

Spring 5-17-2014

Meaning-Making Among Medical Students: Development of a Quantitative Measure of Self- Authorship

Robert Fallar
robert.fallar@student.shu.edu

Follow this and additional works at: <https://scholarship.shu.edu/dissertations>



Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Recommended Citation

Fallar, Robert, "Meaning-Making Among Medical Students: Development of a Quantitative Measure of Self-Authorship" (2014).
Seton Hall University Dissertations and Theses (ETDs). 1954.
<https://scholarship.shu.edu/dissertations/1954>

MEANING-MAKING AMONG MEDICAL STUDENTS:
DEVELOPMENT OF A QUANTITATIVE MEASURE OF SELF-AUTHORSHIP

ROBERT FALLAR

Dissertation Committee

Eunyoung Kim, Ph.D., Mentor
Elaine Walker, Ph.D.
Erica Friedman, M.D.

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Seton Hall University
2014

© 2014 Robert Fallar

SETON HALL UNIVERSITY
COLLEGE OF EDUCATION AND HUMAN SERVICES
OFFICE OF GRADUATE STUDIES

APPROVAL FOR SUCCESSFUL DEFENSE

Doctoral Candidate, **Robert Fallar**, has successfully defended and made the required modifications to the text of the doctoral dissertation for the **Ph.D.** during this **Spring Semester 2014**.

DISSERTATION COMMITTEE

(please sign and date beside your name)

Mentor:

Dr. Eunyoung Kim

Kim Eun Young 2/26/14

Committee Member:

Dr. Elaine Walker

Elaine M. Walker 2/26/14

Committee Member:

Dr. Erica Friedman

Erica Friedman 2/26/14

The mentor and any other committee members who wish to review revisions will sign and date this document only when revisions have been completed. Please return this form to the Office of Graduate Studies, where it will be placed in the candidate's file and submit a copy with your final dissertation to be bound as page number two.

ABSTRACT

Preparation for and application to medical school, as well as the subsequent medical training of matriculating students, can have an important impact on psychosocial development. The premedical baccalaureate is the traditional preparation for medical school, although many medical schools also offer a separate entry path through early assurance programs that provide conditional acceptance in the sophomore year of college. These programs may provide freedom in the remainder of the baccalaureate program to explore the liberal arts, which could be a source of differential development of non-cognitive skills among medical students. Self-authorship is defined as the ability to define one's beliefs, identity, and social relations and provides the basis to operate in a complex, ambiguous environment. Such a capacity is relevant to the medical education, where students face changing roles and expectations as they progress through four years of medical school and then graduate training.

This study investigated the application of a previously validated quantitative measure of self-authorship across the new population of medical students. Principal component analysis identified a 16-item instrument that measured three dimensions and one phase of self-authorship. Content analysis of responses validated the components.

The findings identified the challenges in applying a quantitative measure of self-authorship to medical students. These students may be at a different level of development and require different contexts than those in previous studies. Further research is needed to investigate differences in development across students based upon the entry path to medical school.

ACKNOWLEDGMENTS

While you can read numerous books about the process of writing a dissertation, nothing really prepares you for the actuality when it's time to write your proposal and then follow through with the proposed research. I have to thank some people who have helped make this process less painful and more productive than I thought it could be at times.

First, I extend my biggest thanks to my co-researchers at the Icahn School of Medicine who graciously allowed me to co-opt the data from our survey as the basis for my research. Thanks go to Dr. Basil Hanss, Dr. Craig Katz, and Roberta Sefcik, MS2.

Second, my research committee provided me with invaluable help and perspective as I sought to understand the potential of a new measure of self-authorship. My sincerest thanks go to my mentor, Dr. Eunyoung Kim, as well as my committee members, Dr. Elaine Walker and Dr. Erica Friedman.

I also have to extend thanks to Dr. MaryDee McEvoy, who was a colleague and confidante for many years at The Mount Sinai Hospital. While she has moved on to enjoy retirement, she has continued to provide me with moral support and encouragement in these last few months as this project has come to an end.

Finally, I wanted to thank my partner, my family, and my friends by name for their support during this process and realized I would forget someone unintentionally. Better that I just say a big thanks to each and every one of you. You know who you are.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION	1
Background.....	1
Theoretical Framework.....	12
Conceptual Model.....	16
Justification for the Study	19
Research Context and Motivation.....	20
Purpose of Study and Research Questions.....	23
Significance	23
Overview of Chapters.....	24
CHAPTER 2. REVIEW OF THE LITERATURE	26
Theories of Student Development	26
Self-Authorship and Medical Education.....	34
Quantitative Methods to Assess Self-Authorship	37
CHAPTER 3. METHODS	44
Pilot Study	44
Research Project	56
CHAPTER 4. RESULTS.....	63
Descriptive Statistics	63
Reliability	67
Principal Component Analysis	68
Sub-Group Analyses.....	76
Qualitative Analysis of Open-ended Responses.....	80
CHAPTER 5. CONCLUSION	90
Summary of Research Findings.....	90
Implications for Theory	94
Implications for Practice.....	98
Limitations.....	101
Suggestions for Future Research	102
REFERENCES	106
APPENDIX A. Comparison of Survey Items from Creamer, Baxter Magolda, and Yue (2010) with Study Pilot Phase Items	117
APPENDIX B. Revised Survey Items for the Research Project with Indicators for Proposed Dimension and Phase of Self-Authorship	121
APPENDIX C. Correlation and Covariance Matrix for the Full Data Set	125

LIST OF TABLES

Table	Page
1. Number of Survey Items in Pilot Phase by Developmental Stage and Dimension	46
2. Representative Characteristics as a Percentage of Respondents and Population	51
3. Component Loadings Using Principal Components with Promax Rotation	54
4. Correlation Matrix for Pilot Study Principal Component Analysis with Promax Rotation	55
5. Number of Survey Items in Study Phase by Developmental Stage and Dimension	58
6. Demographic Characteristics of Study Respondents	64
7. Representative Characteristics of Study Respondents as a Percentage of Respondents and Population	65
8. Mean and Standard Deviation by Survey Item and Study Cohort	66
9. Reliability Statistics in Total and by Subgroup	68
10. Component Loadings Using Principal Components Analysis with Promax Rotation	72
11. Component Correlation Matrix after Using Principal Components Analysis with Promax Rotation	73
12. Factor Loadings and Components Using Principal Components Analysis with Promax Rotation – FINAL MODEL	74
13. Component Correlation Matrix after Using Principal Components Analysis with Promax Rotation	75
14. Factor Loadings Using Principal-axis Factor Analysis with Promax Rotation – FINAL MODEL	76
15. Component Loadings Using Principal Components Analysis with Promax Rotation – PREMED STUDENTS	78

16. Communalities for Principal Component Analysis – Early Assurance Students.....	79
17. Frequency Distribution of Responses by 159 Students to Open-end Question	81
18. Summary Table Illustrating Examples of Qualitative Analysis of Responses.....	84

LIST OF FIGURES

Figure	Page
1. Phases and Dimensions of Self Authorship.....	14
2, Relationship between Experience and Meaning-Making among Early Assurance and Traditional Premed Students.....	18
3. Scree Plot for Exploratory Factor Analysis Using Principal Components.....	52
4. Scree Plot for Principal Component Analysis with Promax Rotation Using 19 Variables.....	71

CHAPTER 1
INTRODUCTION

Background

The process of preparation for and application to medical school, as well as the subsequent matriculation into and completion of training, can have an important impact on the psychosocial development of students. The current process of preparation for medical school and the medical school curriculum itself have antecedents dating back to the early 1900s. At that time there were a large number of poorly trained doctors coming from schools that did not provide a proper medical education. In 1910 the Flexner Report, sponsored by the Carnegie Foundation for the Advancement of Teaching, summarized the standards for a good medical education and provided an analysis of each of the medical schools, citing how each performed against these standards (Flexner, 1960). This report, along with subsequent work by the American Medical Association, led to the primacy of a premedical (premed) bachelor's degree as the necessary scientific preparation for medical school (Gross, Mommaerts, Earl, & De Vries, 2008). However, since the emergence of premed programs at the baccalaureate level, there have been concerns over both the curriculum and the high-pressure culture it breeds among students. As part of taking the premed curriculum, students focus on achieving high grades in classes and on the MCAT standardized admissions test in order to improve their chances of admission to medical school, which is decided in their senior year of college (Gunderman & Kanter, 2008). Students may study for and take the MCAT, often up to three times, in order to obtain the

highest score they can. This has become the traditional path upon which students endeavor to apply to medical school.

In addition to the traditional entry path, there are several alternatives offered at some medical schools that provide entry for subgroups of students alongside their traditional premed peers. There are combined BS/MD degree programs, which can run from five to eight years, with the BS taking anywhere from one to four years before the student moves on to medical classes (Eaglen et al., 2012). There are also early assurance programs, which are a loosely defined set of programs that allow a student to be accepted to medical school in his or her sophomore year; however, matriculation into medical school does not occur until after completion of the bachelor's degree.

To date, empirical literature evaluating the purposes or goals of early assurance programs is almost non-existent. An informal interview of the 11 medical schools with the largest classes of early assurance students identified that even after acceptance, many still require the MCAT as part of the requirements for matriculation (S. Shadravan, personal communication, November 27, 2012). A review of online resources at three of the schools reveals that these programs tend to recruit honors students and often have a goal of educating a particular type of doctor, such as rural physicians (TTUHSC School of Medicine, 2013; University of Maryland, College Park, 2013; Upstate Medical University, State University of New York, 2013). Several programs do mention on their web sites that the goals of the programs include the potential to reduce the stresses of the traditional premed program and to allow the students more freedom to explore their undergraduate education. However, no empirical research is available to identify the non-cognitive

effects or benefits of the assumed freedom early assurance students attain during their last two years of college.

In a recent study, Muller and Kase (2010) demonstrated how students from an early assurance program with some unique features (elimination of the MCAT entrance exam and discouragement of a premed undergraduate degree) perform just as well as, if not better than, their traditional premed-trained classmates on many measures of academic performance during medical school. When the findings were highlighted in a front-page article of the *New York Times*, the subsequent online discussion was swift, rather one-sided against the idea of such a program, and somewhat vitriolic in tone (Hartcollis, 2010). A common theme across the comments was that the students were less prepared for medical school and graduates of this particular early assurance program would be less qualified doctors than their peers, incapable of understanding concepts such as prescription drug interactions or advanced testing procedures. Many comments seemed to assume that the basic science in a premed education is the only legitimate source of this knowledge and failed to acknowledge that the same science knowledge is addressed in medical school itself. While admittedly not representative, these comments highlight a view held by the general public about the requirements to become a doctor, as well as the type of physician this particular program graduates.

Though Muller and Kase (2010) provided the first evidence to identify possible differences on cognitive measures between the early assurance students and traditional premed students by the end of medical school, they did not look at possible differences at the start of medical school nor investigate any non-cognitive differences in students either at the beginning or end of medical school. If there are differences in non-cognitive skills

among different groups of medical students, one potential influence could be the different paths these students took prior to matriculation, including removal of the academic pressures of the premed undergraduate program and possible differences in preparing for medical school.

Preparation for Medical School

Historically, students prepare for medical school with a premed undergraduate degree and by taking the MCAT, often more than once during their college years in an effort to achieve the best score they can (Zhao, Oppler, Dunleavy, & Kroopnick, 2010). The premed curriculum includes general chemistry, organic chemistry, biology, physics, and calculus. Critics are concerned that the content of the curriculum places too much or too little focus on science and thereby a premed program might instill some negative behaviors in students, such as competitiveness and a focus on grades (Gross et al., 2008). While no one suggests the premed curriculum should be abolished, the content of a premedical program has received the most consistent criticism over time. The primary arguments against the curriculum relate to the desire to maintain exposure to the liberal arts as part of a student's education (Brieger, 1999; Gross et al., 2008). The development of clinical skills is of primary importance in becoming a doctor, but so is the ability to communicate well, a sense of objectivity, and an understanding of the world from a natural, social, and cultural perspective (Weingartner, 1980). Further, Weingartner suggested that while maintaining some basic science courses, a premed program should require a humanities or social science major or concentration. Some critics also note that the purpose of a premed program is not to make medical school easier, but instead to prepare a student for the rigors of a medical education. Therefore, the actual science of

medical school should not be taught beforehand in a premed program; rather, discipline and a well-rounded education are paramount (Brieger, 1999).

What purpose should an undergraduate education serve for a prospective medical student? An undergraduate degree must provide a foundation in liberal arts for all students. The premed program should not be viewed as just a prerequisite for medical school but as an undergraduate degree that provides a well-rounded, holistic education (Gunderman & Kanter, 2008). The science should be limited to that which is most relevant to medicine, opening up the opportunity to explore the liberal arts (Dienstag, 2008). In particular, students need to learn how to acquire knowledge and understand self-reflection, both of which are regarded as important qualities for a doctor's career.

From a broader perspective, there have been long-standing issues with the definition, value, and purpose of a liberal education itself. A liberal arts education has been a fundamental pillar of American higher education since the colonial period, building on the concepts of the European system (Thelin, 2004). In 1947, the Truman Commission proposed a set of principles for providing access to higher education and guidelines for what every student should learn. The report distinguished liberal arts from a general education, which included the development of basic skills, responsible citizenship, and multicultural relationships in the world (Hutcheson, 2007). The Commission report was instrumental in improving access to higher education with programs like the GI Bill, but it did not effectively lead to initiating change in teaching general education skills. By the late 20th century, an explosion of new technology vastly affected society's ability to automate processes in business and manufacturing, as well as the ways in which we communicate and interact with others. In this more complex work environment, the

curriculum of higher education shifted from a utilitarian focus on specific skills to an interest in the more non-cognitive skills that are desirable for college graduates and the new work force (Baxter Magolda & King, 2012). The “higher learning” skills would help to better prepare students to work in a complex environment where they would have to deal with uncertainty, think critically, and be self-sufficient to find solutions to problems (Hodge, 2009).

In 2005, the Association of American Colleges and Universities began a decade-long project called Liberal Education and America’s Promise (LEAP) in an effort to change the goals of higher education to meet the needs of a more complex world (The National Leadership Council for Liberal Education and America's Promise, 2007). The proposed new outcomes for college graduates included the following:

- Knowledge of cultures and the physical and natural world
- Intellectual and practical skills (critical thinking, problem solving, and teamwork)
- Personal and social responsibility (fostering responsibilities beyond oneself)
- Integrative learning across subject areas

The educational leaders who participated in the creation of LEAP understood the complexity of improving these skills and acknowledged that one program would not be effective across all types of programs/institutions. Rather, programs should be designed to include interventions that are specific to their pedagogy and environment. The proposed skills are, however, somewhat vague and left to the individual institutions to specify and define. LEAP does not identify specific tools to measure these skills but recommends that any assessments be linked to the specific learning situations. They offer broad examples such as portfolios and capstone projects, providing evidence to identify if students have

met the goals or not; but these examples are not specific enough to allow for any measure of progress along a continuum up to and beyond meeting the goal.

In July 2012, the National Academies published *Education for Life and Work*, based upon the work of another expert panel charged with defining “21st century skills” that should be the goal of higher education (Board on Testing and Assessment, 2012). The identified skills fall into three domains: cognitive (reasoning), interpersonal (expressing oneself and interpreting messages from others), and intrapersonal (the ability to regulate oneself when working with others towards goals). This report reflects recent theoretical work on student development, specifically the concept of *self-authorship*, calling for more research to understand the link between these skills and outcomes after college. The reports from the National Academies and LEAP indicate there is a growing public interest to improve the outcomes of higher education and provide accountability towards higher learning skills over and above the accumulation of specific knowledge within an academic discipline.

A common theme to the criticism and discussion thus far points to the need for students to accumulate more non-cognitive skills that can be helpful in both the educational experience and in the students’ eventual careers. The National Academies’ goals apply to all types of higher education, and some critics have particularly specified the benefits for medical students to develop these same types of skills after having exposure to the liberal arts as undergraduates (Brieger, 1999; Dienstag, 2008; Weingartner, 1980). Regardless of whatever efforts may arise from the LEAP and National Academies’ calls for improvement in liberal arts education, the opportunity for

early assurance students to have more exposure to liberal arts than their premed peers may lead to the potential for differences in non-cognitive skill development between the two.

A second issue with premedical education involves the culture it perpetuates among the students in the programs: that the students become too fixated on grades and the likelihood of being admitted to medical school while missing out on a richer, more holistic undergraduate experience (Gross et al., 2008; Weingartner, 1980). Although early assurance programs can be the opportunity to enhance the liberal education of aspiring medical students while removing some of the pressure to achieve the highest grades, there is no empirical evidence to support any potential benefits from these programs.

Assessing Medical School Applicants

Since concerns about the content of the premed curriculum have existed since the early 20th century, it can be assumed that the general nature of the premed curriculum will not change quickly. The focus of criticism then shifts to how medical schools select applicants, with the primary suggestion being to move away from relying on one's GPA and MCAT scores as the most significant indicators of success. Acceptance to medical school is usually determined based on cognitive measures (GPA, MCAT) as well as other aspects of personal development of the student, often assessed based upon interviews and written personal statements. These measures are used to identify the students who are most capable of succeeding in medical school and becoming the kind of doctors the schools seeks to train. There is a large volume of literature on the positive correlation between MCAT scores and GPA with performance in the first two years of medical school (Peskun, Detsky, & Shandling, 2007). For example, the MCAT and GPA have been shown to be valid predictors of written knowledge assessment throughout medical school,

as well as in clinical skills assessments (Kreiter & Kreiter, 2007). However, there is inconsistent evidence in the literature on the relationship between these scores and performance in the last two years of medical school. For instance, research has shown that the MCAT has little predictive validity for clinical performance among minority students (White, Dey, & Fantone, 2009). The undergraduate GPA has also been found to not be predictive of performance in the clinical phase (third and fourth years) of medical school (Silver & Hodgson, 1997). Other researchers have indicated that science indicators (MCAT and GPA) are actually inversely related to non-cognitive performance, suggesting that those with higher MCAT or GPA scores may not be as advanced as others in the domain of non-cognitive skills (Barr, 2010). Furthermore, the issue of bias within these measures has been raised. Standardized tests are designed to distinguish between students and highlight those who are more or less knowledgeable. These types of tests may be biased against minorities, providing inadequate measurement of other desirable skills among these students (Sedlacek, 2004). Therefore, other non-cognitive measures may provide better information to support disadvantaged students in aspects such as gaining access to medical schools.

Admissions decisions can have different purposes, including being a means to identify who is most prepared for the rigors of medical school and/or to select prospective students who will become the type of doctor the school wishes to train. Reliance on standardized cognitive indicators such as the MCAT and the focus on premed training can impact admissions decisions regarding who will be able to be trained to become a physician. Any potential bias introduced by these indicators may also influence who ultimately applies and gets accepted to medical school.

Given the questionable utility of these cognitive indicators in predicting success through various stages of medical training, non-cognitive measures including essays, interviews, and multiple mini-assessments have become more common in the application process as alternative methods of assessing candidates. These non-cognitive measures have been found to be more predictive of academic success in the first two years of medical school than the traditional cognitive measures (Peskun et al., 2007). For example, multiple mini-interviews provide the opportunity for prospective students to react to various situations that measure competencies such as ethical behavior that cannot be measured by the traditional cognitive methods. Mini-interviews have been found to be a fair and engaging method of assessing students without the pressure of traditional assessments (Razack et al., 2009). Despite little consensus on which non-cognitive skills are the best criteria for evaluating applicants, many medical schools have incorporated some type of non-cognitive measure of applicants into their review process (Bardes, Best, Kremer, & Dienstag, 2009). Many of these non-cognitive measures require significant human effort in terms of reviews of written documents or ratings of interactive activities. These are qualitative in nature, requiring significant resources to implement on any large-scale evaluation of students.

Alternatives to the Traditional Entry Path

Along with changes to the applicant review process for the traditional medical school entry path, there are now additional programs to gain entry to medical school. One popular non-traditional entry path to medical school is a combined baccalaureate-MD program. Many of these programs are at medical schools within a university setting, where the early acceptance is for undergraduates at the same institution. As of 2011, 44% (57) of

the medical schools in the United States offered some type of “combined” program (Eaglen et al., 2012). In addition, there are 30 medical schools in the United States that have an early assurance program (Hutcheson, 2007). Early assurance programs guarantee admission to medical school during the sophomore year. There is no standardized format for such programs; but most, if not all, require some type of premedical education while in college and may or may not require the MCAT. These programs guarantee acceptance to medical school prior to the third year of college, thus eliminating some of the traditional premed application steps. More than half of the early assurance programs report that the primary goal of the program is to attract honors students into medicine, and 22% report that the mission is to integrate the liberal arts into medicine (Eaglen et al., 2012).

The types of students applying to medical school are still predominantly biology majors. While the composition of MCAT examinees is not fully reflective of those who attend medical school, the data show that students from non-premed backgrounds are still a minority taking the exam as part of the traditional path into medical school (Association of American Medical Colleges, 2011). No data on the demographic characteristics (race, academic performance, gender) of students nationally who apply to the early assurance programs are available.

How do early assurance programs counter the traditional criticism of medical school admissions? First, they may remove the MCAT scores from the admission decision process and, in so doing, remove the pressure on the prospective student to study for, and take, the exam during their undergraduate education. Second, they provide the student with the opportunity to explore classes of their choice in the last two years of college, free of the need to have their transcript include courses they might think imperative to get

accepted to medical school. Third, the opportunity to broaden the liberal arts education for early assurance students might provide a more humanistic experience than traditional pre-medical programs. Taken together, these differences may create a subgroup of medical students different than their premed-trained peers. Still, there has been little empirical evidence of the effects of an early assurance program. As Eaglen et al. (2012) state, “For those programs that strive to reduce competitive pressures or emphasize humanistic qualities, there have been no studies to date indicating whether students in those programs are different in any meaningful way by the time they begin medical school” (p. 5). The historical arguments regarding premedical undergraduate programs and the criteria used for acceptance to medical school, as well as the number of alternative entry paths to medical school, all suggest the need to investigate non-cognitive development among medical students along the trajectory from baccalaureate through medical school. However, the literature provides no evidence of any broad measure of non-cognitive development among this population. This study sought to validate a quantitative measure that could provide a first effort to more broadly determine potential non-cognitive development among medical students, with analyses to determine validity across sub-groups.

Theoretical Framework

When studying the cognitive and non-cognitive development of students as they enter and progress through medical school, self-authorship provides a useful theoretical lens. Self-authorship, defined as “the internal capacity to define one’s beliefs, identity, and social relations” (Baxter Magolda, 2008, p. 269) seeks to explain how a person is able to consider external influences and persevere in a complex, ambiguous environment through

a threefold set of dimensions: epistemological (what one believes), intrapersonal (what one values), and interpersonal (how one relates with others in mutually beneficial ways) (Baxter Magolda, 2008). The epistemological dimension refers to how a student deals with ambiguity and conflicting information and grows from accepting knowledge as coming from authority figures to the point where they understand that knowledge can be contextually driven and interpreted differently by different people. According to Pascarella and Terenzini (2005), “epistemological skills ... provide the foundation for making choices among truth claims, values, and behaviors” (p. 49). The intrapersonal element refers to how one grows to understand one’s own values and beliefs and becomes comfortable using these to interpret information and deal with others. Similarly, the interpersonal element identifies how one relates with others while maintaining one’s sense of self. These three elements are intertwined and equally important to the development of self-authorship.

Baxter Magolda (2008) identifies three phases in the development towards self-authorship. In the External formulas phase, students base decisions and actions on influences and knowledge provided by authority figures. As they move to the Crossroads phase, students question knowledge and formulas that conflict with their own ideas, yet they do not know how to resolve these conflicts. By the Early self-authorship phase, students can understand their own beliefs and how they coincide or conflict with External formulas. Subsequent decisions can be made by taking into account their own values and beliefs and working with others to achieve mutually beneficial goals. Figure 1 displays the dimensions and phases of self-authorship.

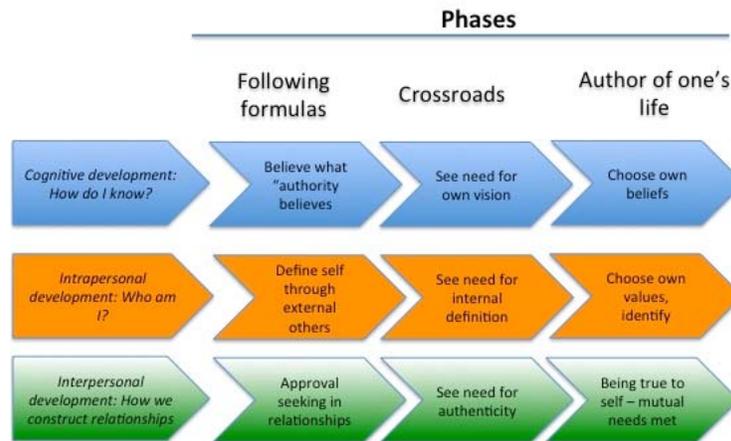


Figure 1. Phases and Dimensions of Self-Authorship. Adapted from Joseph Boehman (2011, April 28). Who are you? Self-authorship defined [Web log post]. Retrieved from <http://imjoeboe.wordpress.com/2011/04/28/self-authorship/>

Self-authorship is relevant to medical education for several reasons. Firstly, premed education relies heavily on external authorities as one learns how to prepare and apply to medical school (Gross et al., 2008). Students who bypass the traditional premed programs may have different perspectives on External formulas and different senses of internal beliefs given the preparation and motivations necessary to apply to medical school as a sophomore.

Next, developmental movement from one phase to another (External to Crossroads) reflects the period in which a person comes across information that conflicts with, or contradicts, the messages from external influences. In the Crossroads phase, a person may not know exactly how to resolve the conflict, but he or she becomes aware of

his or her own beliefs and values and realizes the need to resolve these with the conflicting external messages. Baxter Magolda (2004) found that the break with External formulas was often caused by a perceived conflict regarding the student's current career plans. This conflict is relevant to the pressures of the premed curriculum as well as the experience students have once they matriculate. It also provides a lens to determine both the differences in expectations prior to medical school and the potential conflict between expectations and reality the students face once they matriculate into medical school. These conflicts may be different for early assurance students; they are more likely than traditional premed students to face increased pressures at least for the first year of medical school as they begin studying medical science.

Finally, medical education emphasizes experiential learning, which is key to medical student development and combines basic sciences commonly taught in a classroom setting with clinical learning that takes place in a medical practice setting (Yardley, Teunissen, & Dornan, 2012). Medical education is a progressive path where students' identities and knowledge are formed and revisited over time as students move from undergraduate to resident to practice (Slotnick, 2001). Students face different expectations from teachers and others (e.g., patients) along their education path and need to resolve these changing expectations along with their personal needs and perspectives. Self-authorship can be an important skill necessary to succeed in the arc of experiences during medical school. Self-authorship supports the cognitive as well as sociocultural dimensions within Slotnick's (2001) view of a "life-span" of learning. In light of cognitive development, Slotnick also adds that epistemological development is important in building knowledge and skills in problem solving and in satisfying self-esteem needs. The

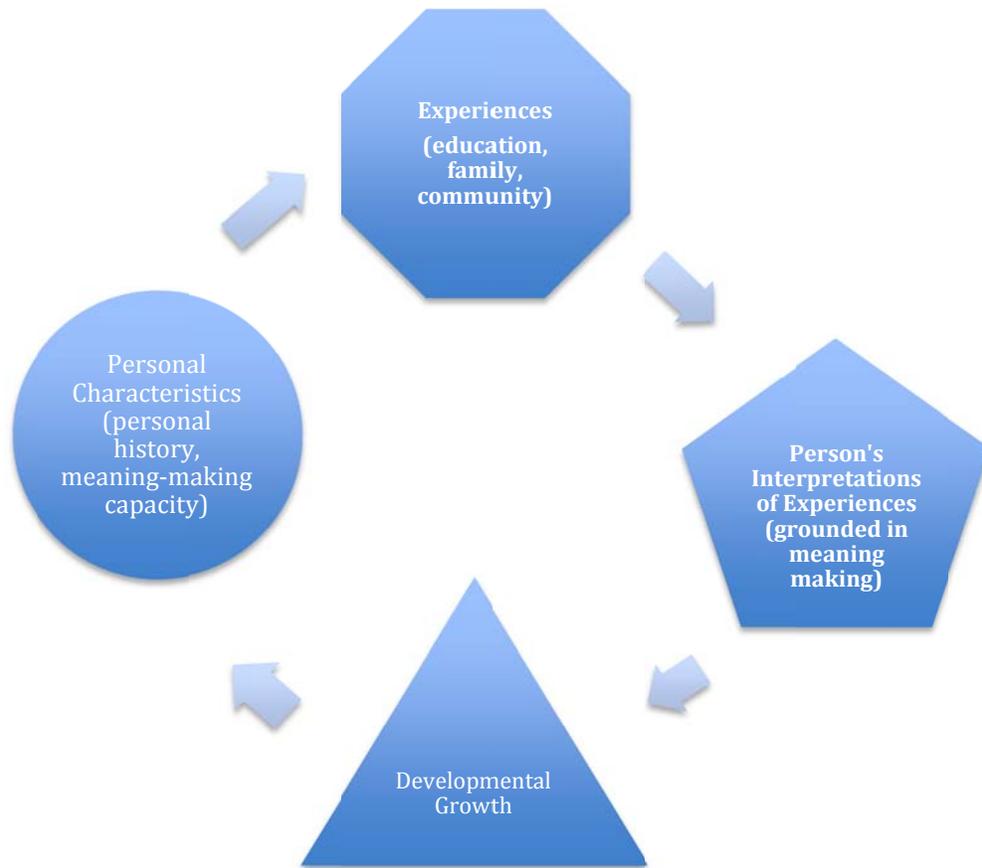
problem-solving skills of medical students parallel Baxter Magolda's epistemologic dimension. The sociocultural dimension relates to the many roles a medical student plays in life, both in school and in practice as a physician after graduation. The ability to maintain relationships with others (family, peers, patients) is key to successfully resolving expectations and meeting one's own needs. This dimension coincides with the interpersonal and intrapersonal dimensions of Baxter Magolda's model.

Conceptual Model

Suggestions to incorporate self-authorship into the curriculum to support different ways of knowing include providing opportunities for self-discipline, as well as providing freedom to exercise choice and develop one's voice. (Baxter Magolda, 1992). The comparative freedom of an early assurance program student's final years of college as compared to a premed student provides an excellent context for the application of self-authorship and testing whether differences in development do exist.

This conceptual model is not intended to explain the possible differences in development but instead to illustrate the potential differences that could occur within the cyclical nature of development towards self-authorship (see Figure 2). The development of self-authorship is a recursive process based upon reflection and revision of beliefs, values, and knowledge when a person is faced with new challenges to his or her current state. These challenges can come from sources such as educators, family, and work situations (Baxter Magolda, 2004). It is hypothesized that the college experiences of early assurance students as compared to those of premeds can create different influences towards self-authorship. These influences can take several potential forms:

1. The freedom to take liberal arts courses for the early assurance students may allow them to grow and develop skills towards self-authorship.
2. The freedom for the early assurance students may remove pressures and challenges to beliefs, values, and knowledge in the traditional premed curriculum that would otherwise spur development toward self-authorship.
3. Once the early assurance students arrive in medical school, they may be faced with additional challenges of feeling less prepared than their peers because of their lack of basic science training during their undergraduate degree program.



*Figure 2. Relationship between Experience and Meaning Making among Early Assurance and Traditional Premed Students. Adapted from “Development of the Ten Positions in the Journey Toward Self-Authorship” by M. B. Baxter Magolda and J. P. Barber, 2012, *Assessing Meaning Making and Self-Authorship: Theory, Research, and Application: ASHE Higher Education Report 38:3*, p. 41. Copyright by John Wiley & Sons.*

Self-authorship is a movement away from External formulas when faced with complexity, ambiguity, and conflict. The early assurance program does not force students to face External formulas but instead removes some of them. The challenge for academics has been how to measure self-authorship on a large-scale method and then apply this knowledge to the educational experience as it relates to student outcomes. This research sought to identify a tool that could be used to measure self-authorship; it sought to allow

for measurement among sub-groups of medical students along the trajectory of their educational paths.

Justification for this Study

The nature of self-authorship lends itself to qualitative methods, as researchers have sought to understand what it is and what factors influence non-cognitive development. Significant training is necessary for a researcher to properly conduct interviews and guide conversations in order to elicit responses that help explain the thought processes that can exhibit self-authorship or a lack thereof (Baxter Magolda & King, 2007). Qualitative studies of self-authorship can be resource-intensive but provide rich data to help understand the phases and dimensions of self-authorship. If, however, a researcher intends to conduct large scale studies of how a population stands in terms of their positioning within the phases and dimensions of self-authorship, qualitative methods are too cumbersome and subject to the potential for variation between reviewers (Baxter Magolda & King, 2007). A quantitative scale that measures self-authorship could be used in large-scale evaluations, providing the opportunity to identify group differences and correlates with development (Pizzolato, 2005). Such a tool could also provide economies of scale over qualitative efforts in allocating resources to study self-authorship.

Very few quantitative tools have been developed to measure self-authorship. Pizzolato (2005) created a scale based upon the skills needed to manifest self-authorship, such as autonomy and relationships with authorities. The scale ultimately was modeled by four skill factors, which Pizzolato called Capacity for Autonomous Action, Problem Solving Orientation, Perceptions of Volitional Competence, and Self-Regulation in Challenging Situations. Creamer et al. (2010) created a scale that attempted to mirror the

phases and dimensions of self-authorship as theorized by Baxter Magolda. This scale identified factors representing the three dimensions and the three phases of self-authorship. In both studies, the sample populations were largely White female undergraduates at a Midwestern university. Neither instrument was directly transferable to the medical student population.

Extant literature fails to provide any evidence of the study of self-authorship specific to the medical student population. Within the last two years, however, there has been discussion regarding the desire for self-authorship to be studied as it relates to medical education curriculum (Jarvis-Selinger, Pratt, & Regehr, 2012; Prober & Khan, 2013). Researchers argue the importance of self-authorship for students and physicians during their education and training. (Jarvis-Selinger et al., 2012). In particular, these discussions relate to measuring phases of self-authorship and how they pertain to medical education. Medical education curricula, as it relates to the experiences of traditional premeds and early assurance students, provide one specific example of the need for a quantitative measure that could be validated for medical students in general and across sub-groups. However, given the potential to apply the concept of self-authorship to medical students, there are no validated quantitative instruments to allow for comparative studies among medical students.

Research Context and Motivation

This study took place at a large (more than 500 students) medical school in the northeast United States that has an early assurance program. The Icahn School of Medicine at Mount Sinai has a limited early assurance admission program and is not affiliated with any university; therefore, the program is available to students from a variety

of undergraduate institutions. Students are able to apply to the medical school in their sophomore year of college. The application process includes a review of the student's college GPA up to that point, as well as personal essays and interviews. Upon acceptance, these students are not required to take the traditional premed undergraduate program, nor are they required to take the MCAT admissions test. Rather, they agree *not* to take a premed undergraduate program and instead choose a major in a liberal arts program. They must take two semesters each of biology and chemistry. The students are also encouraged to take a year off after completing their bachelor's degree for further exploration of interests they might have, such as travel, research opportunities, or volunteering efforts nationally or internationally. Those who are accepted spend part of the summer between their junior and senior years at the medical school doing an immersion class in organic chemistry. In the summer before they start medical school, they are also provided the opportunity to attend an optional six-week preparation program that covers fundamentals in biochemistry and anatomy, as well as sessions on study methods and acclimating to medical school.

The premise of the early assurance program is that it graduates a different type of physician—someone who would be different in a non-cognitive skill such as empathy or morality. A previous study has identified the differences and similarities in cognitive performance between early assurance and premed students during the four years of medical school (Muller & Kase, 2010). Although the early assurance program has been in place for 25 years, there has been no formalized effort to date at this medical school to identify the non-cognitive development of the early assurance (or any other) students, either upon entry to, or at the completion of, medical school. In 2011, a team of three

researchers (including myself) came together for a four-year project to design a method to identify the non-cognitive development of the same student groups. The team reviewed literature on several non-cognitive measures, including empathy, morality, actualization, and self-authorship. We originally viewed the empathy and morality of medical students as very specific characteristics, whereas there may not be much differentiation among students at the start of medical school. Therefore, we decided these were not appropriate for our first analysis. Instead, we selected broader concepts of student development, specifically self-authorship and self-actualization. Self-actualization refers to the self-fulfillment of one's perceived potential (LeClerc, Lefrancois, Dube, Hebert, & Gaulin, 1999). This research primarily focused on the development of the measure of the self-authorship construct because the self-actualization items had been empirically validated previously.

Self-authorship has predominantly been examined qualitatively through in-depth interviews, which can be time-intensive and require skilled interviewers (Baxter Magolda & King, 2012). Previous efforts to create quantitative measures of self-authorship, using both contextual and non-contextual items within the construct, have had mixed results when data from the scale are correlated with qualitative data that is collected concurrently (Creamer, Baxter Magolda, & Yue, 2010; Pizzolato, 2005). In particular, respondents may answer questions in a manner that would indicate self-authorship yet when they explain their reasoning through qualitative measures, they do not actually reflect self-authored behaviors. Still, there is no consensus on whether efforts to build a quantitative measure are worthwhile or not.

The research team decided to use the tool created by Creamer et al. (2010) as the basis for a new quantitative measure of self-authorship among medical students. The challenge was to create construct items that measure the thought process the student goes through, not the actual decision, and apply these concepts to medical students. A pilot study was conducted in 2011 using a test version of a self-authorship scale created by the research team. The analysis of this pilot formed the basis for this research.

Purpose of the Study and Research Questions

The purpose of this study was to perform the analyses necessary to create a valid and reliable scale of self-authorship that can be applied to the medical student population. This study incorporated qualitative data with the scale for contextual verification. Given the scant literature regarding self-authorship in medical school as well as the prior studies of early assurance programs at this site, this study aimed to answer two central research questions:

1. (a) Does data collected for students at the beginning of the 2013-2014 academic year support the validity and reliability of the proposed instrument, and (b) does the reliability and validity extend to the premed and early assurance subpopulations?
2. Does a content analysis of the qualitative responses provide evidence to validate the constructs of self-authorship among medical students?

Significance of the Study

This study focused on the development of an instrument to measure the construct of self-authorship among medical students. This research is significant in several respects. First, it expanded upon the literature of self-authorship by extending efforts to validate the

concept among the medical student population. This analysis explored the dimensions and phases of self-authorship among a new population and highlights differences in the theoretical interrelationships developed among the undergraduate population. Second, most of the previous research has been exploratory and qualitative in defining self-authorship among populations. This research built upon the earlier research into defining quantitative measures of movement within self-authorship. Finally, the potential to identify development among subpopulations of medical students could help in understanding the tool's applicability in evaluating admissions criteria. Early assurance programs in general provide an opportunity to diversify the types of students who enter medical school (Eaglen et al., 2012). As early assurance students are a key subgroup of the student population, this evaluation is the first step towards validating the tool in relation to this population.

Overview of Chapters

This first chapter has provided the background information regarding self-authorship and its relation to medical students. This chapter also included the theoretical framework, justification of the study, the research context, purpose statement, research questions, and significance of the study. In Chapter 2, there is a two-part literature review examining the development of self-authorship theory and its applicability to medical students, followed by a review of the literature regarding efforts at quantitative measures of self-authorship. Chapter 3 describes the methods used in the first wave of data collection, the pilot for this research project, followed by a description of the methods used in the second wave. Next, Chapter 4 presents the results from the second wave. In

conclusion, Chapter 5 discusses the implications of research findings for theory and practice and provides suggestions for future research.

CHAPTER 2

REVIEW OF LITERATURE

Given that little research has been conducted examining the development of non-cognitive skills among early assurance medical students, this chapter provides a review of the literature on the development of self-authorship as a theoretical guide and analyzes the empirical evidence of methods to investigate self-authorship.

Theories of Student Development

Theories of student development explain how students mature and grow toward more complex and sophisticated thinking during their educational years. While some theorists describe development from early childhood on, higher education has been the primary focus of much student development work. There are four general classes of development theory that have been most commonly used with college students since the 1960s (Kuh, Gonyea, & Rodriguez, 2002). In brief, they are the following:

1. Psychosocial Theories – these focus on development of the person and his or her identity, how one builds relationships and becomes self-sufficient.

Conceptually, development can occur in response to pressures from society as one seeks to understand one's identity as it fits in with others. The locus of development is through sociocultural, psychological, and/or biological influences (Pascarella & Terenzini, 2005). Development/regression occurs when one needs to make a choice among alternatives in regard to self-identity. In terms of college students, this relates to how students develop in the face of a new college environment, peer pressure, and academic responsibilities.

2. Cognitive Structural Theories – a class of theories which explain personal development not necessarily from societal pressure but rather from internal questioning when faced with information contrary to one’s beliefs or assumptions. Influential theorists include Kegan (1982) and Perry (1997). Baxter Magolda (2008) refers to how one *makes meaning* of experiences by incorporating beliefs, values, and knowledge. Unlike the psychosocial theorists, cognitive structural theorists include an epistemic factor into development. The movement from accepting knowledge from authority to being able to understand knowledge as contextual and having multiple meanings is an integral part of these theories of development.
3. Person-environment Interaction Theories – these theories do not specifically describe development; instead, they seek to explain how a person goes about having one’s needs match with demands of the environment. These models help to explain the match between student and college and why some succeed and others do not, as well as how one identifies a good match.
4. Typology Models – here, theorists categorize students by how alike or different they are in responding to situations within the college setting. Again, these models do not necessarily describe development but instead reflect inventories that provide the ability to predict how a person might react in various situations. In the case of student development, this can relate to situations such as classroom environment or dormitory living assignments.

Of the four broad theoretical perspectives, cognitive-structural theories would be the most relevant to the study of deeper learning skills because of the inclusion of

epistemological development, which can be related more to learning and introspection. All of the theories include conceptualizations of higher order skills that support the LEAP goals for a liberal arts education, such as building relationships, autonomy, and acceptance of diversity. However, without considering epistemology, many of these theories are more applicable to issues of student life than they are to explaining the growth of non-cognitive skills. This is not to say that the psycho-social theories ignore epistemology, but more that they focus on socialization, acceptance into the higher education environment, and the ability to succeed and continue to develop as adults. A review of key cognitive structural theorists further clarifies the applicability to deeper learning skills.

Perry's Theory of Intellectual Development

When one thinks of development over time, it is helpful to think about indicators of progress or milestones that mark change. Many cognitive-structural theories identify varying numbers of developmental stages, but all see a recursive process of growth within the student's development of cognitive skills. One of the earliest studies into cognitive structural theories of student development consisted of qualitative interviews with White male students during their four years in college (Perry, 1997). Perry proposed that development occurred in nine progressive positions, from what he called *dualism* to *relativism*. In the lowest position (dualism), knowledge is either right or wrong and is handed down from authority. At this stage, there is no regard for others or diversity of thought. By the time students develop to the ninth position (relativism), they are aware that there are a variety of opinions and values and they can analyze and compare these in order to make a decision on what is the best interpretation of an experience or problem. The path through the positions is not linear, nor is it exactly the same for everyone. Perry

explains how people may retreat, or digress, in their development when faced with too difficult situations, but progress may continue again in the future. On the other hand, others may escape, as Perry puts it, and not progress any further or may even revert toward a dualistic state.

Kegan's Theory of the Evolution of Consciousness

Robert Kegan provided a complementary theory that introduced the construct of self-authorship (Love & Guthrie, 2002). Kegan's seminal work relates to growth over a lifetime and does not focus on just the period of academic experiences (Kegan, 1982). His theory is based upon the concept of how we view experiences as either *subject* or *object*. [Emphasis added.] When we are responsible for an experience or can work to fix problems with it, it is considered in our control, or an *object*. Conversely, if there are experiences we are not responsible for, or are not aware of, then these are considered to be *subject*, as in we are subject to them. Moving from *subject* to *object* shows development in knowing and independence. Kegan proposes five stages of development and believes that most people will plateau at the fourth stage. Similar to Perry, he does not see the path through the stages as linear, but more like a helix; movement upward requires reflection on past learning and assumptions and subsequent reconciliation of conflicted material. The ability to reflect is important to these cognitive theories. Of the five stages, three are most relevant to the college years. Stage 2 starts at around age six and goes through adolescence. At this stage, a person's impulses are *object*, but the person is *subject* to his needs and desires. Through awareness and interaction with others, growth into the third stage occurs during post-adolescence and includes movement of one's needs and desires to become *object*, where one is able to control them in relation to one's interactions with

others. On the other hand, at this point people become *subject* to interpersonal relationships and can be influenced by the needs and desires of others (peer pressure), often at the cost of their own needs. As an individual moves into the fourth stage, he or she acquires what Kegan termed self-authorship. In addition to having one's own needs and desires become *object*, at this time interpersonal relationships move into a phase in which the person understands that his or her relationships are mutually created with the other participants. The self-authored individual can mesh his or her own needs with those of the other person to achieve mutually satisfying solutions. The fifth stage, in which one has the ability to have multiple systems of thinking, occurs, if at all, in the adult years after college. One key to understanding Kegan's theory is that it is not necessary for every person to achieve the fifth, or even fourth, stage of development. Instead, he explains how achieving each stage provides a new set of skills that can benefit one as one moves into more complex life situations.

Baxter Magolda's Theory of Self-Authorship

The theory of self-authorship has been further refined by Baxter Magolda based upon her 25-year longitudinal study of a cohort of college students (Baxter Magolda, 2008). Her qualitative research included annual interviews with a cohort of men and women as they went through college, entered careers, and began families. She clarified self-authorship as "a necessary foundation for adults to meet typical expectations they face at work, home, and school, such as the ability to be self-initiating, guided by their own visions, responsible for their experience, and able to develop interdependent relationships with diverse others" (p. 269). One key skill of self-authorship is the ability to learn on

one's own and not be uncritically assimilating from others; another is the ability to reflect on experiences and understand their relation to one's internal needs and desires.

Baxter Magolda's (2008) model can be conceptualized in a 3x3 matrix. The dimensions of self-authorship are epistemological (understanding where knowledge comes from), intrapersonal (understanding one's own values and beliefs), and interpersonal (how one builds relationships with others). The three dimensions are considered equally important to self-authorship, with no priority given for one over the others. Across each dimension there are then three phases of growth: External formulas, Crossroads, and Early self-authorship.

The External formulas phase represents the place where thoughts and behaviors are based upon input from external sources. Knowledge is handed down as rote from authority and relationships are often predicated on trying to please or obey the other instead of understanding one's own internal needs and desires. Moving into the Crossroads phase, one becomes aware of and comfortable with one's own needs, how one is responsible for them, and how relationships need to incorporate the perspectives of both individuals. Knowledge becomes questioned as alternative perspectives are understood in a more complex environment and one realizes there are limits to certainty when faced with problems. In the final phase, self-authorship, one understands the complexity of knowledge and how it is individually constructed. Similarly, the interpersonal and intrapersonal dimensions develop to the point that one becomes comfortable creating one's own sense of values and that, when acting with others, one does not lose sense of oneself but can balance the needs of both parties. One can act on one's own senses and not necessarily just assimilate the knowledge and beliefs of others. Similar to Perry and

Kegan, Baxter Magolda also sees the distinction between the three stages as somewhat fuzzy and the path towards self-authorship as not necessarily the same for every person. The third phase of development (Early self-authorship) has been shown to exist during the college years (Pizzolato, 2004).

One common theme across these cognitive-structural theories is the understanding that conflict of some type usually helps to promote development to a further phase. This conflict could relate to new knowledge that counters what the individual has always believed or that leads to questions about the assumed knowledge of a trusted authority, or a diverse opinion from someone that they had never heard before. Having to react to an affront to one's current comfort levels can provide the opportunity to better understand the conflicting information and how to incorporate it into one's thought processes going forward. The more self-authored the person, the better the person is at taking a challenge to his or her belief systems, comparing and evaluating the information, and using it to learn and grow.

A second common thread is the need for self-reflection if one is to move to different phases of development. These theories focus on the inner voice, being able to understand one's needs and beliefs, and interpreting conflicts to one's prior understanding. Development is not organic or linear. It requires an effort, whether intentional or not, to learn from new information and to reconcile it with past beliefs and values. Self-authorization is thus a thought process and not something that is measured by actions. Through sources of conflict and reflection on one's past experiences, a person learns how to move forward. What has been shown in some of the research is that the ability to reflect is not automatic and needs to be nurtured. Not only must one be able to understand one's

internal voice, but one must also learn to trust it and then reconcile it with the perceived conflicts one experiences with others (Baxter Magolda, 2008).

Third, contextual knowing is key to understanding the development of self-authorship. “Other people’s ideas continue to be important, but no longer simply as a source of ideas that coexist with the learner’s; rather, they are potential elements that, when judged to be valid, can be incorporated into the learner’s own thinking and views.” (Pascarella & Terenzini, 2005) All of the qualitative studies of self-authorship center themselves on the context of the student’s experiences, either in school, work, or with family, and how these particular situations can spur movement towards self-authorship. When faced with a situation or issue, self-authored individuals seek out expert advice but integrate it with their own views in deciding what to think (Baxter Magolda, 2004).

In summary, cognitive-structural theories help to illuminate the growth of the ways individuals construct meaning as they develop skills to question knowledge, understand their own perspectives, and relate to others in a manner that accounts for diverse worldviews. As one develops towards self-authorship, one’s ability to build higher learning skills such as personal responsibility, integrative learning, and critical analysis will be easier. Self-authorship can provide the foundation to foster development of the LEAP goals for the 21st century such as critical thinking and social responsibility. Among undergraduate students planning to enter medical school, the foundation to develop self-authorship may be different for those in an early assurance entry path when compared to the traditional premed student because of their different exposures to liberal arts courses and the requirements of medical school preparation.

Self-Authorship and Medical Education

Literature regarding the study of self-authorship specifically within the medical education setting is negligible, although there is evidence of its utility as a lens for study. Slotnick (2001) argued that the journey through medical education from student in the classroom to student in the clinical setting to graduate in residency requires the individual to model and acquire new roles at each step. Each new step incorporates new expectations upon the individual and a new set of actors (faculty, preceptors) with whom the individual must interact. However, each progressive step can also retain some of the prior expectations/roles as well. His arguments for cognitive and socio-cultural skills to help adapt to these changes were prescient of Baxter Magolda's conception of self-authorship. Once medical students start learning in the clinical arena, the multiple roles that they must progressively master also include that of doctor and student (Shuval, 1975). This creates a conflicting set of messages to the student regarding expected behaviors from teachers and patients, each of whom may have very different expectations of the student as a learner and expert, respectively. Self-authorship has been shown to provide an understanding of how meaning-making can support the perceptions of multiple identities (Abes, Jones, & McEwen, 2007). Abes, Jones, and McEwen did not study medical students, but their conception of the conflict of multiple expectations is relevant to medical education as well, considering the multiple roles students must exhibit.

In addition to medical education requiring the achievement of new roles over time, it is also an experiential process and becomes progressively self-directed as one goes through medical school and residency (Yardley et al., 2012). Yardley et al (2012) argue for an experience-based learning theory that illustrates the relative importance of

understanding the change from classroom-based lectures to the clinical setting and how students experience multiple sources of information and expectations, including the ability to “rehearse the act of a doctor” (p. e111). This interaction creates a process where the student is both learner and student, facilitated by the faculty. Applying the concept of self-authorship highlights the importance with regards to the interpersonal and intrapersonal dimensions of self-authorship as the student incorporates the multiple messages with his understanding of what it means to be a student and doctor.

Understanding the experiential environment can provide a framework to understand how students enter and experience medical school.

Medical education, especially in the clinical years, can also expose students to complicated problems where conflicting information may be presented between all of the different levels of clinicians (or even patients) who may be involved in the teaching experience. The ability to reconcile all of this information within the clinical context may be better achieved by those who are more self-authored.

Finally, within the experiential process of medical education, the literature points to the significant influence of the physician-teachers with whom the students interact throughout the education process. Teachers provide knowledge and exhibit behaviors in the clinical setting that students see and emulate. The behaviors they see can often conflict with their own beliefs and understanding of expected behavior (Shuval, 1975). The influence of physician-teachers is significant to both knowledge and identity formation as the student develops into a physician (Jarvis-Selinger et al., 2012; Yardley et al., 2012). Self-authorship is founded on the idea that a person moves from External formulas to internalized beliefs of one’s own. The significant teacher-student relationship within

medical education highlights the importance of being self-authored as a medical student internalizes his or her own beliefs and values and reconciles that with all of the individuals he or she interacts with during the educational process.

Evidence of the study of self-authorship within medical education is rare but there are compelling arguments for the utility of its study. Medical education is progressive, with each new step (science years, clinical rotations, residency) a potential starting point where the student may feel unprepared and Externally motivated to prove oneself. White, Kumagai, Ross, and Fantone (2009) reported that students who are more developed with regard to self-authorship have been found to complete their clinical rotations while remaining more patient-centered and less subject to compromising their values. The authors provided evidence to support efforts within the curriculum to develop self-authorship but fell short of providing methods to systematically measure this development within the curriculum. Similarly, arguments have been made for widening medical education beyond clinical competencies to incorporate identity formation, calling for a systematic measure of self-authorship to support these specific changes to the curriculum (Jarvis-Selinger et al., 2012). Time-consuming and resource-intensive qualitative methods have been the most common format to study self-authorship. Reliance on qualitative measures will not be sufficient if the medical education system seeks to meet these calls for change. Support to develop self-authorship along the medical education path can help improve the learning experience. The availability of some method to measure self-authorship development in a standardized, economical way can support the appropriate evaluation of curricular efforts.

Quantitative Methods to Assess Self-Authorship

The literature on methods to measure self-authorship is somewhat limited and fairly recent. Given the limited literature, it is helpful to review the empirical evidence regarding the development of assessment methods in a chronological fashion and see how it has accumulated. Prior to Baxter Magolda's introduction in 2001 of her conception of self-authorship, there were some attempts to quantitatively measure concepts that are akin to self-authorship among medical students. Researchers at the Harvard Medical School created a cognitive behavioral survey to measure epistemological beliefs among medical students when comparing three curriculums (Mitchell, 1994). This survey of second-year students included three newly created scales of learning. With 96 responses, a 71% response rate, the findings were appropriate for generalization to the student population at Harvard as part of a program evaluation, but there were not enough responses to allow for validation of the actual scales. Still, the authors were confident in their quantitative methodology to collect useful data on these concepts.

Empathy can be viewed as another characteristic of the self-authored person. One study looked at empathy and how it correlates to the quality of counseling provided by graduate counseling students (Lovell, 1999). Lovell used two previously validated psychological batteries in a national survey in which over half of the respondents had a liberal arts undergraduate degree and found that those in higher cognitive stages of growth showed more empathy. The use of a standardized quantitative measurement tool showed that this methodology could be appropriate for measuring non-cognitive skills (i.e., empathy) as well as cognitive factors. However, his research did not focus on the methodology itself and possible implications for further research.

The National Study of Student Learning (NSSL) was a three-year program from 1992-1995 to survey students' orientation to learning and its relationship to academic experiences. The standardized tools included various constructs that did not explore non-cognitive skills but did include measures on constructs such as acceptance of diversity and critical thinking; both skills could fall into the epistemological and interpersonal dimensions of self-authorship (Pascarella, 2001). However, the purpose of NSSL was to measure changes in general cognitive skills and how they were related to varied activities such as Greek life, diversity training, and volunteer work. By creating a construct on openness to diversity and trending each student over the three years, the researchers were able to identify changes in the student's development within the construct and how that might be related to other academic characteristics. Conceptually, this provides a model for using some quantitative measure of self-authorship and monitoring changes over time within a particular context. The NSSL itself is more applicable to person-environment interaction theories, but the use of a survey to measure skills related to self-authorship supports further research into the applicability of this methodology.

Baxter Magolda's (2008) research was a labor-intensive effort that supported the development of a theory as well as a specific qualitative method in which to collect the data. Starting in 1986, she followed 101 students, conducting annual semi-structured interviews. By 2000, there were still 39 former students participating each year. In her methodology, the interviewer starts the conversation by asking about important learning experiences in the past year, thus giving the interviewee the opportunity to set the context. Baxter Magolda uses a constructivist lens, allowing each person to describe their unique experiences before trying to find commonality in the data to help explain self-authorship.

The method requires significant training and practice for the interviewer to ensure the quality of the data and to ensure the interviewer understands how to properly prompt and lead the interviewee to reflect and express his or her experiences. Her qualitative methodology set the groundwork for other subsequent studies of self-authorship.

Pizzolato (2004) followed Baxter Magolda's method with a qualitative study of marginalized Hispanic students in order to understand the theory and its application to student retention and success. This study of 27 students was exploratory in nature and identified movement towards self-authorship upon entering college but found regression by these students as they initially faced challenges in college similar to the Crossroads stage as identified by Baxter Magolda.

The basis of studying self-authorship using qualitative methods continued in a study to explore ethnic identity and its relationship to self-authorship, which was not part of the original hypothesis but arose out of the coding of the interviews (Torres & Hernandez, 2007). The study supports the idea of a long-term process to measure development of students over the four years of college. In terms of concepts within self-authorship, the researchers identified the need to incorporate students' expectations and how these related to experiences and development of self-authorship. While Baxter Magolda's original work identified the influence of conflict and pressure to spur the development of self-authorship, Torres and Hernandez identified how positive experiences also promoted development. Their research among Latino students followed Baxter Magolda's method, but they did not include any explanation of how they structured their interviews, nor did they explain the challenges of working with this particular population.

Empirical evidence of quantitative measures of self-authorship appeared around this same time. For example, Wawrzynski and Pizzolato (2006) published on the development of a quantitative tool to measure self-authorship because, as they pointed out, the theory of self-authorship was built upon qualitative methods that are not practical for large-scale studies of development. In a study of 368 students at one university, a survey that included a 29-item questionnaire of self-authorship was conducted at the beginning and end of one semester. The authors acknowledged that they did not hypothesize that they would find significant growth in self-authorship in such a short time frame, but instead were looking to see how personal characteristics and school environmental factors might be related to self-authorship. The data identified a link between the students' precollege experiences (high school grades and test scores) and development on subscales of self-authorship. However, this article did not provide any indication of the actual content of the survey, nor was there an explanation of how their concepts may differ theoretically from Baxter Magolda's conception.

Pizzolato (2007) further described the challenges of developing an assessment method for self-authorship that combined a quantitative Self-Authorship Survey (SAS) and a qualitative text exercise called the Experience Survey (ES). Her goal was to develop a measure that would allow for assessment of development towards self-authorship and could support program evaluations of curricular efforts to encourage development. She found moderate ($r=.51$) correlations between scores on the constructs and ratings of the text responses. However, she noted that the text item only asked the student to write about an important experience, not specifically one in which the student was an actor and would need to reflect on his or her thought processes. Thus the language in the responses made it

difficult to identify self-authored reasoning. Pizzolato's constructs are also related to skills necessary for self-authorship, but the items themselves were not situated in any context but instead related to experiences in general.

In a study that included surveys among 467 men and women and follow-up interviews with 40 female college students, there were somewhat conflicting data regarding the epistemological and interpersonal dimensions of self-authorship (Creamer & Laughlin, 2005). The survey asked for sources of information in deciding on a career, but the responses were hard to interpret given the subsequent analysis of the interviews. The study highlighted the difficulty in writing questions that can identify the thought process involved in self-authorship as opposed to behavioral actions that are not representative of self-authorship. The most recent evidence of a quantitative measure of self-authorship is an 18-item scale that was designed to reflect constructs for both the three dimensions of authorship and the three phases of growth (Creamer et al., 2010). It was included as part of a larger survey on career development for women in information technology, and the scale was shown to have acceptable reliability and consistency. The authors included the important point that Baxter Magolda's original qualitative methodology depended on the interviewee identifying a notable experience from the past year and thus directing the conversation toward the topic to ultimately explore self-authorship. A survey of self-authorship would also need a context in which to measure the construct; in this case it was choice of a career, which may not have resonated with all of the respondents. Also, in a 3x3 matrix, the 18 items were limited in measuring each of the nine dimensional stages and self-authorship phases. However, their work did show utility to the scale when the analysis looked more globally at either the three dimensions or the three phases separately.

Although the analysis did not explore the relative importance of any of the dimensions in regard to the others, it gave equal weight to each in the end. Given the acknowledged fuzziness between the three stages of development, it is encouraging that the suggested items did produce evidence of the separate stages.

More recent work using mixed-methods to combine both qualitative and quantitative methods attempted to triangulate the measurement of self-authorship and provided an interesting perspective on assessing outcomes of liberal education.

Researchers at the Wabash National Study conducted a longitudinal mixed methods study using a national sample to measure liberal arts outcomes among students (Seifert, Goodman, King, & Baxter Magolda, 2010). In the first wave of the project, the researchers collaborated on designing the methods for the study; but then the quantitative members worked separately from the qualitative members in terms of data collection and initial data interpretation. They then jointly compared findings, identified common themes, and incorporated the qualitative data to enrich findings from the survey.

The quantitative researchers did not create any new survey items or scales to measure higher learning outcomes, including self-authorship. Rather, they used validated national surveys such as the National Survey of Student Engagement and the Wabash National Study of Student Experiences Survey. The utilized instruments included scales to measure constructs such as reasoning, moral character, and intention for lifelong learning, none of which specifically target self-authorship but many of which related to concepts within self-authorship. On the qualitative side, interviews were conducted using Baxter Magolda's suggested methodology to explore self-authorship. The findings highlighted many correlations between the themes identified in the qualitative and quantitative data,

illustrating how the two methodologies may provide richer information to measure self-authorship than using either one alone. Some of the validated scales may also be helpful in measuring concepts of self-authorship, either alone or incorporated with new scales specific to self-authorship.

Some progress has been made in the move toward creating a quantitative method to measure self-authorship. For the most part, the construct items were validated and/or substantiated through a concurrent analysis of either interviews or reviews of student writing (Baxter Magolda, 2008; Torres & Hernandez, 2007). This study sought to validate a new measure of self-authorship and to build upon these methods through the analysis of survey items and open-ended responses that were captured simultaneously, as described in the next chapter.

CHAPTER 3

METHODS

This study used a mixed-methods approach quantitatively driven with simultaneous qualitative analysis (Morse, 2010). The development and validation of the instrument occurred in two stages. Construct definition, item generation, and a pilot test occurred in the first stage, which was part of an ongoing research project at the study site. At the end of the first stage, the instrument was revised based upon findings in the pilot. Data collection using the revised instrument represents the second stage, which was the basis for this dissertation. This chapter first describes the pilot study in Stage 1, including the steps that were taken to create, analyze and refine the proposed measure of self-authorship, followed by a description of Stage 2, which represents the data collected from the revised instrument and analyzed for this study.

Pilot Study

Research Setting

The setting for this research was a medical school in the northeastern United States. The majority of students self-identify as White (53.2%), Asian (23.2%), Hispanic or Latino (12.1%), and Black or African American (6.6%); this is slightly more diverse than national averages (Association of American Medical Colleges, 2012). Just over half of the students in 2012 were male (53.3%). Similar to national averages, more than three quarters (78.2%) of applicants are from out-of-state; and 69.8% of matriculants are from out-of-state, which is higher than the 38.8% national average. All medical degree candidates from Years 1 through 4 were included in the pilot phase (N = 569). Master's degree students were excluded because they do not have the same undergraduate

requirements, nor do they follow the same curriculum in medical school. Consent to participate in the pilot included acknowledgment that the research database would include identifying information such as email address in order to allow longitudinal tracking of changes in responses over the course of the four years of medical school.

Item Development

For the pilot study in Stage 1, the team consulted telephonically with Baxter Magolda and Creamer in 2009 to discuss revising and utilizing their published tool. (see Appendix A.) Much of the previously reviewed literature accounted for identifying self-authorship qualitatively within a specific context such as career or education experiences (Baxter Magolda, 2004; Creamer & Laughlin, 2005; Seifert et al., 2010). Quantitative research on self-authorship by Creamer et al. (2010) used the context of selecting a career for the survey items, while the tool developed by Pizzolato (2007) is not context bound. We identified selection of a specialty to practice medicine as an appropriate context to develop the construct items.

Using a 3x3 matrix of the three dimensions of self-authorship (epistemological, intrapersonal, and interpersonal) and the three phases of development (External, Crossroads, and Early), the research team aligned the original 18 items from Creamer's scale and then created an additional six items so that there were at least two items per cell (see Table 1). According to Baxter Magolda and King (2007), "...assessing meaning-making structures in the journey toward self-authorship requires unearthing both object and subject aspects of meaning-making (p. 495). Therefore, the research team used item stems already in the scale or based upon the experience of physician-teachers on the research team that would add options related to object or subject where lacking in the

original scale. The team tried to have an equal number of items per dimension of self-authorship under the assumption that each could be equally important. However, it was challenging to create an equal number of items relevant to the interpersonal dimension as compared to the other dimensions. Similarly, without prior knowledge or assumptions of how medical students develop self-authorship, the team tried to allow for an equal number of items representing each phase of development. Again, it was challenging to create items to represent the Crossroads phase. As a result, there were fewer items generated representing this phase. All of the original items from Creamer (2010) were revised to reflect selection of a medical specialty where appropriate (see Appendix A).

Table 1

Number of Survey Items in Pilot Phase by Developmental Stage and Dimension

	Epistemological	Interpersonal	Intrapersonal	Total
External Formulas	4	2	4	10
Crossroads	2	2	2	6
Early Self-authorship	4	2	2	8
Total	10	6	8	24

Item Structure and Scaling

For the pilot, the self-authorship scale consisted of 24 items that were all rated on a 4-point Likert scale of strongly agree, agree, disagree, strongly disagree. The items were contextualized about choice of medical specialty in a manner similar to the original survey by Creamer et al. (2010), with stems that included “To make a good choice about a medical specialty . . . ,” “The most important role of a mentor or advisor . . . ,” and “My

primary role in choosing a medical specialty . . .” There were two additional stems that were not related to choice of specialty. They were “Experts are divided on some scientific issues such as the causes of global warming. I think . . . ,” and “When some people have different interpretations of a book, I think” Order bias occurs when the order in which survey items are presented can affect responses to subsequent questions by creating thought processes based upon initial items (Dillman, 2000). Using an online method, the questions were randomly ordered for each participant in order to remove the potential for order bias.

The pilot version of the instrument was pretested in 2011 with a group of new interns at the medical center since they were recent medical school graduates and could reflect on the process of selecting a specialty. This relates directly to establishing the reliability of the instrument, as the discussion could help identify ambiguous or confusing items that may introduce unintended information to the construct (DeVellis, 2012). Question wording was clarified on several items before data collection began. Appendix A compares the original version by Creamer, Baxter Magolda, and Yue (2010) with the final revised pilot version of the survey.

Procedures

Data collection in the pilot wave included all four current classes. The first wave of data collection began in August 2012. Email invitations with links to the online survey were sent to 569 students, with reminder emails sent to non-responders every 2-3 weeks through September. To increase the response rate, an incentive of a \$5 gift card to a coffee shop on campus was incorporated into the process after the initial three weeks of data collection. The gift cards were also provided to those who had already responded.

Analyses

Analyses after data collection for the pilot phase were twofold. First, reliability testing and exploratory principal component analysis were conducted. Second, focus groups were conducted with students to further refine the instrument. All data were analyzed using SPSS, Version 21.0.

With development of a survey instrument, it is key to determine the reliability of the instrument to ensure consistent data over time and the validity of the instrument to accurately measure its intended concept(s) (Abell, Springer, & Kamata, 2009). Reliability also refers to the extent to which an instrument differentiates among individuals accurately. It is intimately linked to the population being measured (Streiner & Norman, 2008). This was particularly relevant to this study for two reasons: (1) the instrument which was the basis for this study had never before been applied to the medical student population, and (2) the conceptual framework was based upon the hypothesis that early assurance students and traditional premed students may start medical school in different phases of self-authorship. From a quantitative perspective, the reliability was determined using Cohen's alpha internal consistency estimate to measure internal consistency of the construct. In addition, agreement coefficients were calculated using random halves criteria since a test-retest design was not possible. The two random halves were a reasonable proxy for a test-retest (Downing, 2004). If random errors proved to be low, this would be good evidence of score reproducibility.

The construct validity of a measure requires multiple sources of evidence which relate to test content, the response process, internal structure of the construct, its relation to

other variables, and consequences of testing (Downing, 2003). Internal consistency and factor analysis are sources of internal structure evidence.

Regarding factor analysis, there is significant discussion in the literature about the choice between principal component analysis and exploratory factor analysis techniques (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Russell, 2002). Principal components analysis seeks to model all of the variance and focuses on data reduction. Exploratory factor analysis techniques instead model only the shared variance between variables. Exploratory factor analysis is appropriate for instrument development and identifying underlying constructs (Wetzel, 2011). However, in many cases the difference in methods can be negligible, in which case principal components can provide a simpler structure for analysis. Child (2006) in particular recommends starting with a principal components analysis and then using an exploratory factor analysis technique for verification of the initial findings. The pilot study used principal components analysis only, and the research in this study also used principal components.

The set of potential variables should have a good range of response choices which are at equal appearing intervals and are normally distributed (Child, 2006). The data should also be sufficiently correlated to warrant factor analysis as a method to identify an underlying structure. The data from the pilot test were analyzed with the Kaiser-Meyer-Olkin measure of sampling adequacy, which identifies whether the number of data records is factorable in relation to the number of variables, and Bartlett's test of sphericity, where the null hypothesis is that the variables are uncorrelated. Significant results for each test would indicate that the items are sufficiently correlated and factor analytic techniques are justified.

When interpreting principal component or exploratory factor analysis results, a salient variable is defined as one that has a sufficiently high loading and ensures a relationship between that variable and the component (Gorsuch, 1983). Guidelines for interpretation can vary depending on the level of significance one desires and the size of the dataset. Gorsuch included guidelines for interpretation such as using a minimum load of .4 if a sample is 100 and a minimum load of .3 when a sample is at least 175. The decision on an appropriate cutoff can be arbitrary depending on the circumstances of the research (Abell et al., 2009). This study used a more conservative .400 cutoff for the interpretation of principal components.

Initial Assessment of Self-Authorship Component Structure

Principal component analysis was performed using Promax rotation of the component loadings. Promax is an oblique rotation that allows for correlated factors. “American psychologists... believe that most behavioral characteristics of human beings are so interrelated that we should allow for this in any kind of analysis used.... In other words the resulting factors may be correlated” (Child, 2006). Components were identified using two criteria: (1) eigenvalues greater than 1.0 and (2) a visual review of a Cattell’s scree plot. The eigenvalue rule alone might be too loose in allowing factors, especially when discriminating between factors just above or below the 1.0 threshold. Applying the scree test involves some reliance on subjective criteria when reviewing factor interpretability and the visual drop off of the plot (DeVellis, 2012).

The pilot wave of data collection elicited 165 responses (25% response rate). Low response rates alone are not biasing so long as the respondent characteristics are representative of non-respondents (Dillman, 1991). However, Table 2 shows that

respondents' characteristics significantly ($p < .05$) under-represented females and White students. Still, within the responses, there were 42 early assurance students, allowing for some preliminary analysis of early assurance and traditional premed students.

Table 2

Representative Characteristics as a Percentage of Respondents and Population

Characteristic	Survey Respondents ($n = 149$)	School Population ($N = 569$)
Early assurance students	24.2	16.9
Female	55.7	46.7
White	66.0	53.2

Univariate analysis of the responses to each question identified a range of responses to nearly all items from 98% strongly agree or agree to 1% strongly agree or agree over the 24 items, visually confirming that there was variation among the responses. All variables had kurtosis $< |7|$ and skewness $< |2|$, so none were deleted because of significant non-normality (Fabrigar et al., 1999). Using the guideline of a minimum 5 records per variable and 100 total records, the data were determined to be sufficient for principal component analysis methods (Child, 2006; Gorsuch, 1983).

Internal consistency of the construct was measured using Cronbach's alpha internal consistency estimate to check the reliability of the scale and the value was .755, which is acceptable for a formative assessment (Downing, 2004). Removing any of the items would improve the alpha by less than $\pm .005$; thus, no items were deleted. Reliability was tested using split-halves. The correlation between halves was moderate at .569 with Cronbach's alpha of .618 for one half and .609 for the other half.

The Kaiser-Meyer-Olkin index of sampling was .736 for the pilot, indicating that the data were a homogenous collection suitable for factor analysis. Bartlett's test of sphericity was significant ($\chi^2 = 822.694, p < .001$), indicating that the inter-item correlations were significantly different from 0 and suitable for factor analysis.

The principal component analysis with Promax rotation of 24 items produced eigenvalues that showed six factors which met the rule for an eigenvalue greater than 1, although the sixth factor just met that criterion. However, review of the scree plot showed that four factors were more reasonable to interpret. (see Figure 3.)

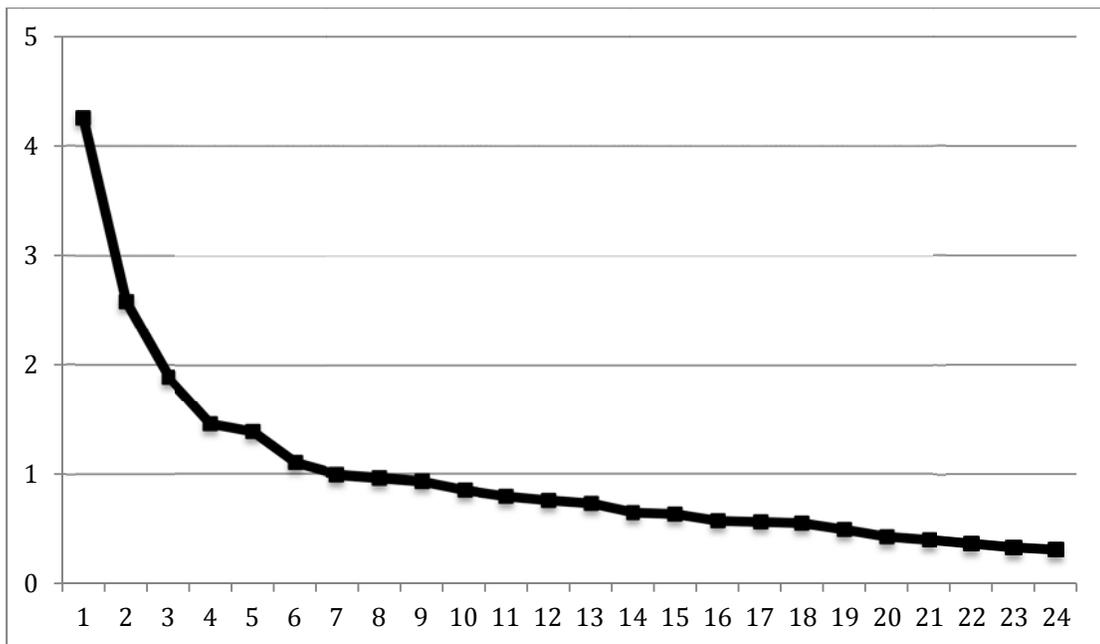


Figure 3. Scree Plot for Exploratory Factor Analysis Using Principal Components

Four components were selected as being most salient; the fifth and sixth components were not selected based upon review of the scree plot and the fact that the variables that loaded on these two components could not be easily interpreted as

constructs. The loadings for the four emergent components are included in Table 3. For interpretation purposes, the more conservative minimum of .40 was used for identifying variables on each component. The four components related to one dimension and the three phases of self-authorship. They were named based upon the proposed structure of the items by phase and dimension. The first component was interpreted as the External phase. The second component was called Epistemological since the items fell under that predesigned dimension. The third component was called Crossroads and the fourth was called Early, both for the respective dimensions represented by the items with high loadings. The analysis did not identify salient factors relating to the intrapersonal and interpersonal dimensions. The inability to identify factors related to the interpersonal and intrapersonal dimensions was most likely related to the lower number of survey items for these dimensions in the pilot instrument. In addition, two items, “When people have different interpretations of a book, I think only one interpretation can be right” and “I am in the best position to know when I’ve found the right specialty for myself,” did not load on any factor. These two items were excluded from the research version of the instrument.

Table 3

Component Loadings Using Principal Components with Promax Rotation

Survey Items	Component Loadings			
	1	2	3	4
Q1	.697	-.081	.118	-.059
Q2	.501	-.195	.045	-.026
Q3	.425	.101	.223	-.406
Q4	.421	.350	.070	.121
Q5	.276	-.214	.140	.012
Q6	-.157	.221	.652	.035
Q7	.210	-.350	.613	.089
Q8	-.032	-.030	.880	-.187
Q9	-.272	.055	.388	.503
Q10	.012	.570	.109	-.083
Q11	.311	.308	.044	-.050
Q12	.218	.108	.284	.339
Q13	.587	-.126	.009	.218
Q14	.024	.345	.145	.076
Q15	.268	-.078	.361	-.049
Q16	.711	.072	-.279	.167
Q17	.302	.703	-.191	-.206
Q18	-.245	.639	.180	-.160
Q19	.345	.218	-.006	.146
Q20	.093	.003	-.005	.518
Q21	.135	-.116	-.147	.798
Q22	-.079	.831	-.113	.114
Q23	.226	.382	-.021	.399
Q24	-.201	.123	.183	.082
Eigenvalues	4.255	2.581	1.886	1.457
% variance	17.73	10.76	7.86	6.07
Cumulative variance		28.49	36.35	42.42

Note. Factor loadings greater than .40 are included in bold.

Table 4 shows moderate correlation among the four factors, substantiating the use of a Promax oblique rotation and the theoretical assumption of correlation between the phases and/or dimensions.

Table 4

Correlation Matrix for Pilot Study Principal Component Analysis with Promax Rotation

Component	1	2	3	4
1	1.000	.187	.238	.124
2	.187	1.000	.293	.226
3	.238	.293	1.000	.328
4	.124	.226	.328	1.000

In addition to the statistical analysis of the data, additional focus groups were conducted after the pilot phase with current fourth year and master's degree students to review the survey instrument and identify any further potential interpretation issues with individual items. While the first, second, and third year students would be asked to complete the survey again in the fall of 2013, the fourth year students were an ideal group to use for this type of information since they were graduating and would no longer be included in the study. The master's students were also appropriate since they were enrolled in a separate parallel program that was not part of this research study. Their review and discussion of the survey would not create any potential bias to future responses.

The focus groups with these two populations highlighted wording changes to clarify items, in particular the change of the stem "My primary role in choosing a medical specialty . . ." to "When choosing a medical specialty, it is most important . . ." The original random order of items in the pilot phase was judged to be confusing as well

because several items would have the same stem and different conditional statements occurring in random order. The randomness of the order made the items seem to be repetitive instead of variations on the stem. The participants suggested that items with a common stem be grouped together to remove the perception that the question had already been asked. Other wording changes included the following:

- Reference to people having different interpretations of a book did not resonate with students; therefore, the referent was changed to “different interpretations of a journal article,” which was considered more relevant to the medical student population.
- The stem “My primary role in choosing a specialty” caused confusion because the students felt every item choice with this stem would be their role. Through discussion, the students agreed that “When choosing a medical specialty, it is most important to . . .” would be a better stem for the same choices.

Research Project

This dissertation research was based on Stage 2 of the original study. After completion and analysis of the pilot study, data were collected in 2013. The methods described here and the analysis in Chapter 4 represent the data collection in 2013, using a revised set of items related to self-authorship and the analysis of responses to an open-ended question to follow-up on choice of a medical specialty.

Sample Validation

For the 2013-2014 academic year, the sample consisted of all medical degree students in Years 1, 2, 3, and 4 (N = 569). Master’s degree students were excluded. There

was no identifying information collected due to changes in the protocol, as described below.

Modification of the Instrument

Based upon the component analysis from the first stage, two items were deleted from the instrument. Because of the limited number of proposed items related to the intrapersonal and interpersonal dimensions, four new items were added to the survey. Taking into account Baxter Magolda and King's (2007) suggestion that items represent the concept of subject and object as theorized by Kegan, items were added in order to ascertain that the scale represented both concepts of subject and object. The items were the following:

- I know what I want to personally achieve from the specialty I choose to practice.
- I have been able to remain true to myself during the past year.
- The most important role of an effective mentor or advisor is to challenge a student to clarify expectations of a medical specialty.
- To make a good choice about a medical specialty, I think that advisors or mentors can provide advice to consider along with my own views.

This instrument consisted of 28 items. This improved the number of potential items per hypothesized factor to the minimum 5-6 item rule suggested by Gorsuch (1983). The distribution of items by dimension and phase of self-authorship is included in Table 5 and the final items are displayed in Appendix B. Two of the items (“When people have different interpretations of a journal article, only experts can say which interpretation is really correct” and “ Experts are divided on some scientific issues, such as causes of

global warming. I rely on the experts to decide”) were negatively worded so that disagreement would imply higher development of self-authoring behavior.

In addition, a revised free text question was added to the survey as follows:

In a couple of paragraphs, please describe the approach you have taken, or plan to take, to reach a decision about what medical specialty you will practice. What was/is your thought process in thinking about a specialty to practice?

One blank 8 ½” x 11” page was included for the response. The large open space encouraged responses by not displaying a limited space to write an answer. Unlike the ES tool used by Pizzolato (2007), this open-ended item situated the respondent in the context where he or she is an actor and should be able to reflect on this or her experiences.

Typically, a trained interviewer who is exploring self-authorship will direct a conversation by often stating, “Help me understand that experience . . .” (Baxter Magolda & King, 2007). The open-ended question sought to do this same prompt through its direction to ask for a thought process.

Table 5

Number of Survey Items in Study Phase by Developmental Stage and Dimension

	Epistemological	Interpersonal	Intrapersonal	Total
External Formulas	6	2	2	10
Crossroads	3	2	2	7
Early Self-authorship	5	3	3	11
Total	14	7	7	28

Survey Protocol

Due to the challenges of achieving a substantial response rate using an online survey in the pilot phase of data collection, the data collection process was changed to a paper-based survey that was administered to all four classes at the beginning of the 2013-2014 academic year. The survey took approximately 15-20 minutes to complete depending on the time spent on the one open-ended question. Data were collected during scheduled meetings when all students were in attendance. For the first-year students, this was during their orientation week; for the second-year students, this was during their first week of classes. The third-year students completed the survey during the orientation week prior to starting their clinical rotations, and the fourth-year students completed the survey during a mandatory class meeting at the beginning of the academic year. A medical student handed out the surveys and collected the responses. The change to a paper format eliminated any identifying information to allow for tracking of individuals over time. However, students could write their name at the end of the survey if they wished to allow the researchers to track their responses over time.

Data Analysis

The methods for data analysis of the data in Stage 2 are similar to those described for Stage 1 previously. Reliability of the construct was determined using Cronbach's alpha internal consistency estimate, using a minimum of .700 as an acceptable level (Nunnally & Bernstein, 1994). Agreement coefficients were also calculated using random halves criteria. The Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity were both calculated to ensure that the data were sufficient for principal component analysis. Once principal component analyses were conducted, both the

eigenvalue greater than one rule and Cattell's scree plot were used to determine the number of components. Following Gorsuch's (1983) guidelines for interpretation, a minimum load of .400 was used as criteria to identify salient items on each component.

Data were collected during scheduled meetings when all students were in attendance; students were handed the surveys before the sessions, and the completed surveys were collected after the respective session. The surveys were data entered and the narrative responses typed by the researcher. Responses were entered on a 1 to 4 scale (1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree). Two items ("When people have different interpretations of a journal article, only experts can say which interpretation is really correct" and " Experts are divided on some scientific issues, such as the causes of global warming. I rely on the experts to decide") were recoded in reverse order (1=strongly agree, 2=agree, 3=disagree, 4=strongly disagree). These two items were negatively worded such that the desirable response to identify development of self-authorship was clearly to disagree.

Reliability and validity testing of a new instrument were the focus of this project. Internal consistency of the scale and subscales was measured through Cronbach's alpha consistency estimate, and reliability was measured through split-half analysis. Data reduction was first performed using principal components analysis and Promax oblique rotation. Exploratory factor analysis was also performed, using principal-axis factoring and Promax oblique rotation as a final verification, as suggested by Child (2006).

Principal component analysis of the data by what were possible subgroups was limited for the early assurance students due to a small number of responses. However, studies have shown that factor analysis can be conducted with as little as 30 records under

certain conditions (MacCallum, Widaman, Zhang, & Hong, 1999; Preacher & MacCallum, 2002). First, the communalities of each variable should be high ($> .6$) and second, the number of factors should be small, but not too few that they will negate the effect of the high communalities. This second condition reflects directly on the ability to properly interpret the factors and was used as the criterion to determine if principal component analysis was appropriate for the early assurance population.

Determining the validity of the instrument provided a more significant challenge. Validity requires multiple sources of evidence to support or refute the interpretation of findings (Downing, 2003). Validity can include a measure of the relationship between the studied measure and other variables. This implies that there is some alternative measure upon which to corroborate the findings from the new conceptual measure in the proposed instrument. However, with vague concepts such as self-authorship, there are no correlated data to substantiate the findings. As an alternative, it was possible to use the responses to the open-ended question in the survey to validate the measure.

Content analysis of the open-ended responses was used as a measure of criterion validity. Analysis of qualitative data can help by validating, interpreting, and illustrating quantitative findings (Miles & Huberman, 1994). Constructive grounded theory was used to analyze the responses, allowing for exploration of how the students constructed meaning (Charmaz, 2006). While Baxter Magolda's work has already described the theoretical dimensions and phases of self-authorship, the theory did not drive the coding of responses. Instead, the theoretical codes were applied in relation to what emerged from the data, not forcing the responses to fit into specific categories (Walker & Myrick, 2007). When identifying phases while reading the comments, the focus was mainly to determine

if any responses identified the extreme (External or Early) phases since these are easier than the Crossroads to identify. Responses were compared and contrasted in order to identify trends and differences among the responses for all respondents (Boyatzis, 1998). This is important given the lack of specificity in identifying one phase of self-authorship from another.

Subsequent to the identification of principal components, component scores were calculated for each student. The univariate distribution of scores was then reviewed to determine the mean scores and standard deviations of each component. Given the lack of specificity in moving from one phase of self-authorship to another, those who scored more than +/- one standard deviation of the mean were used as proxies for the Early and External phases, respectively. Once the text data was coded, the coding was linked to the respondent to determine if they were “scored” as Early or External on a component, allowing for triangulation of the data as a measure of validity.

Summary

This chapter outlines the methods used in this two-stage research study. In Stage 1, which was a pilot test of the proposed instrument, the analysis of reliability and validity led to revisions of the instrument and the collection protocol. Stage 2, which represents the data for this project, included the determination of reliability and validity of the proposed instrument, both quantitatively and through qualitative analysis of text responses. Chapter 4 provides the findings of the analyses of the data in Stage 2.

CHAPTER 4

RESULTS

This chapter summarizes the findings from the data analysis to determine the reliability and validity of a scale to measure self-authorship for medical students. The first part of this chapter provides a quantitative analysis of the survey responses to determine the reliability and validity of the self-authorship scale in response to the first research question. Data are examined in aggregate and separately for traditional premed students and early assurance students. The second part of the chapter integrates a qualitative analysis of the narrative responses to support the validity of the scale as described in the second research question.

Descriptive Statistics

Overall, 271 students completed the survey, representing a 65% response rate based upon the surveys distributed to the four academic classes. Subsequent review of the surveys revealed that most of one class was provided with the incorrect version of the survey and thus had to be eliminated. The final sample was reduced to 199 responses. Table 6 displays the demographic characteristics of the respondents. Of these 199 respondents, 46% were female and 54% male. More than half (54%) of the students self-identified as White (54%), followed by Asian (30%). The most common entry path into medical school was traditional premed (75%), followed by early assurance (18%), and MD/PhD students (8%). The MD/PhD students ($n=15$) were not analyzed as a separate sub-population because of the small size of the cohort.

Table 6

Demographic Characteristics of Study Respondents

Characteristic	Survey Respondents (<i>n</i> = 199)
Entry Path:	
Early assurance students	17.7
Premed Students	74.7
MD/PhD	7.6
Female	45.7
Male	54.3
Age:	
20 – 23	37.7
24 – 26	44.7
27 - 30	14.6
31 or older	3.0
Race:	
White	52.8
Black or African American	8.3
American Indian or Alaskan Native	0.5
Asian	30.1
Native Hawaiian or other Pacific Islander	0.0
Other	4.1
From multiple races	4.1

The respondents are similar to the student population on several key demographic variables. Table 7 provides the demographic characteristics of the respondents and comparative demographic data for the entire student population on those items which were available.

Table 7

Representative Characteristics of Study Respondents as a Percentage of Respondents and Population

Characteristic	Survey Respondents (n = 199)	School Population (N = 569)
Entry Path:		
Early assurance students	17.7	16.9
Female	45.7	46.7
Race:		
White	52.8	53.2

Preliminary review of the data included the frequency distribution of each variable as well as tests for non-normality. None of the variables presented as a bimodal distribution, and skewness and kurtosis values were in acceptable ranges with skewness <2.0 and kurtosis < 7.0 (Kim, 2014). Skewness ranged from -1.1 to .65, M=-.18 and kurtosis ranged from -.79 to 1.91, M=.01.

All items were renamed for tabular display in the research phase of this dissertation research. The name consists of a combination of the dimension and phase of self-authorship hypothesized for each item; similar names were then numbered consecutively. The dimensions were abbreviated as Intra (intrapersonal), Inter

(interpersonal) and Epis (epistemology). The phases were abbreviated as X (External), C (Crossroads) and E (Early). Thus, the first item related to the interpersonal dimension and the External phase was labeled Inter-X1. All names and survey items are included in Appendix B.

For data entry, response options were given numerical values: strongly disagree=1; disagree=2; agree=3; and strongly agree=4. These values implied ordinal data and assumed equal distance between each of the response options. Items Epis-X4 and Epis-X5 were reverse coded (strongly disagree=4; disagree=3; agree=2; and, strongly agree=1) because the nature of the items would suggest disagreement as self-authorization developed. Table 8 shows the mean and standard deviation for each survey item in total and separately for the premed and early assurance students. In total, the means ranged from 1.76 for Epis-X3 to 3.65 for Intra-C2. The range of highest and lowest mean scores represented the same two variables for the premed and early assurance students, with no statistically significant difference in values. Appendix C displays the correlation and covariance matrix for the whole data set.

Table 8

Mean and Standard Deviation by Survey Item and Study Cohort

Item	All Students		Early Assurance		Premed	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Inter-X1	2.34	.64	2.40	.65	2.34	.66
Inter-C1	3.27	.65	3.11	.63	3.36	.61
Inter-C2	3.16	.61	3.17	.51	3.15	.61
Inter-X2	3.24	.67	3.20	.68	3.23	.66

Inter-E1	2.91	.68	2.89	.63	2.95	.72
Intra-X1	3.36	.68	3.26	.70	3.40	.64
Intra-X2	3.19	.65	3.06	.68	3.21	.64
Intra-C1	3.52	.60	3.46	.66	3.52	.61
Intra-E1	3.51	.55	3.43	.61	3.56	.52
Epis-C1	2.95	.66	3.09	.70	2.91	.64
Epis-X1	2.91	.65	2.83	.79	2.94	.63
Epis-X2	2.59	.63	2.46	.61	2.61	.66
Epis-E1	3.52	.59	3.54	.61	3.49	.59
Epis-E2	3.41	.54	3.37	.69	3.44	.50
Epis-C2	2.58	.77	2.71	.62	2.55	.82
Epis-E3	3.24	.58	3.09	.61	3.30	.56
Epis-X3	1.76	.65	1.54	.61	1.78	.63
Epis-X4	3.13	.69	1.74	.66	1.90	.69
Epis-E4	3.13	.55	3.23	.55	3.13	.54
Inter-C3	2.88	.53	2.94	.48	2.87	.55
Inter-E2	3.21	.47	3.06	.34	3.24	.45
Epis-E5	3.10	.67	3.00	.69	3.14	.64
Epis-E6	2.87	.71	2.91	.61	2.87	.69
Epis-X5	2.80	.70	2.20	.63	2.24	.73
Epis-X6	1.89	.67	2.00	.69	1.87	.67
Intra-C2	3.65	.52	3.60	.55	3.66	.52
Intra-C3	2.97	.73	3.17	.75	2.89	.72
Intra-E2	3.18	.65	3.17	.66	3.16	.68

Reliability

Cronbach's alpha internal consistency estimate for the 28 survey items was .658, with item-total statistics indicating that five variables did not contribute to the internal

consistency of the scale. After these five variables (Inter-X1, Epis-X3, Epis-X4, Epis-X5, Epis-X6) were deleted, alpha increased to .724, which is considered acceptable (Nunnally & Bernstein, 1994). The internal consistency of these variables for the subpopulations also was acceptable at .715 for the early assurance students and .711 for the premed students.

Test-retest reliability was measured using Guttman’s split-half method. Values were moderate for these measures and were sensitive to the order in which the variables were entered into the computer analysis. The order of the variables was randomized for each analysis. Table 10 displays the coefficient alpha and split-half statistics for the total population and the subgroups.

Table 9

Reliability Statistics in Total and by Subgroup

	α level	Split-half
Total	.724	.780
Early assurance	.712	.768
Premed	.711	.770

Principal Component Analysis

A principal component analysis with Promax rotation using the 23 variables was conducted as the initial method of data reduction. The Kaiser-Meyer-Olkin (KMO) index of sampling adequacy was .659 for the full sample, which is acceptable but just below the .7 level, which is a more moderate indication of homogeneity among the sample variables (Dziuban & Shirkey, 1974). Subsequent review of the inter-item correlations identified two variables (Epis-C1 and Epis-E6) that had values less than .20 across all of the other

variables. Variables with low inter-item correlation across all other variables may perform poorly in factor analysis; thus, these two variables were deleted from the model (Floyd & Widaman, 1995). In addition, items Intra-C3 and Intra-E2 had low correlations, less than .20 with all other variables except each other. These two variables were also deleted from the model due to their low covariance beyond each other.

After deleting the four items, a principal components analysis with Promax rotation using 19 variables achieved a KMO index of .693; and the Bartlett's test of sphericity was significant ($\chi^2 = 614.7, p < .001$), indicating sufficient correlation among the variables for a factor analysis.

Using the guideline for component selection of including components with an eigenvalue >1 , there were six components identified from the model. However, a review of Cattell's scree plot suggested only two components (see Figure 4). The component loadings after rotation show that four of the components are over-determined with at least three variables loading above .400, which is desirable (MacCallum et al., 1999). The items with large loadings that cluster on the first component suggest that it represents the interpersonal dimension of self-authorship, particularly how it relates to the relationship with a mentor. The items that load on the second component suggest that it represents the Early phase of self-authorship across all of the dimensions. Intra-C2 also loads on this component. While it was theorized to represent the Crossroads phase, this item's concept of accepting that beliefs may change over time because of experiences could also be seen as indicative of the Early phase. Conceptually, the difference would be that the person believes the change could be consciously made after considering the experience versus the change could be made based on external influences without consideration for one's

beliefs. Items that load highly on the third component suggest it represents the intrapersonal dimension. Again, one item (Epis-E4) was originally conceived as representing epistemology but could be better categorized as interpersonal because the item truly focuses on “what makes more sense to me,” or what are the person’s beliefs. Items loading on the fourth factor suggest that it represents the epistemological dimension. However, in this case there is one item (Inter-C1) that does not easily fit into an alternative definition that could apply to this dimension. The fifth and sixth components each had just two variables that loaded on them, falling below the standard for over-determination. Given the fact that neither component provides clear interpretation over and above the previously identified factors, these last two factors were excluded and a final model with four factors was set. A review of the component correlations in Table 11 illustrates that there is moderate correlation among the factors, substantiating the use of an oblique rotation method. In particular, there were positive correlations among the three components representing dimensions of self-authorship, as theorized by Baxter Magolda (2008).

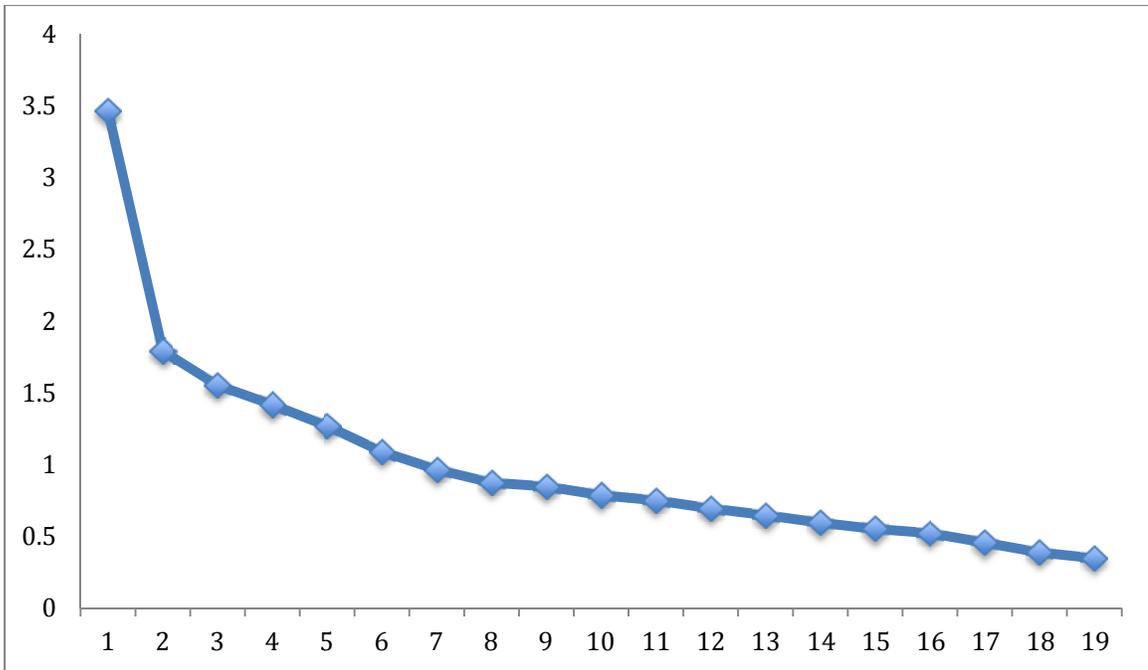


Figure 4. Scree Plot for Principal Component Analysis with Promax Rotation Using 19 Variables

Table 10

Component Loadings Using Principal Components Analysis with Promax Rotation

Rotated Component Loadings						
Survey Items	1	2	3	4	5	6
Epis-C2	-.214	-.095	.207	.009	.047	.906
Epis-E1	.210	.445	-.052	-.384	-.158	.158
Epis-E2	-.014	.681	.032	.161	.116	.080
Epis-E3	.212	.045	-.270	.075	.096	.498
Epis-E4	.107	-.027	-.420	.233	.314	.226
Epis-E5	-.117	.482	-.112	.412	-.085	.054
Epis-X1	.034	-.024	.093	.710	-.314	.163
Epis-X2	-.067	-.030	.186	.658	.185	-.076
Inter-C1	.456	.001	.069	.401	-.183	-.098
Inter-C2	.649	.021	.103	.131	.019	-.113
Inter-C3	-.117	-.057	.187	-.165	.894	.117
Inter-E1	.718	-.086	.114	.038	-.142	-.089
Inter-E2	.122	.451	-.011	.111	.496	-.172
Inter-X2	.834	-.156	.024	-.272	.089	.076
Intra-C1	.071	.266	.600	.074	-.067	.201
Intra-C2	-.219	.770	-.024	-.090	-.035	-.181
Intra-E1	.057	.555	.369	-.141	.038	.030
Intra-X1	.146	-.023	.683	.016	.259	-.036
Intra-X2	.063	-.108	.761	.327	.057	.037
Eigenvalues	3.463	1.791	1.552	1.417	1.268	1.090
% variance	18.23	9.43	8.17	7.46	6.68	5.73
Cumulative variance		27.66	35.83	43.29	49.97	55.70

Note. Loadings greater than or equal to .400 indicated in bold.

Table 11

Component Correlation Matrix after Using Principal Components Analysis with Promax Rotation

		Correlations/Component					
Component		1	2	3	4	5	6
1	1.000	.370	.137	.224	.104	.267	
2	.370	1.000	.090	.121	.147	.213	
3	.137	.090	1.000	-.016	-.092	-.053	
4	.224	.121	-.016	1.000	.169	.074	
5	.104	.147	-.092	.169	1.000	.050	
6	.267	.213	-.053	.074	.050	1.000	

Table 12 shows the final results from a principal component analysis with Promax rotation after removal of the variables in the fifth and sixth components that did not load significantly on any other component (Epis-C2, Epis-E3, Inter-C3). The KMO statistic is .737 in this final model. The four components can be interpreted similarly to the previous model although in different order. The first component loads on items relevant to the Early phase of self-authorship. The second component relates to the interpersonal dimension, and the third component relates to the intrapersonal dimension. The fourth component relates to the epistemological dimension, although it has only three items that load above the .400 threshold.

Review of the component correlation matrix in Table 13 confirmed the use of the Promax oblique rotation given the moderate correlations among the components. Finally, Cronbach's alpha internal consistency estimates for the four components were low to moderate at .614, .621, .534, and .320, respectively.

Table 12

Rotated Components Using Principal Components Analysis with Promax Rotation – FINAL MODEL

Survey Items	Rotated Component Loadings			
	1	2	3	4
Epis-E1	.357	.320	-.064	-.445
Epis-E2	.760	.011	-.003	.104
Epis-E4	.261	.169	-.577	.285
Epis-E5	.524	-.072	-.147	.332
Epis-X1	.005	.105	.028	.648
Epis-X2	.116	-.124	.128	.679
Inter-C1	.003	.454	.020	.367
Inter-C2	.057	.582	.092	.134
Inter-E1	-.127	.696	.103	.030
Inter-E2	.634	.017	-.038	.142
Inter-X2	-.104	.846	-.036	-.223
Intra-C1	.209	.118	.601	.044
Intra-C2	.666	-.274	.090	-.187
Intra-E1	.489	.050	.420	-.186
Intra-X1	.025	.113	.616	.079
Intra-X2	-.094	.020	.742	.355
Eigenvalues	3.384	1.695	1.473	1.306
% variance	21.15	10.59	9.21	8.16
Cumulative variance		31.74	40.95	49.11

Note. Loadings greater than or equal to .400 indicated in bold.

Table 13

Component Correlation Matrix after Using Principal Components Analysis with Promax Rotation – FINAL MODEL

Component	Correlations/Component			
	1	2	3	4
1	1.000	.343	.181	.061
2	.343	1.000	.247	.189
3	.181	.247	1.000	.066
4	.061	.189	.066	1.000

When the research goal is to identify underlying constructs from a battery of survey items, it is recommended to do an exploratory factor analysis (EFA) as well as a principal components analysis (Child, 2006). An EFA seeks to best fit the shared variance among the variables that can be related to the construct as opposed to all of the variance (shared and unique) that is modeled in a principal components analysis. It has been shown that very often the two models can have similar findings, in which case the simpler model can be used (Child, 2006).

A subsequent EFA using principal-axis factoring and Promax rotation identified nearly the same set of factors and factor components. Table 14 displays the results from the EFA model. While the factors were interpreted similarly to the principal component analysis, in all cases the factors were not over-determined. The first factor was interpreted as relating to the Early phase of self-authorship, the second factor related to the intrapersonal dimension, and the third factor to the interpersonal dimension. The last factor regarding the epistemological dimension loaded on just two variables in this model. The PCA model provided a more robust estimate of this scale.

Table 14

Factor Loadings Using Principal-axis Factor Analysis with Promax Rotation – FINAL MODEL

Survey items	Factor Loadings			
	1	2	3	4
Epis-E1	.332	-.047	.201	-.239
Epis-E2	.712	-.039	-.047	.174
Epis-E4	.125	-.326	.054	.273
Epis-E5	.327	-.111	-.033	.287
Epis-X1	-.076	.034	.076	.502
Epis-X2	.012	.105	-.095	.501
Inter-C1	-.013	.023	.313	.331
Inter-C2	.050	.046	.448	.171
Inter-E1	-.096	.084	.543	.075
Inter-E2	.471	-.038	.015	.168
Inter-X2	-.065	-.046	.742	-.124
Intra-C1	.223	.474	.101	-.029
Intra-C2	.503	.010	-.147	-.106
Intra-E1	.501	.328	.032	-.189
Intra-X1	.052	.470	.101	.000
Intra-X2	-.125	.834	-.070	.203
Eigenvalues	3.384	1.695	1.473	1.306
% variance	21.15	10.59	9.21	8.16
Cumulative variance		31.74	40.95	49.11

Note. Loadings greater than or equal to .400 indicated in bold.

Sub-group Analyses

Principal component analysis identified four factors for the entire respondent population. The first research question of this study sought to determine the validity of the measure of the self-authorship construct in total and for two sub-populations: the early assurance students and the traditional premed students. Principal component analyses were

conducted separately for each sub-group starting with the 16 survey items from the overall principal component analysis.

For the traditional premed students (n=142), the KMO statistic was .668, just below the .70 level, and the Bartlett's test of sphericity was significant ($\chi^2 = 392.9$, $p < .001$). A principal component analysis with Promax oblique rotation was conducted using the 16 variables identified from the full model. The model identified four factors using the rule of eigenvalues greater than 1 as well as a visual review of the Cattell's scree plot.

The component interpretation of the results in Table 15 was similar to that for the entire study population. The first component was interpreted as the Early phase of self-authorship. As in the overall analysis, item Intra-C2 ("My choice of medical specialty may change over time based upon my experiences in medical school") was originally hypothesized to represent the Crossroads phase but could also be considered representative of the Early phase, as it suggests an understanding that experiences may be interpreted during the thought process to choose a specialty and not just that the experience would cause a change. The second component represented the epistemological dimension. Item Inter-C1 did not directly relate to epistemology but could be interpreted as a source of knowledge as well as the desired relationship with a mentor. The third component represented the interpersonal dimension. While three items met the stricter .400 cutoff for load interpretation, item Inter-C1 fell just short at .379 and was still influential to the component. Finally, the fourth component was interpreted as the intrapersonal dimension.

Table 15

*Component Loadings Using Principal Components Analysis with Promax Rotation –
PREMED STUDENTS*

Survey Items	Component Loadings/Component			
	1	2	3	4
Epis-E1	.501	-.322	.318	-.066
Epis-E2	.734	.256	-.043	.046
Epis-E4	.161	.447	.134	-.436
Epis-E5	.296	.480	-.027	-.099
Epis-X1	-.117	.665	-.033	.168
Epis-X2	.010	.634	-.151	.128
Inter-C1	-.070	.506	.379	-.040
Inter-C2	.042	.246	.495	.150
Inter-E1	-.104	-.007	.766	.107
Inter-E2	.523	.261	.031	-.014
Inter-X2	-.029	-.142	.846	-.044
Intra-C1	.199	.107	.094	.628
Intra-C2	.722	-.064	-.216	-.127
Intra-E1	.616	-.203	-.007	.402
Intra-X1	-.022	-.067	.111	.630
Intra-X2	-.094	.206	-.043	.790
Eigenvalues	3.107	1.896	1.610	1.360
% variance	19.42	11.85	10.07	8.50
Cumulative variance		31.27	41.34	49.84

Note. Loadings greater than or equal to .400 indicated in bold.

For the early assurance students, there were 35 responses that could be used for factor analysis. Normally, this would violate the suggested guidelines of at least five responses per survey item with a recommended minimum of 100 in total (Gorsuch, 1983). However, research has shown that with just 30 responses it is possible to conduct a factor analysis and have good factor recovery as long as the model seeks to identify fewer factors which are well determined (Preacher & MacCallum, 2002). In particular, another desired

quality for a small sample study is highly over-determined factors having at least three to four variables per factor (MacCallum et al., 1999). The items should also have high communalities, averaging at least .6 to offset the sample size under 100 records (MacCallum et al., 1999). The initial communalities for the early assurance student data are shown in Table 15.

Table 16

Communalities for Principal Component Analysis – Early Assurance Students

Survey Item	Extraction
Epis-E1	.794
Epis-E2	.570
Epis-E4	.661
Epis-E5	.883
Epis-X1	.731
Epis-X2	.728
Inter-C1	.855
Inter-C2	.868
Inter-E1	.901
Inter-E2	.799
Inter-X2	.702
Intra-C1	.859
Intra-C2	.869
Intra-E1	.694
Intra-X1	.679
Intra-X2	.800

A principal component analysis with Promax rotation was again conducted, using all 16 variables for the early assurance students. The KMO statistic was .356, well below the desired cutoff of .700, indicating the data were not appropriate for component analysis. A review of the initial reliability analysis identified that while the reliability coefficient

using all of the variables in the full model was sufficient at .712, the removal of three other variables (Epis-X1, Epis-E1, Epis-E5) raised the reliability coefficient to .748. Principal component analysis was again run on the 13 variables and the KMO statistic was .606, which is still low compared to the desired cutoff. Given the small sample and relatively few construct items to model, no further analysis was done on this sub-population.

Qualitative Analysis of Open-Ended Responses

Overall, 159 (86%) students responded to the open-ended question in the survey. The question specifically asked for “a couple of paragraphs” about one’s thought process in thinking about a specialty, and the level of detail in the responses varied widely from one sentence to several paragraphs. In general, most described some type of behavior(s) they would exhibit over the course of medical school to help them decide on a specialty to practice. However, the majority related to the behaviors without any indication of how these would be integrated into any thought process, as in the following examples:

- “Talk to people in the fields I might be interested in, from residents to attendings”
- “Hoping third-year experiences help make a decision”
- “Talk with physicians in the field. Shadow. Research the specialty”
- “Shadow, clerkships, meet with advisors”

Table 17 lists the frequency distribution of responses by topic. Students may have responded with more than one topic; thus, the percentages add up to more than 100%. The three most common themes regarded ways one can build the knowledge to choose a medical specialty: shadowing, 22%; experiencing the third-year clinical rotations, 19%;

and interacting with others (mentors, upperclassmen, residents, and other physicians), 18%. Considering the match between a specialty and one’s personal interests and values was also mentioned in 18% of the responses. Those categorized as “many methods of gathering information” listed three or more methods of information gathering instead of just one or two specific methods and did not elaborate on the importance of any one method in particular.

Table 17

Frequency Distribution of Response Categories by 159 Students to Open-ended Question

Response Category	Number/Percent of Responses
Shadow physicians	35 (22%)
Wait for third year experience	30 (19%)
Consult with experts (doctors, residents, upperclassmen)	29 (18%)
Consider match with personal interests, values	28 (18%)
Consult with mentors	22 (14%)
Many methods of gathering information	20 (13%)
Conduct research/have experiences in specialties	19 (12%)
Characteristics of specialty	17 (11%)
Reflection, interpretation of experiences	16 (10%)
Lifestyle considerations	10 (6%)
General comments	8 (5%)
Don’t know	4 (3%)
All other	10 (6%)

Shadowing of physicians was referenced both before and during medical school as a source of knowledge. Unfortunately, the comments related mostly to the act of

shadowing and did not explain how what was observed or performed could be incorporated into one's thought processes. In general, without any indication of how these behaviors are integrated into the students' thought processes, it was difficult to analyze the data against the theoretical framework of self-authorship. Still, 63 responses (40%) included specific references to reflection, interpretation of experiences, and their incorporation of personal values that could be related to some dimension of self-authorship.

In terms of the student-mentor relationship, a common theme when seeking advice about specialties was having mentors accessible to discuss why they chose their specialty, what it is like to practice in that specialty clinically, and how that particular specialty allows for work/life balance. From the perspective of the interpersonal dimension, there was rarely any mention of how the student saw this relationship; it often appeared to be a one-way dynamic with the mentor providing perspective but no indication of how the student would respond to the information, especially during the interaction with the mentor. Similarly, many respondents mentioned speaking with others to learn about a specialty. Again, there was no elaboration on the type of interaction they sought to have with these people. The implication was that of a one-sided relationship with the student absorbing information from the other person. Finally, most mentioned the third-year experiences as being important to the thought process. First- and second-year students ($n = 24$) regarded the clinical rotations with anticipation of the exposure but did not express much detail in terms of how they would integrate that exposure with their current thoughts. The third year students ($n = 16$) responded similarly as they began the actual rotations. Freeman (2007) indicated that clinical exposure can be limited in terms of time

and its impact on awareness; however, the students did not acknowledge or seem to be aware of the potentially limited utility in the decision process to choose a specialty.

Given the limited responses some included in their surveys, unless a person provided some sense of how this information was integrated, it would appear that all were externally influenced even though that may not necessarily be the case. The remaining analysis refers only to those comments that did include some reference to a thought process or personal reflection, or an evident lack thereof. The data are discussed in their relation to the identified components representing the three dimensions of self-authorship. Items were compared and contrasted to determine if they related to any particular aspect of self-authorship. Table 18 displays a summary of the dimensions and phases of self-authorship and illustrates how comments were coded. While a comment reflecting Externally driven behavior was straightforward, determining the difference between the Crossroads and Early self-authorship was more difficult because of the relative lack of depth in many of the responses. Thus, the remaining discussion focuses on the dichotomy of External versus more developed self-authorship based upon the scores from the principal components.

Table 18

Summary Table Illustrating Examples of Qualitative Analysis of Responses

Phase	Dimension		
	Epistemology	Intrapersonal	Interpersonal
External	“I’ve been interested in one specialty in high school. I am still interested in it.”	“Ultimately I’m hoping that within one of the few specialties I’ve narrowed down, one will ‘just feel better’ than the others.”	“Most mentors and professionals I met with usually agreed with supporting my decision.”
Crossroads	“Consider my interests (academic) and what excites me. Consider my goals (in and out of academics/ professional life).”	“I strongly believe that I need to know myself to find the specialty that is right for me. As I learn about different specialties, I always consider how appropriate the fit will be for my interests, my needs, and my lifestyle preferences.”	“I would value the advice of my mentor and faculty members the most, over that of my peers, etc.”
Early	“After considering the options, speaking with trusted mentors/ experts and truly understanding myself, I would feel ready to choose a future specialty.”	“I have always been interested in understanding myself – my values, interests, skills and abilities – and I believe that the better I understand myself, the more equipped I will be to choose a specialty.”	“As a woman in science, I plan to shadow and speak with women in many fields to get a sense of how they were able to balance work with raising children.”

One commonly identified theme ($n = 34$) regarded the epistemological dimension along the continuum of self-authorship phases and may or may not have included references to evaluating information against previous knowledge and/or beliefs. Among those who scored relatively high on the epistemological component, some of the

comments exhibited behavior at the Crossroads or Early phases of development. These included the following:

- “I want to choose a specialty that will allow me to continue exploring/pursuing my interests both inside and outside of medicine. I believe that experiences outside of medicine can directly impact medical innovation.”
- “ I understand lifestyle weighs into this decision in ways I may not be able to anticipate or imagine as a first year . . . ”
- “I am mindful about keeping an open mind. I think it would benefit me to go into the different clinical rotations leaving my preconceived notions behind. I don’t want those to alter my perceptions of the experiences.”

These comments illustrated the understanding that knowledge is contextual and changeable. Being open to the possibilities of new information or thinking beyond the traditional sources of information in medical school showed evidence towards self-authored ways of understanding knowledge. Conversely, those that scored lowest on this factor tended to relate knowledge more to external sources, such as “Most mentors and professionals I met with usually agreed with supporting my decision” or “I plan to seek the advice of friends, family, peers and mentors.” Knowledge was confirmed or provided by others but with no evidence of any internal thought process.

There were another group of comments ($n = 41$) that were related to the intrapersonal dimension as well. In general, students may have referred to their “values,” but those comments that were coded for this dimension were more explicit in defining values or beliefs. Those who scored relatively higher on this component exhibited an

understanding of their present values and how they can change, which could be related to the Crossroads or Early phases of self-authorship.

- “Lifestyle is important in terms of what I absolutely need and what I can give up or adapt to.”
- “Although I have a pretty good idea of what I wanna [sic] be, I am fully aware that I can change, so I am looking forward to the experience.”
- “Having a family is very important to me and . . . I won’t be happy if I don’t achieve a good work-life balance.”

The flexibility in values because of internal choices, as opposed to external pressures, showed evidence of self-authorship, as did others who clearly expressed beliefs and needs as being their own. Comments that were more commonly related to those who scored lower on this factor exhibited externally driven senses of value or no awareness of values at the time. In particular, values need to be related to personality or cognitive processes, not just Externally defined measures, to reflect self-authorization.

- “Over the course of the first two years, I plan to pursue a two-track approach. On the one hand, figure out my own preferences in terms of intellectual vs. manual work, lifestyle, finances, etc.”
- “I want to ensure I won’t be bored.”
- “Ask myself frequently if I am fairly interested . . . ”

Additionally, the personality characteristics of a specialty were a common concern. Many of those who scored on the lower end of the intrapersonal factor referred to concepts such as the following:

- “See the personalities of individuals in the specialty to ensure I’d fit in.”

- “Who I want to work with and maybe most important who I want to work for . . . Whatever the specialty, the doctors should be curious and activist.”

There are acknowledged personality types that are attracted to specific specialties (Freeman, 2007). However, in the choice of a specialty, Freeman indicated that the match should be between the student’s personality type and the skills/aptitudes necessary for the type of work in that particular specialty. The comments above referred more to the people the student would work with and not necessarily how they viewed themselves doing the work itself.

Identifying themes that validated the interpersonal dimension was more difficult. Most comments were only specific enough to mention talking to a mentor or professional as opposed to illustrating the anticipated interaction. Without knowledge of self-authorship, students were unaware of the connection that was desired from their text response to some phase/dimension of self-authorship. This may be an example of where qualitative interviews with experienced interviewers, as suggested by Baxter Magolda, might have provided prompts to elicit more information. However, some relatively high scoring students did provide more substantial comments:

- “Consider my strengths and weaknesses as a student/clinician, incorporating the feedback of my mentors/advisors.”
- I would value the advice of my mentor and faculty members the most, over that of my peers, etc.”
- “It helps me to go to influential informed individuals who have a good understanding of who I am.”

Self-authorship in regard to the interpersonal dimension is exhibited by the ability to prioritize feedback and to use it as a resource when needed, instead of an absolute source of information. The last comment exhibited a high awareness of what the student seeks in an interpersonal relationship, but the student was still Externally driven in terms of how he or she would use that information.

Finally, there were student responses that showed evidence to validate the construct related to the Early phase of self-authorship. Unlike the content analysis that related to the three dimensions of self-authorship, the data regarding the Early phase pointed only to development and not the lack thereof. In particular, respondents spoke about the importance of relating knowledge and experiences as they relate to choosing a specialty with their beliefs and values. As opposed to Externally driven needs, these students expressed a strong understanding of themselves, either at present or in the future. As evidence of Early development, they expressed thoughts that consider the give and take necessary to an interpersonal relationship with the specialty itself, willing to respond to pressures instead of the more self-centered behavior of one who is less developed in regard to self-authorship.

- “The foundation of my thinking regarding my future specialty lies in self-knowledge After understanding myself, I believe it is next most important to find mentors and speak with experts After considering the options, speaking with trusted mentors and truly understanding myself, I would feel ready to choose a specialty.”

- “I strongly believe that I need to know myself to find the specialty that is right for me Lifestyle is important in terms of what I absolutely need and what I can give up or adapt to.”
- “I found myself happy and excited. It was a matter of exposure and information for me because prior to that, I was very focused on surgery, probably for the wrong reasons.”

Among the comments that supported more or less development among the three dimensions and the Early phase of self-authorship, there was no clear relationship between year in school and level of development. Students in the first two years of medical school were just as likely as those in the last two years of medical school to have provided supporting responses. However, early assurance students more frequently provided evidence that referenced reflection and development (34%) when compared to premed students (27%).

In summary, while most students provided some response to the question, the response data were not always specific to the thought process behind selecting a specialty. However, those that did provide relevant responses were often somewhat related to how they scored on particular factors within the self-authorship construct. In particular, some students provided information that helped illustrate the dimensions of self-authorship, from both the internally and Externally driven extremes of the development within those dimensions. The analysis of this data provided substance to support the reliability and validity of the survey in general when combined with the constructs identified in the overall models.

CHAPTER 5

CONCLUSION

Summary of Research Findings

Summary of Quantitative Results

This study investigated the reliability and validity of a quantitative measure of self-authorship among medical students. The data were collected as part of an ongoing research project at a large medical school in the Northeast. The survey items were based upon a previously validated quantitative measure of self-authorship developed by Creamer et al. (2010) and adapted to reflect the context of selecting a medical specialty to practice after training. A total of 200 survey responses were collected at the start of the 2013-2014 academic year across all four classes of students. The research objectives were twofold: (1) determine the reliability and validity of the instrument, and (2) conduct a content analysis of textual responses as a method to validate the scale. It was hypothesized that the instrument could prove useful in measuring potential differences in development of self-authorship between traditional premed and early assurance medical students.

Items were developed to represent both a dimension of self-authorship and a phase of development within self-authorship. The original survey instrument included 28 potential items to represent all nine combinations of the three phases and three dimensions (3 x 3) of self-authorship, but it was not known if all phases and dimensions would be identified as separate constructs within the scale. Internal consistency of the scale for the total sample and the premed and early assurance student subgroups was satisfactory at .724, .711, and .712, respectively, after reducing the scale to 16 items.

Using the total sample of medical students and principal component analysis with Promax oblique rotation, the data produced results confirming four of the constructs of self-authorship identified by Creamer et al. (2010). Four components were determined based on a review of the components with eigenvalues greater than one, review of the scree plot, and interpretation of component loadings. Only components that loaded on three or more items were considered appropriate for interpretation. Of the four components, three represented the dimensions of self-authorship (epistemological, interpersonal, and intrapersonal), while this study identified a construct only for the Early phase of self-authorship, not for the External or Crossroads phases. This might be an artifact of the continuous nature of the phases. Baxter Magolda (2008) identified that the distinction between phases may not be clearly delineated, but rather fuzzy. The instrument was able to clearly separate the items proposed to measure development at the Early phase of self-authorship. Low scores within this construct could be interpreted as the opposite of the Early phase, or development within the External phase. This might be an easier definition of the extremes of development phases when compared to having two separate components, one for each phase.

Review of the component correlations identified moderate correlations ranging from .061 to .343, justifying the assumption of correlated components. Values much higher than these would have indicated the potential for collapsing components or second-order principal component analyses (Gorsuch, 1983). The highest correlations were between the interpersonal-Early components (.343) and the intrapersonal-Early components (.247). Baxter Magolda (2008) discussed the unique paths individuals may take to develop self-authorship. The relatively equal correlations of the interpersonal and

intrapersonal dimensions with the Early dimension highlighted that either can be primary to the other as one moves toward self-authorship. However, the relatively low correlations of any component with the epistemological component pointed to the model's inability to fully capture that dimension and/or the relative lack of importance of this dimension in relation to self-authorship among medical students. Internal consistency of the four components was only moderate, ranging from .320 to .621.

The first research question in this study sought to measure the reliability and validity of a quantitative measure of self-authorship among medical students. Sub-group analyses among the traditional premed and the early assurance student populations were the secondary objective of this research question. Repeating the same principal component analysis and limiting the data to premed students identified the same four components and explained 50% of the variance. Since premed students represented more than 70% of the respondents, it was expected that these results would be similar to the total. The early assurance students represented just 18% of the respondents ($n = 35$), which greatly limited the ability to conduct data reduction analyses. Attempts to conduct principal component analysis with this data failed based upon the Kaiser-Meyer-Olkin test for consistency, even though communalities were appropriate based upon prior evidence with small samples (Preacher & MacCallum, 2002). The study did not succeed in validating the instrument for this sub-population.

Summary of Content Analysis

Given the greater evidence of self-authorship development in the literature using qualitative methods, a second research objective was to conduct a content analysis of written free-form text in conjunction with the four identified constructs from the principal

component analysis (the epistemological, interpersonal, and intrapersonal dimensions and the Early phase of self-authorship). These findings paralleled Pizzolato's (2007) prior difficulties with correlating survey constructs and open-ended responses. Although the inclusion and formatting of the open-ended question was intended to elicit rich data, students did not always provide sufficient data for content analysis. As many as 125 students' responses to a question asking for the thought process in their thinking about a medical specialty in which to practice contained brief statements of specific activities, similar to the experience of Creamer and Laughlin (2005). The lack of in-depth responses limited the ability to interpret some comments regarding the level of development of self-authorship. The inherent characteristics of a paper survey and the inability to probe for responses made a non-response or minimal response incapable of being related to any particular type of student. However, more than 60 respondents included thoughtful explanations of their thought processes and experiences.

The responses that included sufficient discussion indicating self-authored behavior, or a clear lack thereof, were used as a source to validate the constructs. Student responses to the self-authorship scale were scored on the four identified constructs as relatively higher or lower on each scale and then the open-ended responses were compared and contrasted to those at the extreme ends of each scale. Among those who provided a response that described a thought process, the evidence supported the higher or lower score on the respective construct. Those who scored at the higher end of the scale tended to describe processes towards the more developed behaviors representing the construct, while those at the lower end tended to describe behaviors that were more Externally driven, or less indicative of self-authored behavior. Thus it was determined that the scale

was a valid measure to separate respondents into the extremes of each dimension. In particular, the data provided evidence to identify themes that separated Externally and internally driven thought processes. Kegan (1982) discussed how movement through the dimensions was fluid and not exact for all. The themes identified did not validate all three phases of development within each dimension but were sufficient to allow for identifying the extremes within each dimension that may be more appropriate, given the theorized fluidity between phases.

Implications for Theory

Baxter Magolda's (2008) work is founded on students' experiences during the college years and into early adulthood as subjects begin careers and family life. In particular, she identified that the Early phase of self-authorship can first become evident during college. The scale devised by Creamer et al. (2010) was the basis for the instrument under investigation in this study; it, too, was originally tested with undergraduate students. This study sought to expand the research regarding self-authorship to the population of medical students, who have a different educational experience and are at a different life stage than previously studied populations. From a theoretical perspective, this study provided new knowledge that highlights unique characteristics of the population and three potential influences on theory.

First, the advanced educational background of medical students and the homogeneity of their career aspirations in one particular field have provided a new window into the exploration of self-authorship. In particular, medical students overall may be at a further point in their educational experience and life span, making comparisons to theory based upon undergraduates difficult. Baxter Magolda (2001) identified career

choice as a pivotal decision that can help develop self-authorship. Upon matriculation to medical school, students have already made one career decision. Analysis of their open-ended responses indicated that the secondary career decision regarding what medical specialty to practice is still uncertain. Principal component analysis identified the three dimensions of self-authorship but identified only a construct of the Early phase, unlike Creamer et al. (2010), who were able to identify the three phases and the three dimensions. In the content analysis of the text responses, the most common source of information for the students was shadowing physicians. Baxter Magolda (2004) suggested that “situating learning in the learners’ experience” (p. 191) is one method to promote self-authorship. Shadowing, working in clinics, and consulting with physicians are all examples of behaviors where the students are already participating in situated activities which may provide more opportunity to develop self-authorship than other types of undergraduate students might experience.

These findings raise the question of whether the study of self-authorship through a quantitative tool can be generic to all populations or whether tools need to be more aligned to particular populations. The instrument in this study built upon a study by Creamer et al. (2010), who focused on the undergraduate population when designing their instrument. While appropriate to the study of self-authorship among undergraduates, the original scales may have been too general for the medical student population. In particular, the instrument in this study may have over-represented the External phase and did not include sufficient numbers of items to distinguish between the Crossroads and Early phases. Questions used in a scale for self-authorship among professional students (medicine or law) may need to reflect a set of processes more advanced than those in the original

instrument. Self-authorship and its context-specific nature may require instruments that are also contextual and necessarily relevant to particular populations. In adapting one instrument for the purposes of this specific student population, there were enough substantive changes and revisions that indicated the need for careful consideration when moving beyond undergraduate students, who were the original foundational population of self-authorship theory. In particular, medical students seem more likely to be approaching the Early phase of self-authorship. Scale items may need to be recalibrated to be more sensitive to higher levels of development for this population.

A second implication for theory is that if the study of self-authorship is contextual, it may not be possible to develop construct items that maintain relevancy across any population. The findings in this study again supported the relative importance of career choice as a context for the construct items when measuring self-authorship (Creamer et al., 2010). Development through the phases of self-authorship is related to perceived conflicts in knowledge, beliefs, and/or values (Baxter Magolda, 2004; Kegan, 1982; Pizzolato, 2007). Survey items that resonated with students and were most important within component analyses related to choice of medical specialty and gathering knowledge about specialties, either on one's own or from mentors. Items that related to knowledge or information in general (interpretations of a journal article or scientific explanations of phenomena) did not provide as much variation or utility to the analyses. The qualitative data from this study further emphasized the importance of mentors and clinical experiences in understanding specialty choice. Measures of self-authorship should focus on the specific context desired in order to measure the thought processes of the population. While the items in the proposed instrument were theorized to represent each of the

dimensions and phases of self-authorship, it is possible that those that did not specifically relate to the medical school context did not resonate with the respondents nor generate suitable variation for analysis. Saliency of the item to the respondent can improve responses (Dillman, 2000). That certain groups of items did not relate to the component analysis might reflect how the theory of self-authorship is applied to specific populations and the need to maintain relevance to the specific context. If one desires to measure self-authorship quantitatively across different student populations, then one context-driven instrument may not be feasible. Pizzolato (2005) argued for a context-free measure of self-authorship; her more generalized scale items might be more appropriate as a standardized measure. It should be noted that researchers need to ensure that all construct items are relevant to the particular population under study.

Finally, the proposed instrument was most successful in identifying the interpersonal and intrapersonal dimensions of self-authorship, as well as somewhat similar correlations between each dimension and the construct of the Early phase. There were also smaller correlations among the interpersonal and intrapersonal dimensions with the epistemological dimension; however, the latter had a statistically insignificant correlation with the Early phase component. Creamer et al. (2010) questioned the importance of having an instrument that measures both the phases and dimensions of self-authorship equally. As discussed earlier, medical students may be at a more advanced life stage and require a different measure of self-authorship. The relative strength of the Early phase construct would indicate that there are a substantial number of students already at some point in this stage based upon their responses. The applicability of this scale to the medical

student population might be more appropriate if it focused on the three dimensions rather than the phases.

This study hypothesized that the development of self-authorship was a recursive process that can be influenced by personal characteristics, experiences, and interpretations of the experiences. Slotnick (2001) argued that the activities experienced in medical training and the trajectory of changing roles influence how physicians-in-training learn. He stressed the need to understand the “social factors . . . , psychosocial development, and epistemologic sophistication” (p. 1025) of individual physicians when evaluating their progress through the educational arc. Physician-teachers exert significant influence on the clinical knowledge and identity formation of medical students (Jarvis-Selinger et al., 2012). The students themselves repeatedly mentioned in their responses the influence of mentors, physician leaders, and other clinicians on their decision processes. Baxter Magolda (2004) argued that altering career choices is a result of growing the interpersonal dimension of self-authorship and understanding one’s concept of “Who am I?” This study identified the relative strength for medical students of both the interpersonal and intrapersonal dimensions of self-authorship as measures of psychosocial development within the self-authorship instrument. This scale can provide data to support Slotnick’s theory and identify potential differences in the growth of medical students.

Implications for Practice

The relationship between exposure to the liberal arts during a baccalaureate and success in the training to be a physician has been argued for many years (Brieger, 1999; Gunderman & Kanter, 2008). A liberal arts education is thought to improve the non-cognitive skills of future medical students. The most common preparation for medical

school in the United States has been the science-centric traditional premed baccalaureate, which may limit the opportunity for a student to explore any liberal arts classes while an undergraduate. Success in medical school, and ultimately as a practicing physician, requires not only technical knowledge and understanding of the role of a physician but also the ability to interact in relationships with diverse others, including patients, mentors, and physician-colleagues. Self-authorship can facilitate mutually beneficial relationships with others and has been described as a skill that can help students succeed during their clinical training without compromising their values (White, Kumagai, Ross, & Fantone, 2009).

The instrument in this study provides one method to assess the level of self-authorship among students when they start medical school and their level of self-authorship development in coordination with their level of exposure to the liberal arts during medical school. Such an analysis could support or disprove arguments for the inclusion of liberal arts in premed education. If those who have a liberal arts background do indeed start medical school with more self-authored behavior, this might suggest the inclusion of a measure of self-authorship such as the one in this study as a non-cognitive measure to assess applicants for medical school.

The findings from this study also provide preliminary evidence of a tool that could eventually be used to measure potential differences between premed and early assurance students, either at the start of medical school or anywhere along the continuum of the educational path. This scale of self-authorship can be utilized to identify the distribution of scores as well as cutoffs to identify significant differences from the mean. The scale would provide a quantitative measure of self-authorship that could be related to the degree of

development. Such an assessment could be used to monitor stages of development and potential changes over time. There have been arguments for changing the application process to medical school to include non-cognitive measures as a method to better assess student potential (Peskun et al., 2007; Sedlacek, 2004). This scale of self-authorship could be one such type of assessment. The data could provide useful information in terms of potential support a student might need to succeed in medical school in comparison to his or her peers. However, admissions officers need to be cautious that the data should not be used to determine acceptance to medical school.

As stated earlier, there is no literature to describe the potential benefits of an early assurance program for medical students, although 22% of schools with early assurance programs indicated that part of the program's mission is to integrate liberal arts into medical training (Eaglen et al., 2012). Self-authorship is being described as a potential tool to help support success in the medical education continuum (Jarvis-Selinger et al., 2012; Prober & Khan, 2013). Having the ability to identify developmental differences, if any, at the start of medical education could provide one piece of evidence on how to improve the undergraduate curriculum to support any population that might start medical school at a disadvantage or develop less than another.

Within the broader field of social research, the evidence from this study supports previous findings such as those by Creamer & Laughlin (2005) and Pizzolato (2007) regarding the potential difficulties in quantitatively measuring self-authorship but still provides encouraging data about these efforts with a new study population. The proposed instrument included new items beyond those in the original tool from Creamer et al. (2010), but many did not improve the internal reliability of the scale. The study did

validate the translation of many of the items to a context relevant to the medical student population. However, it should be noted that the medical student population is conceptually different than the undergraduate populations who have been the subject of previous research. In addition, the literature has shown that the study of self-authorship needs to be context-driven. Many of the previously published studies, such as Creamer and Laughlin (2005) and Creamer et al. (2010), as well as this one, included the scale within a broader survey with other topics. Respondents may not understand the context of the particular scale well enough to truly focus on their responses, particularly in regard to the text responses. Isolating this scale in an instrument specifically about development could make it easier for a student to understand and might provide more-focused responses to inform the study of self-authorship.

Limitations

This study has several limitations. First, the data are limited to one medical school that had a proportionally sizable early assurance student population, even though they were a small sub-sample within this study. The early assurance program is rather liberal in terms of admission requirements when compared to other medical schools with early assurance programs, in particular regarding the program's exclusion of the MCAT entrance exam requirement. Thus the sample may not be representative of students in other medical schools that also have early assurance programs. Also, the data were sufficient to investigate reliability and validity in total and for the premed sub-population but not for the early assurance students. Even with a satisfactory response rate over 60%, upon data cleaning and data entry validation the study was left with just under 200 responses for analysis. This limited the statistical methods that could be used to principal

components analysis and prevented completion of the analysis to answer the research questions regarding the early assurance student sub-population.

A second limitation is related to the relatively small number of items in the test instrument. In this study, 10 items were added to the original 18 in the instrument from Creamer et al. (2010). Reliability analyses and principal component analyses eventually reduced the items in the scale to 16. However, the distribution of items overrepresented the epistemological dimension and the External phase. Future research into this scale should include more items that will represent the intrapersonal and interpersonal dimension in the Crossroads and Early phases of development.

The third limitation regards the methods used. Exploratory factor analysis or principal component analysis is useful in exploring the reliability and validity of a scale. However, the research sub-question of this study was to validate the self-authorship scale separately for early assurance and premed students. To test how the instrument differentiates among specific sub-populations, it is more appropriate to do a confirmatory factor analysis. The amount of raw data from this study precluded any such analysis. Therefore, further research into this area needs to be conducted, subsequent to dealing with the third limitation.

Suggestions for Future Research

The primary suggestion for future research is to build upon the current study sample. One idea would be to repeat the entire study process with all four classes in the next academic year. This, however, would introduce confounding factors by having multiple responses from students who potentially took part in both waves of the study. A compromise would be to survey just the entering class from the next academic year to

build the number of responses without introducing multicollinearity among the data. Unfortunately, it may take up to two calendar years to attain a sufficient sample to investigate the early assurance students separately, focusing specifically on entering medical students to validate the instrument. There is also the potential to expand such a study beyond this one medical school. This would increase the survey population but could also introduce confounding variables if the admission requirements are significantly different at other medical schools. Any differences in programs could be controlled for on a post-hoc basis through multivariate methods when interpreting the results, and the richer data from multiple sites would improve the generalizability of the instrument.

In addition to collection of further quantitative data, future research should consider collecting qualitative data. Collecting meaningful responses without the prompting that can happen in a human interaction has been problematic in this study as well as in previous studies (Creamer & Laughlin, 2005; Pizzolato, 2007). Responses are not always as thorough and thoughtful as hoped. One suggestion is to add a second open-ended item that poses a relevant question/situation to the person answering and asks for the thought process in responding to the situation. For example, in the case of medical students, the situation could relate to the conflict between personal family obligations and an emergency patient situation and how they would respond. A medical student at any level of training could respond to such a question, and the depth of response could provide evidence of self-authorship or the lack thereof. A comparison of responses to this question and the original open-ended question can be also useful in providing additional rich information in conjunction with the quantitative items.

After further validation of the instrument as it applies to the medical student population, the most important research will be to measure the potential differences among students as they enter and progress through their medical education. One specific area of analysis would be a study between early assurance and premed students. Existing literature does not provide any evidence of the benefits of early assurance programs to the students. By using a quantitative tool to measure self-authorship, such a formative assessment in high school and/or college could provide important information to students considering medical school, counselors who advise students, and educators who design curriculum. This information could help identify students who are ready for early assurance programs and also bolster undergraduate curriculum to help develop self-authorship where it might be lacking

Ultimately, this research raises the question as to whether any assessment of self-authorship among medical students needs to be at a more sophisticated and contextual level than those instruments that have been designed for college students. Not all medical students are at a higher developmental phase of self-authorship, but the way these students approach the concept may require different methods of measurement. To better understand how medical students think about self-authorship, qualitative data needs to be collected along the lines of the methods discussed by Baxter Magolda and King (2007). Qualitative research involving interviews with early assurance and premed students could provide rich information to understand how these students relate to self-authorship concepts and identify them within themselves. This research also provides the opportunity to explore the experiences each subgroup had as undergraduates and how each subgroup experienced their college years. The ability to compare and contrast the students' experiences during

their undergraduate college education, including their decisions on which classes to take (Externally driven versus internally driven choices), what their individual goals were in selecting classes, and how this relates to their vision of medical school (values, beliefs) could provide rich information to further explore self-authorship. This information could provide background to further develop the items in the scale and better reflect how self-authorship relates to the medical student population.

Research into self-authorship has been applied mainly to the student affairs literature and how educators can help support the growth of self-authorship (Baxter Magolda & King, 2012). White et al. (2009) argued for the need to have such support specifically for medical students and trainees as they progress through the many stages of medical education. Self-authorship may have the potential to help foster identity formation and better physician-patient relationships for future physicians. Further research to develop and improve this quantitative measure of self-authorship can provide a key method to validate efforts to bring self-authorship into the medical education curriculum in all medical schools.

REFERENCES

- Abell, N., Springer, D. W., & Kamata, A. (2009). *Developing and validating rapid assessment instruments*. New York, NY: Oxford University Press.
- Abes, E. S., Jones, S. R., & McEwen, M. K. (2007). Reconceptualizing the model of multiple dimensions of identity: the role of meaning-making capacity in the construction of multiple identities. *Journal of College Student Development*, 48(1), 22.
- Association of American Medical Colleges. (2011). Medical college admissions test (MCAT) examinees. Washington, DC: Author. Retrieved from <http://www.humanitiesindicators.org/content/hrcoIIIB.aspx#topIII5>
- Association of American Medical Colleges. (2012). FACTS: Applicants, matriculants, enrollment, graduates, MD/PhD, and residency applicants data. (20102). Washington, DC: Author. Retrieved from <http://www.aamc.org/data/facts/>
- Bardes, C. L., Best, P. C., Kremer, S. J., & Dienstag, J. L. (2009). Medical school admissions and noncognitive testing: Some open questions. *Academic Medicine*, 84(10), 4.
- Barr, D. A. (2010). Science as superstition: selecting medical students. *The Lancet*, 376(August 28), 678-679.
- Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco, CA: Jossey-Bass.
- Baxter Magolda, M. B. (2004). *Making their own way*. Sterling, VA: Stylus Publishing.

- Baxter Magolda, M. B. (2008). Three elements of self-authorship. *Journal of College Student Development, 49*(4), 269-284. doi: 10.1353/csd.0.0016
- Baxter Magolda, M. B., & King, P. M. (2007). Interview strategies for assessing self-authorship: Constructing conversations to assess meaning making. *Journal of College Student Development, 48*(5), 491-508.
- Baxter Magolda, M. B., & King, P. M. (Eds.). (2012). *Assessing meaning making and self-authorship: Theory, research, and application: ASHE higher education report 38:3*. Wiley Online Library.
- Board on Testing and Assessment. (2012). Education for work and life: Developing transferable knowledge and skills in the 21st century. Washington, DC: The National Academies
- Boyatzis, R. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: SAGE.
- Brieger, G. H. (1999). The plight of premedical education: Myths and misperceptions, Part II: Science "versus" the liberal arts. *Academic Medicine, 74*(11), 1217-1221.
- Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis*. London, England: SAGE.
- Child, D. (2006). *The essentials of factor analysis* (3rd ed.). London, England: Continuum International Publishing Group.

- Creamer, E. G., Baxter Magolda, M. B., & Yue, J. (2010). Preliminary evidence of the reliability and validity of a quantitative measure of self-authorship. *Journal of College Student Development, 51*(5), 550-562.
- Creamer, E. G. & Laughlin, A. (2005). Self-authorship and women's career decision making. *Journal of College Student Development, 46*(1), 13-27.
- DeVellis, R. F. (2012). *Scale Development* (3rd ed.). Los Angeles, CA: SAGE.
- Dienstag, J. L. (2008). Relevance and rigor in premedical education. *New England Journal of Medicine, 359*(3), 221-224.
- Dillman, D. A. (1991). The design and administration of mail surveys. *Annual Review of Sociology, 17*(1), 225-249.
- Dillman, D. A. (2000). *Mail and Internet Surveys* (2nd ed.). New York, NY: John Wiley & Sons.
- Downing, S. M. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education, 37*(9), 8.
- Downing, S. M. (2004). Reliability: on the reproducibility of assessment data. *Medical Education, 38*(10), 7.
- Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? *Psychological Bulletin, 81*(6), 358-361.
- Eaglen, R. H., Arnold, L., Girotti, J. A., Cosgrove, E. M., Green, M. M., Kollisch, D. O., . . . Tracy, S. W. (2012). The scope and variety of combined baccalaureate-MD programs in the United States. *Academic Medicine, 87*(11), 1600-1608. doi: 10.1097/ACM.0b013e31826b8498

- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.
- Flexner, A. (1960). Medical education in the United States and Canada; a report to the Carnegie Foundation for the Advancement of Teaching. Boston, MA: The Merrymount Press.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286-299.
- Freeman, B. (2007). *The ultimate guide to choosing a medical specialty* (2nd ed.). New York, NY: McGraw-Hill.
- Gorsuch, R. L. (1983). *Factor Analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gross, J. P., Mommaerts, C. D., Earl, D., & De Vries, R. G. (2008). After a century of criticizing premedical education, are we missing the point? *Academic Medicine*, 83(5), 4.
- Gunderman, R. B., & Kanter, S. L. (2008). "How to fix the premedical curriculum" revisited. *Academic Medicine*, 83(12), 1158-1161.
- Hartcollis, A. (2010, July 29). Getting into med school without hard sciences, *The New York Times*. Retrieved from <http://www.nytimes.com/2010/07/30/nyregion/30medschools.html?pagewanted=all>

- Hodge, D. C., Baxter Magolda, M. B., Haynes, C. A. (2009). *Engaged learning: Enabling self-authorship and effective practice*. Paper presented at the The Conference on Liberal Education and Effective Practice, Worcester, MA.
- Hutcheson, P. A. (2007, Fall). The Truman commission's vision of the future. *The NEA Higher Education Journal*, 107-115.
- Jarvis-Selinger, S., Pratt, D. D., & Regehr, G. (2012). Competency is not enough: Integrating identity formation into the medical education discourse. *Academic Medicine*, 87(9), 1185-1190.
- Kegan, R. (1982). *The evolving self: Problem and process in human development*. Cambridge, MA: Harvard University Press.
- Kim, H. (2014). Statistical notes for clinical researchers: Assessing normal distribution using skewness and kurtosis. *Restorative Dentistry and Endodontics*, 38(1), 52-54. doi: 10.5395/rde.2013.38.1.52
- Kreiter, C. D., & Kreiter, Y. (2007). A validity generalization perspective on the ability of undergraduate GPA and the medical college admission test to predict important outcomes. *Teaching and Learning in Medicine: An International Journal*, 19(2), 95-100.
- Kuh, G., Gonyea, R., & Rodriguez, D. (2002). The scholarly assessment of student development. In T. Banta (Ed.), *Building a scholarship of assessment* (pp. 100-128). San Francisco, CA: Jossey-Bass.
- LeClerc, G., Lefrancois, R., Dube, M., Hebert, R., & Gaulin, P. (1999). Criterion validity of a new measure of self-actualization. *Psychological reports*, 85(3f), 10.

- Love, P. G., & Guthrie, V. L. (2002). Kegan's orders of consciousness. *New Directions for Student Services*(88), 65-76.
- Lovell, C. (1999). Empathic-cognitive development in students of counseling. *Journal of Adult Development*, 6(4), 195-203.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84-99.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: SAGE.
- Mitchell, R. (1994). The development of the cognitive behavior survey to assess medical student learning. *Teaching and Learning in Medicine: An International Journal*, 6(3), 161-167.
- Morse, J. M. (2010). Principles of mixed methods and multimethod research design. In R. K. Howard, L. Kennedy-Phillips & C. E. Watt (Eds.), *Qualitative & quantitative research* (3rd ed., pp. 208-224). Boston, MA: Pearson Learning Solutions.
- Muller, D., & Kase, N. (2010). Challenging traditional premedical requirements as predictors of success in medical school: The Mount Sinai School of Medicine Humanities and Medicine program. *Academic Medicine*, 85(8), 1378-1383.
- The National Council for Liberal Education and America's Promise. (2007). *College learning for the new global century*. Washington, DC: Association of American Colleges and Universities.

- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill Inc.
- Pascarella, E. T. (2001). Cognitive growth in college. Surprising and reassuring findings from the National Study of Student Learning. *Change: The Magazine of Higher Learning*, 33(6), 20-27.
- Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students* (Vol. 2). San Francisco, CA: Jossey-Bass.
- Perry, W. G. Jr. (1997). Cognitive and ethical growth: The making of meaning. In P. Altbach (Ed.), *Contemporary higher education: College student development and academic life* (pp. 48-88). New York, NY: Garland Publishing.
- Peskun, C., Detsky, A., & Shandling, M. (2007). Effectiveness of medical school admissions criteria in predicting residency ranking four years later. *Medical Education*, 41(1), 57-64.
- Pizzolato, J. E. (2004). Coping with conflict: Self-authorship, coping, and adaptation to college in first-year, high-risk students. *Journal of College Student Development*, 45(4), 425-442.
- Pizzolato, J. E. (2005). *Creating complex partnerships: A multiple study investigation into self-authorship*. Ann Arbor, MI: Michigan State University. Retrieved from <http://search.proquest.com/docview/305473605?accountid=13793>
ProQuest Dissertations and Theses database. (305473605)
- Pizzolato, J. E. (2007 Spring). Assessing self-authorship. *New Directions for Teaching and Learning*(109), 31-42. doi: 10.1002/tl.263

- Preacher, K. J., & MacCallum, R. C. (2002). Exploratory factor analysis in behavior genetics research: Factor recovery with small sample sizes. *Behavior genetics*, 32(2), 153-161.
- Prober, C. G., & Khan, S. (2013). Medical education reimaged: A call to action. *Academic Medicine*, 88(10), 1407-1410. doi:10.1097/ACM.0b013e3182a368bd
- Razack, S., Faremo, S., Drolet, F., Snell, L., Wiseman, J., & Pickering, J. (2009). Multiple mini-interviews versus traditional interviews: Stakeholder acceptability comparison. *Medical Education*, 43, 993-1000. doi: 10.1111/j.1365-2923.2009.03447.x
- Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in Personality and Social Psychology Bulletin. *Personality and Social Psychology Bulletin*, 28(12), 1629-1646. doi: 10.1177/014616702237645
- Sedlacek, W. E. (2004). *Beyond the big test*. San Francisco, CA: Jossey-Bass.
- Seifert, T. A., Goodman, K., King, P. M., & Baxter Magolda, M. B. (2010). Using mixed methods to study first-year impact on liberal arts learning. *Journal of Mixed Methods Research*, 4(3), 248-267.
- Shuval, J. T. (1975). From "boy" to "colleague": processes of role transformation in professional socialization. *Social Science and Medicine*, 9(8-9), 413-420.

- Silver, B., & Hodgson, C. S. (1997). Evaluating GPAs and MCAT scores as predictors of NBME I and clerkship performance based on students' data from one institution. *Academic Medicine, 72*(5), 394-396.
- Slotnick, H. B. (2001). How doctors learn: Education and learning across the medical-school-to-practice trajectory. *Academic Medicine, 76*(10), 1013-1026.
- Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales. A practical guide to their development and use* (4th ed.). New York, NY: Oxford University Press.
- Thelin, J. R. (2004). *A history of American higher education*. Baltimore, MD: The Johns Hopkins University Press.
- Torres, V., & Hernandez, E. (2007). The influence of ethnic identity on self-authorship: A longitudinal study of Latino/a students. *Journal of College Student Development, 48*(5), 558-573.
- TTUHSC School of Medicine. (2013). Early acceptance program. Lubbock, TX: Author. Retrieved 3/17, 2013, from http://www.depts.ttu.edu/honors/earlyacc_med.php
- University of Maryland. (2013). Early assurance program guideline for University of Maryland, College Park. College Park, MD: Author. Retrieved from [http://www.prehealth.umd.edu/files/prehealth/gwea/EARLY SELECTION PROGRAM - Updated SU 2011.pdf](http://www.prehealth.umd.edu/files/prehealth/gwea/EARLY_SELECTION_PROGRAM_-_Updated_SU_2011.pdf)

- Upstate Medical University, State University of New York. (2013). Special admissions options - early assurance program for college sophomores. Syracuse, NY: Author. Retrieved 3/17, 2013, from <http://www.upstate.edu/com/admissions/options/soph.php>
- Walker, D. & Myrick, F. (2007). Grounded theory: An exploration of process and procedure. In A. Brynam (Ed.), *Qualitative Research 2* (Vol. 4, pp. 47-62). Thousand Oaks, CA: SAGE.
- Wawrzynski, M., & Pizzolato, J. E. (2006). Predicting needs: A longitudinal investigation of the relation between student characteristics, academic paths, and self-authorship. *Journal of College Student Development, 47*(6), 677-692.
- Weingartner, R. H. (1980). Selecting for medical school. *Journal of Medical Education, 55*, 922-927.
- Wetzel, A. P. (2011). *Factor analysis methods and validity evidence: A systematic review of instrument development across the continuum of medical education*. Richmond, VA: Virginia Commonwealth University. Available from Proquest, LLC (3453673)
- White, C. B., Dey, E. L., & Fantone, J. C. (2009). Analysis of factors that predict clinical performance in medical school. *Adventures in Health Sciences Education, 14*, 455-464. doi: 10:1007/s10459-007-9088-9
- White, C. B., Kumagai, A. K., Ross, P. T., & Fantone, J. C. (2009). A qualitative exploration of how the conflict between the formal and informal curriculum influences student values and behaviors. *Academic Medicine, 84*(5), 597-603.

Yardley, S., Teunissen, P. W., & Dornan, T. (2012). Experiential learning: AMEE guide
No. 63. *Medical Teacher*, 34(2), e102-e115. doi:[http://dx.doi.org/](http://dx.doi.org/10.3109/0142159X.2012.650741)

10.3109/0142159X.2012.650741

Zhao, X., Oppler, S., Dunleavy, D., & Kroopnick, M. (2010). Validity of four approaches
of using repeaters' MCAT Scores in medical school admissions to predict
USMLE Step 1 total scores. *Academic Medicine*, 85(10), 564-567.

Appendix A

Comparison of Survey Items from Creamer, Magolda and Yue (2010) with Study Pilot Phase Items

Creamer, Magolda, Yue (2010) Items	Study Pilot Survey Items
The most important role of an effective career counselor or advisor is to be an expert on a variety of career options	The most important role of an effective mentor or advisor is to be an expert on a variety of medical specialties.
The most important role of an effective career counselor or advisor is to provide guidance about a choice that is appropriate to me.	The most important role of an effective mentor or advisor is to provide guidance about a choice of specialty that is appropriate to me.
The most important role of an effective career counselor or advisor is to help students think through multiple options	The most important role of an effective mentor or advisor is to help students think through multiple options [when selecting a specialty].
In my opinion, the most important role of an effective counselor or advisor is to direct students to information that will help them to make a decision on their own.	The most important role of an effective mentor or advisor is to direct students to information that will help them to make a decision on their own.
My primary role in making an education decision is to acquire as much information as possible	My primary role in choosing a medical specialty will be to acquire as much information as possible
My primary role in making an education decision is to seek direction from informed experts	My primary role in choosing a medical specialty will be to seek direction from informed experts
My primary role in making an education decision is to consider my own views.	My primary role in choosing a medical specialty will be to consider my own views.
n/a	My primary role in choosing a medical specialty will be to make a decision considering the available information and my own views.

To make a good choice about a career, I think that it is largely a matter of personal opinion.	To make a good choice about a medical specialty, I think that it is largely a matter of personal opinion.
To make a good choice about a medical specialty, I think the facts are the strongest basis for a good decision.	To make a good choice about a medical specialty, I think the facts are the strongest basis for a good decision.
To make a good choice about a medical specialty, I think that experts are in the best position to advise me about a good choice.	To make a good choice about a medical specialty, I think that experts are in the best position to advise me about a good choice.
To make a good choice about a medical specialty, I think it is not a matter of facts or expert judgment, but of a match between my values, interests, and skills and those of the job.	To make a good choice about a medical specialty, I think it is not a matter of facts or expert judgment, but of a match between my values, interests, and skills and those of the job.
When people have different interpretations of a book, I think that some books are just that way. It is possible for all interpretations to be correct.	When people have different interpretations of a book, I think that some books are just that way. It is possible for all interpretations to be correct.
When people have different interpretations of a book, I think that multiple interpretations are possible, but some are closer to the truth than others.	When people have different interpretations of a book, I think that multiple interpretations are possible, but some are closer to the truth than others.
n/a	When people have different interpretations of a book, I think only one interpretation can be right.
When people have different interpretations of a book, only experts can say which interpretation is really correct.	When people have different interpretations of a book, only experts can say which interpretation is really correct.
If a teacher or advisor recommended a career in a field that I have never considered before, I would explain	If a teacher or advisor recommended a medical specialty that I have never considered before, I would explain my

my point of view.	point of view.
If a teacher or advisor recommended a career in a field that I have never considered before, I would try to understand their point of view and figure out an option that would best fit my needs and interests.	If a teacher or advisor recommended a medical specialty that I have never considered before, I would try to understand their point of view and how it would best fit my needs and interests.
Experts are divided on some scientific issues, such as the causes of global warming. In a situation like this, I would have to look at the evidence and come to my own conclusion.	Experts are divided on some scientific issues, such as the causes of global warming. In a situation like this, I would have to look at the evidence and come to my own conclusion.
Experts are divided on some scientific issues, such as the causes of global warming. In a situation like this, I think it is best to accept the uncertainty and try to understand the principal arguments behind the different points of view.	Experts are divided on some scientific issues, such as the causes of global warming. In a situation like this, I think it is best to accept the uncertainty and try to understand the principal arguments behind the different points of view.
n/a	Experts are divided on some scientific issues, such as the causes of global warming. I rely on the experts to decide.
n/a	Experts are divided on some scientific issues, such as the causes of global warming. I try not to judge as long as different scientists have different opinions on these kinds of issues.
n/a	I am in the best position to know when I've found the right specialty for me.
n/a	My choice of medical specialty may change over time based upon my experiences in medical school.

OPEN-ENDED QUESTION: In a couple of paragraphs, please describe the approach you have taken, or plan to take, to reach a decision about what medical specialty you will practice.

Appendix B

Revised Survey Items for the Research Project With Indicators for Proposed Dimension and Phase of Self-Authorship

	Dimension of Self-Authorship	Phase of Self-Authorship	Survey Item
Inter-X1	Interpersonal	External	The most important role of an effective mentor or advisor is to be an expert on one or more medical specialties.
Inter-C1	Interpersonal	Crossroads	The most important role of an effective mentor or advisor is to provide guidance about a choice of specialty that is aligned with my skills, talents and personality.
Inter-C2	Interpersonal	Crossroads	The most important role of an effective mentor or advisor is to discuss multiple options when selecting a specialty.
Inter-X2	Interpersonal	External	The most important role of an effective mentor or advisor is to direct students to information that will help them to make a decision on their own.
Inter-E1	Interpersonal	Early	The most important role of an effective mentor or advisor is to challenge a student to clarify expectations of a medical specialty.
Intra-X1	Intrapersonal	External	When choosing a medical specialty, it is most important to acquire as much information as possible.
Intra-X2	Intrapersonal	External	When choosing a medical

			specialty, it is most important to seek direction from informed experts.
Intra-C1	Intrapersonal	Crossroads	When choosing a medical specialty, it is most important to consider my own opinions and views.
Intra-E1	Intrapersonal	Early	When choosing a medical specialty, it is most important to consider the available information and my own views.
Epis-C1	Epistemological	Crossroads	To make a good choice about a medical specialty, I think that it is largely a matter of personal opinion.
Epis-X1	Epistemological	External	To make a good choice about a medical specialty, I think the facts (e.g., lifestyle) are the strongest basis for a good decision.
Epis-X2	Epistemological	External	To make a good choice about a medical specialty, I think the faculty is in the best position to advise me about a good choice.
Epis-E1	Epistemological	Early	To make a good choice about a medical specialty, I think it is not a matter of facts or expert judgment, but of a match between my values, interests, and skills and those of the job.
Epis-E2	Epistemological	Early	To make a good choice about a medical specialty, I think that advisors or mentors can provide advice to consider along with my own ideas.
Epis-C2	Epistemological	Crossroads	When people have different interpretations of

			a journal article, I think that some articles are just that way. It is possible for all interpretations to be correct.
Epis-E3	Epistemological	Early	When people have different interpretations of a journal article, I think that multiple interpretations are possible, but some are closer to the truth than others.
Epis-X3	Epistemological	External	When people have different interpretations of a journal article, I think only one interpretation can be right.
Epis-X4	Epistemological	External	When people have different interpretations of a journal article, only experts can say which interpretation is really correct.
Epis-E4	Epistemological	Early	When people have different interpretations of a journal article, I think their ideas should be compared to determine which makes more sense to me.
Inter-C3	Interpersonal	Crossroads	If a teacher or advisor recommended a medical specialty that I have never considered before, I would share my opinion about it.
Inter-E2	Interpersonal	Early	If a teacher or advisor recommended a medical specialty that I have never considered before, I would try to understand their point of view and how it would best fit my needs

			and interests.
Epis-E5	Epistemological	Early	Experts are divided on some scientific issues, such as the causes of global warming. In a situation like this, I would have to look at the evidence and come to my own conclusion.
Epis-E6	Epistemological	Early	Experts are divided on some scientific issues, such as the causes of global warming. In a situation like this, I think it is best to accept the uncertainty and try to understand the principal arguments behind the different points of view.
Epis-X5	Epistemological	External	Experts are divided on some scientific issues, such as the causes of global warming. I rely on the experts to decide.
Epis-X6	Epistemological	External	Experts are divided on some scientific issues, such as the causes of global warming. I prefer not to judge one opinion over another.
Intra-C2	Intrapersonal	Crossroads	My choice of medical specialty may change over time based upon my experiences in medical school.
Intra-C3	Intrapersonal	Crossroads	I know what I want to personally achieve from the specialty I choose to practice.
Intra-E2	Intrapersonal	Early	I have been able to remain true to myself during the last year.

APPENDIX C

Correlation and Covariance Matrix for the Full Data Set

	Inter-X1	Inter-C1	Inter-C2	Inter-X2	Inter-E1	Intra-X1	Intra-X2	Intra-C1	Intra-E1	Epis-C1	Epis-X1	Epis-X2	Epis-E1	Epis-C2
Inter-X1	.412	.098	-.052	-.056	.062	.025	.179	-.072	-.127	.029	.150	.156	-.101	.037
Inter-C1	.041	.428	.357	.194	.127	.231	.218	.228	.143	.117	.230	.199	.033	.022
Inter-C2	-.020	.141	.366	.311	.271	.312	.183	.260	.181	.087	.196	.111	.063	.037
Inter-X-2	-.024	.085	.126	.451	.175	.386	.109	.192	.162	.051	.062	-.025	.185	.092
Inter-E1	.027	.056	.111	.080	.460	.195	.488	.285	.198	-.074	.064	.095	.059	.081
Intra-X1	.011	.103	.129	.178	.091	.468	.249	.180	.154	-.045	.155	.139	.147	.034
Intra-X2	.075	.093	.072	.048	.216	.111	.425	.405	.233	.096	.237	.280	-.071	.066

	Epis-E2	Epis-E3	Epis-X3	Epis-X4	Epis-E4	Inter-C3	Inter-E2	Epis-E5	Epis-E6	Epis-X5	Epis-X6	Intra-C2	Intra-C3	Intra-E2
Inter-X1	-.221	-.079	.124	-.158	-.022	.073	-.045	-.058	-.051	-.116	.122	-.190	.156	.055
Inter-C1	.211	.060	.035	-.044	.060	-.100	.120	.149	-.002	-.075	-.086	-.056	-.058	.069
Inter-C2	.220	.168	-.082	.023	.029	-.003	.228	.099	.110	.003	-.058	.083	.047	.123
Inter-X-2	.160	.174	-.010	-.055	.057	.050	.141	.015	.086	-.088	.046	-.017	.047	.177
Inter-E1	.148	-.113	-.059	-.011	-.070	.148	.157	.089	.054	-.036	.018	.062	.023	.044
Intra-X1	.102	.095	-.016	-.126	-.023	-.073	.093	.169	.029	-.071	.091	.011	.161	.153
Intra-X2	.146	-.024	.036	-.113	-.113	.021	.109	.038	.042	-.031	.012	-.049	.144	.031

(continued)

	Inter-X1	Inter-C1	Inter-C2	Inter-X2	Inter-E1	Intra-X1	Intra-X2	Intra-C1	Intra-E1	Epis-C1	Epis-X1	Epis-X2	Epis-E1	Epis-C2
Intra-C1	-.028	.090	.095	.078	.117	.074	.159	.365	.428	.185	.189	.158	.182	.124
Intra-E1	-.045	.052	.060	.060	.074	.058	.084	.142	.303	.102	.017	.027	.232	.016
Epis-C1	.012	.050	.034	.023	-.033	-.020	.041	.074	.037	.423	.170	-.014	.137	.100
Epix-X1	.063	.098	.078	.027	.028	.069	.101	.074	.006	.073	.427	.284	-.063	.097
Epis-X2	.063	.082	.042	-.011	.041	.060	.115	.060	.009	-.006	.117	.399	-.061	-.036
Epis-E1	-.038	.012	.022	.073	.024	.059	-.027	.064	.075	.053	-.024	-.023	.344	.071
Epis-C2	.018	.011	.017	.048	.043	.018	.033	.058	.007	.051	.049	-.018	.032	.597

	Epis-E2	Epis-E3	Epis-X3	Epis-X4	Epis-E4	Inter-C3	Inter-E2	Epis-E5	Epis-E6	Epis-X5	Epis-X6	Intra-C2	Intra-C3	Intra-E2
Intra-C1	.219	.075	-.033	.186	-.079	.048	.169	.072	.012	.095	-.154	.130	.108	.135
Intra-E1	.394	.090	-.102	.085	.047	.102	.255	.124	.116	.008	.088	.185	.066	.096
Epis-C1	.030	.005	.128	.106	.119	.042	.10	.118	-.047	-.056	-.013	.039	.029	.009
Epix-X1	.093	.086	.032	.107	.106	-.182	.097	.165	.008	.039	.012	-.034	.158	.063
Epis-X2	.160	.070	.134	-.151	.095	.132	.132	.114	.066	.086	.102	-.010	-.028	.038
Epis-E1	.226	.077	-.062	.154	.048	-.050	.136	.021	.087	.010	-.054	.151	.050	.126
Epis-C2	.111	.210	-.254	.102	.067	.144	-.033	.104	-.051	-.027	.122	-.026	.189	.084

(continued)

	Inter-X1	Inter-C1	Inter-C2	Inter-X2	Inter-E1	Intra-X1	Intra-X2	Intra-C1	Intra-E1	Epis-C1	Epis-X1	Epis-X2	Epis-E1	Epis-C2
Epis-E2	-.077	.075	.072	.058	.054	.038	.051	.072	.118	.011	.033	.055	.072	.047
Epis-E3	-.029	.023	.059	.068	-.045	.038	-.009	.026	.029	.002	.032	.026	.026	.094
Epis-X3	.052	.015	-.032	-.004	-.026	-.007	.015	-.013	-.036	.055	.014	.055	-.023	-.128
Epis-X4	-.070	-.020	.010	-.026	-.005	-.061	-.051	.078	.032	.048	.048	-.066	.062	.054
Epis-E4	-.008	.022	.010	.021	-.026	-.009	-.040	-.026	-.014	.043	.038	.033	.015	.028
Inter-C3	.025	-.035	-.001	.018	.053	-.027	.007	.015	.030	.015	-.063	.044	-.016	.059
Inter-E2	-.014	.037	.064	.044	.050	.030	.033	.048	.065	.037	.030	.039	.037	-.012

	Epis-E2	Epis-E3	Epis-X3	Epis-X4	Epis-E4	Inter-C3	Inter-E2	Epis-E5	Epis-E6	Epis-X5	Epis-X6	Intra-C2	Intra-C3	Intra-E2
Epis-E2	.294	.169	-.204	.108	.171	.097	.364	.298	.177	.047	-.106	.300	.098	.149
Epis-E3	.053	.337	-.121	-.012	.213	.058	.163	.124	.175	.013	-.093	.032	.042	.081
Epis-X3	-.072	-.046	.423	-.438	.059	-.009	-.074	-.121	-.057	-.230	.023	-.157	-.027	-.046
Epis-X4	.041	-.005	-.197	.476	.011	.014	.111	-.007	-.040	.214	-.127	.169	.049	.042
Epis-E4	.051	.068	.021	.004	.299	.124	.139	.091	.136	-.028	.052	-.043	.127	.067
Inter-C3	.028	.018	-.003	.005	.036	.281	.310	-.009	.097	.105	.037	.039	-.050	.105
Inter-E2	.092	.044	-.022	.035	.035	.076	.216	.248	.206	.077	-.177	.211	.065	.205

(continued)

	Inter-X1	Inter-C1	Inter-C2	Inter-X2	Inter-E1	Intra-X1	Intra-X2	Intra-C1	Intra-E1	Epis-C1	Epis-X1	Epis-X2	Epis-E1	Epis-C2
Epis-E5	.025	.065	.040	.007	.040	.077	.017	.029	.046	.052	.072	.048	.008	.054
Epis-E6	-.023	-.001	.047	.041	.026	.014	.020	.005	.045	-.022	.004	.029	.036	-.028
Epis-X5	-.052	-.034	.001	-.041	-.017	-.034	-.014	.040	-.003	-.026	.018	.038	.004	-.015
Epis-X6	.053	-.038	-.024	.021	.008	.042	.005	-.063	-.033	-.006	.005	.043	-.021	.063
Intra-C2	-.063	-.019	.026	-.006	.022	.004	-.017	.041	.053	.013	-.011	-.003	.046	-.010
Intra-C3	.073	-.028	.021	.023	.011	.080	.068	.047	.026	.014	.075	-.013	.021	.106
Intra-E2	.023	.030	.049	.077	.019	.068	.013	.053	.034	.004	.027	.016	.048	.042

	Epis-E2	Epis-E3	Epis-X3	Epis-X4	Epis-E4	Inter-C3	Inter-E2	Epis-E5	Epis-E6	Epis-X5	Epis-X6	Intra-C2	Intra-C3	Intra-E2
Epis-E5	.108	.048	-.053	-.003	.033	-.003	.077	.445	.104	.177	.071	.163	.028	.125
Epis-E6	.068	.072	-.026	-.020	.053	.036	.068	.050	.507	-.155	.231	.047	-.076	.004
Epis-X5	.018	.005	-.105	.103	-.011	.039	.025	.082	-.077	.490	-.310	.170	-.022	-.004
Epis-X6	-.039	-.036	.010	-.059	.019	.013	-.055	.032	.110	-.146	.449	-.227	-.018	-.063
Intra-C2	.064	.010	-.053	.060	-.012	.011	.051	.057	.017	.062	-.079	.269	-.001	.211
Intra-C3	.039	.018	-.013	.025	.051	-.019	.022	.014	.040	-.011	-.009	.000	.528	.405
Intra-E2	.053	.031	-.019	.019	.024	.036	.062	.054	.002	-.002	-.028	.071	.192	.425