’s framework for scale validation is provided next followed by an outline of current validity evidence for the PRL and the validity evidence still needed.

**Scale Validation**

Researchers should define the theoretical aspects of the targeted construct during the substantive stage (Benson, 1998). Construct theory is then used to direct the development of scales. Researchers should be mindful to reduce the occurrence of what Messick (1995) refers to as *construct underrepresentation*, or a failure to include all relevant facets of a construct, and *construct irrelevant variance*, or extraneous variance as a result of a broad definition of a construct or methodological issues. A scale that is developed without the use of a strong theory, or the contribution of content experts, may be too narrow or broad in its measurement. Additionally, if a theory indicates that the
construct is multi-dimensional, or consisting of different components, this theoretical scale structure may be tested during the structural and external stages.

The structural stage is marked by analyses of relationships among observed variables (Benson, 1998). The hypothesized interrelations among various dimensions of the construct, defined in the substantive stage, can be tested in the structural stage. Researchers may use empirical evidence to advocate for or against a particular model being evaluated. Common analyses in this stage include exploratory and confirmatory factor analyses. Researchers use these procedures to examine the expected dimensionality among the items on a scale suggested by theoretical foundations.

Finally, the external stage consists of research that examines hypothesized relationships between the measured construct and other variables, or between groups, as predicted by theory (Benson, 1998). If researchers find that the measured construct relates to other variables in theoretically expected ways, evidence is then added to the argument that the scale measures what is proposed (Kane, 1992). Tests of these relationships may involve correlation with other variables or examining known group differences in which a scale indicates a difference between groups that is congruent with expert judgments.

Validity Evidence for the PRL

The educational psychology literature provides little validity evidence for the PRL scale. The evidence that does exist is contained in a few studies. The following sections note existing evidence as organized by Benson’s (1998) framework.

Substantive evidence for the PRL. Zimmerman and Kitsantas (2005) provide limited detail about the theory underlying the PRL items. They indicate that the items
represent three groups of learning outcomes, hereinafter referred to as facets, of responsibility for learning: (a) motivation, (b) deportment (or behaviors), and (c) learning processes. No theoretical basis for these facets was provided by the authors. It is also unclear which items were written to measure each facet. In scale development, we would hope for a much tighter mapping between theory and items. If these facets are to be considered further, the items must be mapped to and analyzed with these potential relationships in mind.

Zimmerman and Kitsantas (2005) did expect one latent factor structure for the PRL. They hypothesized that students should respond to each item similarly as all items were written to tap the same general construct of responsibility for learning. This hypothesized structure was then tested.

**Structural evidence for the PRL.** Zimmerman and Kitsantas (2005) conducted an exploratory principal components analysis (PCA) using responses to 20 items written to measure responsibility for learning. Their sample consisted of an entire student body of a parochial high school for girls ($N = 180$). Three components emerged that accounted for 81% of the variance. The first component accounted for 69% of the variance (eigenvalue = 13.83), and the remaining two components accounted for 7% (eigenvalue = 1.50) and 5% of the variance (eigenvalue = 1.00). The authors do not indicate what type of rotation was used. Using the coefficients from the PCA, the authors removed two items that did not relate to the first component above .70 resulting in a final scale of 18 items. Zimmerman and Kitsantas proposed that, due to the similarity in item content, the scale is unidimensional and item scores may be summed together. Cronbach’s $\alpha$ for the scale was .97 indicating high internal consistency.
It is worth sharing several methodological flaws of the previous study. PCA is an improper technique to model latent factors (Preacher & MacCallum, 2003). PCA creates components that attempt to explain as much variance as possible (including unwanted error variance) using observed variables. On the other hand, an exploratory factor analysis (EFA) attempts to explain common variance by assuming a latent factor is driving responses to items. Because responsibility for learning is a latent variable, and latent variables drive responses on observed variables, an EFA would have been a more appropriate analysis to model latent factors when the factor structure is unknown.

Moreover, when the dimensionality of a scale is hypothesized a priori, a confirmatory factor analysis (CFA) is a more appropriate and stringent analysis to examine the scale structure than an EFA.

Apart from the PCA analysis in the initial study, no other study has investigated the factor structure of the original PRL (B. J. Zimmerman, personal communication, September 17, 2011). However, Magno (2011) developed a new, 30-item measure of perceived responsibility for learning using the same three substantive facets established in Zimmerman and Kitsantas (2005). For this new scale, ten items were written to measure each facet: (a) motivation, (b) deportment, and (c) learning processes. Magno proposed a three-factor structure given the items were written to represent three facets of responsibility for learning. Additionally, a one-factor model was tested as this was the structure championed for the PRL. A series of CFAs were conducted using data from a sample of 2,054 college students from the Philippines: a one-factor model, a three-factor model, and three two-factor models testing a pair of facets as one factor (20 items) and the remaining facet as a factor by itself (10 items).
The one-factor model did not fit the data well, \( \chi^2(405, N = 2,054) = 5,674.94 \), SRMR = .062, RMSEA = .089. The other models did not fit in a global sense; moreover, these models are not relevant to the current study. The results indicated that the three-factor model fit the data best compared to the other models, \( \chi^2(402, N = 2,054) = 4,694.58 \), SRMR = .057, RMSEA = .080. The correlations among the factors were low to moderate (.10, .11, and .65). From these findings, Magno (2011) concluded that the 10 items representing each facet are distinct subscales of perceived responsibility for learning and championed a three-factor model. Although using a different scale, Magno presents a new factor structure for responsibility for learning. As the items from the PRL and Magno’s scale were written to measure the same three facets, the three-factor model is a possible model for the structure of the PRL item responses.

Magno’s (2011) methodology may have some flaws. He did not calculate difference tests between the nested factor structure models. Researchers should conduct difference tests because a more complex model with fewer degrees of freedom will fit better than a simpler model in a comparative sense (e.g., a three-factor model will fit better than a one-factor model). When I calculated this difference test, the one-factor model fit significantly worse than the three-factor model supporting his conclusion (\( \Delta \chi^2(3) = 980.36, p < .001 \)). Additionally, researchers should report comparative fit indices (i.e., CFI) and assess local misfit (i.e., standardized covariance residuals) to support a championed model. No such information was provided. Therefore, caution should be taken before concluding that the three-factor model can adequately explain the relationships among the items. Given Magno’s (2011) findings, it can be presumed that similar relationships among items might occur in the original PRL due to the same
substantive evidence driving the creation of the items; however, this hypothesis has yet to be tested.

**External evidence for the PRL.** Validity evidence for the PRL is limited to the original studies. Recall that higher scores on the PRL indicate that a student perceives student as more responsible for learning than the teacher. In the first study (Zimmerman & Kitsantas, 2005), the PRL was highly correlated with homework quality \( (r = .63) \), quantity of homework \( (r = .74) \), a self-efficacy for learning measure \( (r = .71) \), and end of semester GPA \( (r = .86) \). Zimmerman and Kitsantas proposed that the high correlation between the PRL and GPA and the 22% more variance explained in GPA by the PRL above homework practices provides convergent validity for the measure because “perceived responsibility is clearly an important motive for academic achievement emerging from homework experiences” (p. 410). Although the PRL and self-efficacy for learning measures were highly correlated, both measures explained unique variance in GPA, which Zimmerman and Kitsantas interpreted as discriminant validity evidence.

The aforementioned study on high school students was replicated in a college setting (Kitsantas & Zimmerman, 2009). A sample of college students \( (N = 223) \) provided responses to homework quality, homework quantity, and self-efficacy for learning measures including the PRL. SAT and course grades were also collected for this study. The PRL was positively correlated with course grade \( (r = .40) \), homework quality \( (r = .38) \), homework quantity \( (r = .40) \), self-efficacy for learning \( (r = .50) \), and SAT scores \( (r = .36) \). The authors proposed that the discrepancy between the relative contributions of the PRL to academic outcomes in the 2005 and 2009 study is explained by the types of students in the samples. Students’ homework practices are monitored
more intently by high school teachers than college professors, particularly at the parochial school that participated in the 2005 study; therefore, a greater proportion of their grade was surmised to be dependent on their homework practices.

Given the aforementioned validity evidence, the PRL may have some utility as an outcome measure of academic responsibility. Clearly, further psychometric evaluation must be performed to garner more credibility for the scores. The purpose of this thesis is to gather evidence according to Benson’s (1998) framework; however, it may not be possible to explore every stage given that poor results in one stage may redirect the course of study. For example, if the hypothesized dimensionality of the scale is not supported, it may not make sense to conduct studies in the external stage.

**Research Questions**

Zimmerman and Kitsantas (2005) have provided some substantive validity evidence. Although not derived from a responsibility theory, the researchers wrote the PRL items with extensive knowledge of educational constructs. Because the original mapping of the items to the three facets is unknown, a content alignment will be performed to retroactively map items to facets (Dawis, 1987). After the 18 items are mapped to their respective facets, the structure of the scale may be tested. The structural evidence of the PRL is the primary focus of this thesis. Both a one-factor model and a three-factor model have been proposed in the literature. Neither model has been tested in a confirmatory manner using responses to the PRL. Finally, regarding external validity evidence, Zimmerman and Kitsantas (2005) have correlated the PRL with certain academic outcomes. Given that an acceptable structure of the PRL is established, I will correlate PRL scores with other constructs according to the relevant research.
Hypothesized Structural Models

Competing theories of the dimensionality of the PRL can be addressed using CFA. What follows is a review of the competing hypothesized models as indicated by the literature and the order by which each model will be tested. A study of these models will attempt to add structural validity evidence to the PRL.

One-factor model. A one-factor model will be tested first (see Figure 1). Zimmerman and Kitsantas (2005) championed this model implicitly by the interpretation of the PCA results and explicitly by the summing of the items in their study. If a one-factor model is found to fit the data, this parsimonious model will lend support to the conclusion that the PRL is unidimensional and may be scored by summing the items. I will test alternative models if satisfactory fit is not found.

Three-factor model. Considering the available evidence regarding the item writing (Zimmerman & Kitsantas, 2005), as well as the work of other authors on a similar scale (Magno, 2011), a three-factor model may best explain the relationships among the items. However, it is unknown which items were written to measure each facet of responsibility in the PRL (i.e., motivation, deportment, and learning processes). Therefore, this factor structure will be specified and then tested using the results of the content alignment.

Bifactor model. In addition to the two aforementioned models, a bifactor model may be appropriate for the current study. In a bifactor model, researchers specify a general factor for all items (Reise, Morizot, & Hays, 2007). Additional factors are specified representing common variance among a group of items (see Figure 1). In the current study, the general factor is assumed to be responsibility for learning as proposed
by Zimmerman and Kitsantas (2005). Simultaneously, the item groups representing the three facets will be used to partition unique variance due to method effects. To be clear, these three method effects should not be confused with the original three facets. Recall that the original three facets were hypothesized to be correlated. In a bifactor model, researchers specify all factors to be orthogonal (Chen, West, & Sousa, 2006).

More specifically, in the bifactor model, researchers model common variance among groups of items after controlling for common variance among all items. If the bifactor model fits the PRL, standardized factor pattern coefficients for the general factor may be compared to the standardized factor pattern coefficients of the method effects factors. If the scale is unidimensional, no systematic residual variance should be shared across items representing each method effect (Reise, Morizot, & Hays, 2007). Thus, the factor pattern coefficients associated with the method effect factors should be low. This means that the method effect factors do not need to be modeled and the scale is essentially unidimensional (Reise, Moore, & Haviland, 2010). However, if items correlate above and beyond the general factor, the items are multidimensional and should be modeled as such (with a bifactor model) if the scale is to be used in practice. Additionally, variance shared among method effect factor items may be redundant with the general factor and this may be captured using a bifactor model (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012).

Both global and local misfit, along with replicability, will be assessed when comparing the three competing structural models. Researchers should note that observed model-data misfit may be an idiosyncrasy of a particular sample; therefore, replication
across samples is encouraged (MacCallum, Roznowski, & Necowitz, 1992). If global misfit replicates, then item level misfit will be examined.

**Hypothesized External Relationships**

To provide external validity evidence, the PRL may be modeled with measures of motivation goal orientation given adequate fit from the structural stage. Specifically, students who exhibit student-centered responsibility for learning are hypothesized to score higher on *mastery-approach* measures of goal orientation and lower on *performance-approach* measures. Elliot and Murayama (2008) defined *mastery-approach* as striving to achieve competence according to an intrapersonal standard and *performance-approach* as “focused on attaining normative competence” (p. 614).

Bacon (1993) posited that students who are academically responsible are intrinsically motivated and emphasize greater personal control over their academics. Corno (1992) also proposed that students who are more responsible for their learning differ in the demonstration of their competence. She hypothesized that students who are more responsible for their learning are learning or mastery oriented. That is, these students desire to learn the material deeply, to master the concepts. Students who are low in responsibility for learning are hypothesized to be more performance oriented—striving to do well in relation to others (Elliot & Murayama, 2008). These students tend to be less focused on deep processing. The relationship between motivation and responsibility in education can best be understood by the description of a certain intervention: “the locus of the writing task moved from the teacher to the student, shifting from ‘performance’ toward ‘mastery’ goals” (Corno, 1992, p. 77).
It would be prudent to review the general responsibility literature and how responsibility is manifested in academics prior to delving into the methodology of the current study. The PRL is a potential outcome measure of perceived responsibility for learning. That is, programs designed to increase responsibility should desire to move student scores on this scale to reflect their actual increased level of perceived responsibility for learning. How students determine their level of responsibility is a relevant subject because it is directly related to item responses. Therefore, responsibility theory may help explain how individuals perceive their responsibility for learning in an educational context.
CHAPTER TWO: LITERATURE REVIEW

Based on rhetoric of institutions and initiatives (such as APA, NASPA, and CAS), responsibility for learning is widely endorsed in education. Nevertheless, how students actually conceptualize responsibility is quite nebulous. As this thesis concerns validity evidence for the PRL, particularly in the structural stage, a review of the philosophical roots and definitions of responsibility is merited to establish some substantive background of the construct.

In an attempt to determine how responsibility is distinct from related constructs (e.g., accountability), I review various conceptualizations of responsibility. I then discuss a model that describes the process one takes to ascribe the level of responsibility of an individual. The manifestation of responsibility within education is then considered with a specific focus on the use of the PRL in higher education assessment. In order to establish the appropriate validity evidence for the use of the scale in assessment, a review of the responsibility theory is necessary.

Notions of Responsibility

From its inception in philosophy to its use in modern law and society, the definition of responsibility has eluded unanimity. Modern philosophers have used terms such as accountability and trustworthiness in an attempt to consolidate the facets of responsibility (Baker, 2002). The philosopher Richard McKeon (1990) noted that the word “responsibility” first appeared in English and French (responsabilité) around the year 1787 to describe the political changes during the American and French revolutions. Documents resulting from these revolutions, including the United States Constitution, included the word to establish the responsibilities of government. McKeon attempts to
synthesize various philosophical dissentions regarding the accurate definition of responsibility. He described three aspects of responsibility. First, *accountability* relates to an actor’s obligation to law and punishment. *Imputability* concerns cause and causal agents and *freedom and rationality* relates to the social context with which judgments of responsibility are made.

Whereas McKeon proposed three aspects of responsibility, Neff (1969) instead proposed that there are two general definitions for responsibility: *accountability* and “the voluntaristic act of assuming an obligation” (p. 14). Neff’s notions of accountability are similar to McKeon’s in that his definition of accountability concerns legal liability for which the actor is completely obliged to uphold. Responsibility, on the other hand, must be voluntarily welcomed on the part of the actor. To become responsible, two reactions must take place: “the act of responding…and the assumption of an obligation” (p. 16). Neff proposes that the actor must understand the underlying values of the obligation before it is assumed and that the responsibility must be assumed on the actor’s volition. For example, agreeing to water a friend’s plants would be an assumed responsibility whereas following the laws of a governing body is mandatory and considered to be accountability.

In a review of the responsibility literature, Schlenker (1997) identified six definitions of responsibility. Responsibility as:

1. *Causation* which refers to the connection between the actor and the event;
2. A *mental state* which includes the intentionality of the actor in that they are not deemed responsible for unanticipated consequences of the event;
3. A *mental or physical capacity* considers whether the actor is capable of assuming responsibility. Children and mentally handicapped individuals are good examples of instances where the capacity for responsibility is limited;

4. An *obligation* refers to moral codes or laws that actors are expected or required to follow;

5. *Social roles* are social codes of conduct that hold actors responsible. In the case of a parent of a young child, the child may break an expensive object but due to the parent’s role, he or she is then responsible for the child’s actions.

6. And *answerability* which addresses the judgments of an actor’s conduct and liability.

Schlenker posited that, in accordance with McKeon (1957) and Neff (1969), accountability is conceptually distinct from voluntary responsibility. In his review of the responsibility literature, both obligation and answerability are seen to be conceptually subsumed under accountability. After considering the previous definitions, Schlenker (1997) proposes that the remaining definitions of responsibility may be subsumed in: (a) an individual’s prescribed behaviors or codes of conduct, (b) the event that occurred or is expected to occur, and (c) the identities or roles of the actor.

Up to this point, responsibility has been conceptualized as a construct whereby an actor’s volitions bind him or her to behave in prescribed ways according to his or her complex roles. This definition is conceptually different from accountability and liability as those terms address jurisprudence rather than obligations based on an individual’s role. It appears that a notion as multifaceted as responsibility is best explained by exploring the conceptual processes. In the next section, I will attempt to further clarify the definition of
responsibility as well as explain the process by which an individual’s level of responsibility is determined. These processes are presented to explore what students may be thinking when responding to the PRL.

**Ascription of Responsibility**

When one attempts to ascribe the level of responsibility of an individual, certain information is gathered to make this judgment. Various conceptual processes have been proposed to explain the information required to determine an individual’s level of responsibility. When researchers understand these processes, issues in the measurement of responsibility for learning may be clarified.

Schlenker, Britt, Pennington, Murphy, and Doherty (1994) proposed a model of responsibility comprised of three elements. They argued that an audience, or individual(s) evaluating the level of responsibility, must examine each of these elements to determine the degree to which an actor is responsible. These elements are (a) the *prescriptions* or codes of conduct that guide the actor in the situation, (b) the *event* or consequences that relate to the prescriptions, and (c) *identity images* that indicate the actor’s roles and qualities. These three elements form the Triangle Model of Responsibility (TMR) as seen in Figure 2. All three elements are necessary for responsibility to be ascribed to an actor.

Schlenker et al. (1994) defined *prescriptions* as the prescribed behaviors such as rules or codes of conduct that apply to the actor under evaluation. These prescriptions may include social norms, laws, moral codes, or any other cultural or situational rules that apply to the particular actor. For example, Beth, a student who is given a study guide on Monday for a test on Friday is expected to study the information it contains. Beth has
many prescriptions that may apply to her stemming from various aspects of her identity; however, when evaluating her level of responsibility as a student, the applicable prescriptions are those that relate to her student identity. Only appropriate prescriptions are used to assess responsibility during the relevant occurrence or event.

*Events* refer to the set of occurrences evaluated given an actor’s prescriptions and identity. Events often include the actor’s behaviors, or lack thereof, in a given situation. An evaluator considers only the behaviors that are relevant to the responsibility situation. In our student example, the events under consideration are what Beth does between Monday and Friday given that the test covers material that is completely new to her. To determine Beth’s responsibility, her actions in this time period are the only actual behaviors that are relevant.

*Identity images* are the relevant roles or commitments that pertain to the responsibility situation in evaluation. The applicability of prescriptions is dependent on the actor’s identity images. In our student example, Beth has expectations that are included with the role of being a student. Identity images can also remove prescriptions. Children are immune to some societal norms and laws as they lack the ability to be fully discerning individuals. When children mature to adults, it is assumed that their increased discernment allows them to shift to an adult identity that would make them more responsible for their behaviors.

Schlenker (1997) describes responsibility as the “psychological glue” (p. 241) that adheres the actor to an event and its consequences. An evaluator of responsibility, therefore, is hypothesized to consider the strengths of the three linkages of the TMR in determining the amount of responsibility. That is, the prescriptions must be relevant to
the event (prescription-event link), the prescriptions must be applicable to the actor’s identity (prescription-identity link), and the actor must be connected to the event that has happened or is anticipated to happen (identity-event link).

When an external audience attempts to judge the amount of responsibility of an actor by “looking down” on the links, the TMR becomes the Accountability Pyramid (see Figure 2). The self is also eligible to be an audience in which an individual perceives the level of his or her own responsibility. Schlenker et al. (1994) empirically tested the use of the TMR and garnered support for two conclusions: (1) weak linkages resulted in lower levels of ascribed responsibility and (2) when provided with a myriad of information, people primarily seek out information regarding the strengths of the linkages amidst irrelevant information.

**Responsibility in Education**

The TMR provides an explanation for how responsibility is manifested in the classroom and how students may determine their response on an outcome measure such as the PRL. When ascribing responsibility in academia, the identity of “college student” invokes images of a person who studies and learns from books and courses. Therefore, a college student is expected to have prescriptions that are appropriate for his or her identity as a student. Such behaviors include reading, writing, and synthesizing of ideas. If an event occurs in which the student performs poorly, such as failing a test, the appropriate prescriptions regarding the event (i.e., reading and studying) will be considered. If a student did not prepare for the exam as he or she should (by a lack of reading and studying), the student will be deemed responsible for this poor learning outcome.
When one considers these facets further, it can be surmised that a student’s perceived responsibility may deviate because of misunderstood student identity images. Perhaps students who view their role of student as a “customer” of the higher education system will have different prescriptions they deem appropriate. These academically entitled students do not expect their own actions to determine their success; rather, they believe that for their attendance, tuition, and completion of assignments they are entitled to good grades (Chowning & Campbell, 2009). Interventionists may be interested in increasing responsibility for learning by correcting false perceptions of student identities or prescriptions.

Surprisingly, given the amount of literature on the theoretical construct of social responsibility, these conceptual underpinnings have not been used to develop measures to assess responsibility for learning. The example of the entitled student in the preceding paragraph is just one instance in which theory may inform validity studies. It is prudent that researchers consider these theories in developing or refining existing measures. Specifically, theoretical research literature may be used to explain and direct the collection of appropriate validity evidence.

**Student Affairs Assessment**

Considering the theoretical foundations of responsibility, as well as some students failing to take responsibility for their academics in higher education, administrators and faculty may be interested in increasing student responsibility in the classroom. As mentioned before, the PRL is currently used to assess a student affairs program on the author’s campus. The purpose of the current study is not to discuss interventions; however, the PRL scores may be used to determine the effectiveness of an intervention or
program. The interpretations of those scores must first be validated. That is, without multiple stages of validity evidence, inferences regarding program effectiveness made from student scores on an invalid measure may not be appropriate. The following section reviews the importance of responsibility programming in student affairs, and the role of assessment of these programs.

**Importance of responsibility programs.** A task force at the National Association of Student Personnel Administrators (NASPA) emphasized that through the collaboration of academic units and student affairs, programs should be designed to increase students’ active participation and responsibility for their own learning (Berson et al., 1998). These collaborative programs, designed to increase student responsibility for learning, should occur at the outset of a student’s college career: during orientation and first-year experience programs. Accordingly, the Council for the Advancement of Standards in Higher Education (CAS; 2008) further emphasizes that orientation programs should inform students of their social, civic, and academic responsibilities.

**Assessment of responsibility programs.** Assessment assists programs in determining if their stated objectives are met. Within higher education, both academic and student affairs units design and implement programs with a myriad of purposes such as increasing knowledge, changing attitudes, or influencing behaviors. Assessment provides useful information about how students change as a result of the program. These results can then be used to make substantive improvements (Palomba & Banta, 1999). Because most programs in higher education are administered annually, assessment can be seen to follow a cyclical pattern—informing changes for the next year.
The assessment cycle is comprised of multiple stages or steps completed in a sequential manner. T. Dary Erwin (1991) advocates for an assessment cycle that consists of five stages: (1) establishing program objectives, (2) selecting or designing assessment methods, (3) collecting assessment information, (4) analyzing assessment information, and (5) reporting and using the assessment information. In order to make appropriate inferences about the effectiveness of the program, scores from a scale need to be validated. That is, this validation process provides evidence that scores from the scale are appropriate for its proposed uses. Relative to the assessment cycle, this thesis pertains to the second step: selection of an assessment method.

With regards to the PRL, this measure has been selected for use in an orientation office that facilitates a first-year program during the summer and welcome week program before classes begin. During the welcome week program, students receive programming regarding the academic nature of college and expectations of a student in the classroom. Student affair practitioners designed this program to clarify the students’ understanding regarding responsibility for their academics, which is an objective of the program. In order to assess student gains in responsibility for learning, the PRL was selected during the second stage of Erwin’s (1991) assessment cycle.
CHAPTER THREE: METHOD

Participants and Procedures

Data from a sample of college students at a southeastern university in the United States were collected on a university-wide “assessment day” conducted prior to beginning of the fall semester. During assessment day, classes are cancelled and students are randomly assigned to testing rooms based on their student identification number. The testing rooms are proctored, standardized directions are read aloud to students, and the assessments are completed using paper forms or electronic survey software. A series of cognitive and motivational measures are administered in each session. The testing session lasts approximately three hours, and proctors encourage cognitive engagement with the measures. All incoming first-year students are required to attend the assessment day before the fall semester begins and participation rate is generally high (i.e., 90%) as students are not allowed to register for courses for the spring semester until the assessments are completed.

Measures

Perceived Responsibility for Learning. The PRL was administered to all students during assessment day. As the scale was originally developed to assess responsibility in high school students, the original wording may be confusing to college students. Wording such as “teacher” and “school” may invoke responses pertaining to secondary education settings and not the measurement of responsibility for learning at the collegiate level. I modified the directions of the PRL so that wording would be consistent with a college population. The modified directions read as follows:

How well college students study and learn in college may be partly due to their college teachers and partly due to their own efforts. For each of the activities
listed below, respond with one of the following numbers indicating who is more responsible: the college teacher or the college student. For example, regarding question number 1 below, if you believe that when a college student is unprepared for a test, the college student is slightly more responsible than the college teacher, respond 5; if you believe the college teacher is definitely more responsible than the college student, respond 2. Select the option between 1 and 7 that best represents your belief.

These new directions were written to focus the respondents’ attention to their perceived beliefs about the responsibility for learning in college students. The response options and item content remained the same as the original scale (see Appendix A).

**Achievement Goal Questionnaire-Revised.** To model to the relationships among responsibility and motivation constructs, participants also completed a goal orientation scale. The Achievement Goal Questionnaire-Revised (AGQ-R; Elliot & Murayama, 2008) consists of four subscales: mastery-approach, mastery-avoidance, performance-approach, and performance-avoidance. The two-by-two framework for academic achievement goals has been empirically supported using a CFA (Finney, Pieper, & Barron, 2004). Correlations among the four factors were low to moderate showing additional discriminant validity for four distinct orientations. The AGQ-R is an appropriate scale to measure student motivation due to its established structure and academic context.

**Data Screening**

A total of 3,833 students provided responses the PRL scale. One hundred eighty-eight (4.9%) cases had one or more missing item responses to the PRL scale; these cases were removed resulting in 3,645 cases. I screened the data to detect missing data and multivariate outliers as well as to assess univariate and multivariate normality. Excessive nonnormality may bias the fit indices and standard errors of parameters in structural
equation models (Finney & DiStefano, 2006). Multivariate outliers were identified using a macro written by DeCarlo (1997). Squared Mahalanobis distances were examined and 11 cases with aberrant response patterns were identified and removed. After these 11 cases were removed, the effective full sample was reduced to $N = 3,634$. In this sample, 62% were female, 81.8% were white non-Hispanic, 6.9% were Asian, 4.4% were African-American, 3.2% were Hispanic, 1.2% were American Indian or Pacific Islander, and 2.5% did not specify an ethnicity. The average age was 18.4 years ($SD = .40$).

Finney and DiStefano (2006) have suggested that variables may be considered univariate normal when skewness values are less than |2| and kurtosis values are less than |7|. The PRL items were relatively univariately normal. In the full sample, the largest skewness and kurtosis values were -2.39 (item 8) and 6.34 (item 9), respectively. Multivariate normality was assessed using Mardia’s normalized kurtosis coefficient. For the full sample PRL data, Mardia’s coefficient was 132.21. No universal cutoff for Mardia’s exists (Finney & DiStefano, 2006) although it has been suggested that a value greater than 3 while using maximum likelihood (ML) estimation may produce inaccurate results (Bentler & Wu, 2003). In this context, a value of 132.21 is extremely high suggesting multivariate nonnormality. Due to the nonnormality of the data, models were estimated with ML using the Satorra-Bentler (S-B) scaled $\chi^2$ and robust standard errors (Satorra & Bentler, 1994). Given satisfactory fit, difference tests between the nested models proposed will be conducted among these scaled $\chi^2$ (Satorra & Bentler, 2001).

**Sample Split**

Researchers may randomly split a large sample to test different factor structures on smaller independent samples. One benefit of this methodology is to observe if model
misfit replicates on an independent sample. If global or local misfit replicates across these random samples, the results can be trusted as a more stable reflection of the observed structure of the PRL rather than an idiosyncrasy of a particular sample (MacCallum, Roznowski, & Necowitz, 1992). Using this reasoning, the full PRL sample ($N = 3,634$) was split after the data screening process. Each case was assigned a number using a random number generator. Those numbers were then sorted and the first third of the dataset was split to form sample A ($n = 1,211$), the second third to form sample B ($n = 1,211$), and the final third to form sample C ($n = 1,212$). Item correlations, means, standard deviations, skew, and kurtosis are presented in Table 1 for the full sample, Table 2 for sample A, Table 3 for sample B, and Table 4 for sample C. All item means are above the mid-point on the scale for the full sample indicating that participants, on average, indicated that students were more responsible for all activities listed in the items. Item standard deviations represent the spread of responses about the mean. Interestingly, some items, such as item 9, have high means (6.6) and a small standard deviation (0.66) meaning that most students agree that the content of this item is the student’s responsibility.

**Testing Hypothesized Structural Models**

Consider the item correlations in Table 1 for the full sample. If a unidimensional solution best explains the relationships among the items, all item correlations should be about the same magnitude. However, the correlation between items 2 and 11 ($r = .57$) and the correlation between items 7 and 9 ($r = .07$) are very different in magnitude. There are a number of possible reasons for this discrepancy. Most relevant to a unidimensional solution, dissimilar item correlations may indicate that the items do not
measure the same construct. Because of the number of dissimilar inter-item correlations, it is likely that the unidimensional model will not fit the data well.

If a three-factor solution should be championed, the correlations among the items within each factor representing the three facets of perceived responsibility for learning (i.e., Motivation, Deportment, and Learning Processes) should be approximately the same magnitude. Additionally, the correlations with other items not in the same factor should be relatively lower. However, this is not the case, as the correlations within each facet are quite different (e.g., \( r = .57 \) for items 2 and 11 and \( r = .26 \) for items 2 and 15).

Recall that in a bifactor solution, researchers specify a general factor for all items as well as additional factors representing variance above and beyond the general factor (Reise, Morizot, & Hays, 2007). If a researcher expects that groups of items share variance after controlling for the general factor, the item correlations should reflect these relationships. If a bifactor model best explains the relationships among the items, all items should have correlations of at least a moderate magnitude; moreover, groups of items in method effect factors should have comparatively larger correlations. That is, after controlling for the variance among all items, items in method effect factors will form bloated specifics beyond the general factor. Additionally, there should be little to no correlation of these method effects with other method effects items as these factors are specified to be orthogonal.

**Planned Data Analyses**

The three hypothesized models (see Figure 1) were fit to sample A and the location of any misfit was noted. Then, these models were tested using sample B to assess the stability of misfit. At this point, I considered modifications to the hypothesized
models as suggested by replicated misfit (e.g., items with high residuals with other items across two samples). If a modified model emerged from these analyses, I planned to test this model on sample C. After establishing a satisfactory factor structure for the PRL, further analyses were planned that would correlate the PRL scores with scores from the AGQ-R.

As mentioned previously, Corno (1992) hypothesized that students with higher perceived responsibility for learning should be more mastery-oriented in regards to their academics. Additionally, she hypothesized that students with low perceived responsibility for learning are less focused on learning and should be more performance-oriented in their achievement goals. Mastery-oriented students make learning the material a goal whereas performance-oriented students are focused on performing well compared to a normative standard (Elliot & Murayama, 2008). It would seem reasonable that those who are focused on mastering classroom material believe that learning is their responsibility. That is, students with a mastery goal orientation are more likely to perceive students as primarily responsible for learning. It would also seem reasonable that students who are focused on performing well may easily find others responsible for their performance, especially when it is poor. To test these hypotheses, structural equation modeling (SEM) could be used to correlate the latent variables of responsibility for learning and goal orientation after partitioning measurement error among items (Weston & Gore, 2006).

Recall that the AGQ-R consists of four subscales: mastery-approach, mastery-avoidance, performance-approach, and performance-avoidance. The approach and avoidance valence concerns a student’s view of the goal as a positive (i.e., approach) or
negative (i.e., avoidance) possibility (Elliot & Murayama, 2008). This difference may be seen in comparing item content. Item 4 is a performance-approach item that states: “my aim is to perform well relative to other students” which is different from item 12, a performance-avoidance item that states, “my aim is to avoid doing worse than other students” (p. 617). Corno (1992) does not address goal orientation valence; however, students who score high on the mastery-approach scale should believe that students are more responsible for their learning resulting in a positive correlation between these latent variables. Additionally, students who score high on the performance-approach scale are hypothesized to perceive others, rather than the student, as responsible for learning; therefore, these scores will negatively correlate with the PRL. Given that these hypotheses address latent relationships with responsibility for learning, SEM will be used to correlate these latent variables given a satisfactory factor structure for responsibility. That is, the “responsibility” factors in models 1 or 3 are most appropriate to correlate with goal orientation variables.
CHAPTER FOUR: RESULTS

Item-Facet Mapping

Zimmerman noted that three facets of responsibility for learning influenced the writing of the PRL items: (a) motivation, (b) deportment, and (c) learning processes (B. J. Zimmerman, personal communication, October 5, 2011). Before the three-factor structure could be tested, items needed to be mapped to facets. A content alignment was performed in which raters not involved in the item writing process mapped items to the domains for which they were purportedly written (Dawis, 1987).

A team of eight raters with training in educational assessment and motivation (including faculty, doctoral, and master’s students) conducted the content alignment among the 18 items and the three facets. Zimmerman and Kitsantas (2005) provide the mapping of three items because they were used as examples to explain the content of each facet; therefore, these items were automatically mapped without the consultation of the eight raters (see Appendix B). These automatically mapped items as well as definitions for each facet were provided to all raters. The criterion used to determine the alignment between an item and a facet was agreement among more than four raters. When no clear pattern emerged regarding the facet matching for a particular item, the faculty ratings broke the tie. This tie break rule was evoked only once, for item 15.

The mappings of the items to facets were then presented to one of the original authors and supported (B. J. Zimmerman, personal communication, October 5, 2011). Items 2, 11, 14, 15, 16, and 17 were mapped to the motivation facet, items 3, 4, 5, 8, 10, 13, and 18 were mapped to the deportment facet, and items 1, 6, 7, 9, and 12 were
mapped to the learning processes facet. The results of the content alignment are found in Appendix B.

**Confirmatory Factor Analysis**

**Global fit indices.** When assessing global model-data fit, Hu and Bentler (1998) explained that $\chi^2$ is influenced by sample size in that true models are more likely to be rejected with larger samples. Additionally, $\chi^2$ is a test of exact model-data fit and is difficult to satisfy. Therefore, Hu and Bentler suggested that additional fit indices provide researchers with a more well-rounded perspective of fit. They advocated for an approach using fit indices that are sensitive to model misspecification in factor correlations (i.e., simple model misspecification) and fit indices that are sensitive to misspecification of factor pattern coefficients (i.e., complex model misspecification). Fit indices may be grouped into two additional categories: absolute and incremental (Hu & Bentler, 1999). Absolute indices represent how well the proposed model reproduces that sample data and incremental fit indices compare model fit to a null model in which all observed variables are uncorrelated.

The comparative fit index (CFI) is an incremental fit index that ranges from 0 to 1 with values closer to 1 indicating better fit compared to a null model. A null model specifies that all factor loadings are fixed to zero. This index is sensitive to complex model misspecification (Hu & Bentler, 1998). Complex model misspecification occurs when the model specifies factor loadings as zero when they should be freely estimated.

The standardized root mean square residual (SRMR) is an absolute fit index that is the mean of the absolute values of the correlation residuals. SRMR may range from 0 to 1 with values closer to zero indicating better fit. Hu and Bentler (1998) found that the
SRMR is very sensitive to simple model misspecification meaning that the SRMR will be larger when factor correlations that are set to be zero should be freely estimated.

Similar to SRMR, the root mean square error of approximation (RMSEA) is another absolute fit index with values closer to zero indicate better fit. It differs from the SRMR in that it is more sensitive to complex model misspecification (Hu & Bentler, 1998). When data have high multivariate nonnormality, Yu and Muthén (2002) suggested cutoffs of the robust CFI ≥ .95, robust RMSEA ≤ .05, and SRMR ≤ .07 when N ≥ 250.

**Local model misfit.** Recently, some authors have cautioned against dichotomous decisions regarding global fit indices (Marsh, Hau, & Wen, 2004). Instead, researcher should use fit indices to note areas in which the model does not fit the data suggesting further exploration of why misfit is occurring. Considering this position, I will use the suggested cutoffs for the fit indices with caution and will place a greater emphasis on the standardized and unstandardized residuals when assessing model misfit. Covariance residuals can be assessed to identify local misfit. If the residual between two items is greater than |3|, the model is not reproducing the relationship between those items (Byrne, 1998). Covariance residuals are on a z-score metric and test if the residual is statistically significantly different than zero. A small residual may be statistically significant with a large sample size; therefore, some researchers suggest using correlation residuals as they are not affected by a large sample (Kline, 2011). Correlation residuals between two items greater than |.1| are indicative of local misfit (Kline, 2011).

**Tests of hypothesized models.** Data were submitted to PRELIS 2.80 to generate a covariance matrix analyzed using LISREL 8.80 (Jöreskog & Sörbom, 2006).
Additionally, an asymptotic covariance matrix was generated for each sample to adjust the Chi-square and standard errors for multivariate nonnormality. Global fit indices for the three hypothesized models are presented in Table 5 for sample A and sample B.

The one-factor model resulted in poor fit for both samples; moreover, this model poorly reproduced the relationships among the items with 25 and 20 correlation residuals larger than |.10| in sample A (16% of all inter-item correlations) and sample B (13%), respectively. The unidimensional model proposed by Zimmerman and Kitsantas (2005) was not supported by this study.

The three-factor model did not converge to an admissible solution in either sample. Specifically, the deportment and learning processes factors correlated above 1. Implausible factor correlations over 1 may occur when the specified model is grossly wrong (Rindskopf, 1984). This result does not bode well for the three-factor model.

Consider the correlation residuals for the one-factor model tested using sample A presented in Table 6. These residuals remain after modeling variance shared among all items. In the bifactor model, this is modeled by the general Responsibility factor (see Figure 1). If the bifactor model best explains the relationships among the items, residuals for the one-factor model should remain within the groups of items for each method effects factor. However, this is not true as the correlation residuals seem “scattered” meaning certain items correlate above and beyond the general factor. This pattern is not captured by the proposed bifactor model. Researchers can assume that the global and local fit of the bifactor model will likely be poor.

The bifactor model converged with very similar global fit on both samples. I calculated Satorra-Bentler scaled Chi-square difference tests between the one-factor and
bifactor model in both samples (Satorra & Bentler, 2001). This difference was statistically significant for sample A, $\Delta \chi^2_{S-B} (18) = 490.30, p < .001$ and for sample B, $\Delta \chi^2_{S-B} (18) = 405.16, p < .001$ indicating that the items are likely multidimensional as the bifactor model specifies. Although the bifactor model explains the relationships among the items better than the unidimensional model, global fit indices are still poor with many large correlation residuals.

**Diagnosing Misfit of the Bifactor Model**

Correlation residuals were examined to diagnose areas of local model misfit. A total of 14 correlation residuals were larger than $|.10|$ in sample A (9%) and 13 were larger than $|.10|$ in sample B (8.5%). See Table 7 for correlation residuals for the bifactor model using sample A. Across both samples, the largest correlation residuals for the bifactor model occurred between items 2 and 11 (correlation residual = .313) and items 8 and 9 (correlation residual = .204). Perhaps an idiosyncrasy of the sample data, the correlation residual between items 5 and 6 in sample B was much smaller (i.e., 0.121) than in sample A (i.e., .238) although still notable. The largest residuals are positive indicating that the model is underestimating the relationships between those items.

The largest residual was between items 2 (“Who is responsible for a student being motivated to learn in school?”) and 11 (“Who is responsible for a student being interested in school?”). Given both items represent the same method effect, the large residual in the bifactor model indicates that those items share variance above the general Responsibility factor and the method factor A. Interestingly, the content of items 2 and 11 perhaps elicit school outcomes that are shared between the student and the teacher unlike items 14, 15, 16, and 17 which seem to elicit outcomes or valuations that are within the student’s
influence. When examining the item descriptives, items 2 and 11 have similar means and distributions, closer to the “both equally” response than the student end of the scale (i.e., items 14, 15, 16, and 17) further supporting this hypothesis. Perhaps misfit is occurring because students conceptualize these items differently.

When considering the other correlation residuals in Table 7, one can see that some items still share variance (i.e., positive residuals) after controlling for both the Responsibility factor and the method effects factor (e.g., items 2 and 11 for method effect A, items 3 and 4 for method effect B). This means that the bifactor model is failing to account for the relationships between some items. Additionally, some residuals are negative indicating that the model is overestimating the relationships between some items. These poorly reproduced relationships should be noted in further studies to explore how students are responding to this scale and if the items are functioning appropriately.
CHAPTER FIVE: DISCUSSION

The purpose of this study was to test hypothesized factor structure models of the PRL scale. A unidimensional solution, which was implied by Zimmerman and Kitsantas (2005), was not empirically supported because of poor global and local fit. Of the models tested, a bifactor solution demonstrated the best fit to the data; however, global fit was not satisfactory and local misfit was observed. There are several hypotheses that explain why the model did not fit. One might initially surmise that gross model misspecification was the cause of poor model fit; however, future researchers should consider other potential issues with the PRL.

Theoretical and Methodological Concerns

Zimmerman and Kitsantas (2005) did not use an established theory to create the PRL items. In scale development, researchers are encouraged to create scales that are grounded in sound theory; Benson’s (1998) substantive stage addresses these concerns. When test developers do not address aspects of a theory, it is difficult to conclude why students provide each response. If we use a theoretical model to interpret the results, we find that students may provide disparate responses for two items purporting to measure the same construct.

One theory that we may use is the Triangle Model of Responsibility (TMR) proposed by Schlenker et al. (1994). In this theory, students are to consider the strengths of the linkages between the relevant prescriptions, identity images, and events when ascribing the level of responsibility. Consider the student example I used when describing this theory. Beth, a student, has been told on Monday to study for a test on Friday. Let us say that she studies for the exam for a reasonable amount of time before
Friday yet she was still unprepared and performs poorly. Item 1 in the PRL (“Who is responsible for a student being unprepared for a test?”) represents this scenario. According to the TMR theory, Beth fulfilled her responsibilities (i.e., strong linkages within the TMR). However, it may be that the teacher did not fulfill his or her responsibility to align the test with the material taught in class or the study guide. Or perhaps the tests and the material were properly aligned but the teacher failed to cover all of the material. Responsibility in this scenario may be perceived as shared because a poor grade on an exam is not entirely a lack of responsibility on the student’s behalf.

Items with similar content (i.e., shared responsibility scenarios) are a potential issue with the PRL.

Alternatively, consider how item 9 (“Who is responsible for a student not taking notes in class?”) is a clear example of a situation that is completely in the student’s control. Students completing this scale should all conceptualize this item as solely the student’s responsibility according to the TMR. The student’s prescriptions are clear (i.e., notes should be taken), students are known to take notes (i.e., identity image), thus, if a student does not take notes (i.e., event) then the student is considered responsible. Clearly, students’ responses will differ across two items when one item represents shared responsibility and the other solely student responsibility if the TMR is used for both.

Recall, Zimmerman and Kitsantas (2005) believed the content of the PRL items was relatively similar and thus supposed that the scale was unidimensional. If a scale is to measure responsibility for learning as a unidimensional construct, all items must be conceptualized similarly, thus manifesting in similar responses. When students
conceptualize items differently, their responses will be dissimilar and a unidimensional
model will result in poor fit with the data.

In addition to conceptual issues, negative wording may contribute to model misfit. PRL items 1, 3, 4, 7, 8, 9, 13, 14, 16, and 18 all contain the word “not” or have a negative valence. Negative wording is problematic when assessing the factor structure of a scale. According to Podsakoff, MacKenzie, Lee, and Podsakoff (2003), negative items may form artificial factors. These factors can form when as few as 10% of the sample do not recognize the wording change (i.e., from positive to negative). In the current study, it is unlikely that a negative method factor will help to address model misfit for the one-factor or bifactor model as the negative valence items are not the only items with residuals. However, in future studies, it may be wise for researchers to avoid negative wording in the items to prevent student confusion.

Zimmerman and Kitsantas (2005) did not indicate that any of the items should be reversed-coded when scoring the PRL. Perhaps these items were not initially considered to be conceptualized negatively. Zimmerman and Kitsantas may have included negative words with the intention that students conceptualize these scenarios as the student’s responsibility; however, students may be confused by the negative valence. Consider item 9 once again (“Who is responsible for a student not taking notes in class?”). When one removes the “not,” students should still consider the student in the scenario as the responsible agent. That is, if students are asked “how responsible is a student for taking notes in class” they may then rate the level of perceived responsibility for the student in that scenario. Notice the shift of focus from “who is responsible” to “how responsible is a student.” I suggest a focus on the student solely because the purpose of the PRL is to
measure student perceived responsibility for learning not the perceived responsibility of the teacher or professor.

**Future Directions for Measuring Responsibility for Learning**

Clearly, a scale that measures perceived responsibility for learning is needed in higher education assessment; however, strong validity evidence supporting any scale is elusive. The current study has leant some insight into the performance of the PRL, yet more work is needed. It seems that there are two possible directions for the furtherance of this work: (a) the PRL could be revised to address some of the aforementioned issues or (b) a new scale could be designed based on responsibility theory. The researcher may consider the cost of each direction.

**Suggested revisions for the PRL.** Future development of the PRL should address issues pertaining to item content and the response scale. Specifically, researchers may want to include items that represent academic outcomes in which the student is solely responsible, the teacher is solely responsible, and scenarios in which responsibility is shared. Most items in the current PRL represent these three categories. Some items in the PRL are clearly the student’s responsibility (e.g., “Who is responsible for fooling around in class?”) yet more items should be written to be the teacher’s responsibility. Responsibility theory may help guide future researchers to determine the academic outcomes that are the student’s responsibility. When writing these items, researchers are also advised to avoid including negatively worded items as negative wording factors may form that will add noise to future structural studies, although this was not an issue in the current study.
Response scale concerns. Two new response scales are proposed for the PRL. The current response scale presupposes that responsibility is always on a continuum between the student and the teacher with an equal level of responsibility mid-scale. It is unclear if students are selecting “both equally” because they truly believe the item reflects a situation in which responsibility should be equally shared or if they are averaging the level of responsibility between teacher and student (i.e., high responsibility for student, high responsibility for teacher). Perhaps students may rate responsibility as “high” or “low” for both the teacher and the student depending on the outcome. Without the ability to provide responses to both aspects, students respond with “both equally.” Additionally, both the student and teacher may not be responsible for certain academic outcomes. With the current scale, students are unable to indicate when neither the teacher nor the student is responsible.

For these reasons, I suggest that researchers provide a scale that allows students to select the degree to which they perceive the student is responsible and one to which they perceive the teacher to be responsible for one item. For example, students may be provided with a 7-point Likert scale from “Not At All” to “Completely” to rate the student’s level of responsibility and a second scale to rate the teacher’s responsibility. This would eliminate the issue of respondents becoming potentially confused on how to respond when considering one scale for each scenario.

Suggested analyses for the revised PRL. With these modifications, new data may be collected for analysis. Future researchers should first test a unidimensional factor structure. Zimmerman and Kitsantas (2005) supposed that the PRL was unidimensional however indicating that they used three types of items: motivation, deportment, and
learning processes. Additionally, the TMR dictates that, as long as the linkages among
the prescriptions, identity images, and event are strong, responsibility should be ascribed
similarly. If a unidimensional model does not fit, and due to the lack of theory to suggest
a factor structure, perhaps an exploratory methodology is appropriate. Researchers may
attempt to unearth a factor structure for the PRL using an EFA. It is suggested that this
new factor structure be replicated on an independent sample to reduce the possibility of
capitalizing on chance variations in one sample (Bandalos & Finney, 2010). However, if
these methods still prove unsuccessful or seem arduous for the researcher, perhaps a new
scale may be selected from the literature. Researchers should be aware that validity
evidence for this new scale may be limited and therefore needed before the scale may be
used for assessment purposes.

**Designing a New Scale**

Assessment practitioners may not be interested in revising the PRL or selecting a
new scale due to the lack of a theoretical foundation for the items. Future researchers
should consider responsibility theory, such as the TMR, when writing items. I do not
suggest that researchers write items to assess each component of the TMR theory as this
is the process by which students determine their responsibility. I do suggest that
researchers split the response scale to assess the degree to which students perceive each
scenario as the student’s responsibility or the teacher’s responsibility. Most importantly,
scores from the student response scale may provide researchers with rich information
about student’s perceived responsibility for learning.

Researchers should not edit the items to be scenarios in which student responses
have no variability (i.e., most answer that the student is responsible) because these
scenarios provide researchers with limited information. For example, the largest item mean and smallest standard deviation in the current study was item 9 of the PRL (“Who is responsible for a student not taking notes in class?”). Although item 9 is a good item by itself, researchers are cautioned against providing too many items that may cause a restriction of range.

The suggestions I made regarding the item content and response scale format for the revised PRL should be applied to this new scale. Benson (1998) suggests that, during the substantive stage, one must “ensure the operational definition (specific set of items and response format) adequately reflects all the aspects of the theoretical domain of the construct” (p. 13). Scale response format should be developed in tandem with new items. However, some of the scenarios in the current PRL items may be used with the new response scales. Again, students may respond to items twice using a scale that assess the perceived responsibility of the student and another scale that assess the perceived responsibility of the teacher.

After researchers complete the substantive stage for the new scale, further validity evidence should be collected in Benson’s (1998) structural and external stages. The theory used to write items for the new responsibility for learning scale should dictate the expected factor structure. Researchers may test this hypothesized factor structure using a CFA. After a sufficient factor structure is achieved, the external stage collects evidence regarding the external relationships with the purported construct. One such empirical analysis may be a comparison of scores between two groups known to differ on the construct. For example, one might consider seniors in college to be more responsible for their academics than freshman (Schmelzer, Schmelzer, Figler, & Brozo, 1987).
Therefore, seniors in college should have scores indicative of higher student perceived responsibility for learning than freshman on the new scale.

**Limitations of the Current Study**

Several limitations are apparent in the current study. Recall that eight people participated in a content alignment whereby they determined the item-facet mapping for the PRL scale. The primary concern with this procedure is the ability of motivation students and faculty to make the appropriate mappings. Perhaps the true mapping of the scale items could be accomplished with experts in responsibility literature. Moreover, the content alignment procedure was performed assuming that the three facets mentioned by Zimmerman and Kitsantas (2005) were the only ways that the items may be grouped together.

Related to the concerns with the item-facet mapping, it can only be speculated what each of the method effects in the bifactor model represent after controlling for the general factor of responsibility. Although this model was the one most explored in the current study, it should be noted that this was simply because the fit indices were closest to the desired values. Recall that a more complex model will fit the data better than a simple model, but this does not mean that it is the true model. The fit indices for the bifactor model were still poor with many correlation residuals.

The directions of the PRL may be another concern relative to the sample. Recall that the directions for the PRL were changed to address college students in a college classroom. The students who completed the PRL in the current study were entering college freshman who had not taken a college course. Without experience in the college classroom, these students are either considering the responsibility of high school students
and teachers or how they imagine a college student would be responsible when completing the PRL. With issues pertaining to the validity of the content alignment and the sample, Light, Singer, and Willett (1990) advise researchers in this situation that “you can’t fix by analysis what you bungled by design” (p. viii).

**Summary and Concluding Remarks**

Unfortunately, validity evidence for the PRL is limited. The structural validity evidence collected in the current study does not support the suggested factor structure of the PRL. Future researchers and assessment practitioners should be aware that the PRL item responses may not be summed or averaged to provide a single score. Students may conceptualize some of the PRL items differently leading to the hypothesis that the lack of validity evidence in the substantive stage explains disparate item responses.

Zimmerman and Kitsantas (2005) should be commended for creating the PRL as the construct is clearly important. Responsibility for learning is considered by Davis and Murrell (2003) to be foundational to other constructs learned in college. That is, if students increase in responsibility for learning they may become more successful in other academics endeavors. This drives university officials to design interventions to increase responsibility for learning. In order to assess the effectiveness these interventions, we need a scale with valid interpretations of scores.

Future research in this area may provide exciting new developments to better understand the manifestation of responsibility in academia. Student success is a primary goal of higher education. Therefore, it is our responsibility as researchers in higher education to design and develop the best ways to teach our students to be more
responsible for their learning. This new research should make better measurement of responsibility for learning a primary goal.
Appendix A

Perceived Responsibility for Learning Scale

How well students study and learn in school may be partly due to their teacher and partly due to their own efforts. Next to each of the activities listed below, write one of the following numbers indicating who is more responsible: the teacher or the student. For example, regarding question number 1 below, if you believe that when a student is unprepared for a test, the student is slightly more responsible than the teacher, put a 5 in the space next to the question.

<table>
<thead>
<tr>
<th>The Teacher</th>
<th>Both Equally</th>
<th>The Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>mainly 1</td>
<td>definitely more 2</td>
<td>slightly more 3</td>
</tr>
</tbody>
</table>

Who is more responsible:

1. for a student being unprepared for a test?
2. for a student being motivated to learn in school?
3. for a student not finishing homework assignments?
4. for a student being unprepared to participate in class?
5. for a student writing assigned papers well?
6. for a student understanding assigned homework readings?
7. for a student not understanding a class discussion?
8. for a student fooling around in class?
9. for a student not taking notes in class?
10. for a student doing homework assignments correctly?
11. for a student being interested in school?
12. for a student remembering information from assigned readings?
13. for a student not concentrating in class?
14. for a student not valuing good grades in school?
15. for a student giving extra effort when needed?
16. for a student just going through the motions without really trying in class?
17. for a student seeing school as important to his or her future success?
18. for a student receiving poor grades in school?
Appendix B

Mapping of Perceived Responsibility for Learning Scale Items to Facets

M = Motivation; D = Deportment; L = Learning Processes

<table>
<thead>
<tr>
<th>Facet</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>1. for a student being unprepared for a test?</td>
</tr>
<tr>
<td>M</td>
<td>2. for a student being motivated to learn in school?</td>
</tr>
<tr>
<td>D</td>
<td>3. for a student not finishing homework assignments?</td>
</tr>
<tr>
<td>D</td>
<td>4. for a student being unprepared to participate in class?</td>
</tr>
<tr>
<td>D</td>
<td>5. for a student writing assigned papers well?</td>
</tr>
<tr>
<td>L</td>
<td>6. for a student understanding assigned homework readings?</td>
</tr>
<tr>
<td>L</td>
<td>7. for a student not understanding a class discussion?</td>
</tr>
<tr>
<td>D*</td>
<td>8. for a student fooling around in class?</td>
</tr>
<tr>
<td>L*</td>
<td>9. for a student not taking notes in class?</td>
</tr>
<tr>
<td>D</td>
<td>10. for a student doing homework assignments correctly?</td>
</tr>
<tr>
<td>M</td>
<td>11. for a student being interested in school?</td>
</tr>
<tr>
<td>L</td>
<td>12. for a student remembering information from assigned readings?</td>
</tr>
<tr>
<td>D</td>
<td>13. for a student not concentrating in class?</td>
</tr>
<tr>
<td>M</td>
<td>14. for a student not valuing good grades in school?</td>
</tr>
<tr>
<td>M</td>
<td>15. for a student giving extra effort when needed?</td>
</tr>
<tr>
<td>M*</td>
<td>16. for a student just going through the motions without really trying in class?</td>
</tr>
<tr>
<td>M</td>
<td>17. for a student seeing school as important to his or her future success?</td>
</tr>
<tr>
<td>D</td>
<td>18. for a student receiving poor grades in school?</td>
</tr>
</tbody>
</table>

Note. Items denoted with an asterisk (*) were given as examples in Zimmerman & Kitsantas (2005) and thus were not mapped by the raters.
### Table 1. Correlation Matrix and Descriptive Statistics for Full Sample

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>11</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>8</th>
<th>10</th>
<th>13</th>
<th>18</th>
<th>1</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>.570</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>.273</td>
<td>.293</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>.257</td>
<td>.265</td>
<td>.484</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>.274</td>
<td>.316</td>
<td>.447</td>
<td>.475</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>.302</td>
<td>.338</td>
<td>.445</td>
<td>.396</td>
<td>.410</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.251</td>
<td>.176</td>
<td>.345</td>
<td>.372</td>
<td>.322</td>
<td>.208</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.259</td>
<td>.194</td>
<td>.329</td>
<td>.338</td>
<td>.307</td>
<td>.224</td>
<td>.488</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.268</td>
<td>.379</td>
<td>.244</td>
<td>.244</td>
<td>.230</td>
<td>.244</td>
<td>.251</td>
<td>.263</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.163</td>
<td>.170</td>
<td>.349</td>
<td>.311</td>
<td>.325</td>
<td>.251</td>
<td>.219</td>
<td>.298</td>
<td>.183</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>.235</td>
<td>.267</td>
<td>.182</td>
<td>.238</td>
<td>.250</td>
<td>.205</td>
<td>.230</td>
<td>.227</td>
<td>.379</td>
<td>.186</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>.331</td>
<td>.367</td>
<td>.415</td>
<td>.394</td>
<td>.455</td>
<td>.313</td>
<td>.351</td>
<td>.333</td>
<td>.257</td>
<td>.342</td>
<td>.286</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>.214</td>
<td>.270</td>
<td>.297</td>
<td>.296</td>
<td>.324</td>
<td>.260</td>
<td>.317</td>
<td>.275</td>
<td>.284</td>
<td>.177</td>
<td>.324</td>
<td>.365</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.251</td>
<td>.224</td>
<td>.179</td>
<td>.171</td>
<td>.217</td>
<td>.165</td>
<td>.288</td>
<td>.288</td>
<td>.259</td>
<td>.164</td>
<td>.247</td>
<td>.239</td>
<td>.348</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.289</td>
<td>.282</td>
<td>.138</td>
<td>.134</td>
<td>.175</td>
<td>.166</td>
<td>.121</td>
<td>.155</td>
<td>.243</td>
<td>.104</td>
<td>.309</td>
<td>.236</td>
<td>.268</td>
<td>.275</td>
<td>.485</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SD**: 1.337 1.310 0.809 0.822 0.993 1.089 0.625 0.818 1.023 0.833 1.076 1.022 0.924 0.961 1.264 1.221 0.659 0.930  
**Skew**: -.0212 -.0182 -1.809 -1.845 -1.336 -1.109 -1.577 -1.300 -.339 -.291 -.788 -.346 -.135 -.074 .0193 -2.330 -.798  
**Kurtosis**: -.4755 -.2893 3.6074 4.4591 1.6540 0.4873 3.0832 2.0980 -.3021 6.1746 -.2510 0.1540 -0.3560 -0.4970 -.0830 0.1360 6.3420 0.6940  

*Note.* Correlations are presented in groups based on the content alignment results (see Appendix B). Presenting the correlations in this way allows for readers to easily note where misfit may be observed in the hypothesized models.
Table 2. Correlation Matrix and Descriptive Statistics for Sample A

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>11</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>8</th>
<th>10</th>
<th>13</th>
<th>18</th>
<th>1</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>.583</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>.293</td>
<td>.300</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>.252</td>
<td>.245</td>
<td>.435</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>.254</td>
<td>.323</td>
<td>.437</td>
<td>.439</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>.275</td>
<td>.338</td>
<td>.427</td>
<td>.380</td>
<td>.429</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.248</td>
<td>.181</td>
<td>.395</td>
<td>.386</td>
<td>.326</td>
<td>.231</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.228</td>
<td>.230</td>
<td>.337</td>
<td>.334</td>
<td>.310</td>
<td>.237</td>
<td>.488</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.250</td>
<td>.379</td>
<td>.231</td>
<td>.265</td>
<td>.201</td>
<td>.258</td>
<td>.272</td>
<td>.287</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.184</td>
<td>.186</td>
<td>.346</td>
<td>.264</td>
<td>.316</td>
<td>.259</td>
<td>.296</td>
<td>.320</td>
<td>.169</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>.241</td>
<td>.266</td>
<td>.183</td>
<td>.239</td>
<td>.241</td>
<td>.243</td>
<td>.229</td>
<td>.231</td>
<td>.379</td>
<td>.179</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>.352</td>
<td>.403</td>
<td>.425</td>
<td>.362</td>
<td>.443</td>
<td>.278</td>
<td>.391</td>
<td>.354</td>
<td>.279</td>
<td>.370</td>
<td>.310</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>.187</td>
<td>.264</td>
<td>.275</td>
<td>.275</td>
<td>.314</td>
<td>.260</td>
<td>.316</td>
<td>.298</td>
<td>.237</td>
<td>.170</td>
<td>.310</td>
<td>.376</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.261</td>
<td>.234</td>
<td>.182</td>
<td>.139</td>
<td>.183</td>
<td>.132</td>
<td>.287</td>
<td>.281</td>
<td>.259</td>
<td>.177</td>
<td>.239</td>
<td>.233</td>
<td>.339</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.287</td>
<td>.266</td>
<td>.135</td>
<td>.099</td>
<td>.169</td>
<td>.189</td>
<td>.174</td>
<td>.204</td>
<td>.440</td>
<td>.095</td>
<td>.332</td>
<td>.203</td>
<td>.283</td>
<td>.311</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.301</td>
<td>.302</td>
<td>.132</td>
<td>.114</td>
<td>.166</td>
<td>.180</td>
<td>.170</td>
<td>.141</td>
<td>.283</td>
<td>.080</td>
<td>.318</td>
<td>.263</td>
<td>.295</td>
<td>.304</td>
<td>.499</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>.216</td>
<td>.201</td>
<td>.394</td>
<td>.325</td>
<td>.351</td>
<td>.234</td>
<td>.439</td>
<td>.413</td>
<td>.204</td>
<td>.502</td>
<td>.227</td>
<td>.425</td>
<td>.229</td>
<td>.213</td>
<td>.083</td>
<td>.061</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

SD  1.364  1.332  0.808  0.855  0.988  1.080  0.605  0.813  1.051  0.847  1.112  1.014  0.922  0.955  1.225  1.192  0.633  0.932
Skew -0.258 -0.182 -1.677 -1.979 -1.280 -1.088 -1.567 -1.126 -0.342 -2.372 -0.337 -0.760 -0.286 -0.073 -0.146 -0.170 -2.334 -0.803
Kurtosis -0.442 -0.351 2.595 5.467 1.387 0.577 2.488 0.893 -0.366 6.098 -0.069 0.037 -0.563 -0.604 -0.039 0.208 6.634 0.560

Note. Correlations are presented in groups based on the content alignment results (see Appendix B). Presenting the correlations in this way allows for readers to easily note where misfit may be observed in the hypothesized models.
Table 3. Correlation Matrix and Descriptive Statistics for Sample B

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>11</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>8</th>
<th>10</th>
<th>13</th>
<th>18</th>
<th>1</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.566</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.266</td>
<td>0.291</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.258</td>
<td>0.264</td>
<td>0.473</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.281</td>
<td>0.322</td>
<td>0.430</td>
<td>0.470</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.343</td>
<td>0.341</td>
<td>0.448</td>
<td>0.410</td>
<td>0.382</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.244</td>
<td>0.169</td>
<td>0.293</td>
<td>0.326</td>
<td>0.296</td>
<td>0.159</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.304</td>
<td>0.176</td>
<td>0.316</td>
<td>0.335</td>
<td>0.299</td>
<td>0.204</td>
<td>0.476</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.271</td>
<td>0.380</td>
<td>0.238</td>
<td>0.240</td>
<td>0.249</td>
<td>0.206</td>
<td>0.268</td>
<td>0.272</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.165</td>
<td>0.135</td>
<td>0.341</td>
<td>0.322</td>
<td>0.308</td>
<td>0.236</td>
<td>0.310</td>
<td>0.275</td>
<td>0.196</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.237</td>
<td>0.269</td>
<td>0.147</td>
<td>0.219</td>
<td>0.220</td>
<td>0.173</td>
<td>0.260</td>
<td>0.207</td>
<td>0.380</td>
<td>0.154</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.303</td>
<td>0.342</td>
<td>0.425</td>
<td>0.413</td>
<td>0.447</td>
<td>0.334</td>
<td>0.291</td>
<td>0.293</td>
<td>0.250</td>
<td>0.315</td>
<td>0.267</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.211</td>
<td>0.287</td>
<td>0.303</td>
<td>0.303</td>
<td>0.339</td>
<td>0.253</td>
<td>0.298</td>
<td>0.239</td>
<td>0.327</td>
<td>0.175</td>
<td>0.348</td>
<td>0.353</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.230</td>
<td>0.173</td>
<td>0.162</td>
<td>0.166</td>
<td>0.218</td>
<td>0.161</td>
<td>0.300</td>
<td>0.291</td>
<td>0.224</td>
<td>0.145</td>
<td>0.247</td>
<td>0.218</td>
<td>0.313</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.361</td>
<td>0.332</td>
<td>0.158</td>
<td>0.225</td>
<td>0.207</td>
<td>0.195</td>
<td>0.202</td>
<td>0.200</td>
<td>0.356</td>
<td>0.115</td>
<td>0.378</td>
<td>0.245</td>
<td>0.273</td>
<td>0.236</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.246</td>
<td>0.247</td>
<td>0.146</td>
<td>0.201</td>
<td>0.188</td>
<td>0.174</td>
<td>0.129</td>
<td>0.183</td>
<td>0.194</td>
<td>0.090</td>
<td>0.291</td>
<td>0.227</td>
<td>0.248</td>
<td>0.253</td>
<td>0.480</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.183</td>
<td>0.144</td>
<td>0.399</td>
<td>0.365</td>
<td>0.325</td>
<td>0.234</td>
<td>0.403</td>
<td>0.375</td>
<td>0.246</td>
<td>0.509</td>
<td>0.235</td>
<td>0.336</td>
<td>0.261</td>
<td>0.168</td>
<td>0.155</td>
<td>0.078</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.315</td>
<td>0.408</td>
<td>0.331</td>
<td>0.382</td>
<td>0.336</td>
<td>0.220</td>
<td>0.344</td>
<td>0.323</td>
<td>0.306</td>
<td>0.225</td>
<td>0.358</td>
<td>0.349</td>
<td>0.345</td>
<td>0.257</td>
<td>0.295</td>
<td>0.247</td>
<td>0.322</td>
<td>1</td>
</tr>
</tbody>
</table>

SD 1.327 1.288 0.801 0.785 0.991 1.099 0.636 0.828 1.012 0.834 1.045 1.006 0.918 0.960 1.269 1.180 0.659 0.924
Skew -0.164 -0.155 -1.680 -1.614 -1.319 -1.112 -1.408 -1.373 -0.385 -2.526 -0.321 -0.776 -0.395 -0.143 -0.206 0.180 -2.196 -0.819
Kurtosis -0.490 -0.320 3.020 3.156 1.739 0.431 2.292 2.776 -0.151 7.460 -0.057 0.138 -0.263 -0.454 -0.023 0.286 5.391 0.985

**Note.** Correlations are presented in groups based on the content alignment results (see Appendix B). Presenting the correlations in this way allows for readers to easily note where misfit may be observed in the hypothesized models.
Table 4. Correlation Matrix and Descriptive Statistics for Sample C

|     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | 11    | 12    | 13    | 14    | 15    | 16    | 17    | 18    | 19    | 20    | 21    |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 2   | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 11  | .561  | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 14  | .258  | .288  | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 15  | .262  | .287  | .543  | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 16  | .288  | .305  | .474  | .516  | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 17  | .287  | .337  | .460  | .398  | .420  | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 3   | .260  | .178  | .349  | .405  | .344  | .237  | 1     |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 4   | .243  | .176  | .334  | .346  | .312  | .230  | .500  | 1     |       |       |       |       |       |       |       |       |       |       |       |       |
| 5   | .285  | .379  | .263  | .239  | .268  | .218  | .227  | 1     |       |       |       |       |       |       |       |       |       |       |       |       |
| 8   | .140  | .188  | .359  | .351  | .352  | .259  | .351  | .301  | .183  | 1     |       |       |       |       |       |       |       |       |       |       |
| 10  | .226  | .267  | .213  | .254  | .290  | .198  | .205  | .244  | .379  | .224  | 1     |       |       |       |       |       |       |       |       |       |
| 18  | .244  | .262  | .311  | .312  | .320  | .266  | .339  | .288  | .288  | .187  | .316  | .365  | 1     |       |       |       |       |       |       |       |
| 1   | .262  | .265  | .193  | .207  | .248  | .200  | .281  | .291  | .291  | .169  | .254  | .263  | .389  | 1     |       |       |       |       |       |       |
| 6   | .332  | .314  | .197  | .199  | .247  | .220  | .174  | .239  | .419  | .155  | .424  | .249  | .305  | .313  | 1     |       |       |       |       |       |
| 7   | .317  | .297  | .135  | .092  | .171  | .144  | .070  | .140  | .248  | .139  | .317  | .219  | .262  | .267  | .476  | 1     |       |       |       |       |
| 9   | .184  | .194  | .409  | .405  | .342  | .256  | .457  | .378  | .244  | .527  | .266  | .341  | .305  | .205  | .192  | .062  | 1     |       |       |

SD: 1.318 1.312 0.818 0.825 1.002 1.086 0.632 0.812 1.002 0.817 1.071 1.047 0.933 0.966 1.297 1.287 0.683 0.935
Skew: -0.212 -0.208 -2.063 -1.887 -1.414 -1.132 -1.758 -1.402 -0.275 -2.269 -0.211 -0.828 -0.360 -0.192 -0.176 0.200 -2.441 -0.778
Kurtosis: -0.490 -0.194 5.178 4.281 1.860 0.475 4.428 2.614 -0.420 4.989 -0.661 0.276 -0.230 -0.416 -0.186 -0.074 6.902 0.567

Note. Correlations are presented in groups based on the content alignment results (see Appendix B). Presenting the correlations in this way allows for readers to easily note where misfit may be observed in the hypothesized models.
Table 5. Fit Indices of Three Hypothesized Models of Perceived Responsibility for Learning

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2_{ML}$</th>
<th>$\chi^2_{S-B}$</th>
<th>df</th>
<th>RMSEA$_{S-B}$</th>
<th>SRMR</th>
<th>CFI$_{S-B}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. One-Factor</td>
<td>1717.67</td>
<td>1403.48</td>
<td>135</td>
<td>0.10</td>
<td>0.08</td>
<td>0.87</td>
</tr>
<tr>
<td>2. Three-Factor</td>
<td>Did not converge to an admissible solution.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Bifactor</td>
<td>1088.62</td>
<td>895.67</td>
<td>117</td>
<td>0.08</td>
<td>0.06</td>
<td>0.93</td>
</tr>
<tr>
<td>Sample B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. One-Factor</td>
<td>1619.86</td>
<td>1256.72</td>
<td>135</td>
<td>0.09</td>
<td>0.07</td>
<td>0.87</td>
</tr>
<tr>
<td>2. Three-Factor</td>
<td>Did not converge to an admissible solution.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Bifactor</td>
<td>1048.45</td>
<td>825.37</td>
<td>117</td>
<td>0.08</td>
<td>0.06</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Note. $\chi^2_{ML}$ = Maximum Likelihood Chi-square. $\chi^2_{S-B}$ = Satorra-Bentler scaled Chi-square. RMSEA$_{S-B}$ = robust root mean square error of approximation. SRMR = standardized root mean square residual. CFI$_{S-B}$ = robust comparative fit index. The bifactor fit statically significantly better than the one-factor model in sample A, $\Delta \chi^2_{S-B}$ (18) = 490.30, $p < .001$ and sample B, $\Delta \chi^2_{S-B}$ (18) = 405.16, $p < .001$. $n = 1,211$ for sample A and sample B.
Table 6. One-Factor Model Correlation Residuals for Sample A

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>11</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>8</th>
<th>10</th>
<th>13</th>
<th>18</th>
<th>1</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td><strong>0.296</strong></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-0.025</td>
<td>-0.036</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>-0.040</td>
<td>-0.063</td>
<td>0.087</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>-0.057</td>
<td>-0.003</td>
<td>0.074</td>
<td>0.096</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.011</td>
<td>0.053</td>
<td><strong>0.106</strong></td>
<td>0.088</td>
<td><strong>0.121</strong></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.053</td>
<td><strong>-0.140</strong></td>
<td>0.028</td>
<td>0.059</td>
<td>-0.023</td>
<td>-0.073</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.070</td>
<td>-0.085</td>
<td>-0.014</td>
<td>0.017</td>
<td>-0.026</td>
<td>-0.059</td>
<td><strong>0.164</strong></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.002</td>
<td><strong>0.148</strong></td>
<td>-0.065</td>
<td>-0.014</td>
<td>-0.088</td>
<td>0.002</td>
<td>-0.010</td>
<td>0.005</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-0.067</td>
<td>-0.080</td>
<td>0.054</td>
<td>-0.012</td>
<td>0.030</td>
<td>0.008</td>
<td>0.025</td>
<td>0.044</td>
<td>-0.067</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-0.003</td>
<td>0.003</td>
<td><strong>-0.107</strong></td>
<td>-0.024</td>
<td>-0.040</td>
<td>0.002</td>
<td>-0.052</td>
<td>-0.034</td>
<td><strong>0.148</strong></td>
<td>-0.043</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.005</td>
<td>0.033</td>
<td>0.009</td>
<td>-0.019</td>
<td>0.035</td>
<td>-0.068</td>
<td>0.002</td>
<td>-0.034</td>
<td>-0.050</td>
<td>0.045</td>
<td>-0.009</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>-0.083</td>
<td>-0.021</td>
<td>-0.034</td>
<td>-0.017</td>
<td>0.007</td>
<td>-0.011</td>
<td>0.013</td>
<td>0.004</td>
<td>-0.011</td>
<td>-0.073</td>
<td>0.066</td>
<td>0.023</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.046</td>
<td>0.005</td>
<td>-0.078</td>
<td>-0.094</td>
<td>-0.071</td>
<td>-0.081</td>
<td>0.046</td>
<td>0.049</td>
<td>0.059</td>
<td>-0.021</td>
<td>0.041</td>
<td>-0.046</td>
<td><strong>0.123</strong></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.077</td>
<td>0.045</td>
<td><strong>-0.108</strong></td>
<td><strong>-0.130</strong></td>
<td>-0.070</td>
<td>-0.022</td>
<td>-0.068</td>
<td>-0.027</td>
<td><strong>0.245</strong></td>
<td>-0.097</td>
<td><strong>0.141</strong></td>
<td>-0.071</td>
<td>0.079</td>
<td><strong>0.140</strong></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.098</td>
<td>0.088</td>
<td><strong>-0.107</strong></td>
<td><strong>-0.112</strong></td>
<td>-0.071</td>
<td>-0.022</td>
<td>-0.065</td>
<td>-0.086</td>
<td>0.091</td>
<td><strong>-0.108</strong></td>
<td><strong>0.130</strong></td>
<td>-0.002</td>
<td>0.095</td>
<td><strong>0.137</strong></td>
<td><strong>0.341</strong></td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-0.074</td>
<td><strong>-0.108</strong></td>
<td>0.061</td>
<td>0.010</td>
<td>0.016</td>
<td>-0.058</td>
<td><strong>0.101</strong></td>
<td><strong>0.102</strong></td>
<td>-0.067</td>
<td><strong>0.241</strong></td>
<td>-0.043</td>
<td>0.051</td>
<td>-0.063</td>
<td>-0.019</td>
<td><strong>-0.136</strong></td>
<td><strong>-0.165</strong></td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.028</td>
<td>0.066</td>
<td>-0.018</td>
<td>-0.016</td>
<td>-0.028</td>
<td>-0.032</td>
<td>0.009</td>
<td>-0.002</td>
<td>0.019</td>
<td>-0.040</td>
<td>0.021</td>
<td>0.025</td>
<td>0.002</td>
<td>-0.028</td>
<td>0.019</td>
<td>0.028</td>
<td>-0.034</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note. Correlations are presented in groups based on the content alignment results (see Appendix B). Presenting the correlations in this way allows for readers to easily note where misfit may be observed in the hypothesized models. Correlation residuals above .1 are bolded indicating poor representation of item-pair relationships. Positive correlation residuals indicate that the model underestimates the relationships between the items whereas negative residuals indicate that the model overestimates the relationships between the items.
Table 7. Bifactor Model Correlation Residuals for Sample A

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>11</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>8</th>
<th>10</th>
<th>13</th>
<th>18</th>
<th>1</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td><strong>0.313</strong></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-0.025</td>
<td>-0.045</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>-0.040</td>
<td>-0.072</td>
<td>0.029</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>-0.057</td>
<td>-0.019</td>
<td>-0.002</td>
<td>0.037</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.004</td>
<td>0.032</td>
<td>0.003</td>
<td>-0.020</td>
<td>-0.001</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.053</td>
<td><strong>-0.128</strong></td>
<td>0.048</td>
<td>0.078</td>
<td>-0.006</td>
<td>-0.043</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.061</td>
<td>-0.075</td>
<td>0.001</td>
<td>0.031</td>
<td>-0.013</td>
<td>-0.025</td>
<td><strong>0.104</strong></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.005</td>
<td>0.011</td>
<td>-0.053</td>
<td>0.008</td>
<td>-0.069</td>
<td>0.046</td>
<td>0.021</td>
<td>0.029</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-0.058</td>
<td>-0.071</td>
<td>0.068</td>
<td>0.002</td>
<td>0.042</td>
<td>0.041</td>
<td>-0.033</td>
<td>0.001</td>
<td>-0.033</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-0.003</td>
<td>0.016</td>
<td>-0.095</td>
<td>-0.013</td>
<td>-0.031</td>
<td>0.027</td>
<td>0.008</td>
<td>0.010</td>
<td>0.011</td>
<td>0.010</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.020</td>
<td>0.055</td>
<td>0.046</td>
<td>0.016</td>
<td>0.075</td>
<td>-0.013</td>
<td>-0.030</td>
<td>-0.058</td>
<td>-0.040</td>
<td>0.022</td>
<td>-0.009</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>-0.067</td>
<td>0.003</td>
<td>-0.021</td>
<td>0.008</td>
<td>0.040</td>
<td>0.039</td>
<td>0.013</td>
<td>0.004</td>
<td>-0.042</td>
<td>-0.073</td>
<td>0.008</td>
<td>0.034</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.053</td>
<td>0.013</td>
<td>-0.065</td>
<td>-0.082</td>
<td>-0.050</td>
<td>-0.052</td>
<td>0.028</td>
<td>0.023</td>
<td>0.049</td>
<td>-0.033</td>
<td>0.022</td>
<td>-0.056</td>
<td><strong>0.123</strong></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.089</td>
<td>0.063</td>
<td>-0.088</td>
<td><strong>-0.111</strong></td>
<td>-0.045</td>
<td>0.015</td>
<td>-0.068</td>
<td>-0.037</td>
<td>0.238</td>
<td><strong>-0.107</strong></td>
<td>0.134</td>
<td>-0.071</td>
<td>0.079</td>
<td>0.011</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td><strong>0.110</strong></td>
<td><strong>0.106</strong></td>
<td>-0.087</td>
<td>-0.082</td>
<td>-0.046</td>
<td>0.009</td>
<td>-0.065</td>
<td>-0.086</td>
<td>0.091</td>
<td><strong>-0.118</strong></td>
<td>0.130</td>
<td>-0.002</td>
<td>0.095</td>
<td>-0.004</td>
<td>-0.001</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-0.085</td>
<td><strong>-0.108</strong></td>
<td>0.061</td>
<td>0.010</td>
<td>0.016</td>
<td>-0.029</td>
<td>0.075</td>
<td>0.043</td>
<td>-0.098</td>
<td><strong>0.204</strong></td>
<td>-0.071</td>
<td>0.020</td>
<td>-0.080</td>
<td>0.014</td>
<td>0.006</td>
<td>-0.005</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.044</td>
<td>0.099</td>
<td>0.022</td>
<td>0.022</td>
<td>0.004</td>
<td>0.017</td>
<td>-0.009</td>
<td>-0.015</td>
<td>0.019</td>
<td>-0.053</td>
<td>0.001</td>
<td>0.025</td>
<td>0.002</td>
<td>-0.039</td>
<td>-0.016</td>
<td>0.001</td>
<td>-0.051</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note. Correlations are presented in groups based on the content alignment results (see Appendix B). Presenting the correlations in this way allows for readers to easily note where misfit may be observed in the hypothesized models. Correlation residuals above .1 are bolded indicating poor representation of item-pair relationships. Positive correlation residuals indicate that the model underestimates the relationships between the items whereas negative residuals indicate that the model overestimates the relationships between the items.
Appendix D

Figures

*Figure 1.* Three Hypothesized Models for PRL Scale

**Model 1: One-Factor Model**

![Model 1 Diagram]

**Model 2: Three-Factor Model**

![Model 2 Diagram]

**Model 3: Bifactor Model**

![Model 3 Diagram]
Figure 2. The Triangle Model of Responsibility (left) and the Accountability Pyramid (right)

References


