

5-2-2014

Creativity-Based Music Learning: Modeling the Process and Learning Outcomes in a Massive Open Online Course

Nicholas Michael Stefanic
University of South Florida, nstefanic@mail.usf.edu

Follow this and additional works at: <https://digitalcommons.usf.edu/etd>



Part of the [Other Education Commons](#)

Scholar Commons Citation

Stefanic, Nicholas Michael, "Creativity-Based Music Learning: Modeling the Process and Learning Outcomes in a Massive Open Online Course" (2014). *USF Tampa Graduate Theses and Dissertations*. <https://digitalcommons.usf.edu/etd/5315>

This Dissertation is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact digitalcommons@usf.edu.

Creativity-Based Music Learning:
Modeling the Process and Learning Outcomes in a Massive Open Online Course

by

Nicholas M. Stefanic

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Music Education
College of The Arts
University of South Florida

Co-Major Professor: Clint A. Randles, Ph.D.
Co-Major Professor: Peter R. Webster, Ph.D.
C. Victor Fung, Ph.D.
Jennifer A. Bugos, Ph.D.
David A. Williams, Ph.D.
Robert F. Dedrick, Ph.D.

Date of Approval:
May 2, 2014

Keywords: Composition, Engagement, Motivation, MOOC, Musical Experience

Copyright © 2014, Nicholas M. Stefanic

TABLE OF CONTENTS

List of Tables	vi
List of Figures	viii
Abstract	x
Prologue: Visions and Re-visions	1
Chapter One: Introduction	3
Creativity in American Music Education	4
Creativity-Based Music Learning	7
Indicators of Learning	8
Operationalizing and Measuring Components	10
Specifying Relationships Between Components	10
Technical and Logistical Issues	11
The Proposed Solution	12
Initial Purpose and Rationale for the Study	13
Initial Research Questions	15
Revised Research Questions	16
Delimitations	18
Chapter Two: Review of Literature	20
Perspectives on Creativity	20
Creativity Theories	21
Theories of Musical Creativity	23
Cognitive models	23
Socio-cultural models	24
Systems models	25
Empirical Model Evaluation	27
Model Specification	28
Enabling Skills	30
Aptitude	30
Conceptual understanding	33
Enabling Conditions	34
Motivation	34
Self-determination theory	34
Motivation and creativity	40
Motivation and musical creativity	41

Measuring motivation.....	43
Modelling motivation	44
Personality	45
Big Five factors	46
Measuring personality	49
Self-efficacy.....	50
Context	53
Past experience	57
The Creative Process	58
Musical content analysis	59
Verbal content analysis	60
Creative process and learning.....	66
Engagement	66
Divergent and convergent thinking.....	71
The General Specified Model of Creativity-Based Learning	74
Global level	75
Contextual level	75
Situational level	76
Sub-Models	77
Chapter Three: Method	81
Research Context.....	81
Participants.....	84
Measurement Process	84
Primary Measures	86
Demographics	88
Age	88
Education	88
Past experience in music	89
Music aptitude	89
Personality	90
Contextual support.....	92
Musical creative self-efficacy.....	93
Situational motivation.....	93
Situational cognitive engagement.....	94
Learning outcomes.....	95
Self-evaluation of creative product	95
Perceived learning	98
Additional Measures	98
Project participation	99
Welcome to Canvas Network survey.....	99
Primary reason	99
Type of learner	100
English	101
Geographic region	101

Plan A: Proposed Data Analysis	101
Data screening	102
Identification.....	102
Model estimation	102
Model comparison	103
Equivalent models.....	103
Model interpretation.....	103
Model modification.....	104
Plan B: Proposed Data Analysis.....	104
Plan C: Actual Data Analysis for Revised Research Questions	105
Instrument development data analysis overview	105
Primary study data analysis overview.....	107
Chapter Four: Instrument Development.....	110
Initial Item Pools	111
Expert Review	112
Pilot Study #1	113
Pilot Study #2.....	114
Revisions to PEM index.....	114
Calculating the PEM total variable.....	115
“Frequency/recency” questions	117
“Number of years” questions.....	117
“Listening” question	117
Rationale for indexing to 1	118
PEM total calculation.....	118
PEM latent class analysis	119
Unconditional models.....	121
Conditional models.....	124
Musical creative self-efficacy.....	127
Convergent validity.....	130
Chapter Five: Results	140
Description of the Final Dataset.....	140
Descriptive Statistics for Each Variable	141
Demographics	143
Past experience in music	145
Musical creative self-efficacy.....	147
Music aptitude	150
Personality	151
Motivation	153
Perceived learning.....	155
Project participation	155
Primary reason	157
Expected hours.....	161
Type of learner and expected hours per week	161
Logistic Regression Analyses	163

Type of Learner Group Differences	167
Continuous variables	168
Categorical variables	170
The General Specified Model Explored	172
Statistical significance and missing data	173
A rough sketch	173
Demographic variables	174
Past experience in music	175
Musical creative self-efficacy	175
Personality	175
Music aptitude (PROMS)	176
Contextual support	176
Motivation	178
Situational engagement and perceived learning	179
Chapter Six: Discussion	180
Contributions and Issues with Measurement	180
Past experience in music	180
Musical creative self-efficacy	185
Other considerations and future measurement research	186
Measurement invariance and the PEMI	186
Measurement invariance and the MCSES	187
Issues with the latent class measurement model and the PEMI	188
Latent class analysis with type of learner	189
Moving beyond implicit measurement and structural modeling	190
Creative Music-Making Project Participation	191
Validity of the type of learner framework	192
Past experience in music and participation	193
Group Differences for “Active Participants”	196
Other Avenues for Exploring Participation	199
More Questions	201
Creativity-based learning	202
Why should we care about participation/engagement in a MOOC?	203
What about MOOCs as a research tool?	210
Feasibility issues	210
Unique opportunities	212
What about the model?	214
Epilogue: Future Directions	217
Try It Again	217
Advance the Research on the PEMI and MCSES	218
Institutional Collaboration	218
Creativity-Based Learning Research	219
Onward	220
References	221

Appendix A: The General Specified Model of Creativity-Based Learning	238
Appendix B: The General Specified Model of Creativity-Based Learning (Full Structural Model)	239
Appendix C: Measurement Items	240
Examples of Questionnaire Appearance in Canvas.....	240
Welcome to Canvas Network Survey.....	241
Pre-Course Questionnaire	243
Musical Creative Self-Efficacy	247
Personality.....	248
Contextual Support.....	249
Situational Motivation	251
Situational Cognitive Engagement	252
Perceived Learning.....	253
Appendix D: Equations For Combining Frequency and Recency Variables	254
Appendix E: Profile Plots from PEM Latent Class Analysis (Pilot Study #2)	255
Appendix F: Profile Plots for PEM Latent Class Analysis (MOOC study data).....	257
Appendix G: Intercorrelations (Kendall's Tau) Between All Non-Nominal Variables	261
Appendix H: Plots for Possible PEM Frequency and Recency Comparisons.....	266
Appendix I: IRB Approval Letter	269

LIST OF TABLES

Table 3.1. Initial Measurement Schedule	86
Table 3.2. Summary of Measures for All Variables	87
Table 3.3 Education Variable Coding Scheme	88
Table 3.4 Expected Hours per Week Response Options and Coding	101
Table 4.1 Listening Experience Coding Scheme	118
Table 4.2 Unconditional PEM Models for Latent Class Analysis (Pilot Study 2)	122
Table 4.3 Conditional Models (Age and Gender as Covariates) for Latent Class Analysis (Pilot Study 2).....	125
Table 4.4 Classification of Individuals Based on Their Most Likely PEM Class Membership.....	126
Table 4.5 Reliability Estimates and Confidence Intervals for MCSE Sub-Scales (Pilot Study #2).....	129
Table 4.6 Validity Coefficients for PEM Class and Musical Creative Self-Efficacy.....	131
Table 4.7 Logistic Regression MIMIC Model Results for MCSE Predicted by PEM Class	133
Table 4.8 Fit Statistics and Model Comparison Statistics for MIMIC Models for MCSE on PEM.....	134
Table 5.1 Descriptive Statistics for All Measured Variables	142
Table 5.2 Unconditional PEM Models for Latent Class Analysis.....	145
Table 5.3 Conditional (Age, Gender, and Education as Covariates) PEM Models for Latent Class Analysis.....	146
Table 5.4 Classification of Individuals Based on Their Most Likely PEM Class Membership.....	147

Table 5.5 Reliability Estimates for Musical Creative Self-Efficacy Sub-Scales and Overall Scale	148
Table 5.6 Descriptive Statistics for Musical Creative Self-Efficacy Sub-Scales and Overall Scale	148
Table 5.7 Paired Sample T-Tests and Correlations for MCSE Sub-Scales.....	150
Table 5.8 Descriptive Statistics for Scores on the PROMS Sub-Scales	151
Table 5.9 Inter-Item Correlations for Scores on the PROMS Sub-Scales	151
Table 5.10 Descriptive Statistics for Scale Scores for Each Personality Factor	152
Table 5.11 Reliability Estimates for Scale Scores for Each Personality Factor	152
Table 5.12 Descriptive Statistics for Motivation	153
Table 5.13 Correlations Between Motivation Sub-Scales.....	154
Table 5.14 Descriptive Statistics for Situational Engagement	154
Table 5.15 Descriptive Statistics for Perceived Learning (General)	155
Table 5.16 Correlation Matrix for Perceived Learning Items	156
Table 5.17 Frequencies and Proportions for Type of Learner	159
Table 5.18 Type of Learner by Any Project Participation Cross Tabulation	159
Table 5.19 Frequencies and Proportions for Expected Hours/Week	161
Table 5.20 Proportions of Individuals Expected Hours/Week Participation by Type of Learner (Active/Non-Active)	162
Table 5.21 Results from Univariable Logistic Regression (Any Project Participation)	165
Table 5.22 Logistic Regression Results from Step 2 (Any Project Participation)	166
Table 5.23. Results from Independent T-Tests for “Active Participant” Group Differences	169
Table 5.24 Results from Chi-Square Tests for “Active Participant” Group Differences	171

LIST OF FIGURES

Figure 2.1. Path model for motivation implied by Vallerand's (1997) Hierarchical Model of Intrinsic and Extrinsic Motivation.....	45
Figure 2.2. MCSE Sub-Model 1.....	77
Figure 2.3. MCSE Sub-Model 2.....	78
Figure 2.4. Contextual Support (Self-Determination Theory) Sub-Model.....	79
Figure 2.5. Situational Need Satisfaction Sub Model.....	80
Figure 2.6. Situational Autonomous Motivation Sub-Model.....	80
Figure 3.1. Proportions of Participants by Geographic Region.....	85
Figure 3.2. Self-Evaluation of Creative Product (SECP) measurement models.....	97
Figure 4.1. Possible Response Formats for MCSE Items.....	112
Figure 4.2. MIMIC model specification and results for testing PEM group/class differences on three MCSE factors.....	135
Figure 5.1. Complete and Missing Data by Measure.....	141
Figure 5.2: Distribution of Age of Participants.....	144
Figure 5.3: Distribution of Education of Participants.....	144
Figure 5.4. Boxplots for Scores on MCSE Sub-Scales.....	149
Figure 5.5. Percent Creative Music-Making Project Participation.....	157
Figure 5.6. Percentages of each Primary Reason for Taking an Open Online Course.....	158
Figure 5.7. Any Project Participation by Type of Learner.....	160
Figure 5.8. Proportions of Project Participants (and Non-Participants) by Type of Learner.....	160
Figure 5.9. Differences in Scores for MCSE Factors Between "Active Participants" and Not "Active Participants".....	170

Figure 5.10. Percentages of "Active Participants" and Not "Active Participants" by Education.....	172
Figure 6.1. Proportions of "Active Participants" and Non-"Active Participants" by PEM Class	195

ABSTRACT

While developing creativity is an important goal of many educational endeavors, creating music, from a music education perspective, is a powerful pedagogical tool. Beyond comparing the relative creativity of individuals' musical creative products (e.g., melodies, songs, lyrics, beats, etc.), research in musical creativity must consider how engaging in the creative process can be an effective teaching tool, what I have termed creativity-based music learning. If music teachers are to develop students' abilities "to experience music as meaningful, informed by sensitive discernments and broad understandings, in each particular musical role engagement in which one becomes involved" (Reimer, 2003, p.214), then we must gain a better understanding of how different aspects of the person and context interact during the creative process.

Based on the available literature, Webster (1987a, 2002) conceived the *Model of Creative Thinking in Music* as a conceptual model for understanding the importance of various components that are at work in the musical creative process. Since, generally speaking, learning results from thinking of some sort, Webster's model represents a reasonable starting point from which to examine how musical creative thinking leads to musical learning. There is much research in music education and the general creativity literature that has investigated how these various component parts (e.g., music aptitude, personality, motivation, previous experience, context) relate to creativity, but there has yet to be any substantive attempt to understand how all of these various elements simultaneously interrelate during a given musical creative process. More importantly, there is limited research on how creativity-based music learning contributes to

important learning outcomes such as students' perceptions of learning from the process, students' self-evaluations of creative products (e.g., songs they have written), the development of conceptual understandings, and the development of musical creative self-efficacy.

The initial primary purpose of this study was to develop and identify a statistical model that best represents the nature of the various interrelationships of components of the musical creative process, as identified in Webster's (2002) model, and as they relate to learning outcomes. Understanding how all of these components relate and ultimately impact various learning outcomes has important implications for how we educate our music students.

Data were collected from students taking a Massive Open Online Course entitled "What is Music?: Finding Your Song," which was designed, developed, and taught by the researcher, and offered in January 2014 through the Canvas Network. In the course, the question "what is music?" was approached from several perspectives, including Music as Human Activity, Music as Emotion, Music as Physics, and Music as Form. While learning about each perspective, students were encouraged to engage with and complete various musical creative projects (e.g., creating a representative playlist, writing lyrics, writing a melody, writing a song). Such an educational context in which creativity is used as a pedagogical tool provided an opportunity for studying the educational outcomes of such an approach. Embedded within the course were measures of several predictors of learning (based on Webster's model), including past experience in music, personality, music aptitude, contextual support, musical creative self-efficacy, motivation, and situational engagement.

Initial analysis plans included the use of structural equation modeling to (1) compare and contrast the statistical fit of competing models; and (2) examine how each of these constructs not only relate to each other, but also how they each contribute (uniquely and in combination) to

various learning outcomes, including perceptions of learning, self-evaluations of creative products, and musical creative self-efficacy. However, a sufficient number of students did not engage in and complete the creative projects, nor did a sufficient number of students complete all of the research items, in order to examine the full structural model. When it became apparent that sufficient data would not be available, the study was re-envisioned to examine questions about why students chose to participate or not participate in the creative music-making projects.

Data were collected from 281 students, and although missing data was quite extreme for variables measured late in the course (e.g., motivation), large amounts of data were available regarding students' past experience in music, their expectations regarding participation as MOOC learners, and demographic information (e.g., age, gender, education, language, geographic region). The available data were used in an exploratory manner to derive a model for predicting creative project participation in the course.

The sole important predictor of project participation was whether students identified themselves as an "active participant" at the beginning of the course, although this variable explained only a small amount of variability in project participation. Follow-up analyses for group differences in Active Participant (individuals who identified themselves as "active participants" versus all other Types of Learners) found that "active participants" had significantly higher levels of Musical Creative Self-Efficacy, greater perceptions of the learning context as challenge-supportive, and higher scores on the Openness personality factor. Notably, students' Past Experience in Music appeared to be unrelated to both whether they *intended* to participate in the creative music-making projects and whether they *actually* participated in the projects.

In addition to the primary MOOC study, the development and initial validation procedures and results for two new research instruments utilized in the MOOC study, the Past Experience in Music Inventory (PEMI) and the Musical Creative Self-Efficacy Scale (MCSES), are described in detail. The latent class measurement model utilized for measuring Past Experience in Music is a unique and potentially valuable approach for measuring this important variable in music research of all kinds.

Finally, an exploratory analysis of all zero-order rank-order intercorrelations of all non-nominal variables indicated some initial support for the General Specified Model of Creativity-Based Learning. It was not possible to take the next step with the model: to prune it, alter it, or reject it altogether, but when viewed as a very large-scale pilot study, this study did provide enough evidence to warrant investing the considerable amount of resources necessary to take that next step.

Implications for creativity-based music learning and the significance of MOOCs and MOOC research are discussed. In particular, music MOOCs represent an opportunity to fill in some much needed space for lifelong learning. However, if we are to promote lifelong musical engagement, then the pedagogy within a MOOC should also promote engagement. As such, questions and further research regarding such engagement, especially within a creativity-based learning framework, are central to better understanding how to promote and facilitate lifelong musical engagement and musical learning.

PROLOGUE: VISIONS AND RE-VISIONS

I have chosen to begin this dissertation with a section that is common in literary works, not because what follows is a literary work, but because in addition to the standard components of a dissertation (e.g., introduction, review of literature, results, discussion), there is also a story to be told with this research. I begin with a prologue because the story that follows has several twists and turns and it may be helpful to be able to anticipate those twists and turns. What began as a study about creativity-based music learning, which happened to take place in a Massive Open Online Course (MOOC), eventually became a study about a MOOC, which happened to utilize a creativity-based music learning approach.

In my pre-MOOC world (i.e., prior to actually teaching a MOOC), I wanted to examine how engaging in creative music-making resulted in learning. I was operating under the premise that there are things that can be learned when one creates music that are either different or cannot be learned from a performance-focused or listening-focused approach to music education. I saw the MOOC as a convenient way to mitigate several technical and logistical issues related to carrying out my research. In a perfect world (at least a world perfect for a Ph.D. candidate), I would have designed a study that was well within my grasp, and then carried out that study as planned, analyzed the data as planned, and drawn conclusions that had implications within the context for which the study was designed. In my real world, I designed a study that required very large numbers of individuals to participate, but that took place in a context that is well-known to have very low retention rates, the Massive Open Online Course. About midway

through the course it became apparent that the vast majority of students were not engaging in the creative projects (at least not in a way that was observable by me), and the number of individuals completing the research items was drastically decreasing over time. It is difficult to study learning and creative music-making if the creative music-making is not occurring. Rather than conceding to a failed study, I re-envisioned the study to uncover potential reasons why students did or did not engage in the creative music-making projects and to explore the characteristics of the different types of learners in my course.

I have chosen to keep (with revisions) much of what I had written prior to beginning my study for two reasons. First, the data I actually collected and the analysis I actually carried out on those data would not have made sense when viewed in isolation from the initial study design and study intentions. Second, I view the preparatory work, including the research concept, the review of literature, the General Specified Model of Creativity-Based Learning that resulted from the review of literature, the review and selection of appropriate measures, and the development of my own measures, as valuable contributions to the field of music education, even in the absence of any results that may have come from the research.

Finally, because my research did not follow a clear linear trajectory, but does indeed have a story bound up with the design and the analysis, there are portions of my writing that are in a more narrative style, which is possibly different than what would be expected of a clearly quantitative-focused research study. I trust that in doing so, a human dimension to this research will emerge amid the much less “colorful” tables, statistics, and analyses, but will also not detract from perceiving my sincere efforts at producing high quality research. A dissertation on creativity would be at least somewhat hypocritical if I did not grant myself the use of *my own* creativity.

CHAPTER ONE: INTRODUCTION

Creativity and creative thinking are fundamental characteristics of human functioning that underlie virtually all aspects of activity in daily life. Creative thinking allows one to do things as simple as carrying on a conversation with another individual or making a sandwich, to things as complex as designing or constructing a building, deriving theories of physics, writing a computer program, scripting a screenplay, or composing a piece of music. We use creative thinking to both solve problems and to identify problems. The central nature of creative thinking to human functioning has made the development of creative thinking an important goal for educational institutions. Because we exist within complex social environments, the creations (ideas and products) generated by our creative thinking vary considerably in the degree to which they are appreciated by others as being unique, useful, well-made, aesthetically pleasing, or elegant, and this is dependent on what or how many creations upon which one bases judgment (e.g., comparison within a small group at one point in time, comparison within an individual over time, comparison across a society over its history). Understanding this variability has been the subject of a substantial body of research in both educational and organizational contexts.

Research in creativity has tended to focus on between-person differences. Researchers have studied how different aspects of the person, process, and context impact some creative product. The general goal in this type of research is to better understand what makes one person more creative than another, or what components of the creative process make one creative product more creative than others. In an education setting, these research aims make sense when

the goal of an educational institution is to develop students' abilities to solve problems and innovate. While critical thinking and reasoning abilities have long been (within the last few centuries) the objective of public and private educational institutions, the development of creative abilities has only recently (within the last few decades) become an explicit educational objective.

The music education profession in the United States has followed a similar trajectory, although much more condensed in time. Researchers in music education have sought to understand the notion of creativity as it is manifest in music. But there is an important difference between the purpose of creativity in music education and the purpose of creativity in general education. In general education, the purpose of including curricular components that involve creativity is to develop innovators, individuals that can solve the world's (and our country's) problems, create new products and services, and function as productive members of society. In such a context, differences between more and less creative people or more and less creative products are very important. However, in a music education context, the purpose of including curricular components that involve creative thinking, as I see it, has very little to do with developing musical innovators (although this might be a beneficial result). Instead, engaging in creative thinking is a means to an *individual educational* end more than it is a means to a *societal* or *national innovational* end.

Creativity in American Music Education

For a very long time in American music education, the primary purpose of educating children in music was either functional (i.e., we need singers and musicians that can make music for others to enjoy and use) or cultural, to develop in individuals an appreciation for music as an art (i.e., aesthetic education) (Mark & Gary, 2007). For much of the history of American music

education, this was accomplished by either direct instruction in listening to music or direct instruction in learning how to perform music, or more specifically, how to replicate music created by someone else. Over the last (less than) half a century, the profession has recognized the importance of other musical behaviors, namely composition and improvisation. The concept of comprehensive musicianship was developed, and the implicit goal with comprehensive musicianship was to develop a more well-rounded musician (CMP & MENC, 1965). One might consider this a focus on *breadth* of knowledge and ability.

While some changes have been made, music performance still holds the most prominent position in many US music education curriculums (Orman, 2002; Williams, 2007), and composition teaching tends to be prescriptive and notation-based (Morris, 2010). With composition and improvisation still relegated to the position of “we’ll do that if we have time once the concert music is ready,” scholars and researchers in the profession have honed their argument: there is music learning that can happen when one creates new music that either cannot happen or does not happen to the same degree when one only re-creates someone else’s music (Hickey & Webster, 2001; Kratus, 1990b; Randles, 2012; Randles & Stringham, 2013; Reimer, 2003; Stringham, 2010; Webster, 1987a, 1990, 2002). As the argument goes, there is a *depth* of understanding that is possible when one works through a musical creative process that is either not possible or qualitatively different than when one works through a musical (re-)productive process.

When depth of understanding is the goal, and not musical innovation, then understanding the creative process is important insofar as it represents a manner in which new knowledge and understanding is constructed and negotiated during generative processes (Lehmann, Sloboda, & Woody, 2007; Sloboda, 2000). The vast majority of research in general creativity (and musical

creativity) tends to focus on the former goal: innovation. Researchers have studied aspects of the person (e.g., personality, intelligence) (Batey & Furnham, 2006; Kim, 2005), process (Claxton, 2006; Collins, 2007; Mumford, Medeiros, & Partlow, 2012), and context (Amabile, 1983, 1996) to see how those aspects relate to differences in creative thinking abilities (e.g., divergent thinking) (Collins, 2005; Dietrich & Kanso, 2010; Torrance, 1966; Webster, 1987b, 2002) or differences in relative creativity of actual created products (Baer, Kaufman, & Gentile, 2004; Cropley, Kaufman, & Cropley, 2011; Hickey, 2001; J. C. Kaufman, Baer, Cole, & Sexton, 2008; Priest, 2006b). Implicit in much of this research is the belief that we should want all people to be *more* creative, that is, be able to generate products and solutions that are more novel, innovative, elegant, or useful than their predecessors.

However, in music education (and probably most arts education), I am not convinced our role as music teachers is to develop students' musical creative abilities such that they may continue to innovate the art form. Historically, the innovation of an art form seems to carry on its merry way regardless if artists are actually trained to innovate. To follow Reimer (2003) and Webster (2013), I believe our primary role as music educators is to improve the musical intelligence of our students, that is, to expand their abilities "to make increasingly acute discriminations, as related to increasingly wide connections" (Reimer, 2003, p. 204) within music. Or put another way, "to experience music as meaningful, informed by sensitive discernments and broad understandings, in each particular musical role engagement in which one becomes involved" (p. 213). Webster (2013) contends:

The ability to make discriminations (differentiations) in increasingly more subtle ways is a clear example of gaining a finer understanding of patterns of experience. Relating this to wider realization of connections broadens and deepens this understanding. (p. 7)

The musical creative act, defined as bringing sounds into existence that are organized to be inherently meaningful (Reimer, 2003), requires one to delve into this messy world of fine discriminations and broad connections and engage with it in ways that are perhaps different than performing or listening to music, which in turn possibly leads to different (not necessarily better) understandings. This represents a focus on *depth* of knowledge and ability.

Creativity-Based Music Learning

I use the term *creativity-based music learning* (or also *creativity as pedagogy*) to refer to the notion of using creative music-making for the purposes of gaining deeper and broader understandings of music. Framed in this way, creative music-making can be seen as the educational context within which learning of some sort occurs. The rationale for teaching creative thinking is therefore somewhat analogous to the long-held rationale for teaching students critical thinking (Moore & Stanley, 2010), that it improves academic achievement; a critical and creative mind is a well-functioning one. While there is much research to back up the claim regarding critical thinking, the connection between musical creativity as pedagogy and actual learning outcomes still requires more empirical support.

What makes studying this connection difficult is the complexity of the creative process. A large body of research (discussed in more detail in Chapter Two) has established connections between individual factors (e.g., intelligence, personality) and differences in creative achievement, creative thinking, and problem-solving styles. Research in motivation has demonstrated the pivotal role that motivation plays in learning outcomes. What are missing are the connections between the individual factors and the learning outcomes. Again, this gap in the literature is primarily a result of focusing on creative thinking for promoting innovation as opposed to creative thinking for promoting learning.

One way we attempt to make sense of the complexity of our world is by creating models. Models can vary in their degree of specificity, breadth of explainable contexts, and degree of empirical support. Webster (1987a, 2002) has advanced a model of the creative thinking process in music that identifies factors that come to bear on the musical creative process, and which forms the basis of the present study. The model, as a basis for the workings of the creative process, can then be used to investigate how the components of the creative process contribute to learning outcomes in the context of creative music-making. In order to do this, several matters must first be addressed, including what is considered evidence of learning, how the components should be operationalized and measured, how the individual relationships between components should be specified, and several other technical and logistical issues.

Indicators of Learning

Webster's model includes several products that result from the creative process, most of which are traditionally considered as examples of musical creativity (e.g., compositions, improvisations, recorded performances, and even written analyses). While the model does not explicitly include learning as a creative product, it does include "mental representations of the music heard." *New* mental representations are evidence of learning, by definition, but mental representations are *internal*, so it is necessary to determine what is considered as *external* evidence of learning. Evidence of learning could be found with an objective measure, such as a formal test of new conceptual understandings. There has also been much research that uses perceived learning as an indicator of learning (Grudnitski & Krentler, 2011; Rovai, Wighting, Baker, & Grooms, 2009; Rowley & O'Dea, 2009), although not without some controversy (Sitzmann, Ely, Brown, & Bauer, 2010). Changes in self-efficacy, one's belief in one's capabilities, can also be an affective indicator of learning (Esfandagheh, Harris, & Oreyzi, 2012;

Sitzmann, Brown, Casper, Ely, & Zimmerman, 2008). If one feels more capable, then it is feasible that some learning or growth of some sort has occurred. Finally, I will argue that it might be possible to conceptualize students' self-evaluations of their created products as indicators of learning, but with several caveats, to which I now turn.

The creativity literature has come to some general agreement on defining something as creative based on the extent to which it is both novel and useful/appropriate (Runco & Jaeger, 2012), although there are other single-criteria (Runco, in press) and three-criteria definitions (Boden, 1991, 2004; Simonton, 2012) in the literature. Based on the available research, students' self-evaluations of the *creativity* of their creative products tend not to be in agreement with ratings of those same products made by experts in the relevant domain (e.g., J. C. Kaufman, Evans, & Baer, 2010; Priest, 2006a). However, as previously discussed, the extent to which a composition is creative in relation to other students is secondary in importance, particularly if novelty is part of one's definition of creativity. What *is* important is the extent to which the created product is useful to the student, or put another way, the extent to which the student sees the musical creation as being appropriate within the musical cultural context. Other components of a self-evaluation might include the extent to which the product is aesthetically pleasing or well-crafted (Amabile, 1983, 1996; Hickey, 2001), or maybe even elegant (Cropley & Cropley, 2008; Cropley & Kaufman, 2012).

In considering these various facets of one's created product, a student makes a judgment regarding the value of that product (a composite of all these facets) for that student. Since we tend to value things more highly the more engaged (i.e., more investment of time and physical, emotional, and cognitive effort) we have been with their creation, value can be seen as an indirect indicator of engagement. Research has made clear that higher levels of engagement lead

to greater and deeper learning in both classroom (Pike, Smart, & Ethington, 2012; Schunk & Pajares, 2005) and online environments (Shea & Bidjerano, 2009; Shea et al., 2012). Therefore, it could be argued that the more engaged a student has been in the creative process, the more likely they will have learned, and the level of engagement will be manifest in the extent to which the student values the product he/she has created. This connection between engagement and product evaluation has scarcely been researched, and although initially included in the present study's design, it remains unexamined.

Operationalizing and Measuring Components

The second matter to be addressed when considering Webster's model is how each component can be operationalized and therefore measured. For example, motivation is a remarkably complex construct, and is operationalized differently depending on the particular motivation theoretical framework in which one chooses to work. While all theories of motivation address the role of motivation in human behavior, a theory that views motivation in regards to one's expectations and value placed on a desired goal measures something different than a theory that views motivation as a product of one's perceptions of autonomy or competence. So it is not enough to simply discuss the role of motivation in the creative process, but it is necessary to choose some motivation theory framework, which in turn determines the type of measurements that one makes. I address this issue in more detail in Chapter Two and Chapter Three.

Specifying Relationships Between Components

A third matter to be addressed is how the individual components of the model relate to each other, when measured quantitatively. Again, the motivation theory to which one ascribes has important implications because the hypothesized antecedents and consequents of motivation

vary by theoretical framework. Expectancy-value theory (Eccles, 1983) holds different assumptions and operationalizes motivation differently than Self-Determination Theory (Deci & Ryan, 1985) or Achievement-Goal Theory (Ames, 1984; Deshon & Gillespie, 2005; Dweck & Leggett, 1988; Vandewalle, 1997). Each theory has different implications for the relationship between motivation and other components. Personality and motivation are both components of the model, and while both are believed to have an impact on the process, it is necessary to specify relationships between *all* components as they are operationalized and measured in the actual study. This specification should be firmly rooted in the research literature from the sub-field to which the component belongs (e.g., creativity and intelligence).

Technical and Logistical Issues

The final set of considerations have to do with technical aspects of carrying out the research. Once measures have been chosen, it is obviously necessary to find some participant sample that can be observed. The sheer number of constructs involved in the model creates logistical issues in terms of collecting that many measurements from the same individuals in a sample, as opposed to utilizing separate samples and following a piece-meal approach. In retrospect, this is a much bigger and more difficult problem to manage than initially anticipated.

There are also technical considerations regarding the statistical analysis of the data. To analyze the data using a technique like Structural Equation Modeling (SEM), very large sample sizes are required. So on top of measuring a large number of constructs, these many measurements must be carried out on a large number of individuals. This introduces yet another logistical issue: finding the time to carry out the measurements and also gaining access to the same large number of individuals in order to carry out those measurements.

Finally, the proposed research must also occur within an educational context in which students actually engage in the creative process and *do* creative thinking. It might be possible to gather a large enough sample by combining many different classrooms from various different schools, but this presents further statistical issues regarding independence of observations. There is no way to control for differences that exist between schools, between teachers, and between individual classrooms. Although this could be managed statistically in a multilevel modeling or multilevel SEM approach, it is difficult to account for the many other confounding variables that could contribute to between-group differences. A situation in which hundreds (or thousands) of students could be brought together under one instructor, in the same educational environment, and at the same time, would help alleviate some of these issues.

The Proposed Solution

The logistical and technical issues just mentioned are by no means trivial and represent a substantial challenge. The solution I chose was to move the research to a web-based environment in which issues of time, place, and space can be managed. As for creating a web-based environment in which thousands of individuals can be educated, the recent phenomenon of the Massive Open Online Course (MOOC) has advanced a model for doing just this.

In 2008, researcher Stephen Downes of the National Research Council of Canada and professor George Siemens of Athabasca University offered a course entitled Connectivism and Connective Knowledge, which enrolled several dozen University of Manitoba students and several thousand online (non-paying) students from around the world. The pedagogical model for the course was based on the tenets of an emerging learning theory, connectivism (Downes, 2012; Siemens, 2005). The course was later dubbed a Massive Open Online Course. Since then,

MOOC's have been developed by many different institutions, although most do not follow the connectivist model.

Without getting into the sticky details of a connectivist pedagogy and concerns whether connectivism warrants being considered a "learning theory," I suggest that the MOOC as a general phenomenon offers a possible solution to the logistical and technical issues. In particular, MOOCs are open; they are free to enroll and open to anyone who chooses to enroll. The open aspect is what tends to generate large enrollment numbers. In addition, most MOOCs have some central Learning Management System (LMS), which is used primarily to distribute course content and facilitate communication with and between students. A MOOC provides a means with which to gather a large number of individuals into one common context and connect/communicate with them via the LMS. Given the logical potential of this measurement technique, I created a MOOC offered through the Canvas Network, and sanctioned by the University of South Florida, which sought to engage students in creative activities for the purposes of learning about "what is music?" Embedded within the course were various means to measure the constructs of interest in Webster's model, or more specifically, the constructs in the General Specified Model of Creativity-Based Learning, the model I devised based on both Webster's model and the literature. The course itself was actually a modified version of an undergraduate online course that I had previously developed and taught, entitled "What is Music?" (explained in more detail in Chapter Three).

Initial Purpose and Rationale for the Study

As previously mentioned, this research study had to be revised and re-focused midway through completion. This section describes the initial purposes and rationales for the original study design. The revised purposes and rationales are described in a later section.

This study had four initial primary purposes and four respective rationales. The first purpose of this study was to develop and identify a statistical model that best represents the nature of the various interrelationships of components of the musical creative process, as identified in Webster's (2002) model. This is important from the perspective of basic research in that it could help us to better understand the musical creative process when the component parts are integrated into a coherent whole. This would help us to be more precise about not just *what* is involved in the process, but *how* these components interrelate.

Beyond identifying an appropriate model that sufficiently represents the observed data (and is explainable via relevant theory), a second objective was to determine the relative importance of each component as they each relate to various learning outcomes. This objective has important implications for connecting theory and research to practice. If, for example, it was found that the extent to which students perceive autonomy in the learning context is the single greatest predictor of all learning outcomes (such a clean result is always unlikely), then teachers must develop strategies for enhancing learner autonomy, as conceptualized within this particular theoretical framework. Perhaps engagement would be strongly predicted by one or two personality traits. Again, such a result would require teachers to consider ways in which they can better engage students who do not exhibit the requisite personality traits.

Third, while an a priori model was developed, there is always the potential of finding something new and interesting. So another purpose was to generate new ideas (something particularly appropriate given the topic of creativity for the study). Because such a quantitative systems-level approach has not been reported in the literature, to my knowledge, it is possible that certain peculiarities or previously unnoticed relationships might become apparent. Remaining open to new possibilities and spawning new ideas is essential to advancing any model

of human activity, but particularly when examining activities like creative music-making in new contexts (e.g., web-based environments).

Finally, my proposed method of data collection (i.e., many measures embedded within a MOOC) is somewhat new, especially for music education. There are many questions regarding the practicalities of collecting the data and the validity of the data once collected. The method has yet to be proven feasible, let alone useful. Therefore, the final purpose of this research was a sort of proof-of-concept.

Besides what I have proposed here, there are many other interesting possibilities for data collection when the environment is web-based. For example, although it was not possible for this study, it is conceivable to record (anonymously, but not surreptitiously) a wide range of individuals' actions in a web-based music composing application, which would allow for an extremely data-rich examination of the musical creative process (as it happens in such an environment). The fields of Learning Analytics, Academic Analytics, and Educational Data Mining have seen sweeping developments and improvements over the last few decades in order to account for, manage, and attempt to understand the massive amounts of data that have exploded with the advent of the internet. Incorporating these modern methods and techniques into research would offer significant potential for important research on the musical creative process, particularly as it has evolved (and continues to evolve) in the 21st century. As such, one final rationale for this study was that it represents a first step into this ocean of new possibilities.

Initial Research Questions

The following research questions were considered in this study:

1. Which statistical model best fits the observed data?
2. How much variance is explained by each component of the model?

3. What are the most important predictors of learning outcomes, and are they the same or different for each outcome?
4. Given suggested modifications of the model from the SEM analyses, are these modifications plausible, rational, and meaningful?
5. What are the challenges and unique aspects of conducting research via a MOOC?

Revised Research Questions

After it became clear that students were, on the whole, simply not participating in the creative music-making projects, the focus of the research shifted for two reasons. First, there simply was not going to be enough data to answer the first four research questions. While there was not enough data to answer the initial questions, there still was large amounts of data available for answering some new questions. Second, the premise of the research was that students would learn by way of their participation in the creative projects, but it was clear that students were choosing not to actively participate in those projects. Somewhat necessarily, I shifted to a much more exploratory approach, seeking to better understand *why* students chose not to participate. Below I present several research questions that I considered, although because of the exploratory nature of the analysis, the questions expanded and arose throughout and in response to the results of the ongoing analysis. For this reason, they are presented more in narrative form, as opposed to being presented all at once, up front.

Following an exploratory approach, I focused on uncovering relationships in the data. Strictly speaking, the new dependent variable was Project Participation (the extent to which any given individual participated in the projects). I define several versions of this variable in more detail in Chapter Three: Method. With Project Participation as the new dependent variable of interest, the first two new research question were the following:

1. Which variables (of those available in the data) are important in predicting Project Participation?
2. How well do those variables predict Project Participation?

The results of the analysis for the first research question uncovered only one statistically significant predictor: Type of Learner. More specifically, individuals who identified themselves as being an “active participant” at the beginning of the course were more likely to participate in the projects. Upon discovering this, a series of analyses were conducted to answer the following general research question:

3. Are there group differences in any of the other characteristics for the different Type of Learners (“active participants” vs. non-“active participants”)?

Finally, although the sample size did not warrant any examination of the full SEM model, I decided it was worth using the data I had to at least describe what picture appeared to be emerging when considering all of these variables, even if this picture is potentially highly influenced by idiosyncratic characteristics of the small sample. The final research question was the following:

4. What are the noteworthy relationships in the data that warrant further investigation?

While the first three research questions focused on a single dependent variable (project participation) using the lens of prediction, the fourth research question widened the focus to all variables (especially those for which there were a small number of responses) and the General Specified Model, using the lens of simple correlation. Planning, designing, developing, and teaching a MOOC requires a huge amount of time and resources, and completing the questionnaires required a sizeable investment of time by the participants. It would be irresponsible of me to not utilize the data from this experience to the fullest extent possible, that

is, within the limits of the specific characteristics of the data (e.g., considering sample size and measurement quality). Therefore, this fourth question was about looking at the data from more of a large-scale pilot study perspective. For the purpose of designing the next study (not for making any substantive or generalizable conclusions), what relationships appear to be strong enough to warrant more focused attention?

Delimitations

This research was not designed to be able to make claims about general creativity or creativity in other domains, but rather limited claims to the domain of music. The age of participants in this study was not be experimentally controlled, although age was included as a covariate for statistical control. Nevertheless, it would not be possible to generalize results to specific age groups, unless sufficient numbers of students for each age group were enrolled such that sub-samples of data could be analyzed for age group differences. This was certainly not the case. The research context was a massive open online course, which clearly has different characteristics than a face-to-face course (at any level), so generalization beyond the context of an online environment was limited. I say limited (as opposed to not possible) because the activities (including the technologies and software utilized in the course) could also be conducted via a face-to-face format. However, it is important to consider that differences in social context, physical resources, time constraints, etc. would almost certainly have an impact on results in a different educational context. Finally, the research was correlational in nature. While structural equation modeling allows for a greater comfort of making causal-type statements regarding results, because the design was not a true experiment, some may wish to limit the scope of conclusions that are drawn to be more relational than causal in nature. Furthermore, because the design is not experimental, it was not possible to make comparisons between creativity-based

music learning and other teaching models such as performance-focused and listening-focused pedagogies.

CHAPTER TWO: REVIEW OF LITERATURE

This dissertation approaches creativity from a systems perspective (explained in more detail below), which means it involves research literature in a wide variety of areas including general creativity, music education, motivation psychology, personality psychology, organizational psychology, cognitive psychology, music aptitude, and online learning, among others. The purpose of this chapter is the following: (a) review the literature in each of these areas that is most relevant to this study; (b) synthesize the literature as it relates more specifically to Webster's *Model of Creative Thinking in Music*; (c) integrate the reviewed research into at least one general path model depicting the hypothesized relationships between the variables of interest for this study. Other rival models could then be constructed by making modifications to the general model.

Perspectives on Creativity

The modern-day study of creativity is often said to have begun with J. P. Guilford's (1950) inaugural address as president of the American Psychological Association in which he rebuked the profession for its apparent neglect of creativity as a research topic, a state of affairs he considered to be "appalling" (p. 445). Since then, the topic of creativity has expanded into a field in its own right, and the topic has been addressed from a wide variety of perspectives in an ever-expanding number of fields including music education, marketing and design, engineering, leadership, business, and virtually all sub-fields of psychology (e.g., organizational psychology). Rhodes (1961) was perhaps the first to concisely (and cleverly) classify the various perspectives

on the study of creativity into the four P's: Person, Process, Product, and Press/Place. Other authors (e.g., Runco, 2007b) have since expanded the four P's to include two more P's: Persuasion (Simonton, 1990) and Potential (Runco, 2003).

The first P, Person, refers to research on the various internal characteristics or traits of an individual (e.g., personality, motivation). Research that has investigated the Process of creativity attempted to explain what is happening during the act of creativity. Product creativity research has sought to describe, analyze, and compare the creativity of objects that are manifest in the real-world as a result of some creative process. Press/Place refers to research on the context in which creativity occurs, which seeks to identify and explain the many pressures that impact a creative process (e.g., environmental, historical, political, socio-cultural, economic). The Persuasion perspective has used primarily historiometric research methods to examine the manner and extent to which creative individuals are persuasive in convincing others of their ideas and inducing a change in others' thinking. Finally, researchers studying Potential have sought to disentangle creative potential from creative performance (Runco, 2008), particularly as it pertains to identifying and nurturing potential within institutions of education.

Creativity Theories

With the many different ways from which to approach the study of creativity, many researchers have attempted to synthesize research findings into broader theories of creativity. Kozbelt, Beghetto, and Runco (2010) identified ten types of creativity theories including Developmental, Psychometric, Economic, Stage & Componential Process, Cognitive, Problem Solving & Expertise-based, Problem Finding, Evolutionary (Darwinian), Typological, and Systems. Kozbelt et al. (2010) explained that the differences in these many different types of

theories lie in which aspect of creativity they attempt to explain (i.e., which of the Six P's) and at what "level of magnitude" creativity is observed (Mini-c, Pro-c, Little-c, Big-c).

Traditionally, researchers have distinguished between Little-c creativity (subjective, relative to small, local groups) and Big-c creativity (objective, genius-level, eminent, relative to an entire domain, field, or humanity in general). Kaufman and Beghetto (2009) drew a further distinction between these two types of creativity. Mini-c represents creativity characterized by "novel and personally meaningful interpretation of experiences, actions, and events" (p. 3), also known as personal creativity. Pro-c creativity is creativity that rises above Little-c creativity, but not to the same heights as the revolutionary sort of creativity represented by Big-c creativity. It is roughly analogous to the difference between amateur and professional (but not world-renowned) performance.

Earlier I mentioned that I take a systems perspective, an idea that should be clarified at this point. The most well-known systems theory in the general creativity literature is Csikszentmihalyi's (1988) Systems View of Creativity, which tends to focus on Big-c creativity, but from a very broad perspective. According to Csikszentmihalyi, creativity involves a complex interaction between the field (experts in a discipline, the gatekeepers), the domain (the entirety of what is known in a discipline), and the individual (one who interacts with the field and acquires knowledge in the domain). Less emphasis is placed on the psychological, cognitive, or neurological characteristics and processes of the individual, but rather emphasizes the role of environment (context, culture, history) and others (e.g., experts, consumers, etc.). This is but one of several systems approaches to creativity in the general creativity literature (c.f., Gruber & Wallace, 1999). More generally speaking, Kozbelt et al. (2010) describe systems theories as those that "take the view that creativity is best conceptualized not as a single entity,

but as emerging from a complex system with interacting subcomponents – all of which must be taken into account for a rich, meaningful, and valid understanding of creativity” (p. 38). It is this notion of a complex, interactive, multi-faceted system that I ascribe when I say this study of creativity follows a systems perspective.

Theories of Musical Creativity

Theories in the general creativity literature tend to approach creativity from a domain-general standpoint, although most recognize the importance of domain-specific factors. That is, they attempt to explain the general notion of creativity, whether it be Big-c/Little-c, or whether it be explained in terms of development, economics, cognition, sociology, strategies, typologies, etc. Researchers in music and music education have taken two general approaches as they relate to theories in the general creativity literature: (a) explain musical creativity in relation to an existing theory in the general creativity literature; or (b) develop a theory specific to musical creativity. I focus on the latter approach for the remainder of this section.

Cognitive models. Swanwick and Tillman (1986) developed a spiral model of musical development based on analysis of children’s compositions. While not a theory of musical creativity, per se, it does offer a way in which to conceptualize children’s development of musical understanding when viewed through the lens of creative activity.

By analyzing data from a think-aloud protocol with four composers, Younker and Smith (1996) developed a model of the thought processes utilized by the four composers in composing a 16-bar piece of music. The model identifies several types of input (tactile, visual, aural), a somewhat nebulous depiction of the processing of that input, and then several types of output, including verbal descriptions, aural output (playing, humming), and visual output (notations). As a cognitive model, the model is lacking in its ability to explain the actual cognitive processing

that occurs during the compositional process. However, an important aspect of the model is the notion that output can become input, thus representing a recursive process.

Through their comparative analysis of data from four different data banks of students of different ages who completed various different composition tasks in three different countries (Canada, Australia, and the UK), Burnard and Younker (2002) identified three general pathways by which students approached the compositional process. Depicted as models, these pathways illustrate the paths that students took through Wallas' (1926) four stages of the creative process (preparation, incubation, illumination, and verification) and are conceptualized as existing along a continuum from "linear" to "recursive" to "regulated." Each model portrays different ways in which students established and worked within different degrees of freedom and constraints.

Working from a computer science and artificial intelligence perspective, Wiggins, Pearce, and Müllensiefen (2009) have developed a computational model of musical creative behavior, which improved upon previous work in computational modeling of musical creativity. What is most interesting about this approach to studying creativity is that a model of music perception must be explicitly embedded into the process. In other words, for a computer to model musical creativity, it must also have a model defining how the process receives input (i.e., perception).

Socio-cultural models. While much of the research on creative thinking in music has focused on the individual, from a systems perspective a cognitive model fails to account for context-specific aspects that come to bear on the process, including social, cultural, and environmental components. Research on the nature of group creativity has been an important addition to the literature because while examining the mental processes of the individual, group creativity research also inherently examines the social aspects of creativity. Wiggins (1994)

found evidence of a creative thinking sequence that moved from a focus on whole to part to whole again. More recently, several researchers have also begun to explore more informal online learning environments and their role in musical learning (Waldron, 2009, 2011a, 2011b), an area that has received limited research as it relates to musical creativity.

Systems models. Individual researchers have tended to focus on specific aspects of musical creativity when developing models of musical creativity (e.g., cognitive processes, social processes, technology-mediated processes). However, only one researcher, to my knowledge, has attempted to articulate how all of these various aspects relate to each other from a systems perspective. Webster's (2002) *Model of Creative Thinking in Music* identifies three main aspects of musical creative thinking: product intentions, thinking process, and creative products. In other words, every individual comes to every potential music situation with some sort of intention. According to Webster's model, they intend to compose, perform for others, listen (once or repeatedly), or improvise. This intention is then acted upon through some creative process (discussed in more detail below). Given some intent of the musical activity, the creative process results in some creative product, which can be in the form of a musical notation, audio recording, written analysis, or internal mental representation.

The core of the model is the depiction of the thinking process itself. At the center of this thinking process is a multi-directional path through four stages (preparation, time away, working through, and verification), which are re-named from Wallas' (1926) four stages (preparation, incubation, illumination, and verification). This process is characterized by alternations between convergent and divergent thinking. Divergent thinking is generally viewed as the process by which one generates possibilities, while convergent thinking is the process by which one works

toward a singularity, bringing together disparate pieces to reach a single answer, response, or decision.

Supporting (or at least influencing in some way) these processes are enabling skills and enabling conditions. Webster (2002) includes the following enabling skills in his model: aptitudes, conceptual understanding, craftsmanship, and aesthetic sensitivity. Each of these skills has an implied causal effect on convergent and divergent thinking. In addition to enabling skills, Webster recognized that any creative process resides within a specific context, which he accounted for with the inclusion of enabling conditions. These enabling conditions can be divided into two sub-categories, personal and social/cultural (roughly analogous to the two P's Person and Press). Enabling conditions specific to the individual person include constructs such as motivation, personality, and subconscious imagery. Enabling conditions that arise from social/cultural dimensions include context, task, peer influence, and past experience. As with enabling skills, each enabling condition is depicted as having some causal effect on convergent and divergent thinking.

What is apparent from the model is that the creative process is extremely complex. This complexity is precisely what has restricted the model from advancing from a conceptual model to an empirical model, a task that is central to this dissertation. Advancing the model also requires a more fine-grained level of detail. For example, in Webster's model all enabling skills and enabling conditions are implied to have a causal effect on convergent and divergent thinking, although details such as magnitude of the effect, direction of the effect (positive/negative) and possible interaction effects (e.g., higher values of one particular skill have a different magnitude of effect than lower values) are not specified. As will become apparent later, in order to empirically evaluate the model, these details needed to be specified both theoretically and

mathematically. Decisions regarding these more specific hypothesized relationships were drawn from the research literature, a task to which I now turn.

Empirical Model Evaluation

Before specifying the individual relationships in the model, it is necessary to consider how one actually quantitatively empirically evaluates a model as complex as this one. To begin, virtually any quantitative statistical analyses involves the evaluation of a model. To calculate a correlation coefficient between two variables is to develop a model in which two variables (among the infinite number of other possible variables that one could measure) are assumed to have some relationship to each other, that is, they co-vary in a linear fashion, but no assumptions are made about any causal relationship. The strength of the correlation coefficient is an indication of the extent to which this model (the specified relationship between the two variables) matches reality (the observed data). A low correlation coefficient suggests our model (that these two variables are related) does not reflect reality very well (i.e., these two variables are actually unrelated). The notion of evaluating how well our specified model fits reality is central to all quantitative research.

In addition to evaluating how well a model reflects reality, we also often wish to understand how much of reality our model explains. To return to the simple correlation model, squaring the Pearson correlation coefficient gives us the statistic known as R^2 , which we interpret as the proportion of variability in the data that the model explains. For example, if the correlation between the two variables is .4, then $R^2 = .16$, which means the symmetric relationship between these two variables explains 16% of the variability in the data. In order to explain a larger proportion of variability in observed data, we can specify additional relationships (with additional variables) or be more specific about those relationships (e.g., a unidirectional

causal relationship as in regression vs. a non-directional symmetric relationship as in a correlation).

If every variable we wished to measure and include in our model were directly observable, then the model can be evaluated with some form or some subset of multiple regression or path analysis. However, much of what we wish to measure in social sciences are constructs or traits that cannot be measured or observed directly. Rather, we take measurements we believe to be either manifestations or causal indicators of an underlying trait (e.g., motivation) and we specify how this latent variable relates to other variables (which may be latent or observed). When one wishes to specify the effects of one latent variable on another variable, traditional multiple regression and path analysis are no longer sufficient, and one must move into the world of Structural Equation Modeling (SEM). I have explained elsewhere (Stefanic, in press) the logic behind SEM and have provided examples of using an SEM technique in a music education context. Therefore, in what follows I will focus less on explaining the technique itself and more on the first step, model specification, and how the related literature informed this step for the purposes of this dissertation.

Model Specification

The first step in SEM is to specify a model, which involves identifying the relevant variables, operationalizing those variables, and indicating the specific relationships between those variables. In other words, a measurement model is specified to indicate how any latent constructs are measured, and a path model is specified to indicate the relationships between the constructs. Combining the measurement and path models results in a structural model, which implies the respective structural equations that form the mathematical model. In this section I will concentrate on identifying relevant variables and hypothesizing relationships between those

variables. The manner in which these variables were operationalized and measured is discussed in Chapter Three.

It is common in SEM studies to evaluate not one model, but several competing models to determine which model best fits the data. This was one initial goal of this dissertation. Several models could have been developed based on both the over-arching structure of Webster's model as well as information from the relevant literature. Even if it were as simple as turning Webster's model into a path model, there are several different ways to interpret the relationships depicted in the model, so several competing models could have been constructed based on different interpretations. For example, one major question regarding the variables listed under enabling skills and enabling conditions is whether they are modelled as co-varying with each other or independent of one another (or some combination of both), or whether a more complex mediation or moderation relationship exists.

In addition, by incorporating information from the relevant literature, individual paths were specified based on the relationship having been verified in the research literature. For example, there is substantial research on the relationships between personality traits (e.g., Openness) and other variables (e.g., creative self-efficacy). Decisions about how to specify these relationships were grounded in the literature. The next section will review the literature that informed the specification of the General Specified Model of Creativity-Based Learning.

The literature to be reviewed is organized according to categories of variables in the model, including enabling skills, enabling conditions, and creative process. For practical and logistical reasons, it was not possible to include every aspect of Webster's model (e.g., subconscious imagery) in this research. Components that I attempted to incorporate into the model included the following: aptitude, conceptual understanding, motivation, personality,

context, and past experience, in addition to several learning outcomes and a proxy for measuring the creative process (engagement). Each will be addressed in turn below.

Enabling Skills

Webster's model identifies four enabling skills: aptitudes, conceptual understanding, craftsmanship, and aesthetic sensitivity. Of these four areas, the vast majority of available research has investigated aptitude. After a discussion of the literature related to aptitude and creativity, I will briefly address the role of conceptual understanding.

Aptitude. At the top of the list of enabling skills is aptitude. Aptitude can be considered in at least three ways, each representing a higher level of specificity: general aptitude (e.g. intelligence or g-factor), musical aptitude, and musical creative aptitude.

The relationship between intelligence and creativity has been the subject of a long and on-going debate in psychology and creativity research. There is much research that suggests the two constructs have a modest relation (Batey & Furnham, 2006; Getzels & Jackson, 1958), which suggests they are distinct from each other, and yet still related. It is also possible that creativity is but one component of intelligence. The Cattell-Horn-Carroll (CHC) theory identifies 16 different intelligence factors (McGrew, 2009). Many have argued that creativity is largely a function of the Gf factor, also known as fluid intelligence (A. S. Kaufman, 2009; Kim, Cramond, & VanTassel-Baska, 2010). Sawyer (2012) has claimed that "most factor analytic studies suggest that creativity loads on the factor Glr, long-term storage and retrieval" (p. 55). Glr (long-term storage and retrieval) is the factor that incorporates many of the more narrow abilities traditionally associated with creativity (McGrew, 2009; Sawyer, 2012).

The traditionally low correlations between intelligence and creativity can be explained in another way. Silvia (2008) suggested that previous research has potentially underestimated the

relationship between intelligence and creativity because of the manner in which the factors have been modelled statistically. Silvia (2008) found low correlations with three individual factors of intelligence (fluid reasoning, verbal fluency, and strategy generation), but found a much larger effect ($\beta = .42$) when the three factors were modelled as effect indicators of a higher-order latent variable (g-factor). However, this effect was reduced to a more moderate effect ($\beta = .26$) when the confounding variable Openness was included in the model. He also cautioned that other personality factors (e.g., Openness) may account for at least a portion of this relationship (an issue addressed in the Personality section below).

Several researchers have developed measures of musical creative aptitude (also known as creative thinking in music or musical creative potential), including the Measure of Creative Thinking in Music – Version II (MCTM-II) (Webster, 1987b), the Measures of Creativity and Sound in Music (MCSM) (C. Wang, 1985), and the Measure of Musical Problem Solving (MMPS) (Vold, 1986). Generally speaking, measures of musical creative aptitude have been found to have a fairly low correlation with measures of music aptitude, such as the PMMA (Schmidt & Sinor, 1986; Webster, 1987b; Willing, 2009) or IMMA (Baltzer, 1990).

Since it was not be feasible to include a measure of musical creative aptitude (e.g., MCTM-II) in this research due to the logistical issues of administering this measure via the internet, the aptitude component of Webster's model was measured by general musical aptitude. There are conflicting results from the literature regarding the relationship between music aptitude and musical creativity.

Webster (1979) identified a consistent relationship between music achievement (measured using Colwell's Music Achievement Tests) and three different criteria measures of musical creativity (composition, improvisation, and analysis) for high school students. However,

a significant correlation was not found between analysis nor the compositional creativity measures and music aptitude (as measured by the Tonal Imagery subtests of Gordon's Music Aptitude Profile [MAP]). One potential explanation for this result is that Colwell's Music Achievement Tests (especially the Auditory-Visual Discrimination test) and several of the composition and analysis activities that formed the basis for the creativity measurement were heavily dependent on music notation. It is possible that these observed relationships are more of a function of students' notational understanding because students with a better grasp of notation would score higher on measures dependent upon notation. Since total scores on each of these measures (as opposed to sub-test scores) were the variables included in analyses, it is not possible to tease out possible different relationships when looking at notation-based vs. non-notation-based tests. That being said, the correlation between improvisation creativity and music achievement cannot be explained by this logic.

On the other hand, the MAP is an entirely aural-based test, so one's ability to interpret notation does not come into play. Interestingly, a significant correlation was found between improvisation and music aptitude. A similar argument can be made that any dependency on notation is removed in both the music aptitude and improvisation measure, and thus the correlation is more prominent.

Also using the Musical Aptitude Profile, Laycock (1992) found statistically significant correlations between the various subtests of the MAP and 10 out of 15 different characteristics (e.g., originality, cohesiveness) of the original compositions of 56 high school students. Unfortunately, the analysis did not move beyond examining the zero-order correlations, so an understanding of the structure of these relationships is not possible. Furthermore, given the number of significance tests calculated on the correlations, it is highly likely that several

correlations deemed statistically significant were spurious due to the failure to control for type 1 error inflation.

Willing (2009) found moderate correlations (.24 to .32) between music aptitude (measured by the PMMA) and a three judges' ratings of the creativity of kindergartners' tonal patterns using a researcher-created measure. The correlations were not statistically significant, but this could be due to a small sample size of 21 students.

The research relating music aptitude to characteristics of the created products (not to measures of creative thinking) are somewhat inconclusive. The differences in results are very possibly a consequence of the use of different instruments, different creative tasks, and different operationalizations of creativity. Because no clear causal relationship between music aptitude and creativity exists in the literature, music aptitude was treated as a control variable, which means it was modelled to have an effect on all contextual and situational variables. This would have allowed for an examination of its relative contributions to other components of the model, and also allow for those components effects to be interpreted as the effect of X on Y "after controlling for music aptitude."

Conceptual understanding. Virtually every model of creativity incorporates conceptual understanding into the model, although it may it may have different names, such as domain expertise, crystallized intelligence, or knowledge. For example, Mednick (1962) contended that creative thinking was an associative process between remote elements (ideas, concepts, words, etc.) in the mind, and more remote associations generated more creative results. As such, creativity is an extension of intelligence, or more specifically, crystallized intelligence (Batey & Furnham, 2006). In creative thinking, the efficiency with which someone makes associations (i.e., fluid intelligence) is somewhat dependent on the amount of knowledge available with

which to associate (i.e., crystallized intelligence). In the context of this study, conceptual understanding is conceived of as both an antecedent to and a consequent of the creative process. As an antecedent, it was treated as a control variable similar to musical aptitude discussed above. As a consequent, change in conceptual understanding was modeled as a learning outcome of the process.

Enabling Conditions

Webster (2002) identifies several enabling conditions, including elements of the person (subconscious imagery, motivation and personality) and socio-cultural elements (context, task, peer influence, past experience). This study attempted to incorporate motivation, personality, context, and past experience into the models to be evaluated.

Motivation. I believe motivation is one of the most central and important aspects of any model of creativity. I view motivation as the gateway of human behavior. Thoughts and actions require effort, and the extent to which effort is exerted on some task or activity is a direct result of one's motivation. The type and amount of motivation one has toward a task or activity controls the floodgates of effort, and as such, stands squarely in the center of any investigation of human behavior. For this reason, I dedicate a significant portion of this literature review to this construct.

Self-determination theory. Motivation represents an entire field of psychology in and of itself, which means there have been and continue to be a wide variety of theories attempting to explain the construct. Researchers attempting to investigate the relationship between motivation and creativity have tended to co-opt one of the existing theories to provide a framework in which to conduct their investigations. Because of the variety in theoretical frameworks of motivation, it can be difficult to synthesize research that utilizes different, sometimes incompatible, theoretical

frameworks because each maintains different assumptions, definitions, and operationalizations of constructs.

Self-Determination Theory (SDT) is an extensively researched macro theory of human behavior that focuses on the role of motivation in human activity and life course. In short, SDT posits that our behavior can be motivated in three qualitatively different ways, intrinsically motivated, extrinsically motivated, or amotivated, with the former tending to be associated with more positive outcomes (Deci & Ryan, 2012a). One becomes more intrinsically motivated toward an activity to the degree that one's basic psychological needs of autonomy, competence, and relatedness are met.

SDT researchers have treated motivation as both a trait-like variable in which between-individual differences are of interest, and as a state-like variable in which within-individual differences are of interest. Vallerand (1997) proposed a Hierarchical Model of Intrinsic and Extrinsic Motivation. In his model, intrinsic, extrinsic, and amotivation exist at three levels of generality: the global level, contextual level, and situational level. The global level is analogous to a trait-like treatment of motivation, similar to aspects of personality, and is believed to be relatively stable. Deci and Ryan (2012b) refer to this global level of motivation as a causality orientation. The contextual level relates to the motivation that is more consistent for specific contexts or domains. For example, one's motivation in a school context might be different than one's motivation in a sports context. Within every context, specific aspects of the situation at hand can also have an effect on motivation, as represented by the situational level. For example, the way instructions are presented by a teacher can impact one's motivation for a particular task.

Importantly, within any given level, one's perceptions of the three basic psychological needs mediate the effects of social factors on motivation. As such, Vallerand (1997) maintains,

“Social factors that do not impact these three types of perceptions should have minimal effects on motivation” (p. 274). This is particularly important in terms of specifying a model because it suggests that any effects from social factors (those not internal to the individual) on motivation should be fully mediated by individuals’ perceptions of autonomy, competence, and relatedness.

Finally, the three levels are also posited to have a top-down effect, such that one’s global motivation may impact one’s contextual motivation, which in turn impacts one’s situational motivation. However, the social factors within any particular level are expected to have no direct effect on motivation at another level. The impact of one level on another is by way of the motivation at one level (e.g., contextual level) impacting motivation at another level (e.g., situational level). Because of the top-down effect though, social factors at the contextual level can have an indirect effect on motivation at the situational level by way of their effect on contextual motivation. Therefore, one’s situational motivation for a particular activity at a given point in time should be primarily a result of their contextual motivation for those sorts of activities in general and the situational factors present during the time in which the activity is occurring.

Some SDT researchers have advanced the concept of intrinsic motivation beyond an undifferentiated singular type of motivation (in contrast to the various forms of external motivation). Vallerand and his colleagues (Corbonneau, J., & Lafreniere, 2012; Vallerand, 1997; Vallerand & Bissonnette, 1992; Vallerand et al., 1992) have proposed that intrinsic motivation can be further distinguished between three major types (the Tripartite Model of Intrinsic Motivation), IM to know, IM toward accomplishment, and IM to experience simulation. IM to know involves engagement in some activity “because of the pleasure and satisfaction of learning, exploring, and trying to understand something new” (Corbonneau et al., 2012, p. 1149). That is

distinct from IM toward accomplishment in which an activity is pursued for the pleasure and satisfaction of improving one's self, accomplishing something, or creating something.

Importantly, with this type of IM, the satisfaction and pleasure is derived from the process itself, and not necessarily the end result. Contrary to the previous two types of IM, when one “engages in an activity for the stimulating sensations, excitement, or aesthetic enjoyment associated with it” (p. 1150), one is said to have IM to experience stimulation.

The aforementioned distinctions between levels of generality (global, contextual, and situational) and between types of intrinsic motivation imply a need to both model and measure motivation at a level of complexity beyond a simple unidimensional, single-level construct. Unfortunately, there is an inherent measurement problem when treating motivation from this perspective. Consider the notion of situational motivation. By definition, it relates to the specific context within which a participant is involved, so individual items in a measurement instrument must be specific to that situation. Guay, Vallerand, and Blanchard (2000) developed and validated the Situational Motivation Scale (SIMS), which has been used in hundreds of studies after being modified for the specific situation of the individual researcher. The measure is constructed in such a way that all items have the generic stem “Why are you currently engaged in this activity?” followed by 16 items that differentiates between situational intrinsic motivation, identified regulation, external regulation, and amotivation. The measure can also be modified slightly to measure situational motivation *after* the task has been completed.

While modification of a situational motivation measure seems somewhat straightforward, contextual motivation measurement is a bit more complex for a few reasons. Whereas a situational motivation measure can be, ironically, somewhat generic, a contextual measure requires items to be much more specific to the domain. Also, it is at the contextual level that

intrinsic motivation is believed to be more differentiated (or at least this is the level at which this differentiation has been researched). The Academic Motivation Scale (AMS) (Vallerand, Blais, Briere, & Pelletier, 1989 in French; Vallerand et al., 1992 in English) was developed to differentiate between seven factors, the four types of extrinsic motivation and the three types of intrinsic motivation, with four items per factor. The scale is designed for measuring individual's motivations to pursue college, and no similar scale has been developed, to my knowledge, within the domain of music.

At this point it is important to consider another important distinction between motivation as construed as a goal and motivation as construed as the energizing force during an activity or over the course of some longer period of time. Vansteenkiste and his colleagues (Vansteenkiste, Lens, & Deci, 2006; Vansteenkiste, Simons, Lens, & Sheldon, 2004; Vansteenkiste, Timmermans, Lens, Soenens, & Van den Broeck, 2008) refer to a difference between intrinsic/extrinsic goals (the goal contents) and autonomous/controlled motivation (goal motive). The goal contents are the intended ends of an activity. An extrinsic goal is one focused on obtaining something outside of one's self, such as "fame, financial success, and physical appearance," but intrinsic goals involve ends that are "satisfying in their own right," such as "community contribution, health, personal growth, and affiliation" (Vansteenkiste et al., 2006, p. 22). Intrinsic goals, when pursued, provide direct satisfaction of the three basic psychological needs in SDT (autonomy, competence, and relatedness).

On the other hand, goal motives are the reasons why one pursues a particular goal content. For example, an individual may choose to learn a particular instrument in order to gain popularity (an external goal content) or because he/she wishes to learn something new (an internal goal content). However, one may choose to learn an instrument in order to gain

popularity because one is feeling pressured by one's peers (controlling motivation) or because one values the social connectedness that may result from popularity (autonomous motivation). Similarly, an individual may choose to learn an instrument solely because he wishes to learn something new, but the value of learning something new is something he knows his parents think is important. In this situation, the value of learning something new has not been fully internalized, so it represents a less-autonomous form of motivation. Conversely, if he chooses to learn the instrument in order learn something new because he personally values learning new things, then this represents a more autonomous form of motivation. Goal contents are the "what" of behavior and goal motives are the "why" (Vansteenkiste et al., 2008).

This distinction between goal contents and goal motives can be somewhat subtle, but nonetheless important. Although most of the research on goal contents has focused on psychological well-being, some more recent research has indicated more positive educational outcomes associated with intrinsic goal contents as opposed to extrinsic goal contents, including greater and deeper conceptual learning and greater persistence (Vansteenkiste et al., 2004; Vansteenkiste et al., 2008). SDT posits that intrinsic learning goals are, by their very nature, more likely to lead to behavior that satisfies the basic psychological needs, which in turn should support learning by way of improving optimal functioning (Vansteenkiste et al., 2008). In contrast, extrinsic learning goal are less likely to promote deep engagement with a learning activity because engagement is contingent on the extent to which the effort results in achieving the extrinsic goal. Furthermore, there is evidence of an interaction between goals and autonomy-support such that deepest involvement in a learning activity occurs when the learner aspires an intrinsic goal (as opposed to an extrinsic goal) in an autonomy-supportive (as opposed to controlling) environment, and autonomous motivation functions as a mediator on the relationship

between interpersonal context (autonomy-supportive vs. controlling), goal content, and their interaction on learning outcomes (Vansteenkiste et al., 2004).

It should also be noted that, similar to autonomous-controlled motivation, goal contents have also been conceptualized as a trait-like, global variable and as a more situation-specific variable (as in the research just discussed). From either perspective, the basic tenets of SDT are still upheld, intrinsic goals promote positive learning outcomes either along with or through an impact on autonomous motivation.

The research reviewed so far paints a very complex picture of motivational constructs, even within the fairly coherent and well-articulated SDT. Several decisions needed to be made as to how to approach motivational constructs in this study. In particular, this included decisions regarding the level of generality (global, contextual, situational), level of detail (distinct factors for each type of motivation vs. a unidimensional autonomous-controlling continuum), and motivational perspective (goal contents vs. goal motives). Before explaining these decisions and addressing model specification, I will first consider motivation research that is more specific to the contexts of creativity and music.

Motivation and creativity. Hennessey (2010) has noted that there are two bodies of research that have addressed the relationships between motivation and creativity, each focusing on different contexts (schools vs. organizations), yet each finding similar or analogous results. One body of research includes research with Expectancy Value Theory (Eccles, 1983), Self-Determination Theory (Deci & Ryan, 1985), and social psychology of creativity perspectives (Hennessey, 2004), while the other research tradition is that of Achievement Goal Theory (Ames, 1984; Deshon & Gillespie, 2005; Dweck & Leggett, 1988; Vandewalle, 1997). Both

research traditions have followed almost entirely separate, but still parallel paths, and have only somewhat recently begun to converge and cross-pollinate (Hennessey, 2010).

What initially seemed clear from these two parallel strands of research is that external rewards (along with other external constraints such as time limits) undermine intrinsic motivation (Deci, Koestner, & Ryan, 2001). Some research has found that certain types of rewards in certain situations may actually enhance intrinsic motivation (Deci & Ryan, 1985). Furthermore, intrinsic motivation can only be undermined if there is some initial level of motivation in the task to begin with. As Hennessey (2010) put it, “innate levels of interest in the target creativity task mark one crucial difference between empirical studies showing negative and positive effects of reward” (p.354).

Motivation and musical creativity. One important question regarding motivation in a music creativity context (particularly in the classroom context) is whether teachers should place limits or constraints on the compositional or improvisational tasks. In support of the hypotheses of SDT, Koestner, Ryan, Bernieri, and Holt (1984) found that certain types of constraints could be placed on a painting task with first- and second-graders without lowering intrinsic motivation. In particular, an informational style of communicating constraints (limits phrased as information about the task) did not undermine intrinsic motivation or decrease quality and creativity of the children’s paintings as did a controlling-style of communicating constraints (limits phrased as requirements, “things that you will have to do” [p. 239]). Wiggins and Medvinsky (2013) echo the need to avoid controlling constraints in their suggestion to frame compositional problems in relation to what J. Wiggins (2009) calls the metadimensions of music. Similar to a constraint framed as information, the metadimension provides structure within which a student can work, but without controlling the actual musical material within that structure. Although Wiggins and

Medvinsky do not refer to SDT for a theoretical rationale, which supports their contention, they base their recommendation on actual experience with students composing in the classroom, an equally relevant piece of evidence.

A related question to placing constraints on the creative task is how teachers should proceed with students during the actual process. Wiggins and Webster (in J. Wiggins, 2005) have both advocated for a collaborative approach to revision, which occurs during the process and not just after the student has decided the work is complete. Furthermore, Webster advanced this notion with the following:

...if the teacher does not actively teach children how to reconsider initial gestures, how to extend and rework their first ideas, and how to question how musical elements combine to create a whole artistically, I believe that real understanding about music does not often occur and the real power of doing composition as a teaching strategy is lost. (in J. Wiggins, 2005, p. 37)

The question that arises from this statement, in light of SDT, is how a teacher engages with the student without undermining intrinsic motivation. In SDT, internalization is the process by which external regulations become internal regulations. Webster's comments on the necessity to actively teach students to reconsider, rework, etc. suggests that there is something about this process of revision that is not intrinsically interesting. If it was, students would not require a teacher to question their initial musical decisions; they would do so because they are intrinsically motivated to do so. Therefore, the power of the compositional process as music pedagogy rests on the extent to which the revision process is motivationally internalized by the student. If it is internalized, then the student sees intrinsic value in the process, and as such should be motivated to continue engaging with his/her music, which in turn leads to additional

learning about and within music. This premise is supported by research in SDT that has shown the more autonomous forms of motivation (intrinsic, integrated) to be associated with persistence in coursework (Vallerand & Bissonnette, 1992), meaningful, deep, or intensive cognitive engagement (Assor, Kaplan, Kanat-Maymon, & Roth, 2005; Walker, Greene, & Mansell, 2006), and deeper conceptual learning (Grolnick & Ryan, 1987), among other beneficial outcomes.

The extent to which motivation is autonomous is impacted by the degree to which the three basic psychological needs (autonomy, competence, and relatedness) are met. Therefore, in order to support intrinsic motivation (and more highly internalized extrinsic motivations), the teacher should approach the composition and revision process in a way that maximizes the students' perception of these three needs being met. In other words, when a student feels a sense of control (autonomy), a sense of competence for the task, and sense of relatedness in the learning context then he/she is more likely to internalize his/her motivation. For example, Deci, Schwartz, Sheinman, and Ryan (1981) found that students who's teachers they perceived to be more autonomy supportive were also more intrinsically motivated and perceived themselves to have higher self-worth and cognitive competence.

Wiggins (2005; and in J. Wiggins & Medvinsky, 2013) has described a supportive, collaborative environment that reflects the principles of SDT quite nicely. In this environment described by Wiggins, students are treated as experts amongst each other (supporting perceptions of competence), encouraged to collaborate with each other (supporting relatedness), and given significant control over compositional decisions (supporting autonomy).

Measuring motivation. In regards to model specification, the next question is how these related constructs of needs fulfillment, different types of motivation, and various outcomes should be measured. Vallerand (1997) has urged that motivation must be measured

independently of both its determinants (needs fulfillment) and its consequences (outcomes), thus allowing for conceptual clarity and also a more fine-grained account of the differential effects of different types of motivation. For this reason, self-report questionnaires have been the measurement of choice because questions can be constructed to model the hypothesized types of motivation. The extent to which an individual endorses the items for each type of motivation provides an indication of the individual's motivation. Many different measures have been developed for various different contexts, including the Academic Motivation Scale (Vallerand et al., 1992) and the Academic Self-Regulation Scale (SRQ-A) (Ryan & Connell, 1989) for school contexts, the Treatment Self-Regulation Questionnaire (Levesque et al., 2007) for health treatment contexts, and the Intrinsic Motivation Inventory for experimental contexts (for validation evidence see McAuley, Duncan, & Tammen, 1989), to name just a few.

Modelling motivation. Taken together, the empirical research and theoretical hypotheses of SDT suggest a general path model as displayed in Figure 2.1. What is immediately apparent from this model is that it is extremely complex and therefore likely to suffer from SEM identification problems (discussed in Chapter 3). Vallerand (1997) has noted that many researchers have chosen to combine the various subscales (intrinsic, integrated, identified, etc.) of a motivation measure to form a unidimensional self-regulation index, a notion that makes sense conceptually because each type of motivation is believed to be along a continuum from autonomous to controlled motivation. A unidimensional self-regulation index derived from subscales would help solve identification issues related to SEM, particularly with complex models, by reducing the number of parameters to be estimated. Studies that have utilized this approach have found the unidimensional index to have sufficient reliability and validity (Ryan & Connell, 1989; Vallerand & Bissonnette, 1992; Vansteenkiste et al., 2008). The path model in Figure 2.1

could then be simplified by replacing the three types of motivation at each level with a unidimensional variable that represents the relative autonomy (with larger numbers indicating a more intrinsic, internalized motivation).

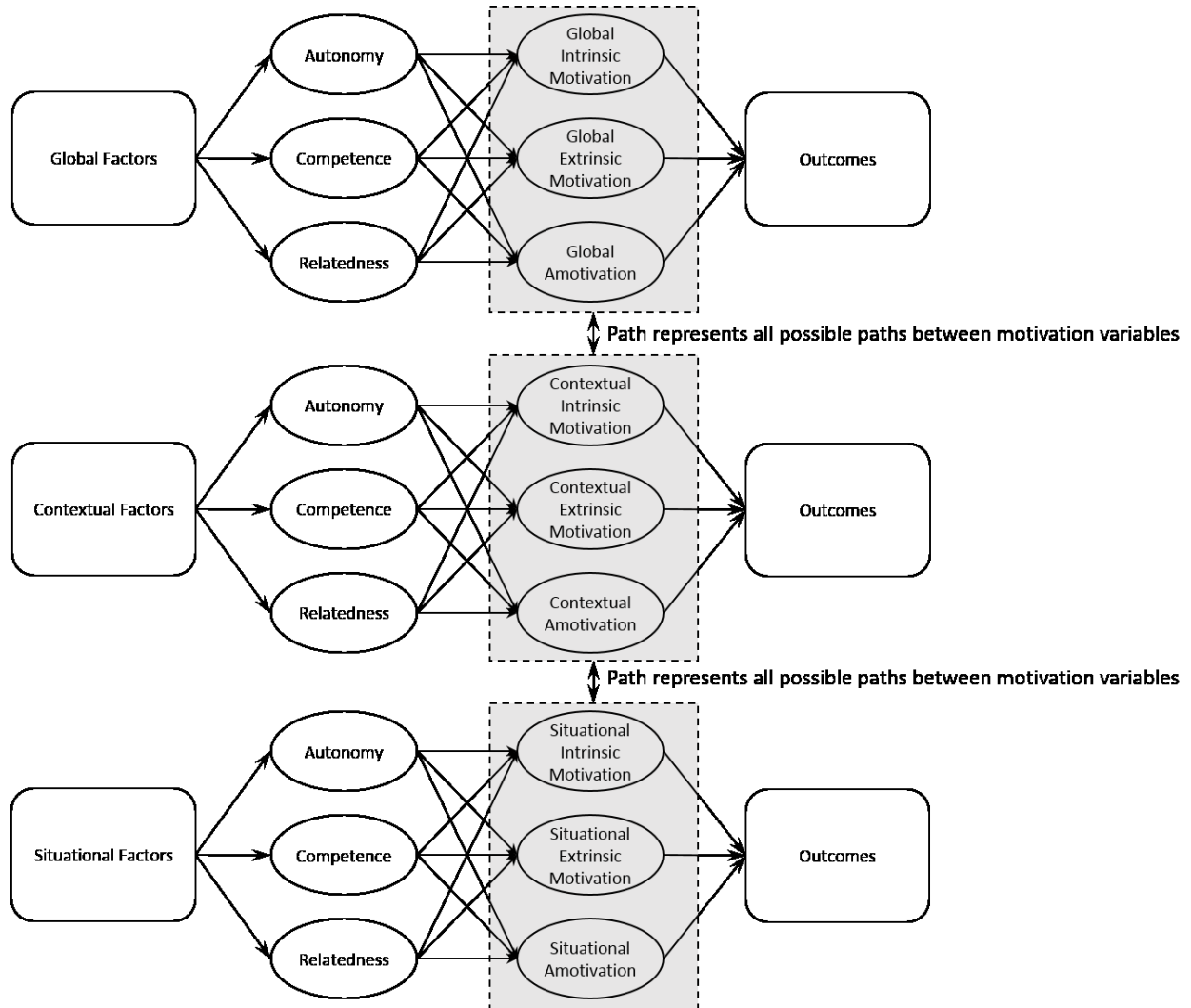


Figure 2.1. Path model for motivation implied by Vallerand's (1997) Hierarchical Model of Intrinsic and Extrinsic Motivation.

Personality. The field of personality research has approached creativity from several different perspectives. Recently, Feist (2010) characterized the varying effects on creative thought and behavior in his Functional Model of the Creative Personality. The model includes

six primary latent variables: genetic and epigenetic influences, brain qualities/characteristics, cognitive personality traits, social personality traits, motivational-affective personality traits, and clinical personality traits.

The singular exogenous variable in his model is genetic and epigenetic factors, indicating the fundamental role of both nature (genetics) and the environment's influence on how genes are expressed (epigenetics) in developing personality development. According to the model, genes make their impact on the latter four traits by way of brain qualities/characteristics, a component of the model that was not possible to include until the last decade or two given the vast improvements in brain imaging technology and research in neuroscience. In regards to brain characteristics, the picture that has emerged from research highlights, generally speaking, the importance of the frontal lobe in creative thought (although the entire brain is active in creative thought) and the importance of not just more activity, but greater interconnectivity between various associative areas of the brain (Dietrich, 2004; Heilman, Nadeau, & Beversdorf, 2003; Sawyer, 2011, 2012).

The remaining four latent variables in Feist's (2010) recently revised model are the types of traits that we tend to more typically think of in regards to personality, although Feist separates them into four categories of cognitive, social, motivational-affective, and clinical traits. For the purpose of this dissertation, I focus on the first two, although the third (motivational-affective) is discussed in the section on motivation.

Big Five factors. One of the most ubiquitous (and hotly debated) frameworks of modern personality research is the Big Five Factor model, which generally identifies Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism as the five broad personality factors. A large amount of research has continually indicated a significant relationship between

Openness and creativity. Feist's (1998) meta-analytic study provided perhaps the strongest argument for the importance of this relationship, particularly in regards to the difference between creative versus non-creative scientists and artists versus non-artists. More recent research continues to confirm this relationship (Griffin & McDermott, 1998; Prabhu, Sutton, & Sauser, 2008). Importantly, Prabhu and colleagues (2008) found evidence of a partial mediating effect of intrinsic motivation between Openness and self-reported creative behavior.

Batey and Furnham (2006) proposed that Openness could be the result of two underlying factors, attitudinal Openness and perceptual Openness. Perceptual Openness represents an inability to disregard or inhibit irrelevant stimuli, which makes individuals high on this trait more open to new ideas simply because they are limited in their ability to filter out such seemingly irrelevant information. Others (e.g., DeYoung, Quilty, & Peterson, 2007; Woo et al., 2013) have suggested Openness is comprised of two sub-factors, one related to aspects of intellect and one related to more experiential aspects of Openness. Woo et al. (2013) call these sub-factors Intellect and Culture.

Differences in personality traits have also been found to be associated with differences in the way people choose to use music. Chamorro-Premuzic and colleagues (2007; 2009) have twice found that individuals who score high on Openness tend to “use music to create cognitively enriching experiences” (Chamorro-Premuzic et al., 2009, p. 152) as opposed to using music for emotional regulation or for playing in the background. Researchers have also found a strong relationship between Openness and IQ, so much so that some have suggested considering Openness as an ability factor and not a personality trait (see Chamorro-Premuzic & Furnham, 2005; Connelly, Ones, & Chernyshenko, 2013; DeYoung, 2010 for reviews of this literature).

On the other hand, there is also an argument that can be made for considering intelligence as one of many facets of the Intellect sub-factor of Openness (DeYoung, 2010)

Taken together, an individual scoring high on Openness would be expected to be more intelligent and more creative, and also approach a musical creative task as an intellectual, cognitive task. This suggests the individual who is open to experience will potentially develop more conceptual understandings of music, particularly given an anticipated higher intelligence.

While Openness is the personality factor that has received the most attention in creativity research, other personality factors have been found to have a very important impact in other outcomes such as academic performance. Given this study took place in a quasi-academic setting, it was worth considering the literature on the relationship between academic performance and personality. Perhaps the most comprehensive review and statistical analysis of this relationship is Poropat's (2009) meta-analysis, which included aggregated sample sizes of over 70,000. Results from this meta-analysis indicated significant correlations between academic performance and personality factors, with Conscientiousness having an effect size only slightly less than intelligence. The next smaller effect sizes (Cohen's d) were for Openness ($d = 0.24$) followed by Agreeableness ($d = 0.14$). After correcting for intelligence, the sample-weighted correlation (corrected for scale reliability) with academic performance was reduced the most (.12 to .09) due to it having the highest correlation with intelligence, but Conscientiousness actually increased (.22 to .24). All correlations remained both statistically and practically significant for Agreeableness, Openness, and Conscientiousness, which Poropat contended is evidence that personality factors make their impact on academic performance not simply by way of their shared correlations with intelligence, but somewhat independently. Poropat suggested "the FFM [Five Factor Model] dimensions appear to be part of the set of factors that contribute to

performance by affecting students' willingness to perform" (p. 331). The notion of some effect by way of "willingness to perform" suggests some sort of motivation-related mechanism by which personality makes an impact on academic performance, an important implication for model specification. I shall return to this issue after first addressing the social personality traits.

One final note should be made regarding Openness and creativity. A recent meta-analysis has added yet another perspective to the already somewhat confusing, conflicting, and murky waters of the Openness factor. Connelly, Ones, Davies, and Birkland (2013) have uncovered four "true facets" (independent of other Big Five factors) of Openness, including aestheticism, openness to sensations, nontraditionalism, and introspection. These facets are unique to the Global Openness factor, as opposed to what Collins and colleagues call "compound facets," those that share a relationship with some other Big Five factor than Openness. Of particular interest to creativity research is the identification of the innovation (creativity) facet as being comprised of Global Openness and two more emotion-related (as opposed to intellect-related) factors, Global Extraversion and Global Emotional Stability (another conceptualization of Neuroticism). This relationship with Extraversion is somewhat to be expected, given previous research has found Extraversion (Dollinger, Urban, & James, 2004), particularly the sub-component reflecting independence and confidence (Feist, 1998), to be associated with highly creative people and scores on self-report creative personality scales.

Measuring personality. In choosing a personality measure, there are several issues to consider. The first is whether to utilize a five factor, six factor, or some other number of factor model. While the Big Five model is by far the most well-known, recent research has suggested the existence of a sixth factor, often referred to as Honesty-Humility (Ashton & Lee, 2007; K.

Lee & Ashton, 2008). Several six-factor measures have been developed, including the 100-item HEXACO-PI-R (K. Lee & Ashton, 2004), and a shorter 60-item version (Ashton & Lee, 2009).

A second issue to consider is length of the measure. Big Five scales range from full-length scales like the 240-item NEO Personality Inventory-Revised (Costa & McCrae, 1992) and the 100-item International Personality Item Pool (IPIP) (Goldberg, 1999), to medium length scales like Saucier's (2002) 40-item Mini-Markers and the 44-item Big-Five Inventory (John, Donahue, & Kentle, 1991), to even shorter scales like the 20-item Mini-IPIP (Donnellan, Oswald, Baird, & Lucas, 2006), to very brief scales like the 10-item BFI-10 (Rammstedt & John, 2007). Generally speaking, there are tradeoffs with reliability for shorter length scales, but these tradeoffs must be considered in relation to practical demands of the research context (Credé, Harms, Niehorster, & Gaye-Valentine, 2012).

Self-efficacy. To return to Feist's (2010) model, the social personality traits include terms such as "norm-doubting, nonconformity, independence, extraversion-introversion, aloofness, hostility, coldness, and dominance/self-confidence/arrogance" (p. 121). Creative people tend to be confident and independent (non-conforming) (Chavez-Eakle, Lara, & Cruz-Fuentes, 2006; Feist, 1998), a trait that has also been investigated in the form of self-efficacy (Bandura, 1977). Self-efficacy can be defined as the set of beliefs one has about one's abilities and is believed to "vary across activity domains and situational conditions rather than manifest uniformly across tasks and contexts in the likeness of a general trait" (Bandura, 2012, p. 13). It can be developed through mastery experiences, social modeling, social persuasion, and changes in physical and emotional states.

In general creativity literature, the more specific construct of creative self-efficacy has been proposed and subsequently researched for its antecedents and consequents in various

different contexts (Beghetto, 2006; Beghetto, Kaufman, & Baxter, 2011; Gong, Huang, & Farh, 2009; Karwowski, 2011, 2012; Lemons, 2010; P. Tierney & Farmer, 2002; Pamela Tierney & Farmer, 2011). Kaufman, Evans, and Baer (2010) found that students' predictions of their creativity in four domains (visual art, math, writing, science) were not consistent with expert's ratings of their created works (as measured using Consensual Assessment Technique [Amabile, 1996]). However, it is difficult to evaluate whether the researchers were measuring students' perceived abilities in a domain, students' perceived self-efficacy in a domain, or even something more akin to a personality trait. For example, one item asks students "How creative in XXX do you think you are?" (p. 6), which potentially blurs the line between self-perceptions of past achievement, domain competence, and current capability. In addition, it is possible for such a question to be interpreted in two ways, whether one *can* create in a given domain or whether one is *more* creative in relation to some standard or some group of people (e.g., fellow classmates). Self-efficacy measures are generally worded in terms of the degree of confidence one has in being able to do something (self-efficacy strength) and confidence is usually assessed for varying degrees of difficulty/challenge. While Kaufman and colleagues did not claim to be measuring self-efficacy, the point is that there is a need for conceptual clarity because their results could be construed as suggesting self-efficacy (or some similar but differently named construct) does not predict actual creative performance. In reality, I believe they were actually measuring differences in students' self-evaluations of their creative products (prior to creating them) and experts' ratings of those products. Thus, it is a study about evaluation and measurement, not motivational processes.

In music, Martin (2012) found middle school band students to identify four general influences on musical self-efficacy that were in line with those proposed by Bandura (2012) (see

above). Ritchie and Williamon (Ritchie & Williamon, 2011a) have developed two self-efficacy measures, one for self-efficacy for music performance and one for self-efficacy for music learning. Factor analysis indicated the two constructs to be distinct from each other. In further research they found musical performance self-efficacy to predict quality of music performance in university music students (Ritchie & Williamon, 2012), a notion that had been previously identified with grade-school-aged students (McPherson, 2006). In primary school children, experience with formal music instruction was found to predict children's self-efficacy for musical learning for both males and females (Ritchie & Williamon, 2011b). Zelenak (2011) found evidence for a potential mediating effect of musical performance self-efficacy on the relationship between music aptitude and music achievement. As for creative self-efficacy specific to music, I was unable to locate any research.

Relations between personality, self-efficacy, and motivation. The theoretical work and empirical research supports the general notion that self-efficacy is domain-specific and is influenced by prior experience as well as factors specific to the environment and task at hand. Self-efficacy also has an impact on various outcomes including musical performance, creativity, and motivation. Therefore, in terms of model specification, musical creative self-efficacy should function as a central mediator between all or most antecedent variables (e.g., past experience, aptitude) and outcome variables (e.g., product creativity, change in conceptual understanding, change in self-efficacy). In the context of social cognitive theory, Bandura (2012) contends that self-efficacy influences motivation and perseverance, a notion that matches with the placement of perceived competence between global/contextual/situational factors and motivation in Self-Determination Theory (Vallerand, 1997).

The role of personality in the context of self-efficacy and motivation remains to be specified. Research and theory consistently supports the notion that personality characteristics will influence the way one views the world, and as such, one's self-efficacy (c.f., Bandura, 2012; Jackson, Hill, & Roberts, 2012). For example, in organizational psychology, meta-analytic researchers have found Big Five traits to predict self-efficacy for task or job performance (Judge & Ilies, 2002). In later research, Judge and colleagues suggested the role of self-efficacy on actual job performance has been drastically over-stated because the effect of self-efficacy on job performance is relatively small after controlling for other individual-difference factors such as personality and general mental ability (Judge, Jackson, Shaw, Scott, & Rich, 2007). Together, this research suggests a partial mediation role of self-efficacy between personality factors and either motivation or specific outcomes, or both.

Context. Context is a difficult component to narrow down because it is difficult to delineate precisely what comprises or what defines a context, let alone how it is measured. There is an almost limitless number of ways in which it can be defined. For the purposes of this study, I chose to focus on one specific aspect of the context in which the study took place (an online environment) and which was also compatible with the motivation theoretical framework for this study. That aspect was contextual support as conceived through SDT.

According to SDT, the more autonomy-supportive an environment is, the more likely individuals are to perceive their basic psychological needs are being met, which will in turn increase their intrinsic motivation. Several studies have supported this basic hypothesis of SDT in various face-to-face environments (Standage, Duda, & Ntoumanis, 2006; Tsai, Kunter, Lüdtke, Trautwein, & Ryan, 2008; Van Nuland, Taris, Boekaerts, & Martens, 2012). Beyond the connection to intrinsic motivation, autonomy supportive environments have also been found to

foster engagement (Assor et al., 2005; Assor, Kaplan, & Roth, 2002; Reeve, Jang, Carrell, Jeon, & Barch, 2004), an important variable in the current study. A wealth of evidence also exists regarding various other positive outcomes of an autonomy-supportive environment related to engagement, learning, development, academic performance, and psychological well-being (Reeve, 2009). Additional research indicates that children, adolescents, and pre-service teachers in a face-to-face environment are able to perceive a distinction in autonomy support as opposed to controlling behavior from the teacher (Assor et al., 2002; Reeve & Jang, 2006), with the latter promoting anxiety and anger, and thus amotivation and the less autonomous extrinsic motivation (Assor et al., 2005).

Reeve (2009) explained that teachers (particularly in K-12 settings) tend to adopt a more controlling motivating style (as opposed to autonomy-supportive) because of a wide variety of pressures, including pressures from “outside agents, such as school policies, administrators, parents, societal expectations, or cultural norms” (p. 163); aspects of the moment-to-moment dynamics of a classroom environment; personality traits; and the beliefs teachers hold regarding the nature of motivation. Incorporating work from Pelletier, Seguin-Levesque, and Legault (2002), Reeve distinguished these pressures as being pressures from above (e.g., administrators, state requirements), pressures from below (e.g., students), and pressure from within (e.g., “beliefs, values, and personality dispositions” [p. 164]).

Given the benefits of an autonomy-supportive environment, Reeve (2009) identifies three enabling conditions (of the teacher) that are necessary to achieve an autonomy-supportive environment. These include the teacher’s willingness to “adopt the students’ perspective, welcome students’ thoughts, feelings, and actions, support students’ motivational development and capacity for autonomous self-regulation” (Table 1, p. 160). Given these requisite conditions,

Reeve also identifies five behaviors that promote an autonomy-supportive environment: (a) nurture inner motivational resources; (b) provide explanatory rationales; (c) rely on non-controlling and informational language; (d) display patience to allow time for self-paced learning; and (e) acknowledge and accept expressions of negative affect (Table 1, p 160).

It is important to point out that there is also a difference noted in the literature between a teacher providing autonomy support versus providing structure. Jang, Reeve, and Deci (2010) had five raters observe teachers' and students' behavior for 133 classroom visits. Using confirmatory factor analysis they found that a two-factor model (ratings from items for structure were distinct from ratings from items about autonomy-support significantly) fit the data quite well (SRMR = .04, CFI = .98, NFI = .97), and significantly better than a single-factor model. The researchers also compared the relationship between teachers supporting autonomy and providing structure. One might expect that providing autonomy might result in lack of structure, but Jang and colleagues found the two behaviors to be positively, linearly related, which means they are neither antagonistic to each other nor independent of each other. In addition, while autonomy-support predicted both measures of student engagement (observer-rated and student self-reported), structure only predicted the observer-rated measure of engagement. Taken together, what these researchers recommend is that while structure is important for promoting engagement, the structure must be provided in an autonomy-supportive manner in order for students to subjectively perceive a sense of engagement in the learning.

The notion of autonomy-supportive environments have also been investigated in music settings as well. In the applied music lessons context, Troum (2010) found further support for the SDT framework in that students' level of perceived autonomy support significantly predicted their task persistence (self-perceived), and this relationship was partially mediated by perceived

competence. In the high school band setting, Legutki (2010) found evidence that students' perceptions of teacher autonomy support are related to their perceptions of musical competence, interest and enjoyment in band, and intrinsic motivation.

It is possible that the ways in which individuals perceive autonomy support and satisfaction of basic psychological needs are different in online environments. Several previous studies have found the SDT framework to be useful for examining motivation in an online environment (Chen & Jang, 2010; Hartnett, St. George, & Dron, 2011; Roca & Gagné, 2008; Xie, Debacker, & Ferguson, 2006). Hartnett, St. George, and Dron (2011) suggested a need to recognize the situated nature of motivation and found complex interactions between the aspects of the online environment in which the study participants (pre-service teachers) were operating and the different types of motivation they experienced (e.g., intrinsic, extrinsic, amotivation). Also working within the SDT framework, Chen and Jang (2010) found that contextual support (autonomy and competence) had a strong impact on students' perceptions of psychological needs satisfaction, which in turn, predicted autonomous motivation. The level of autonomous motivation did not significantly predict students' satisfaction with the course, but contextual support did. The authors viewed their results as supporting the general theoretical hypotheses of SDT in an online environment.

Shroff and Vogel (2009) sought to identify differences in factors that contributed to students' intrinsic motivation between online and face-to-face course discussion environments. They examined six possible factors: perceived competence, perceived challenge, feedback, perceived choice, perceived interest, and curiosity. Interestingly, perceived competence and perceived choice were the only factors that differed between course settings, with online students perceiving greater competence and greater choice than their face-to-face counterparts. While

statistically significant, the differences were only .17 and .26 points on a 5-point Likert scale, and perhaps represent little practical significance.

The research within the SDT framework reviewed above supports the importance of students' perceptions of autonomy support (a component of the context) in promoting intrinsic motivation, satisfaction of the basic psychological needs, and other important learning outcomes, namely engagement. These findings seem to hold true regardless of whether the learning environment is face-to-face or online.

Past experience. From a theoretical standpoint, virtually every creativity theory acknowledges the importance of prior knowledge and experience (Mumford et al., 2012; Sawyer, 2012). Experience in improvisation can increase one's musical creative thinking abilities (Koutsoupidou & Hargreaves, 2009) as can an entire music education program that focuses on creative thinking in music (Corakli & Batibay, 2012). The questions for this study are how to define and quantify past experience, and how it relates to other variables in the model. As previously discussed, one of the four ways in which self-efficacy can be developed is through mastery experience (Bandura, 2012). People evaluate their capabilities in relation to their experiences of exercising those abilities in the past. This suggests a direct effect of past experience on self-efficacy. However, it is also feasible that past experience, even if self-evaluated as being unsuccessful, still results in domain knowledge, which can be brought to bear on the current task. As such, past experience should also have a direct effect on outcomes.

The vast majority of music research tends to operationalize music experience as a dichotomy between "musicians" and "non-musicians," typically as a function of having or not having formal training. As Chin and Rickard (2012) point out, these operationalizations tend to focus exclusively on music production (i.e., performance) at the expense of other components of

musicianship (e.g., listening). In addition, frequency and duration of participation in musical activities is distinct from the importance or value that one places on those activities, but both are important facets of one's prior engagement with music.

Scales have been developed to gain a broader and deeper measurement of musical experience, musicianship, or musical engagement. For example, Werner, Swope, and Heide (2006) developed the *Music Experience Questionnaire* to include 53 items that measure six dimensions of musical experience (commitment to music, innovative musical aptitude, social uplift, affective reactions, positive psychotropic effects, and reactive musical behavior), which represent both reactions/responses to music and active involvement with music. Chin and Rickard (2012) developed the *Music Use Questionnaire* (MUSE) to provide a measurement of both frequency/duration of music experience and differences in style of engagement with music (i.e., the way one uses music or how music functions for an individual).

The Creative Process

Researchers have taken different approaches to studying the musical creative process, which can be divided into at least two broad categories: musical content analysis and verbal content analysis. In musical content analysis, the object of analysis is a *musical artifact*, such as an audio recordings, MIDI data, or notations. The researcher attempts to reveal characteristics of the overall musical creative process by searching for indicators of different sub-processes that are manifest in the music itself. Alternatively, with verbal content analysis, the object of analysis is the *words* an individual uses to talk about the musical creative process. The researcher searches for evidence of decisions, strategies, beliefs, or patterns of thought that are manifest in the words an individual uses when talking about what he/she thought while engaged in the process.

Musical content analysis. Kratus (1989, 1990a, 1994, 2001) has studied the musical creative process by having judges listen to recordings of music created by grade-school students and then rate whether students were engaged in either exploration, development, repetition, or silence for every five-second interval of the recording. In addition to other variables such as length of time working on recording and length of recording, this method provides a means to quantitatively measure the process such that the extent to which an individual spent time exploring, as opposed to developing, repeating, or making no sound (silence), could then be related to other quantitative variables such as musical aptitude and characteristics of the process (e.g., tonal cohesion, use of repeated rhythmic patterns, etc.) and compared across different grouping variables (e.g., age, sex, tonality, pitch availability, etc.).

One potential weakness of Kratus' approach is that the operationalization of the different behaviors (exploration, development, repetition) is too simplistic. For example, Kratus (1994) defined exploration as the following: "The music in a 5-second interval sounds unlike music played in earlier 5-second intervals. No specific references to music played earlier can be heard" (p. 121). Whether a 5-second interval sounds similar or different to previous 5-second intervals does not necessarily imply exploratory behavior, especially given the participants were novices on the keyboard. Differences in musical content between 5-second intervals could be divergent in nature, an attempt to find something new or different. But these differences could also be unsuccessful attempts to replicate what an individual had previously played or was hearing in his head, which would indicate a more convergent or confirmatory behavior that is less exploratory than it would appear. What is apparent from these dual interpretations is that the individual's intentions are not known, and therefore the same behavior can imply qualitatively different cognitive processes. So even while independent judges may be consistent in applying the

operationalization of the constructs, the operational definition itself leaves room for conflicting interpretations. Nevertheless, Kratus's research represents an important contribution to the literature and highlights the difficulties with any research that attempts to analyze musical content in isolation from verbal content.

Verbal content analysis. Within the broad category of verbal content analysis, researchers have taken different approaches to gathering data on individuals' thoughts during the creative process. Researchers have used qualitative methods such as ethnographic observation, discourse analysis, interviews, and talk aloud protocols (Burnard, 2000a, 2000b, 2000c; Burnard & Younker, 2002, 2004, 2008). Younker (2000a) used data from audio recordings a think aloud procedure, each collected during a different 30-minute session with each of six university students, during which the students created a song (with their voice) based on a poem. The data were used to create profiles of the six students' composing processes, each of which demonstrated varying degrees of expert and novice behavior. In particular, more expert composers exhibited a greater ability to generate an overall framework for the piece at the start and followed a recursive path between the whole and the parts, remaining flexible and making adjustments throughout the process. The more novice composers did not appear to create an initial holistic vision of the song and followed a much more disjoint, note-to-note approach.

In a similar study using think aloud protocol and interview data, Younker (2000b) searched for differences in the "patterns of thought" (p. 25) of 8-, 11-, and 14-year-old students while composing using a music software program (*Musicshop*) over the course of four separate sessions. Younker found that while all students did some amount of "exploring, recording, listening, evaluating, and editing" (p. 30), they seemed to differ in the depth and amount of time for which they engaged in these processes. In particular, the researcher noted that "the

differences exhibited by the student appeared to be due to their perceived abilities and comfortableness, as opposed to age” (p. 30). This seems to harken back to the importance of self-efficacy in the musical creative process, although familiarity with the musical software is almost certainly a mitigating factor and needed to be considered in the present study.

Taking a different approach, Burnard (2000a, 2000b) used an ethnographic methodology whereby, as a participant-observer, the researcher observed and interviewed (individual and focus group) 18 middle school students over six months as they engaged in various creative music-making activities in their weekly school music class. Children ascribed meaning to the concepts of composition and improvisation in three different ways: (a) the two are independent activities; (b) the two are interdependent upon each other; and (c) the two are indistinguishable from each other; and these different ways of experiencing the two concepts did not appear to be related to whether students had formal musical training. Burnard urged music educators to avoid making assumptions about how children experience composition and improvisation, and more importantly, support and encourage creativity in music learning. As she put it, “we need to encourage children's natural capacity and thirst for creating music in ways that are authentic and risky, unrepeatabe and unpredictable as well as the repeatabe and predictable” (Burnard, 2000a, p. 243).

Although already mentioned previously in the section on musical creativity theories, Burnard and Younker’s (2002, 2004) is worth mentioning again because it represents another way in which the creative process has been studied. By re-analyzing data from their previous studies using the constant comparative method, Burnard and Younker (2002) uncovered six different “pathways” or strategies that children followed when composing. They did this by organizing their data by events and searching for moments that represented decision points

(evidence of some choice being made). By mapping individuals' decisions over a composing task the different strategies (linear, recursive, and regulated) emerged from the data.

In a more descriptive study of the creative process, Auh (1999) had seventh-grade students complete a Compositional Process Questionnaire after they had completed composing with an instrument and notating their compositions. The questionnaire included questions like "How did you get the first idea for your composition?" and "How did you develop musical ideas?" (p. 58). The judges then rated the extent to which the students' answers exhibited enactive or reflective thinking (along a 5 point continuum). The answers were also analyzed qualitatively to search for specific characteristics of enactive and reflective thinking in the students' responses. Enactive thinkers are characterized by a tendency to "first, act out and then think about what they are doing," while reflective thinkers "first, think about what and how they are going to do, and then act out following their strategies" (p. 58).

Biasutti and Frezza (2009) developed the Improvisation Process Questionnaire (IPQ) and Improvisation Abilities Questionnaire (IAQ) to study the relationships between cognitive processes and ability perceptions of 76 musicians with at least two years of improvisation experience. Results from exploratory factor analysis of the IPQ indicated the presence of five factors. Drawing on previous research from Kenny and Gellrich (2002), the researchers interpreted the five factors as the following: *anticipation* (e.g., "When I am improvising, I can anticipate the rhythmic development of the whole solo."), *emotive communication* (e.g., "While improvising, all the emotions I can communicate depend on rhythmic features."), *flow* (e.g., "While improvising nothing distracts me."), *feedback* (e.g., "While improvising I turn the errors I make to my own advantage."), and *use of repertoire* (e.g., "While improvising I modify licks and phrases I have heard from other musicians.").

There are several weaknesses related to research that relies on participants verbal descriptions of their process that should be mentioned. First, because the data to be analyzed in this sort of research are the words of the participants, individuals' ability to verbalize what they think and how they think would certainly have an effect on how one interprets the nature of their creative thinking. In other words, think aloud procedures (during a task) and interviews (after a task) both require a certain degree of metacognitive ability (e.g., monitoring and evaluating) and verbal fluency.

Second, it is very possible that the method itself (i.e., thinking aloud) alters the thinking. More specifically, in summarizing their research on think aloud methods, Ericsson and Simon (1998) were explicit that such methods will only result in "direct expression" (i.e., verbalization of actual thoughts that do not in turn affect the thinking process itself) under certain specific preconditions, namely that "the participants are allowed to maintain uninterrupted focus on the completion of the presented tasks" and "participants are explicitly instructed to focus on the task while thinking aloud and merely to verbalize their thoughts...rather than describe or explain them to anyone else" (p. 181). They go on to explain that people (especially children) often feel naturally compelled to describe/explain their thoughts instead of simply verbalizing them. Because of this, a training procedure is recommended before conducting the actual think aloud procedure (Willis, 1999). Unfortunately, details regarding precisely how the think aloud procedure was conducted are rarely included in this research I have reviewed.

Finally, the research described above provides important insights into the creative thinking processes of specific individuals, but is not able to explain how the different thought processes relate to specific musical learning outcomes or how different individual characteristics come to bear on the thinking process (e.g., one's motivation for completing the composition,

one's self-efficacy beliefs, etc.). This is not to say the research is not important or has no value, but simply to note that, like all research, there are limitations.

Donin and colleagues (Donin, 2009; Donin & Féron, 2012; Donin & Theureau, 2007) have expressed a deep cynicism regarding previous research into the musical creative process. They espouse a belief that one must study “a real, and not a realistic nor a likely, creative process” because “[s]uch a process does not come into being as a result of a research project: its autonomous existence is an essential component of the composer's creative course and the history of music as a whole” (Donin & Féron, 2012, p. 264). In order to overcome some limitations of previously used methods to study the creative process, the researchers utilized a “situation simulation interview,” or what they have also referred to as an “interview within situation simulation through material traces” (Donin & Theureau, 2007, p. 235). With this method, over the course of many interviews (e.g., eleven) and over a lengthy time period (e.g., 6 months to two years), an individual composer re-enacts the compositional process in a room that simulates the composer's work environment (e.g., includes materials such as his computer, sketches, documents, etc.) and verbalizes his thoughts.

The researchers believe that this approach overcomes limitations of in-the-moment think aloud protocols, experimental designs, and other methods they view as not ecologically valid. They argue a need to be physically present to observe the process as it unfolds, although these researchers are present for a *re-enactment* of the process, which has its own limitations. Other researchers (Collins, 2005, 2007; Folkestad, Hargreaves, & Lindström, 1998; Nilsson & Folkestad, 2005; Seddon & O'Neill, 2003) acknowledge the potential problems of a “surveillance effect” and have used alternative means such as computer-based data collection to intentionally

distance themselves from the process and allow it to unfold as naturally as possible. Both observational approaches have their limitations.

Collins (2007) sought to overcome these limitations by combining both types of analysis (musical content and verbal content). As an analogue to the think aloud method, the researcher collected MIDI “save-as” files at specified points during each session and at points when the composer deemed a significant change was made. These files provided a record of the composer’s work as it changed and progressed over the course of composing the work. In addition, Collins implemented a pre-process verbal protocol to gain insights into the composer’s plans for each session and an immediately post-process verbal protocol for each session. From this research, Collins concluded the following:

...immediately retrospective verbal reporting with computer-based data collection techniques could provide substantially rich data to postulate a time-based hypothetical model of compositional cognition for this particular composer. The methodological procedure, which underwent several layers of careful refinement, clearly allowed the tracking in real-time of stages of creative thinking in an individual composer and how music is structured over time. Moments of creative insight which reflected Gestalt theorist’s notions of problem restructuring were observed throughout the emerging compositions, and the recognition of such fleeting moments was only achievable through the use of the ‘essentially human methodology’ of verbal protocol techniques allied to digital media. (p. 253)

While this methodological procedure appears to be quite valuable, it has since only been reported in one additional study (Collins & Dunn, 2011). Nonetheless it remains a potential avenue for future research.

Creative process and learning. As far as I can tell, none of the research I have just reviewed conceptualized the creative process specifically as a mode of learning. At base, this research has attempted to reveal the *structure* of the musical creative process: what and how do people think when engaged in musical creative activity? What does not appear to have been sufficiently addressed in the empirical research literature is the *function* of the musical creative process within an educational context. In educational contexts, the function of process has been investigated by way of the construct known as engagement.

Engagement. In educational and organizational research, engagement has been conceptualized and measured in many different ways. In an extensive review of the engagement literature, Fredricks, Blumenfeld, and Paris (2004) identified three types of school engagement: behavioral, emotional, and cognitive. These are virtually identical to those identified by Jimerson, Campos, and Greif (2003) (behavioral, affective, and cognitive) in a similar, but far less extensive review. Behavioral engagement is indicated by behaviors such as effort and attention (Wellborn, 1992) (Wellborn, 1992), as well as involvement and choosing challenging tasks (Skinner & Belmont, 1993). Emotional engagement is marked by positive emotions such as enthusiasm, optimism, curiosity, and interest (Skinner & Belmont, 1993), and the absence of negative emotions such as boredom, sadness, and anxiety (Fredricks et al., 2004). Much of the earlier research on school engagement defined the construct in terms of both behavioral and emotional components. Cognitive engagement is often defined in two different ways, either in terms of psychological investment in learning or in terms of “being strategic or self-regulating” (Fredricks et al., 2004, p. 64). Still others have noted that Csikszentmihalyi’s (1990) concept of flow can be seen as a form of engagement (Rupayana, 2010).

From Fredricks and colleagues' review, the authors found a good deal of conceptual overlap with the definitions of the different types of engagement and other constructs in the more general motivation literature (e.g., motivation to learn, learning goals, etc.). This overlap is potentially beneficial because the construct attempts to integrate concepts that are often studied in isolation from each other. However, the overlap can also result in a loss of conceptual clarity. From a measurement standpoint, the engagement measures utilized in the broader literature tend to operationalize engagement using only a few items for each sub-component (e.g., effort, persistence, value, etc.), each of which is typically measured with a finer degree of conceptual distinction (more items, more facets) when the sub-components are measured in their respective theoretical frameworks (e.g., goal theory). Nevertheless, the same authors suggest that "any disadvantages of using only a few items to tap each construct may be offset by the increased predictive strength of a streamlined single measure" (p. 70).

The research on engagement can also be separated into school-level versus classroom-level analyses. Much of the educational research has approached engagement from the contextual level (i.e., not specific to a given task or situation) by focusing on broad levels of engagement over time (e.g., over a school year, over an entire course) because the outcomes of interest were at a similarly broad level (e.g., dropout rates, academic success in college, overall academic achievement). As such, many of the measures of engagement operationalized the construct as trait-like in nature.

Other researchers have approached engagement at the situational level. Research on the classroom context has indicated that factors such as teacher support, peer groups, classroom structure, autonomy support, and characteristics of the task have varying degrees of influence on behavioral, emotional, and cognitive engagement (Fredricks et al., 2004). For example, Rotgans

and Schmidt (2011) hypothesized that cognitive engagement is dependent upon autonomy, which is task-specific. They define the construct as “the extent to which students’ are willing and able to take on the learning task at hand. This includes the amount of effort students are willing to invest in working on the task...and how long they persist” (p. 467). Because no measure of situational cognitive engagement had been reported in the literature, they developed a measure in order to test their hypotheses. Their initial validation studies indicated their four-item instrument had very good fit with a single-factor model and a reliability estimate (Hancock’s *H*) of .93 for the exploratory study and .78 for the cross-validation study.

Unfortunately, Rotgans and Schmidt (2011) did not test the impact of autonomy on situational cognitive engagement directly. Working within a Problem-Based Learning (PBL) framework, they measured students’ situational cognitive engagement at different stages of the learning process. In PBL, students begin with an initial discussion phase, followed by student identification of learning goals and then a period of self-directed learning. After the self-study stage, students share what they have learned and confirm their understanding, which is followed by the identification of new learning goals, and the cycle continues. The researchers hypothesized that different stages would inherently be perceived as being more or less autonomous by the students (e.g., the initial discussion stage would be less autonomous than the self-directed learning stage). The student perceptions of autonomy during the different phases were never measured, so while previous levels of engagement predicted levels of engagement in subsequent phases, it was not possible to determine the extent to which autonomy truly played a role in encouraging engagement.

Engagement as a construct has received fairly limited attention in music education research, at least not as defined above. As discussed in the section on past experience, music

engagement has been treated as what might be considered a global or contextual level variable, but not a situational variable. Chin and Rickard's (2012) defined musical engagement as "an individual's level of active participation in music activities, measured by the frequency and regularity of participation, and the value assigned to the music activity" (pp. 430-431). This use of the term engagement is somewhat similar to the way engagement has been defined in research that examines engagement at the school-level (not the classroom or situation level).

Working from the positive psychology framework, Lamont (2011b) approached engagement as a component of happiness or psychological well-being, and defined it as "the pursuit of gratification through absorption in a given task or activity" (p. 230). The researcher analyzed participants' descriptions of "Strong Experiences in Music" for evidence of (flow-like) engagement. Although more situational in nature, this approach to engagement in music was primarily descriptive.

Focusing on music listening, several researchers have utilized an *Experience Sampling Methodology* (Czikszenmihalyi & Lefevre, 1989) as a means of accounting for the influence of the specific context on individuals' listening experiences (Greasley, 2008; Greasley & Lamont, 2011; North, Hargreaves, & Hargreaves, 2004; Sloboda, O'Neill, & Ivaldi, 2001). With the *Experience Sampling Methodology* participants typically carry some device (e.g., pager, cell phone, etc.) and receives notifications at different times (determined by the researcher) to complete a short questionnaire regarding their experience at the time of the notification. In the music listening research, the participant typically answers questions about the music they are listening to at that moment. The questions may address affective components (e.g., feelings and reactions), functions or uses of the music (e.g., reasons for listening), and details related to the situation and listening environment (e.g., where the listening is occurring, what type of music).

While this research approaches engagement from a situational perspective, it has focused almost exclusively on music listening.

There has also been some research into engagement during creative activity (e.g., composition). In summarizing their previous research, Brown and Dillon (2012) identified five attributes that characterize meaningful engagement during act of composition, which are “those aspects of practice that lead to a satisfying involvement in the act of composition” (p. 94). The five attributes are challenge, motivation, involvement, sensitivity, and virtuosity). These five attributes also relate to each other in specific ways. In particular, Brown and Dillon claim the following:

Our research suggests that composers who are meaningfully engaged choose a sufficiently complex task (the attribute of challenge) that will maintain their interest (the attribute of motivation) through all the composition and production stages. By paying attention to the material (the attribute of involvement), they notice a great range of opportunities and are able to make effective choices between those possibilities (the attribute of sensitivity). Their skill in using the available tools and medium (the attribute of virtuosity) provides them with the capacity to realise any musical idea that emerges. (p. 95)

Most of the research reported by Brown and Dillon involves networked music-making using the Jam2Jam software. On the Jam2Jam website (<http://explodingart.com/jam2jam/jam2jam/Research/Research.html>), over 40 research publications are listed, indicating a keen interest in the notion of developing software environments that promote meaningful engagement. While I do not wish to dismiss the importance of this research, not one of the articles appeared to connect meaningful engagement

with actual learning outcomes. The vast majority of the articles (based on a review of the titles, abstracts, and a careful reading of several of the actual articles) focus on educational *potentials* and affordances of meaningful engagement (e.g., Brown & Dillon, 2007; Hirche, 2011), while evidence of such engagement or actual learning that results from this potential for engagement appeared to be absent. Even the five attributes that characterize meaningful engagement in composition (see above) are said to lead to “satisfied involvement,” which is certainly not the same as an increase in knowledge (i.e., learning).

I suggest that this lack of evidence related to learning outcomes in the research on meaningful engagement in music composition is a direct result of a seemingly exclusive use of qualitative, exploratory, and descriptive research methods. Engagement is described, but not necessarily measured. Potentials are explored and identified, but not necessarily tested and confirmed. For this reason, it is necessary to attempt to make those connections from environment and engagement to learning as an observable and measureable outcome.

In all of the literature I searched and reviewed I was unable to locate any research that attempted to quantify situational engagement during creative activity (as has been done with other research fields such as motivational and organizational psychology), let alone relate it to antecedents (motivation, perceptions of support/context, self-efficacy, etc.) or consequents (learning, perceived learning, changes in self-efficacy, etc.). This represents a significant gap in the research literature related to musical engagement and learning, especially as related to creative musical activity like composition and improvisation.

Divergent and convergent thinking. Central to Webster’s (1987a, 2002) conception of creative thinking in music is the notion of convergent and divergent thinking. Specifically, he contends that “creative thinking is a dynamic process of alternation between convergent and

divergent thinking, moving in stages over time, enabled by certain skills (both innate and learned), and by certain conditions, all resulting in a final product” (2002, p. 26). Therefore, the absence of convergent and divergent thinking from the General Specified Model is quite conspicuous. This was neither a trivial, nor a thoughtless, nor an inconsequential decision on my part. It was, however, made on defensible grounds for theoretical and practical reasons, which I will expand upon below.

There is a very rich history of divergent thinking research in the general creativity literature, much of which is built upon work of Guilford (1968) and his Structure of Intellect (SOI) model, Mednick’s (1962) theory regarding remote associations, Torrance’s notions of fluency flexibility, originality, and elaboration that formed the basis of the Torrance Tests of Creative Thinking (TTCT) (e.g., Torrance, 1966, 1998), Wallach and Kogan’s (1965) work with children, among many others (for a review of this literature see Runco, 2010). The vast majority of this research has treated divergent thinking (DT), measured by DT tests, as what might be called creative potential (Runco, 2007a), something more akin to a trait-level variable like personality or IQ. This research has related DT almost exclusively to creative production (i.e., relative creativity of products or relative creative achievement in a given domain).

In the domain of music, measures analogous to DT tests have also been developed, such as the Measure of Creative Thinking in Music – Version II (MCTM-II) (Webster, 1987b), the Measures of Creativity and Sound in Music (MCSM) (C. Wang, 1985), and the Measure of Musical Problem Solving (MMPS) (Vold, 1986). All of these measures examine creative thinking in music as an ability, as something that some individuals have more or less of, but which can potentially be developed. Divergent and convergent thinking are certainly a component of these measures (either explicitly or implicitly), but they purport to measure

between-individual differences in general (context-independent) creative thinking abilities, not within-individual differences for a given creative task. Put simply, they measure a trait, not the process. In addition, much of the research on the DT test analogues in music has related musical creative thinking abilities to other global or contextual level variables such as musical aptitude or music achievement. To my knowledge, none of this research has related musical creative thinking ability to more situation-specific learning outcomes. If such research were conducted, it would still need to relate musical creative thinking ability to other components of the process and the context in order to provide meaningful information on how such ability manifests its effects on learning outcomes.

There is one additional reason I have not included divergent and convergent thinking in the statistical models. While both types of thinking are likely part of the process, I argue that anything short of an analysis of these processes over time while a person is engaged in creative activity will fail to provide meaningful information about the relative importance and contributions of these modes of thinking during the process. In other words, it is conceivable that one might devise a self-report measure that includes items designed to gauge the extent to which an individual exhibited behavior that is characteristic of divergent or convergent thinking (e.g., “I searched for as many possible types of sounds as I could when choosing the instruments in my song,” rated on a Likert scale). Such a measure would allow a researcher to see how individuals differ in regards to their overall divergent or convergent creative thinking behaviors for a specific task, and these varying levels could be related to different learning outcomes. I am not entirely certain what value such an approach would have in determining teaching strategies. Perhaps we would find out that more divergent thinking behavior leads to better learning outcomes. But we still believe convergent thinking is an important part of the process. How

does the teacher determine what ratio of divergent to convergent thinking is “ideal?” More importantly, such a prescriptive approach would almost certainly undermine the ebb-and-flow nature of the creative process, and likely affect students’ motivation and sense of autonomy. Even if determining relative amounts of convergent/divergent thinking was deemed important, these levels would almost certainly be dependent upon the extent to which an individual engages with the process. Given the literature reviewed above regarding engagement, I suggest that an appropriate operationalization of “the process” component of the model is to measure cognitive and affective engagement, not relative levels of convergent or divergent thinking. Such an operationalization seems to follow Webster’s (2002) definition of creativity in music: “the engagement of the mind in the active, structured process of thinking in sound for the purpose of producing some product that is new for the creator.”

The General Specified Model of Creativity-Based Learning

After taking into consideration all of the literature that has been reviewed, I have constructed a base model, which I refer to as the General Specified Model of Creativity-Based Learning (or the General Model for short), which could also function as the basis for other competing, but still related, models. The General Specified Model of Creativity-Based Learning can be found in Appendix A. Alternative models could include slight variations of the General Model based on different interpretations of the research literature. The full structural model, which shows the path model (relationships between constructs) and the measurement model (relationships between the items and the unobserved factor/construct), including all constructs actually measured in this study, can be found in Appendix B.

The General Model identifies three levels of generality at which variables reside (following the general notion of Vallerand’s (1997) Hierarchical Model): global, contextual, and

situational. These three levels of generality are also roughly analogous to Webster's (2002) enabling skills and two components of enabling conditions (context and task). I have chosen this hierarchical organization of variables in order to account for the different ways in which the different components of Webster's model have been conceptualized and measured in the literature.

Global level. At the global level are factors that are viewed as trait-like in nature (relatively stable over time) and are hypothesized to have a general impact across all contexts and situations. As such, they are modelled as impacting all latent variables below the global level. From a SEM standpoint, they are all exogenous variables and are allowed to covary. Global factors include Personality, Musical Aptitude, Past Music Experience, Conceptual Understanding (prior to beginning the course), Age, and Education Level. In the General Model all global factors were also treated similar to control variables (covariates), that is, they have an impact on all other endogenous variables (all variables at the contextual and situational level). The result of this approach is that any variance attributable to these global factors is partitioned out when considering the effects of the remaining contextual and situational variables on each other. Since they are treated as control variables, the General Model can also be seen as the least constrained model because all parameters related to the global factors' effects on lower level variables are freely estimated. Alternative models could place further constraints on the effects of global factors on lower level variables.

Contextual level. At the contextual level are variables identified in the literature as being more context- or domain-dependent, but not necessarily specific to a given situation. There are two outcome variables at the contextual level, Perceived Learning and Conceptual Understanding (at the end of the course). While both of these outcomes have been treated as

both contextual and situational in the literature, they were intended to be measured in relation to the MOOC course as a whole, and therefore are not situation-specific. Conceptual Understanding was not actually measured in this study, but I have left it in the model for the sake of completeness. Other variables at the contextual level are Musical Creative Self-Efficacy (MCSE) and Contextual Support. Following self-efficacy theory (Bandura, 1991, 2006, 2012), MCSE is domain-specific (music) and even somewhat task-specific (creating music as opposed to listening, performing, etc.). However, every musically creative task has characteristics that are specific to the situation, which is not accounted for by MCSE, which is why it is not viewed as a situational factor. Similar to other research under the SDT framework, particularly in online environments (e.g., Chen & Jang, 2010), Contextual Support resides at the contextual level because it is a general perception of autonomy-support over an entire course. As such it is not a trait (global factor) and not specific to just one assignment or activity in the course (situational factor).

Situational level. Finally, the situational level includes variables that are related to a specific task (e.g., writing a song). These variables include Situational Need Satisfaction, Situational Autonomous Motivation, Situational Cognitive Engagement, and the learning outcome Self-Evaluation of Creative Product. Following the tenets of Self-Determination Theory, the extent to which someone is autonomously motivated is dependent upon the extent to which their basic psychological needs are being met (Deci & Ryan, 1985; Deci & Ryan, 2012a, 2012b). An autonomy-supportive environment and an individual's self-efficacy for a task should then also help meet the basic psychological needs in a given situation. Furthermore, when one is more autonomously motivated, they are often more deeply cognitively engaged in a task (Walker

et al., 2006). It is this deeper engagement in the task that is hypothesized to result in positive learning outcomes.

When viewed as a whole, the General Model attempts to explain how some of the various components of the creative process (taken from Webster’s (2002) *Model of Creative Thinking in Music*) interact and eventually result in learning outcomes. What remains to be seen is whether the General Model (or any alternative models based on the General Model) will adequately represent data from a real-world creative learning context. The following chapter will explain how I attempted to collect and analyze these data in order to address the initial research questions for this study as well as the revised research questions.

Sub-Models

The General Model specifies relationships between a large number of components in Webster’s model, but the General Model could also be looked at as a combination of several smaller component mediational models. For example, Musical Creative Self-Efficacy, a

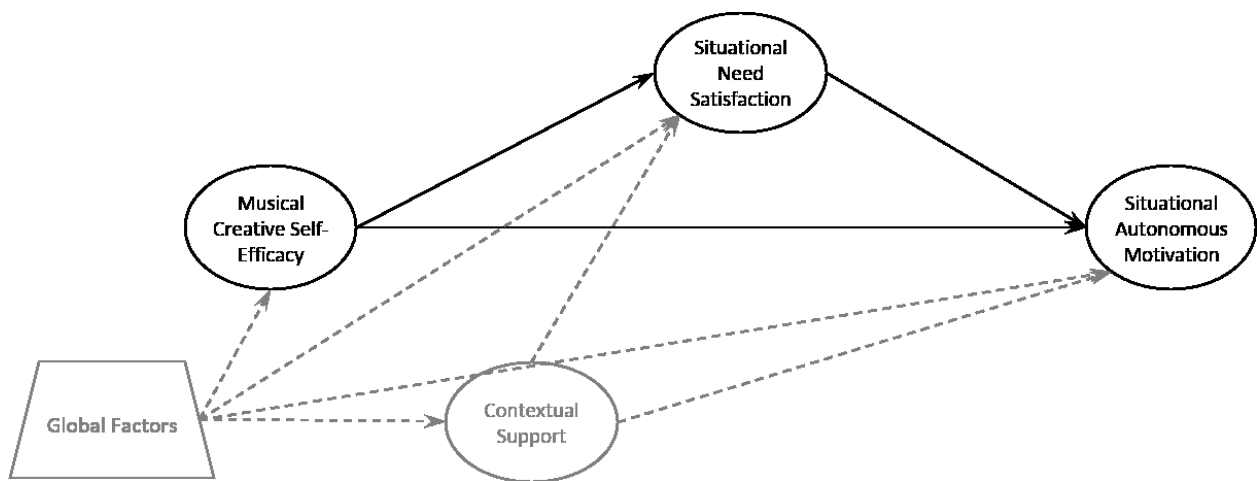


Figure 2.2. MCSE Sub-Model 1.
Effects of Self-Efficacy on Situational Autonomous Motivation, as mediated by Situational Need Satisfaction, controlling for all Global Factors, and controlling for Contextual Support effects on Situational Need Satisfaction and Motivation.

contextual level variable, is hypothesized to predict Situational Autonomous Motivation, but this effect is mediated by Situational Need Satisfaction, also a situational level variable (see Figure 2.2).

Similarly, Musical Creative Self-Efficacy is hypothesized to have an effect on Situational Cognitive Engagement, and this effect is transmitted by Situational Need Satisfaction and Situational Autonomous Motivation (see Figure 2.3). When seen within the context of the larger model, these effects can be examined after controlling for Global level factors and other related factors (e.g., contextual support).

Another sub-model that can be examined is what might be considered the Self-Determination Theory component of the model, which is the effects of Contextual Support on Situational Autonomous Motivation as mediated by Situational Need Satisfaction. Figure 2.4

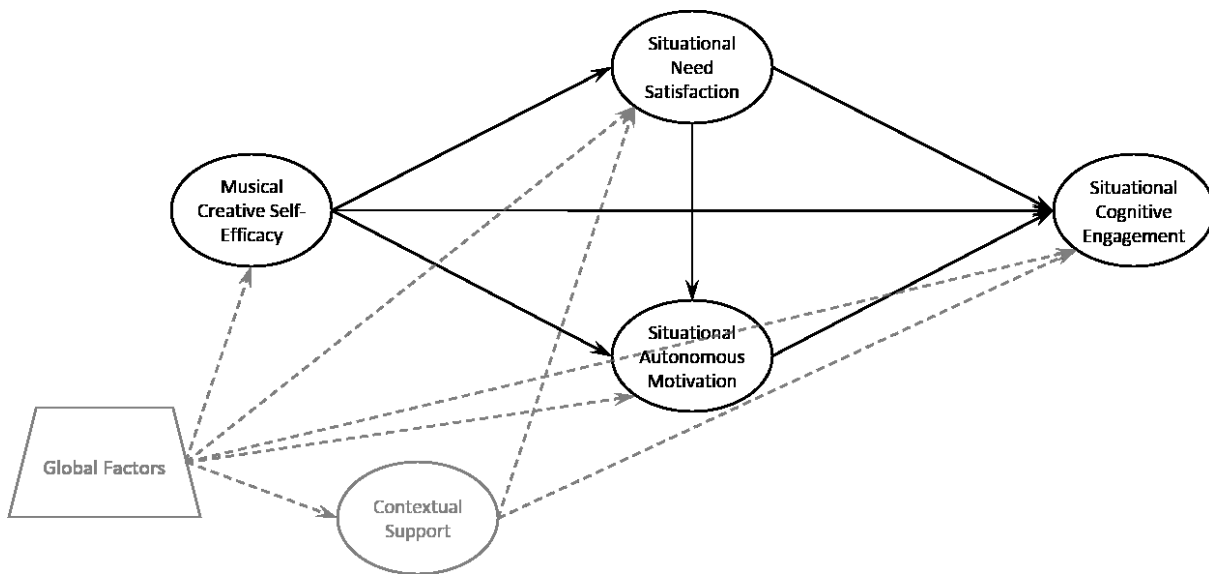


Figure 2.3. MCSE Sub-Model 2.

Effects of Self-Efficacy on Situational Cognitive Engagement, as mediated by Situational Need Satisfaction and Situational Motivation, controlling for all Global Factors, and controlling for Contextual Support’s effects on Situational Need Satisfaction and Situational Cognitive Engagement.

illustrates how these effects can be examined after controlling for the various Global level factors and also Musical Creative Self-Efficacy.

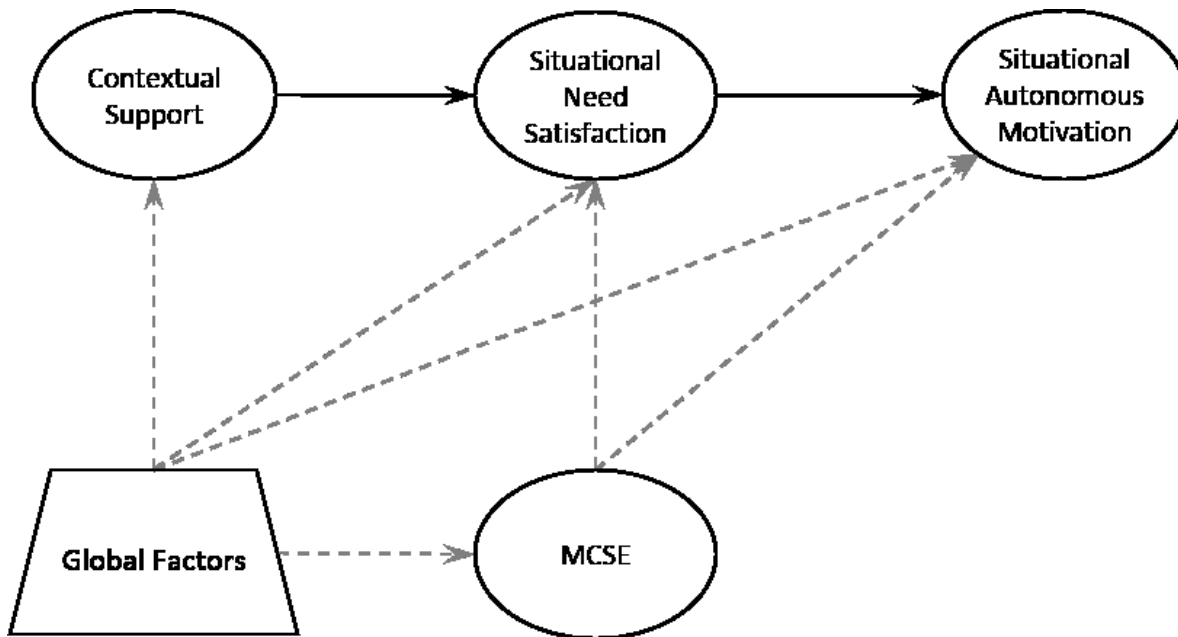


Figure 2.4. Contextual Support (Self-Determination Theory) Sub-Model. Effects of Contextual Support on Situational Motivation as mediated by Situational Need Satisfaction, controlling for all Global Factors and MCSE.

We can also look at the way in which Situational Autonomous Motivation transmits the effects of Situational Need Satisfaction on Situational Cognitive Engagement, once we have controlled for differences in factors at the Global level and Musical Creative Self-Efficacy (see Figure 2.5). Finally, as we reach the final outcome of the model, Learning Outcomes, Figure 2.6 shows how we can examine the extent to which motivation exhibits its effects on the process (as measured by Situational Cognitive Engagement), and how those effects lead to differences in Learning Outcomes. Again, these effects can be viewed when controlling for Global level differences. Each of these component sub-models could represent individual research studies by themselves, but when placed in the context of the entire model, there is great potential for integrating the different component theories.

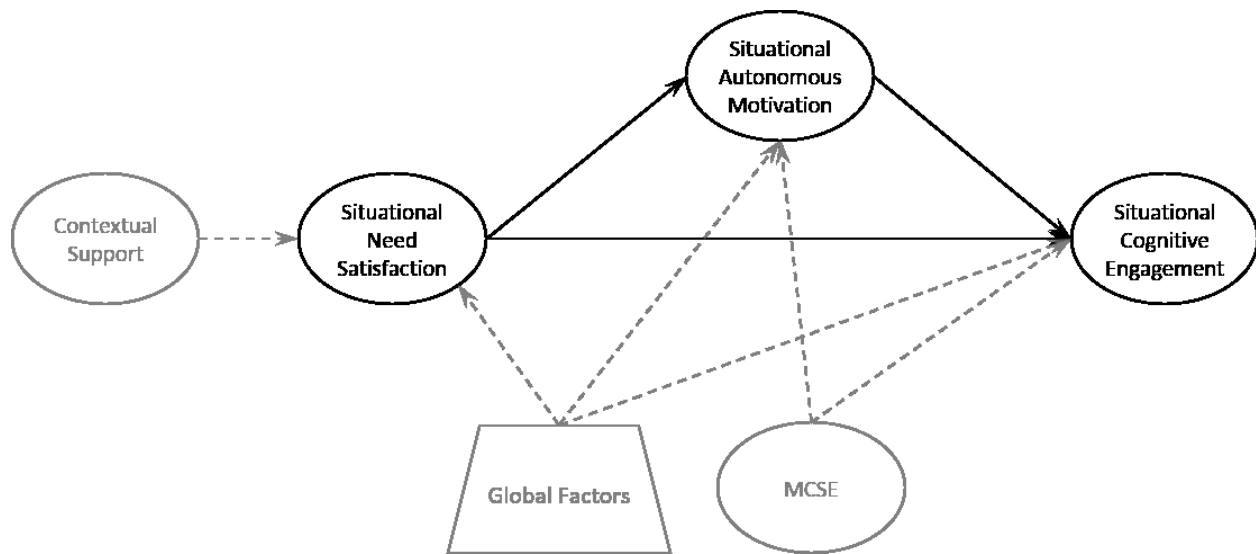


Figure 2.5. Situational Need Satisfaction Sub Model.
Effects of Situational Need Satisfaction on Situational Cognitive Engagement, as mediated by Situational Motivation, controlling for all Global Factors and Musical Creative Self-Efficacy.

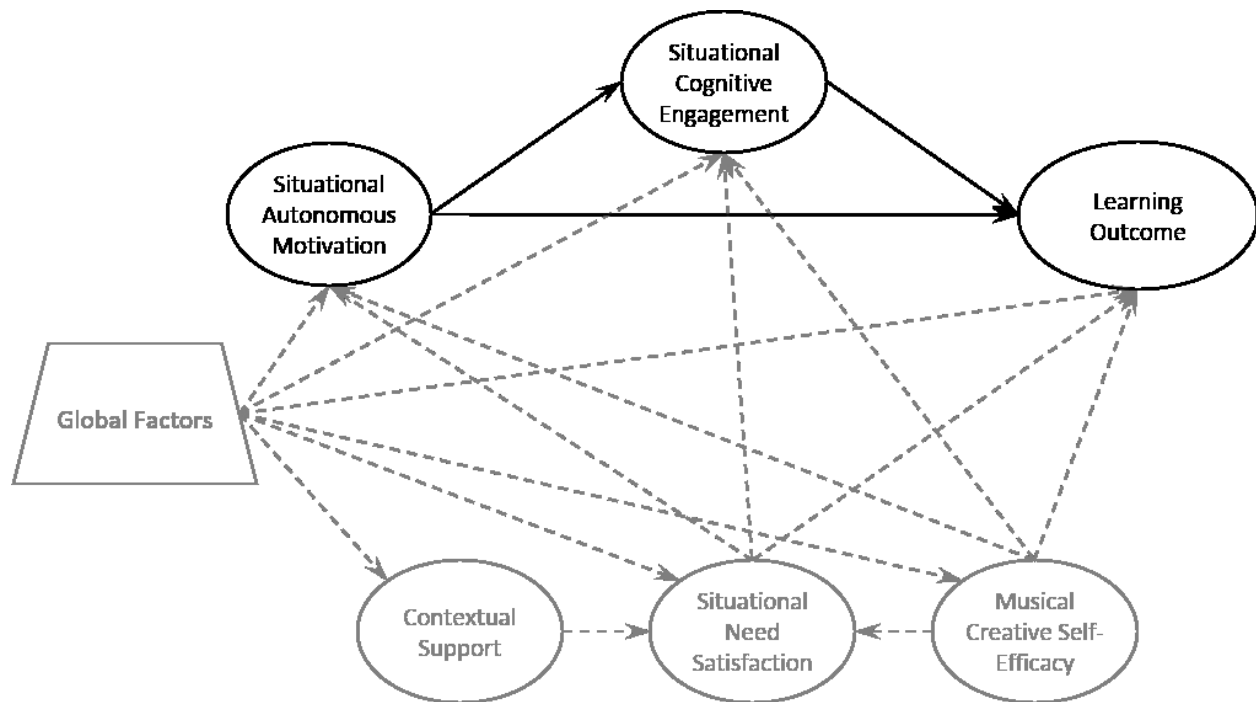


Figure 2.6. Situational Autonomous Motivation Sub-Model.
Effects of Situational Motivation on Learning Outcomes, as mediated by Situational Cognitive Engagement, controlling for all Global Factors, Situational Need Satisfaction, and Musical Creative Self-Efficacy.

CHAPTER THREE: METHOD

In this chapter I outline the design of the study, including a description of the research context and participants, as well as the measurement procedures and measurement instruments. I present the details of the design as originally conceived, followed by an explanation of the procedures I had intended to follow for data analysis. This is followed by a description of the actual data analyses that were carried out on the data that was actually collected. The overall initial design could be simply stated as follows: (1) create a creativity-based learning context, (2) enroll students from around the world, (3) embed measurements of relevant constructs throughout the course, and (4) analyze the relationships between those constructs in an attempt to identify a model that best fits the observed data. I begin with a description of the course I created, which was specifically designed to allow students to engage in creative music-making for the purpose of learning about music.

Research Context

The research was conducted within the context of a Massive Open Online Course (MOOC) entitled “What is Music?: Finding Your Song” offered in January 2014 via the Canvas Network, and sanctioned by the University of South Florida. As a MOOC, the course was not offered for college credit, so it was open to anyone in the world with an internet connection and a desire to enroll. MOOC’s have generated enrollment numbers anywhere from a few hundred to hundreds of thousands of students, although only a small percentage of those initially enrolled complete a MOOC. A recent MOOC offered by the University of Rochester through Coursera

entitled “History of Rock Part 1” saw an enrollment around 44,000, of which about 63% actively participated (Covach, 2013). Because MOOC’s are inherently geared toward larger enrollment numbers, the course design of most MOOC’s follows a generally self-directed format. This course followed a similar self-directed format with the instructor (the researcher) functioning more as a facilitator and course administrator during the actual running of the course. This was possible because all instructional content was developed in advance and the majority of interactions in the course were between students, as opposed to with the instructor. The instructor’s interaction with the students was done primarily by way of course design (e.g., planning of student activities, course content, development support materials and resources, inclusion of opportunity for student autonomy).

The course was based on an online course I had previously designed and taught at the University of South Florida, intended for non-music major undergraduate students. I designed the course around three important premises:

1. It should be accessible to anyone, regardless of prior experience with music (so knowledge of standard music notation was not required, nor was it used)
2. Individuals with different levels of experience and musical abilities should be able to adapt the course according to their individual interests and abilities.
3. Beyond learning about music through viewing the content I created, students would learn by doing, or more specifically, by creating. My belief was that the concepts would come alive more and gain more meaning when applied to the students’ own musical creations.

The course consisted of five modules, each offering some answers to the question “What is Music?” from a different perspective, including Music as Human Activity, Music as Emotion, Music as Physics, Music as Form, and Music as Culture. Within each module students viewed

content I created (text and videos), which presented the fundamental concepts for the module. The video presentations were fully interactive and mobile-compatible. They included a combination of standard PowerPoint-style slides with annotations, student-controllable interactive media content, transcripts for the narrations, fully searchable text, animations, interactive graphs, and videos.

After viewing the main content for the module, students then learned more about those concepts by way of creative music-making projects, which they shared with fellow students in a discussion forum. For example:

- In the Music as Human Activity, students created a “Top 10 Best Music Playlist”, which they shared using services like Spotify, and then discussed the various functions of music within their lives.
- In the Music as Emotion module, students wrote lyrics for a potential new song, or wrote new lyrics for a song they were already familiar with. Students then analyzed those lyrics using the concepts covered in the module.
- In the Music as Physics module, students wrote a melody using any number of resources that were suggested or provided to them, and then analyzed that melody with a free visualization software (Sonic Visualiser).
- In the Music as Form module, students built upon writing lyrics and writing a melody to move toward writing a song, and then examining the large-scale structure of the song.

Students shared their projects via a discussion board and were encouraged to review and comment on the projects of their fellow students.

Participants

Total enrollment by end of the course was 824 students, but of those 334 never logged into the course (“no-shows”). Students interested in participating in the research component of the course were directed to a course webpage that explained the research and requested students indicate their consent to participate (via a survey within the course). This research was approved by the Institutional Review Board at the University of South Florida. A total of 281 students participated in at least some component of the research. Students enrolled from all over the world.

Figure 3.1 shows the proportions of the research participants by geographic region. Nearly half of the participants were from North America (primarily the US and Canada). Participants ranged in age from “13-18” to “65 and older”, but the largest proportion of students were within the age range of “25-34”. Participants in the age range of “13-18” were removed from further analyses because parental permission could not be acquired. Approximately 25% of the participants held graduate degrees and another roughly 25% had completed a four-year college degree. For only 55.4% of my participants (of those that answered the question about primary language) was English actually their primary language.

Measurement Process

Clearly, there was a large number of constructs to be measured in this study. As such, it was important to balance issues related to validity and reliability with practical issues of obtaining data on each construct. Given the number of constructs to be measured, it was possible that participants would become over-whelmed, annoyed, or deterred from participating (or continuing to participate) if the measurement process was not carefully managed. One way I attempted to address this issue was to frame the completion of each measure as a way for the

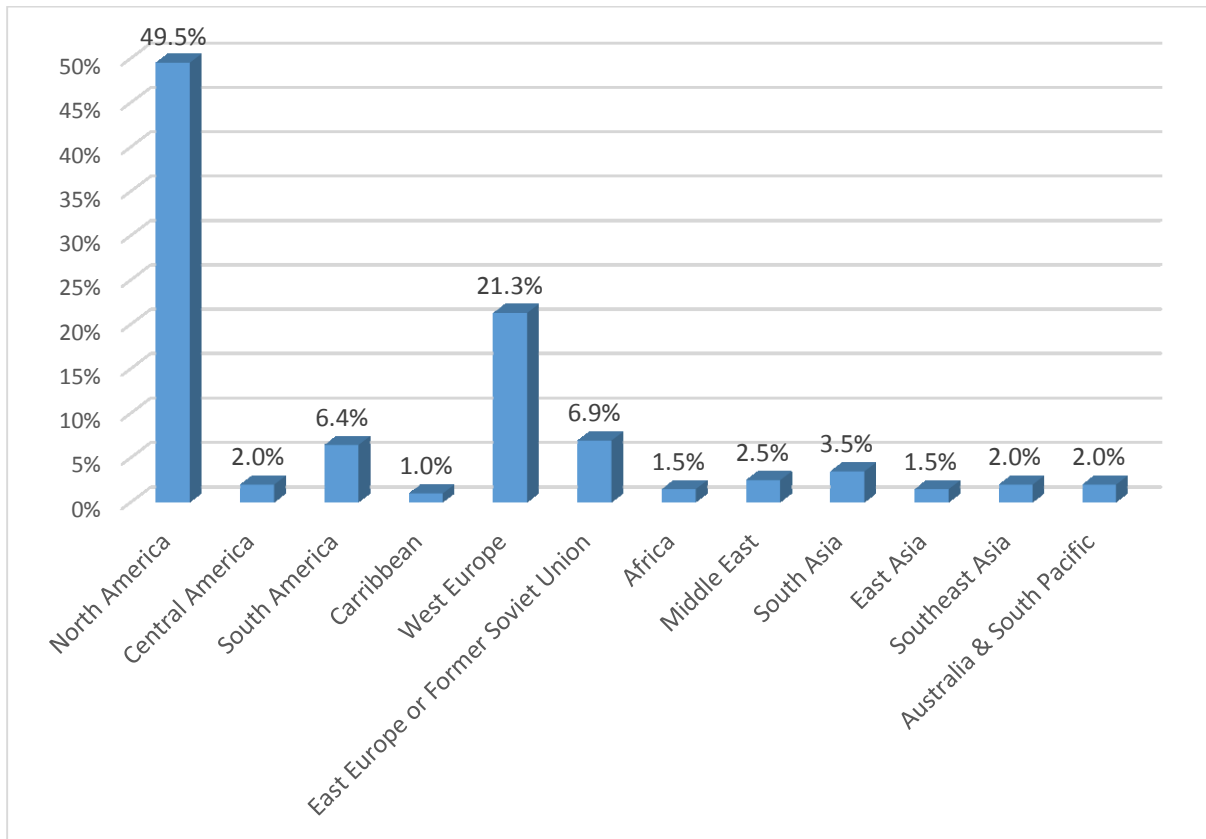


Figure 3.1. Proportions of Participants by Geographic Region

Note. $N = 202$. Proportions are based on participants for whom data about geographic location was available, not the total course enrollment.

participant to learn more about one's self. For example, many people have never completed a legitimate personality profile or music aptitude test. I hoped that the desire to know more about one's self would motivate students to complete measures.

A second way I attempted to deal with the measurement overload problem was to spread out the measurements throughout the course. For example, constructs that are conceptualized as more trait-like, and therefore more persistent over time and across contexts (e.g., personality) could be measured at any point in the course. Other constructs were more situation or task-specific (e.g., situational motivation) and therefore needed to be measured at specific points. The

initial measurement schedule is displayed below in Table 3.1, although this was altered for several reasons. First, Pilot Study #2 was delayed for reasons beyond my control, so I decided to wait on administering the Musical Creative-Self Efficacy Scale as part of the pre-course questionnaire. Instead, it was not administered until week 3, which in retrospect was rather unfortunate because this was around the same time when I witnessed a fairly drastic decrease in course activity from the students. Not included in the measurement schedule is the information collected through the Welcome to Canvas Network survey. This survey was conducted by Canvas Network, but was embedded into the course. It is administered in all Canvas Network courses at the beginning of the course. It is discussed in more detail below.

Table 3.1.

Initial Measurement Schedule

Level	Variable	Time of Measurement	# Items
Global Level	Demographics (Age, Gender, Education)	Pre-Course Questionnaire	3
	Past Experience in Music (PEM)	Pre-Course Questionnaire	11
	Personality	Week 2: Music as Human Activity Module	24
	Music Aptitude (MA)	Week 4: Music as Physics Module	72
Contextual Level	Musical Creative Self-Efficacy (MCSE)	Pre-Course Questionnaire	24
	Contextual Support (CS)	Week 5: Music as Physics Module	24
Situational Level	Situational Autonomous Motivation (SAM)	Weeks 6-7: Music as Form Module	12
	Cognitive Engagement (CE)	Week 7: After submitting Song Project	4
Learning Outcomes	Self-Evaluation of Creative Product (SECP)	Week 7: After submitting Song Project	4
	Perceived Learning (PL)	Week 8: Music as Culture Module	15
Total Items			193

Primary Measures

The following section describes which specific measures were used for this study and how the constructs were operationalized. Table 3.2 displays a summary of the measures that

were used, including whether the measure was based on a measure previously used in the literature, or a newly developed measure. The specific items for each measure can be found in Appendix C.

Table 3.2.

Summary of Measures for All Variables

Level	Variable	Measure/Scale/Index	Variable Type
	Gender	-	Observed
	Age	-	Observed
	Education Level	-	Observed
Global Level	Past Experience in Music (PEM)	Researcher-developed measure (PEMI)	Latent
	Personality	Mini-IPIP6 (Milojev, Osborne, Greaves, Barlow, & Sibley, 2013)	Latent
	Music Aptitude (MA)	Brief PROMS (Law & Zentner, 2012)	Observed
Contextual Level	Contextual Support (CS)	Items revised from the Intrinsic Motivation Survey (Schroff & Vogel, 2009)	Latent
	Musical Creative Self-Efficacy (MCSE)	Researcher-developed measure (MCSES)	Latent
Situational Level	Situational Autonomous Motivation (SAM)	modified Situational Intrinsic Motivation Scale (SIMS) (Guay, Vallerand, & Blanchard, 2000)	Latent
	Cognitive Engagement (CE)	Situational Cognitive Engagement scale (Rotgans and Schmidt, 2011)	Latent
Learning Outcomes	Self-Evaluation of Creative Product (SECP)	researcher developed (based on Amabile, 1996)	Latent
	Perceived Learning (PL)	CAP Perceived Learning Scale (Rovai et al., 2009)	Latent

Demographics. In the General Model, all global factors are treated as control variables. Three of the global-level variables are demographic variables: Age, Gender, and Education Level. Demographic information was collected at the beginning of the course via a combination of the Pre-Course Questionnaire and the Welcome to Canvas Network survey. Items from both of these sources can be found in Appendix C.

Age. Age data were collected via two different sources: (1) the Pre-Course Questionnaire; and (2) the “Welcome to Canvas Network” survey. However, age data from the Pre-Course Questionnaire were continuous (i.e., actual age in years), whereas the age data from the Canvas survey were categorical (ordinal age categories). In order to ensure there was an age variable for the maximum number of participants, both sources of data were combined into one ordered categorical variable with the following categories: 13-18, 19-24, 25-34, 35-44, 45-54, 55-64, and “65 and older.” While this resulted in a less finer-grained measure of age, it allowed for a substantial increase in the number of individuals for whom there was an age variable (compared to using only one of the two data sources).

Education. Similar to Age, participants’ Education level was collected by either the Pre-Course Questionnaire or the Welcome to Canvas Network survey. Responses were coded according to the coding scheme in Table 3.3, resulting in an ordered categorical variable.

Table 3.3

Education Variable Coding Scheme

Codes	Education Level	Codes	Education Level
0	Compulsory Education Not complete	5	Some graduate school
1	High School or College Preparatory School	6	Master's Degree (or equivalent)
2	Some college, but have not finished a degree	7	Ph.D., J.D., or M.D. (or equivalent)
3	Completed 2-year college degree	Missing	None of These
4	Completed 4-year college degree		

Past experience in music. Initial plans for measuring this variable included using three sub-scales from the Music Use (MUSE) Questionnaire (Chin & Rickard, 2012): the Index of Music Listening (IML), the Index of Music Instrument Playing (IMIP), and the Index of Music Training (IMT). However, after further consideration, I decided to develop a more comprehensive measure, the Past Experience in Music Inventory (PEMI), that accounted for the many ways of engaging with music that an individual may have over the course of his/her life, any of which may or may not come to bare on their creative music-making experiences in the course. This instrument is discussed in more detail in Chapter Four: Instrument Development. Items for the instrument can be found in the Pre-Course Questionnaire in Appendix C.

Music aptitude. Music aptitude (MA) was measured with the Profile of Music Perception Skills (PROMS) (Law & Zentner, 2012). A demo of the PROMS can be found at the following URL: <http://www.york.ac.uk/res/musicality/>. Although this is a newly developed measure, it offers several advantages over other well-known measures such as Gordon's Advanced Measures of Music Audiation (AMMA) or Music Aptitude Profile (MAP). First and foremost, the PROMS is available to be administered via the internet. At the time of the study, GIA, the company that distributes the AMMA and MAP, did not have any licensing options for administering the AMMA via the internet. So while the AMMA has a long tradition in music education research, it was not a feasible option for this study, if for nothing other than logistical reasons. In addition, the PROMS has been developed to current validity and reliability standards. The researchers who developed the PROMS to measure musical skills with more specificity, thus avoiding the conflation of multiple skills in single items, something that is particularly problematic if individual items measure different combinations of multiple skills. The PROMS is comprised of nine sub-tests: melody, standard rhythm, rhythm-to-melody,

accent, tempo, pitch, timbre, tuning, and loudness. Factor analysis of the items revealed two general factors, which are labeled as “sequential processing” (melody, rhythm-to-melody, accent, and standard rhythm) and “sensory processing” (loudness, tuning, tempo, pitch, and timbre) (p. 11).

Given the practical demands of the number of measures in this study, it was decided to use the Brief PROMS, which consists of two “sensory” subtests (tuning and tempo) and two “sequential” subtests (melody and accent) and takes about 20 minutes (compared to 60 minutes for the full battery) to complete. Law and Zentner (2012) reported Cronbach’s alpha estimates of reliability of .84 and test-retest reliability estimates of ICC = .82, Pearson’s $r = .84$, and Spearman’s $\rho = .84$ for the Brief PROMS.

Personality. The final global level variable to be measured was personality. In order to limit the time demands on participants and to incorporate the possibility of a sixth personality construct, the Mini-IPIP6 (Milojev, Osborne, Greaves, Barlow, & Sibley, 2013) was used to measure personality. The Mini-IPIP6 includes four additional items (24 in total) compared to the original 20-item Mini-IPIP and has demonstrated excellent stability over time (Milojev et al., 2013). Although the Mini-IPIP6 is a very new measure, there has been research on the overlapping 20 items from the Mini-IPIP. While Donnellan et al. found the Mini-IPIP to be less reliable ($\alpha = .65-.82$ for each scale) than a much longer 50-item IPIP, the Mini-IPIP was found to have acceptable convergent and criterion-related validity. One advantage of the Mini-IPIP is a reduction in intercorrelations between factors. Ideally, Big Five personality factors should be uncorrelated, indicating the distinctiveness of each factor. The average absolute scale intercorrelation for the Mini-IPIP was .19, compared to .25 for the 50-item IPIP.

Another important consideration for this research is the notion of measurement invariance. Since variables such as gender and age were included as control variables, it is important that the scale does not function differentially between groups (e.g., males/females). In one study the Mini-IPIP exhibited “strict invariance” (i.e., factor loadings, intercepts, and residual variances are identical between groups) across age, gender, and even across two independent samples (Laverdière, Morin, & St-Hilaire, 2013). The study also found significant differences in latent factor means across gender and age, which confirms previous research that Big Five traits do indeed vary across age (Donnellan & Lucas, 2008) and gender (Costa, Terracciano, & McCrae, 2001), and also supports the construct validity of the measure. Another study reported evidence for either partial metric or full metric invariance across gender and race (between Blacks and Whites) for each of the five factor scales (Baldasaro, Shanahan, & Bauer, 2013).

One final consideration regarding the choice of personality scale is the format of the scale. Scales differ in the extent to which they use descriptive adjectives as opposed to statements about behavior. For example, Saucier’s (2002) Mini-Markers is a simple listing of single-word adjectives, which the participant rates in terms of accuracy for describing one’s self. The Big Five Inventory utilizes a sentence format that combines items that are adjective-based phrases (e.g., “I am someone who is reserved”) with items that are more descriptive of behavior (e.g., “I am someone who starts quarrels with others”). The Mini-IPIP6 includes items that are almost entirely descriptive of behavior with limited use of adjectives (e.g., “I sympathize with others’ feelings” as opposed to an adjective-based approach like “I am sympathetic”). Since the participants in this study could be from almost anywhere in the world, I felt that a format that relied heavily on adjectives could potentially bias the instrument, depending on students’

familiarity with the adjective and culture-specific understanding of it. Instead, a format that utilizes descriptions of behavior is potentially less susceptible to this bias because the items represent more concrete behaviors than abstract adjectives.

The mini-IPIP6 uses a 5-point response scale (*Very Inaccurate, Moderately Inaccurate, Neither Inaccurate nor Accurate, Moderately Accurate, Very Accurate*) for the following prompt: “Please use the rating scale below each phrase to describe how accurately each statement describes you.”

Contextual support. The extent to which the course environment supports autonomous motivation was measured using Schroff and Vogel’s (2009) Intrinsic Motivation Survey, with minor revisions to items to more accurately match the study context (e.g., “I felt I was competent in my performance in the online discussions” was changed to “I feel I am competent in my performance in the course activities”). Schroff and Vogel’s measure was developed specifically within the Self-Determination Theory framework, but also incorporates related research on curiosity. The instrument measures six constructs (Perceived Competence, Perceived Challenge, Feedback, Perceived Choice, Perceived Interest, and Perceived Curiosity), with four items per construct. The researchers suggest a higher order factor structure such that several first order constructs can be seen as measures of higher order constructs. Those higher order constructs are the 3 basic psychological human needs posited by SDT. The modified items, grouped by hypothesized factor, can be found in Appendix C.

Ultimately, I chose the instrument because it offered a large degree of flexibility in specifying a measurement model for Contextual Support. With 24 items and 6 hypothesized factors, I would be free to choose between higher-order measurement models or first-order measurement models, but with several first-order constructs from which to choose. I felt it was

important to have some flexibility with this variable because, aside from Musical Creative Self-Efficacy, it was the only other contextual-level variable. Therefore, it would be important to have sufficient information to create an appropriate measurement model at the contextual level.

Musical creative self-efficacy. While there are measures of music performance self-efficacy (e.g., Martin, 2012; Schmidt, 2007; Zelenak, 2011) and measures of creative self-efficacy (Abbott, 2010; Karwowski, 2011; P. Tierney & Farmer, 2002), I was unable to locate a previously developed measure of Musical Creative Self-Efficacy (MCSE). Since self-efficacy scales “must be tailored to activity domains and assess the multi-faceted ways in which beliefs operate within the selected activity domain” (Bandura, 2006, p. 310), a self-efficacy measure was constructed specifically for the construct of MCSE. Development of the MCSE scale is discussed in further detail in Chapter Four: Instrument Development. The full set of 24 items can be found in Appendix C. After factor analysis, the final solution included three factors, with items 1-3 measuring the General MCSE factor, items 7-12 measuring the Component MCSE factor, and items 13, 17, 18, and 19 measuring the Improvise MCSE factor. An individual’s summary score for each factor was calculated by taking the mean of all items on the factor. Only one individual (of those that completed the measure) did not complete every single item, and that individual had only one item missing. As such, the summary score for the Component MCSE factor was the mean of items 8-12 (5 items, instead of 6 items) for that individual.

Situational motivation. Situational motivation was measured using a modified version of the Situational Motivation Scale (SIMS) (Guay et al., 2000). Since the measure was administered retrospectively (after completing the song-writing project), items were initially modified to reflect having completed the activity (e.g., changed to past tense). In addition, rather than referring to the generic term, “activity,” items were initially modified to refer to the specific

task, “the songwriting project.” However, when I realized that very few participants were going to complete the songwriting project, I revised the items to reflect motivation related to the projects in general. For example, in response to the main question stem “Why did you engage in any of the discussions/projects?” I modified the item “There may be good reasons to do this activity, but personally I don't see any” to “There may have been good reasons to do the projects/discussions, but personally I didn't see any.”

Situational cognitive engagement. There are several scales that have been developed to measure engagement (Appleton, Christenson, Kim, & Reschly, 2006; Greene & Miller, 1996), but most of these measures focus on general engagement over an entire course or general context, and not specific to a particular situation or task. Recognizing the need for an instrument to measure cognitive engagement in a particular task or situation, Rotgans and Schmidt (2011) developed the a short four-item self-report instrument to measure situational cognitive engagement. The instrument was cross-validated and showed very good fit ($\chi^2(df \text{ not reported}) = .02, p = .94, CFI = 1.00, RMSEA = .00$) with a single-factor model, and reliabilities of .93 and .78 (estimated with Hancock's H) for the exploratory and cross-validation studies, respectively (Rotgans & Schmidt, 2011). The four items from Rotgans and Schmidt used to measure situational cognitive engagement are displayed below. The five-point response scale ranged from 1 (not true at all for me) to 5 (very true for me).

Situational Cognitive Engagement

I was engaged with the topic at hand

I put in a lot of effort

I wish we could still continue with the work for a while

I was so involved that I forgot everything around me

Learning outcomes. This study considered two learning outcomes. At the contextual level (the course overall), the outcome was perceived learning. Initial research plans included a measure of conceptual understanding (which could have been administered at the beginning of the course and at the end of the course) as a contextual level learning outcome. However, development of an appropriate measure, particularly given the need to not use western standard music notation, was not feasible in the time frame available prior to the course beginning. At the situational level (specific to the task of writing a song), the intended learning outcome of interest was Self-Evaluation of Creative Product. Because of lack of participation in the final project for which this variable was intended to be a measure (the write a song project), this variable was not measured. It is discussed below because it still remains a viable avenue to pursue with future research.

Self-evaluation of creative product. The term self-evaluation is used intentionally, as opposed to self-assessment, because it represents a judgment of value of the product, beyond a descriptive assessment. Another name for this variable could be self-perceived value of created product. This variable is treated as a situational variable because it is specific to a particular task (i.e., writing a song). Ideally, Self-Evaluation of Creative Product (SECP) would be modelled as a formative latent variable with four formative indicators. There are issues with treating SECP as a latent variable in this model because of its function as a dependent (endogenous) variable. I will return to this issue shortly. The indicators are usefulness, novelty, aesthetic appeal, and craftsmanship, each measured using a 7-point scale. Because each indicator is seen as a different facet of the construct, and not interchangeable items (as is assumed with the typical specification of the much more common reflective latent variable model), a formative model is necessary.

Aesthetic appeal and craftsmanship are often used as indicators of discriminant validity in some measures of product creativity (Amabile, 1983, 1996), which means they should correlate only modestly with a measure of creativity. Their inclusion as discriminant validity variables suggests they are a component of evaluations of a creative product, and as such, should be included as indicators of value. The indicators novelty and usefulness are the two components of the most commonly agreed upon definition of creativity (Runco & Jaeger, 2012). As such, they each represent different facets of an evaluation of a creative product, according to the literature.

There are potentially other facets that could be included as well. Researchers that investigate functional creativity (especially in engineering and design fields) have developed scales that are much more multi-faceted than the four facets already mentioned. For example, the Creative Solution Diagnosis Scale (Cropley & Cropley, 2005, 2008) includes four primary criteria (relevance/effectiveness, novelty, elegance, and genesis) and several indicators of each (30 in all). Some indicators of elegance include “Pleasingness (beholder finds the solution neat, well-done), Completeness (the solution is well worked out and ‘rounded’), gracefulness (the solution well-proportioned, nicely formed), and harmoniousness (the elements of the solution fit together in a consistent way)” (Cropley & Kaufman, 2012, p. 125), to name a few. The inclusion of additional facets will need to be considered in further research.

To return to the issue of treating SECP as a formative latent variable, there has been much recent criticism of the improper use of endogenous formative latent variables (Cadogan & Lee, 2013; Hardin, Chang, Fuller, & Torkzadeh, 2011; N. Lee & Cadogan, 2013; Rigdon, 2014). Cadogan and Lee (2013) have shown the conceptual, logical, and statistical issues (and impossibilities) related to an endogenous formative latent variable. In particular, because a formative latent variable is defined by its indicators, any antecedents to a formative latent

variable cannot produce effects on the latent variable directly. Instead, they must affect the latent variable by way of an effect on the formative indicators of the latent variable. This means that all variables modeled as having an effect on SECP should actually be modeled to impact the indicator variables (usefulness, novelty, aesthetic appeal, and craftsmanship) instead of directly impacting the latent variable. The difference between these two different ways of specifying the model are illustrated in Figure 3.2. What is apparent from Figure 3.2 is that Panel A requires four paths to be estimated from each antecedent variable while Panel B only requires one path to be estimated from each antecedent variable. The more parameters (paths) included in the model, the more likely there will be problems with model identification or problems with convergence upon estimation. An alternative solution is to treat SECP as a formative composite variable, which is a simple weighted linear combination of the four indicator variables. Without going further into these issues, especially since SECP was not even measure in this study, the point here is that more work needs to be done on both the theoretical notion of SECP as an indicator of learning, and on the measurement issues.

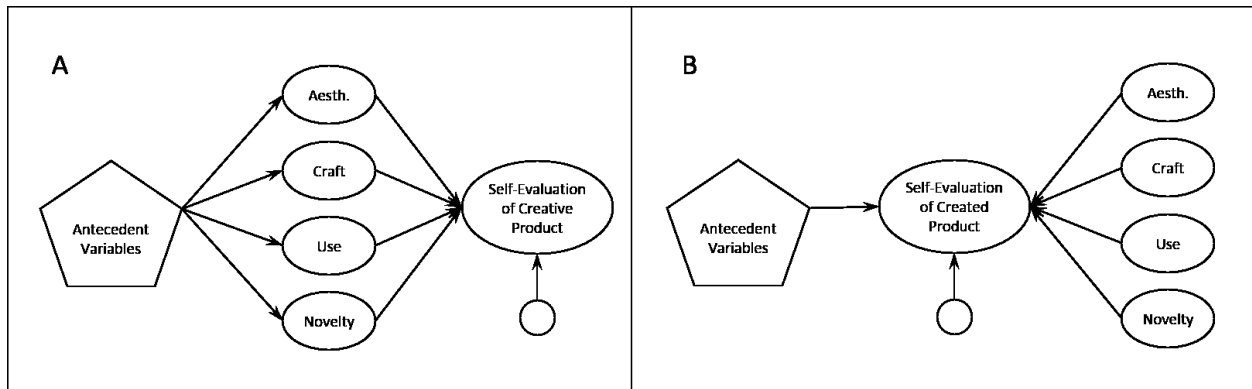


Figure 3.2. Self-Evaluation of Creative Product (SECP) measurement models. Both measurement models of SECP are formative latent variable models with four cause indicator observed variables. Panel A shows antecedent variables making their effect by way of the SECP indicator variables. Panel B shows antecedent variables making their effect directly on the SECP latent variable.

Perceived learning. Perceived Learning (PL) was initially planned to be measured using a slightly modified version of the CAP Perceived Learning Scale (CAP PLS) (Rovai et al., 2009). The CAP PLS is a 9-item self-report scale that measures both overall perceived learning and three sub-scales, cognitive, affective, and psychomotor. However, after further consideration, I decided that the actual items from the CAP PLS could not all be modified in a way that made sense with the course context and course activities. Instead, I created a measure based on Rovai and colleagues' basic premise of measuring different aspects of perceived learning. I wrote 15 items based on the CAP PLS, Caspi and Blau's (2011) work and Lo's (2010) work. Five items were intended to measure a general perceived learning factor (e.g., "I know more things"). An additional 10 items were included to measure the cognitive, affective, and psychomotor components of perceived learning. All 15 items can be found in Appendix C. Participants who completed the Perceived Learning measure responded using a seven-point scale regarding the degree to which they agreed or disagreed with the statements about their learning. Only items from the general perceived learning factor were actually used in the final data analyses. A summary score for Perceived Learning (General) was calculated by taking the mean of all five items.

Additional Measures

In addition to the variables initially planned to be included in this study, several additional variables were added after I realized that the sample size would not be sufficient and it was clear that the majority of students were not participating in the creative projects. In particular, the new dependent variable of interest was Project Participation. In addition, several variables from the Welcome to Canvas Network survey were added to the revised design.

Project participation. To determine which students actually participated in each of the creative projects, I systematically reviewed the discussion pages for each project and assigned a value of 1 for every individual that made a submission for that project. While this measure does not provide any indication regarding students who did work on the projects but never made an actual submission in Canvas, it is a perfectly reliable measure of actual observable participation in the projects. From these data, I created two project participation variables. First, anyone that participated in at least one project was assigned the value 1 for the variable Any Project Participation. Second, anyone who participated in a project beyond the first project (Top 10 playlist) was assigned the value 1 for the variable Project Participation Beyond Top 10. This resulted in two dichotomous variables that capture two different aspects of project participation. Any Project Participation is a variable that identifies those individuals that were willing to at least try one of the projects. Project Participation Beyond Top 10 is a variable that identifies individuals who, after completing the Top 10 Playlist project, continued to engage in the projects. This means that they participated in one of the projects that actually required one to make music, or at least a component of music (e.g., lyrics, melody, or a song).

Welcome to Canvas Network survey. At the beginning of the course, students were asked to complete the Welcome to Canvas Network Survey. This survey was created by Canvas Network and is included in every MOOC they offer. In addition to basic demographic questions (age, gender, education, geographic region, primary language [English or not English]), the survey also included several questions regarding the students' reasons for taking the course, expectations regarding participation in the course, and previous MOOC experience.

Primary reason. This item included nine response options for answering the question "What is your primary reason for taking an open online course?" Being a nominal variable,

eight dummy variables were created for use in regression analyses. Response options are shown below.

- I like the format (online)
- I enjoy learning about topics that interest me
- I enjoy being part of a community of learners
- I hope to gain skills for a new career
- I hope to gain skills for a promotion at work
- I am preparing to go back to school
- I am preparing for college for the first time
- I am curious about MOOCs
- I want to try Canvas Network

Type of learner. This question is based on research on MOOCs that has identified four different types of learners in MOOCs: the observer, the drop-in, the passive participant, and the active participant (Hill, 2013). The four categories were dummy coded into three dummy variables for use in regression analyses. The available responses to the question, “What type of learner are you?” are shown below.

- An observer. I just want to check the course out. Count on me to “surf” the content, discussions, and videos but don’t count on me to take any form of assessment.
- A drop-in. I am looking to learn more about a specific topic within the course. Once I find it and learn it I will consider myself done with the course.
- A passive participant. I plan on completing the course but on my own schedule and without having to engage with other students or assignments.
- An active participant. Bring it on. If it’s in the course, I plan on doing it.

Expected hours per week. This item asked the question, “How many hours a week are you planning to spend on this course?” The six possible responses and the coding are shown in Table 3.4. While the response categories do not represent truly equal time intervals, it was believed the coding scheme provided a close enough approximation to warrant its treatment as a continuous variable.

Table 3.4

Expected Hours per Week Response Options and Coding

<u>Response</u>	<u>Code</u>
Less than 1 hour	1
Between 1 and 2 hours	2
Between 2 and 4 hours	3
Between 4 and 6 hours	4
Between 6 and 8 hours	5
More than 8 hours per week	6

English. This item asked “Is English your primary spoken language?” Answers were either “yes” or “no.” “Yes” was coded as 1, and “no” coded as 0.

Geographic region. This item asked “where do you live?” Below the question was a map of the world, and beneath the map were the following options: North America, Central America, South America, Caribbean, West Europe, East Europe or Former Soviet Union, Africa, Middle East, South Asia, East Asia, Southeast Asia, Australia & South Pacific.

Plan A: Proposed Data Analysis

The following two sections (in smaller font) were included in my dissertation proposal. I have chosen to include these sections (as opposed to removing them) because they outline the

process with which I had intended to analyze the data, had I been able to collect all of the necessary data. Normally these sections would have been moved to my Method chapter, reworded to present tense, and expanded to describe what actually occurred. I have chosen to leave them “as is” because this analysis never took place. I view it as an integral part of the overall dissertation because it is helpful in terms of illustrating how I had initially conceived the study, and how I intended to answer the initial primary research questions. If the reader wishes to skip this section, you may continue on to the description of how I analyzed the data that I was actually able to collect, which begins in the section entitled “Plan C: Actual Data Analysis for Revised Research Questions”.

Data screening. Data analysis will begin with screening the data, which includes examining data for meeting the assumptions of statistical analyses (e.g., normality), searching for and deciding how to handle missing data, and identifying potential outliers in the data.

Identification. After data screening, each model must be examined to determine if it is properly identified, that is, it is mathematically possible to estimate the parameters in the model. As a simple example, if there are five observed variables, then there are 15 possible covariances (and variances) between these variables, each representing a piece of information that is used to estimate model parameters. If there are more than 15 parameters (path coefficients, error variances, latent variable variances and covariances, etc.) to be estimated in a given model, then there are not enough pieces of information to estimate the parameters, so the model is said to be underidentified. There are several other issues to be addressed with model identification, which for the purposes of this proposal, I will save for later. Both components of a structural model (measurement model and path model) must be identified.

Model estimation. Assuming each model is identified, the next step is to estimate the parameters of each model. Syntax will be written for each model so the model parameters can be estimated using the Mplus software. The specific estimation method to be used will be dependent on the characteristics of the data. For example, although maximum likelihood (ML) is the most common method, if the data are severely non-normal, then other estimation methods like Weighted Least Squares (WLS) will need to be considered. It is also possible that some of the variables should be treated as categorical instead of continuous (given the reliance on Likert scale items), which would require other estimation methods.

Model estimation will include two steps. The first step is an estimation of only the measurement component of the model. In this step all observed variables are specified to the latent construct they are intended to measure and then all latent variables are allowed to co-vary. No paths are specified between latent variables. This is similar to a typical CFA approach. This step allows the researcher to determine whether the measurement model sufficiently fits the data such that the direct effects are worth considering. If the measurement model does not adequately fit the data, then it will be necessary to consider re-specifying the model.

Once there is evidence for a satisfactory measurement model, then the path model can be added by specifying the paths for all direct effects and limiting only the exogenous latent variables to covary. This two-step process is only necessary for the first structural model. Once the measurement model has been deemed satisfactory, it remains the same for all other models to be estimated because the only differences between the competing structural models are in path component of the model. In other words, the measurement model stays the same for all structural models.

Model comparison. To answer the first research question, assuming there are no issues with convergence (the software is unable to produce a solution for any number of reasons), the next step will be to compare the models to determine which model best fits the data. For nested models (which there will be some), this is accomplished by a Chi-Square difference test. For non-nested models, alternative fit indices that can be compared are the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), or sample-size adjusted Bayesian Information Criterion (ssBIC). There is also potential that two or more models will fit the data well, but not exhibit any statistically significant differences in terms of model fit. In this situation, the more parsimonious model is preferred.

Equivalent models. After deciding which model to retain as the best-fitting model, it is also necessary to consider equivalent models. Equivalent models are mathematically identical models to the researcher-specified model, but represent alternative specifications of relationships between variables. Since these alternative models are mathematically identical (and therefore have identical fit), it is necessary to determine which model makes the most sense according to theory.

Model interpretation. Once an appropriate model has been retained, the specific parameter estimates can be examined and interpreted. The fit indices provide an indication of how well the overall model fits the models, but it is also important to understand how much of the variance in each component of the model is explained by other components that are believed to predict (or impact) those components. The R^2 value for each latent factor can be calculated by subtracting the standardized variance of the disturbance term from one. Since the disturbance term represents all other factors not incorporated in the model that contribute to variance of a specific latent factor, then one minus the variance of the disturbance represents the proportion of variance that is explained by factors that are included in the model (and predict the specific latent factor). Therefore, the R^2 values for each endogenous latent variable will be examined to answer the second research question (How much variance is explained by each component of the model?).

The third research question seeks to determine the relative importance of each component of the model in terms of predicting learning outcomes. The individual path coefficients (interpreted as regression coefficients) will be examined to answer this question. Each learning outcome has a direct effect from individual components in the model (e.g., direct effect of music aptitude on perceived learning) and a series of indirect effects (e.g., the indirect effect of music aptitude on perceived learning via contextual support). The direct effects and the indirect effects can be added together giving a total effect on each learning outcome for each individual component in the model.

In addition to the effects of components on the learning outcomes, it will also be important to consider the intermediary effects (e.g., the total effects of music aptitude on cognitive engagement). If, for example, situational cognitive engagement is a strong predictor of learning outcomes, then it is also important to understand the relative importance of the components of the model that impact cognitive engagement.

Model modification. This portion of the analysis could potentially occur at one of two points. If none of the proposed models adequately fit the data, then the interpretations of the model discussed above will not be all that meaningful. In this situation, it would be necessary to consider modifications of the model based on the empirical data, which is accomplished by examining the modification indices. Modification indices provide an estimate of the improvement in the Chi-Square model fit statistic if a specific parameter is altered. For example, the path from Situational Cognitive Engagement to Situational Motivation is currently constrained to be zero (depicted by the absence of an arrow pointing from Situational Cognitive Engagement to Situational Motivation). The modification indices might suggest that this path be relaxed and freely estimated, which would improve the overall model fit. However, the researcher runs the risk of capitalizing on chance characteristics of the data (resulting from this specific participant sample) if such suggestions are followed without considering the logical and theoretical implications of such changes. Nevertheless, such changes are worth considering if a plausible explanation can be made or if previous research and theory might support such a change. I expect the potential for uncovering potentially useful modifications to the models is quite high given the complexity of the model. Because there are so many variables in the model, the number of parameters that could potentially be modified is quite large.

On the other hand, if one or more of the models adequately fits the data, then it is not necessary to consider model modifications upfront. Instead, the retained model can be interpreted as proposed above. But after these interpretations have been made, it is still potentially useful, in terms of theory development, to consider modifications to the model based on the empirical data.

Plan B: Proposed Data Analysis

If it is not already apparent, it is worth saying that there are a great number of moving parts to this study, all of which need to come together in order for any meaningful and trustworthy conclusions to be drawn from the data. I recognize that it is quite plausible that I will not achieve a sample size large enough to analyze such a complex model. It is also very possible that

there will be large amounts of missing data, especially given the number of constructs I am attempting to measure. It is also possible that I will have large numbers of participants at the beginning and smaller and smaller numbers towards the end of the course. Without going into too much detail, I have conceived of several “Plan B” options so that I can still conduct a meaningful research study in the event that things do not go as planned. Generally speaking, each of the sub-models I discussed in Chapter Two represent potential smaller scale studies in and of themselves.

In the worse-case scenario, if I have nowhere near enough data to conduct an analysis using structural equation modeling, it will still be possible to utilize a traditional multiple regression approach, but latent variables will need to be converted to composite scores. This is not ideal because measurement error is not taken into account with such an approach. Nevertheless, relationships between the various components (for which I do have usable data) can still be examined. An evaluation of overall model fit cannot be accomplished in a traditional regression-based approach, but the relative importance of individual components could still be examined.

If prior to beginning the course, it is clear that enrollment numbers are much lower than will be necessary, it might also be possible to transition to a more longitudinal, within-subjects approach. This would involve focusing on a much smaller number of variables (e.g., musical creative self-efficacy), but investigating how they change over time as a function of aspects of the course (e.g., type of creative activity, engagement, etc.). The point here is that I have considered the possibility that things will not go entirely as planned and the design is flexible enough that I have several options depending on how things play out.

Plan C: Actual Data Analysis for Revised Research Questions

The following section provides an overview of the analyses that were actually carried out on the data that was actually collected. This includes analyses of the pilot study data that were conducted to develop several research instruments, as well as analyses of the primary study data. Specifics for each analysis can be found in chapters four and five. Prior to conducting all analyses, the data were screened for outliers and anomalies (e.g., miscoded scores) and were examined to determine the extent to which relevant assumptions were met for each particular analysis.

Instrument development data analysis overview. A somewhat novel approach (at least for the study of PEM) was used for analysis of the pilot study and primary study data from the

Past Experience in Music Inventory (PEMI). Rather than using traditional factor analysis, which assumes the underlying factor is continuous, a latent class analysis (LCA) was conducted. The purpose of LCA is similar to cluster analysis or discriminant function analysis (DFA), but with some notable differences. In all three analyses, the general goal is to use a set of variables to separate individuals into different groups. With DFA, the groups are known in advance, so the goal is to identify which specific variables are best at predicting group membership. With cluster analysis, the goal is to determine whether individuals tend to cluster or group together on a pre-determined set of variables. LCA is more similar to cluster analysis, although LCA operates within a latent variable framework and decisions regarding the number of groups/clusters/classes are based on statistical evidence.

In the case of PEM, it was assumed that individuals may vary in overall level of PEM (i.e., more or less experience), but individuals may also have qualitative differences in PEM. For example, some individuals may have a large amount of experience in composing, improvising, and performing in small groups and little experience with private lessons, large ensembles, or professional experience, while other individuals may be the exact opposite. This would represent two different *types* of experience, which would be represented by different classes in a LCA. As such, LCA was used, as an exploratory technique, to determine whether there were indeed different types of PEM.

As for MCSE, a more traditional approach using a series of both exploratory and confirmatory factor analyses was used to determine the underlying factor structure of the responses to the Musical Creative Self-Efficacy Scale (MCSES) for both the pilot study and primary study data. Contrary to the LCA used with the PEMI, the factor analyses of the MCSES scores assumed the underlying factor(s) to be continuous. Methods for determining the number

of factors to extract in exploratory factor analyses were Kaiser's criterion (eigenvalues > 1), parallel analysis (O'Connor, 2000), the Minimum Average Partial (MAP) test (Velicer, Eaton, & Fava, 2000), and a comparison of model fit statistics (e.g., RMSEA, CFI, SRMR). The analysis also resulted in the removal of several items from the MCSES for several reasons, including low factor loadings and cross loadings, each of which had a corresponding possible explanation (e.g., confounding of instrumental performance experience with self-efficacy).

To examine convergent validity in both the pilot study and primary study data, correlations between PEM and MCSE were calculated. Following self-efficacy theory, relevant past experience, especially mastery experiences, should improve one's self-efficacy, which suggests PEM and MCSE should be correlated. In addition to examining correlations, which utilized *observed* variables, structural equation modeling (namely in the form of a MIMIC model) was utilized to explore the relationship between PEM and MCSE in a *latent* variable framework. As such, measurement error could be taken into account when examining this relationship.

Primary study data analysis overview. The first step in analyzing the data from the primary study was to examine descriptive statistics for central tendency (e.g., mean, median) and distribution (SD, Skewness, kurtosis), as well as visual plots (e.g., histograms) for each variable. This helped provide a clearer picture of the sample, particularly in regards to demographic variables.

To answer the first question (Which variables are important in predicting Project Participation?), a series of logistic regression models were explored. Logistic regression was used because the dependent variable was a binary outcome (did participate/did not participate). In building a logistic regression model, I followed the "purposeful selection" process described

by Hosmer, Lemeshow, and Sturdivant (2013), which involves first examining the univariable logistic regression model for each potential predictor variable (akin to simple linear regression), and then following an iterative procedure for adding and removing variables from the full logistic regression model (akin to multiple linear regression). After arriving at the final model, Nagelkerke R^2 , which is really a pseudo- R^2 , was calculated to provide some indication of the predictive power of the model, thus answering the second research question (How well do those variables predict Project Participation?).

The third research question (Are there group differences in any of the other characteristics for the different Type of Learners?), which arose from the answers to the first two research questions, was examined using the appropriate test for group differences according to the level of measurement of each variable. For example, for continuous variables (i.e., interval or ratio level of measurement), the independent sample t-test was used, but for categorical variables (nominal or ordinal), the chi-square test of independence was used. After identifying variables with statistically significant group differences, a point biserial correlation was calculated as an effect size measure for continuous variables. Kendall's tau correlations were calculated as an effect size for ordinal variables.

Finally, answering the fourth research question (What are the noteworthy relationships in the data that warrant further investigation?), I calculated the Kendall's tau correlation for all pairwise combinations of non-nominal variables. I chose to use Kendall's tau, as opposed to the more common and familiar Pearson correlation, for several reasons. First, there were over 30 variables that I intended to correlate, and they varied in terms of scale of measurement from ordinal to interval to true ratio scale. The lowest common denominator between them all was to assume that all variables represented at least an ordinal level of measurement, and as such, a rank

order correlation would be appropriate. After assuming ordinal data, I chose Kendall's tau instead of Spearman's rho for the following reasons: (1) Kendall's tau has a much more intuitive interpretation than Spearman's rho (the proportion of concordant to discordant pairs); (2) Kendall's tau will generally be smaller than Spearman's rho; (3) Kendall's tau is generally more accurate with small sample sizes and a better estimate of the population parameter; (4) Kendall's tau is much less sensitive to large, but rare discrepancies in the rank order of two variables. Taken together, these four characteristics make Kendall's tau a much more conservative measure of association, which I felt was important, given the small sample size for many of the correlations and the large number of variables being considered.

The present chapter has described the initial design of the study, the revisions that were made in response to the realities of the data that was collected, and an overview of how those data were analyzed. The following chapter describes the development of two research instruments used in this study: the Past Experience in Music Inventory and the Musical Creative Self-Efficacy Scale. Two pilot studies were conducted as part of the development process. The development process is described in detail.

CHAPTER FOUR: INSTRUMENT DEVELOPMENT

It was necessary to develop several instruments for this research, including measures of Past Experience in Music, Musical Creative Self-Efficacy, and Perceived Learning. The first two underwent extensive development, although Perceived Learning remains fairly underdeveloped due to limitations on acquiring a large enough sample to pilot test and examine the psychometric properties of the measure. The reader is referred to my discussion of this measure in Chapter Three for more information. In this chapter, I describe the instrument development process for the Past Experience in Music Inventory (PEMI, pronounced pe-mee) and the Musical Creative Self-Efficacy Scale (MCSES, pronounced mik-ses). Although I did not initially envision this process as being a substantial component to this dissertation, I believe the results of my efforts in this area represent a significant contribution to the research fields of music education and musical creativity, and therefore warranted a chapter dedicated to a more comprehensive presentation of the process and the results.

The two instruments were developed in parallel with each other, and subsequently researched in parallel with each other. From a theoretical standpoint, the two measures should be mutually supportive. Self-efficacy, as a construct, “is concerned with people’s beliefs in their capabilities to produce given attainments” (Bandura, 2006, p. 307), and it can be developed (i.e., increased) via four sources (Bandura, 2012): (a) mastery experiences (success in a domain that results from surmounting obstacles, persevering, and sustaining effort); (b) social modeling (observing individuals similar to one’s self being successful; also known as vicarious

experience); (c) social persuasion (becoming convinced to believe in one's self through persuasion by others); and (d) physical and emotional states (e.g., lowering anxiety, increasing physical strength). It could be argued that the first three sources (mastery experience, social modeling, and social persuasion) are more likely to happen in any domain as one's gains more experience in that domain. More overall experience in a domain should, although perhaps indirectly, result in higher self-efficacy. Therefore, a measure of PEM should be at least moderately related to a measure of MCSE.

Initial Item Pools

The initial pool of items for the MCSES was generated by (a) reviewing existing measures of musical and general self-efficacy; (b) reviewing measures of creative self-efficacy in the creativity literature; (c) consulting recommendations made by Bandura (2006) for constructing self-efficacy scales; and (d) attempting to account for the many ways in which one can be creative with music in the 21st century (e.g., including things like making mashups of songs). The initial pool of items and general framework for the PEMI was taken from some of my own previous work (Stefanic, 2011), but additional sources of input were other measures of PEM in the literature (e.g., Chin & Rickard, 2012). I had two main goals in terms of construct coverage (i.e., representing all facets of the construct). First, I needed to account for the many different ways one may engage with music over the course of a lifetime. Second, I wanted the measure to be applicable across all ages and across all levels of ability and knowledge of music. This meant that items needed to be devoid of jargon and also allow an individual to provide some detail regarding the extent to which he/she has experience within a particular way of engaging with music.

Expert Review

Prior to conducting the first pilot study, I consulted several of my music education doctoral student colleagues to get some initial feedback on the items for both instruments. I then submitted the full questionnaire containing both instruments and items for demographic information, as it would be administered via the online survey platform (SurveyGizmo), for expert review by two music education researchers/professors. Experts were asked to provide feedback regarding construct coverage as well as comments about the clarity, relevance, reading ability, and subject-matter knowledge required of the items.

Sliders (grouped by factor)

14. I am _____% confident that I can...

(Cannot do at all) (Moderately confident I can do) (Highly confident I can do)

0 10 20 30 40 50 60 70 80 90 100

create a simple melody with an instrument.	34
create a simple melody with my voice (with or without lyrics).	68
create lyrics for a song.	5
create a chord progression.	84
create a "beat" or background accompaniment for a song.	50

Sliders (individual, not grouped by factor)

29. I am _____% confident that I can

create a simple melody with my voice (with or without lyrics).

(Cannot do at all) (Moderately certain I can do) (Highly certain I can do)

30. I am _____% confident that I can

create an entire music album or collection of songs.

(Cannot do at all) (Moderately certain I can do) (Highly certain I can do)

Table (Number Entry)

31. I am _____% confident that I can...

	Confidence (0-100)
create new music.	<input type="text"/>
improvise on an instrument.	<input type="text"/>
create new music "on-the-spot" or "in the moment".	<input type="text"/>
create music that would gain international popularity and high praise from music critics.	<input type="text"/>
create a re-mix of another song or songs.	<input type="text"/>
improvise a melody to match a chord progression.	<input type="text"/>
create an entire music album or collection of songs.	<input type="text"/>

Drop-Down Menus

32. I am _____% confident that I can

create an entire music album or collection of songs.

-- Please Select --

33. I am _____% confident that I can

create a simple melody with an instrument.

-- Please Select --

Multiple Choice (only 11 response categories)

41. I am _____% confident that

I can create new music.

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Figure 4.1. Possible Response Formats for MCSE Items

This expert review was administered within the actual online survey. Comment boxes were added for each item so reviewers could give specific feedback for individual items. In addition to reviewing the items, expert reviewers were also asked to review the question format and response format of the items. Several different options were provided as examples for the reviewers to see. Examples of the various response formats that were considered are shown in Figure 4.1. Feedback from the experts was used to re-word items if necessary, add additional items, and ultimately decide on the individual sliders (not grouped by factor) format. After incorporating the comments from all sources of input I consulted regarding items, I administered the questionnaire in Pilot Study #1.

Pilot Study #1

For Pilot Study #1, I solicited responses to the Music Experience Questionnaire (a questionnaire that contained both the PEMI and MCSES) from undergraduate students ($N = 79$) in three online music courses that were offered primarily for non-music majors, although music majors commonly take them. Similar to the expert review survey, participants were encouraged to use the comment boxes comment on the questions, clarify their responses if they felt it necessary, or add general comments at the end of the questionnaire. The primary purpose of this pilot study was to test the administration of the instrument and gain additional feedback regarding the items and content coverage. Because the observations were not independent (students were nested in three different courses) and because the sample size was quite small for factor analysis (given the number of items), these responses were not further analyzed for underlying factor structure. Bivariate correlations were examined to look for any potentially poorly behaving items. This was done on the entire sample and also for the sub-samples by course. No items appeared to perform consistently poorly, although item 4 (“create a simple

melody with an instrument” and item 5 (“create a simple melody with my voice (with or without lyrics”) were somewhat erratic between the sub-samples. These items were flagged for further analysis in the second pilot study.

Pilot Study #2

For the second pilot study, a second sample ($N = 132$) of volunteer undergraduates enrolled in an online course entitled “Survey of Jazz” completed a revised version of the Pilot Study Survey, again using the SurveyGizmo platform. The “Survey of Jazz” course regularly has 1000 or more students enrolled each semester, including students from a variety of different majors (many of which are not music majors). The course instructor posted a brief announcement via the course LMS (Canvas) to solicit participation and briefly explain the purpose of the survey. Student participants accessed the survey via a hyperlink in the announcement.

Revisions to PEM index. The revised version of the survey included the following changes:

- An additional Past Experience in Music question was added (“On average, for how many minutes during any given day would you say that you intentionally listen to music?”). This was done for two reasons. First, it addressed an aspect of musical engagement that was not included, but almost certainly has at least some bearing on the sum total of one’s PEM. Second, it allowed individuals who did not have experience in any other types of musical engagement (e.g., performing, composing, producing) to potentially have a non-zero PEM score. It was believed this could potentially help the positive skew in the PEM total score that was observed in the first pilot study.

- Demographics questions were moved to the end of the survey. This was done so that the most important questions (PEM and MCSE) were completed first, in the event that a participant chose not to complete the entire survey.
- Formatting improvements were made to better allow participants to complete the survey on a mobile device (e.g., cell phone, tablet). Logic was also added to the survey entry page to help the participant determine whether he/she could effectively complete the survey on his/her specific mobile device.
- Performance Experience was expanded to include three questions, one for large ensemble experience (>5 people in ensemble), one for small ensemble experience (2-5 people in ensemble), and one for solo performance experience (e.g., solo pianist).
- An additional question was added at the end of the PEM section asking participants the following: “Are there any other ways that you engage with music that you feel were NOT accounted for with this survey?”
- Instead of displaying subsets of the MCSE items on separate pages, all items were displayed on a single page, but in the same order as the pilot study. All participants saw all items in the same order. This was done to minimize the possibility of an additional factor arising from the separate-page item presentation format.

Calculating the PEM total variable. In order to create a single variable to function as a total measure of one’s Past Experience in Music, I examined two different approaches. The first approach treated PEM as existing along a unidimensional continuum (from no experience, through low experience, through moderate experience, through high experience, etc., but on a continuous scale). However, given data on an individuals’ frequency and recency with which they engage in a particular musical activity (e.g., improvising music), there is no readily apparent

or self-evident way in which to combine these two dimensions. They each exist on different scales (frequency is on a count-type scale that is indexed to time, but recency is on a time-based scale that is indexed to the present moment), so a simple summing of these dimensions does not make intuitive sense.

One solution would be to multiply the two measures (frequency score \times recency score; $F \times R$), which mathematically places the resulting score in a two-dimensional space. This is roughly analogous to the difference between variance and standard deviation. Variance is a squared value, and represents distance in two-dimensional space, whereas standard deviation represents distance along a single dimension. But while variance has some advantageous statistical properties, it is a value that is difficult to interpret because it exists in two-dimensional space, hence the need to transform variance into standard deviation. I attempted to apply a similar logic to the combination of frequency and recency. Multiplying the two scores together creates a nice summary of the two scores that incorporates the difference in scales, but the resulting value is uninterpretable, much like variance. However, if the square root of this value is taken, then the value is moved back into single dimensional space, and now the score is interpretable (and least theoretically interpretable) because it exists somewhere along a unidimensional continuum. The unit of this scale is admittedly a bit ambiguous¹, but it represents a combination of the units from the initial scales for recency and frequency. Using this general approach, the scores from the individual PEM items were transformed and combined as describe below.

¹ We can think about the unit of this scale by looking at what happens to the original units when recency and frequency are combined in this way. This is shown in Appendix D. The resulting unit is the square root of number of times. Again, this is somewhat ambiguous.

“Frequency/recency” questions. Five questions (create, improvise, sing, instrument, and record/produce) had a “recency” component and a “frequency” component. The following steps were taken to transform those components into a single score for each question:

1. The frequency score was multiplied by the recency score, and then the square root was taken of that value (e.g., $PEM_{create} = \sqrt{Create_{frequency} \times Create_{recency}}$). Each of the PEM questions are structured in different ways.
2. The resulting score was then indexed to 1 by dividing it by the square root of 30. The highest possible score for these recency questions was 6 and the highest possible score for these frequency question was 5. Therefore, the highest possible score after the transformation described above is $\sqrt{6 \times 5} = \sqrt{30} \sim 5.477$.

“Number of years” questions. Six questions asked individuals how many years they had participated in a particular type of activity (formal lessons, large ensemble, small ensemble, solo performance, professional, and semi-professional). The number of years was indexed to 1 by dividing it by the maximum score for that question. For example, the maximum score from this sample for the small ensemble participation question was 30 years. All individuals’ scores for this variable were divided by 30. This resulted in a maximum possible score of 1 for each of these questions (with the individual who had 30 years of experience receiving a score of 1).

“Listening” question. The question for listening experience utilized ordinal categories (e.g., “between 1 and 2 hours per day”), which were transformed into an ordinal scale using the coding scheme shown in Table 4.1. Once transformed to this ordinal scale, the scores were indexed to 1 by dividing by 13, thus resulting in a maximum possible score of 1 on this question.

Table 4.1

Listening Experience Coding Scheme

Less than 30 minutes	.5	Between 6 and 7 hours	7
Between 30 and 60 minutes	1	Between 7 and 8 hours	8
Between 1 and 2 hours	2	Between 8 and 9 hours	9
Between 2 and 3 hours	3	Between 9 and 10 hours	10
Between 3 and 4 hours	4	Between 10 and 11 hours	11
Between 4 and 5 hours	5	Between 11 and 12 hours	12
Between 5 and 6 hours	6	More than 12 hours	13

Rationale for indexing to 1. The purpose of the indexing procedures described above was to make each dimension of the PEM scale be worth the same proportion of the PEM Total score. Each dimension had a total possible score of 1. There were five frequency/recency questions, four “number of years” questions, and one listening question, each with a total possible score of 1.

PEM total calculation. The individual scores for each question were then summed to create a PEM Total score, which had a maximum possible value of 12. This PEM Total score was used in subsequent analyses as a single value representing one’s sum musical experience over his/her lifetime, accounting for a wide variety of ways in which one might engage with music. It should be noted that this approach makes several assumptions. First, by indexing each of the PEM variables to 1, the resulting total PEM score places equal weight on each component. So while the approach results in a number that is fairly straightforward in terms of interpretation

between individuals, it might not be the best way to capture the way in which each component likely differentially contributes to one's overall PEM.

Second, because the resulting total PEM score is a simple weighted linear combination of the components (also known as a formative composite variable), it is still potentially littered with measurement error. This is because the composite does not partition the variance into shared and unique/error variance. All of the variance from each component variable is treated as true variance, which is likely not the case. The composite is an observed variable, not a latent variable. The advantage to latent variables is they take measurement error into account when examining relationships between constructs. But treating PEM as a continuous reflective latent variable is probably not appropriate given that the individual items are not really interchangeable. The item for improvisation experience cannot be viewed as exchangeable with any other item. This implies PEM might best be conceived as a formative latent variable. With a formative latent variable, the latent construct is defined by the items, and not measured by them. However, exploratory techniques like exploratory factor analysis cannot be done with formative latent variables. This is because items included in a formative measurement model are included because they are necessary components of the definition of the construct, not because of their correlation with other items that define the construct. But if the latent variable is treated as a categorical variable instead of a continuous variable, there is a way in which PEM items can be conceptualized as belonging to a reflective latent variable, and therefore be submitted to exploratory-type latent variable analyses, which is discussed next.

PEM latent class analysis. Another approach to examining the latent structure of the PEM observed variables is to assume individuals are actually sampled from different sub-populations of Past Experience in Music. Using Latent Class Analysis (LCA), it is possible to

uncover these sub-populations (i.e., groups/classes) by modelling the discrete classes as a latent variable. To compare LCA to CFA, with CFA the latent variable is assumed to be continuous and all observed variables load on one or more factors. The factor score provides a measurement of each individual's score on the latent variable on a continuous scale. However, if the latent variable is assumed to be categorical, then the factor score represents different categories. In essence, an individual's responses to the observed variables are assumed to be the result of him/her belonging to one (and only one) category/class/group. So while CFA is a variable-centered approach, LCA is a person-centered approach because it focuses on categorizing individuals (J. Wang & Wang, 2012).

The procedures I followed for conducting the LCA were largely based on procedures recommended by Wang and Wang (2012) in their chapter on mixture models and based on recommendations made by Bauer and Curran (2004). LCA is a fundamentally exploratory analysis, somewhat similar to EFA. The first question to address with LCA is the number of underlying classes in the latent variable (analogous to choosing the number of factors to extract in EFA). To determine the number of classes, a series of models are fit to the data and a collection of fit statistics and model comparison statistics (in conjunction with theory and interpretability) are examined to choose the model that best fits the data. This process is usually done in two separate steps. First, one set of unconditional models (models with only the dependent variables included) are estimated to determine the number of classes. Second, another set of conditional models are estimated with the inclusion of relevant covariates. The second set of models allow the probability that an individual is assigned to a particular class to be conditioned upon these covariates (i.e., after controlling for the scores on the covariates).

The algorithms used for LCA are well-known to be particularly susceptible to converging on local maxima of the likelihood function, instead of the global maximum (J. Wang & Wang, 2012). To counter this problem, researchers often run the analysis with many different random starting values and verify that the best log likelihood value is replicated more than once. All LCA analyses were run with 1000 random starts for initial stage optimizations and 250 random starts for final stage optimizations, with a maximum number of iterations of 20 for each optimization. Unless otherwise noted, the best log likelihood function was replicated for all reported results. All LCA analyses were run using Mplus 7.

Unconditional models. For the Pilot Study #2 PEM data, a series of unconditional models (models for 2 through 5 classes) were estimated with the 11 PEM dichotomous variables as dependent variables of the class latent variable. That is, for each PEM item, participants first answered whether they had or had not created their own music, improvised, learned to play an instrument, etc. For all 11 PEM items, a “yes” answer to this question was coded as 1, and a “no” answer was coded as 0. Thus, there were 11 dichotomous (binary) variables to be used in the LCA analysis. The dichotomous variables (instead of some combination of frequency and recency responses to these items) were used because I felt the yes/no differentiation provided a cleaner, perhaps more reliable, variable that did not rest upon any assumptions of how to combine recency and frequency responses.

Results from each of the unconditional models are displayed in Table 4.2. For LCA, while AIC, BIC, and sample-size adjusted BIC can be used to identify the best-fitting model², the BIC has been found to be the best indicator (Nylund, Asparouhov, & Muthén, 2007). The

² AIC and BIC values are uninterpretable in isolation. They gain meaning when their values are compared to the AIC and BIC values of other models. The lower AIC or BIC value is the better-fitting model.

unconditional model with the lowest BIC value is the 2-class model. However, a second piece of information to consider are the model comparison tests. The Vuong-Lo-Mendell-Rubin Likelihood Ratio Test (LRT), the Lo-Mendell-Rubin LRT, and the Parametric Bootstrap LRT all test whether the more complex model (i.e., the model with more classes) is statistically significantly different than the model with one fewer number of classes. Of the three tests, Nylund and colleagues (2007) found the Bootstrap LRT to function the best. In regards to the unconditional models, the Bootstrap LRT suggests that moving from a 2-class to a 3-class model is statistically significant, even though the other two LRTs are only marginally significant.

Table 4.2

Unconditional PEM Models for Latent Class Analysis (Pilot Study 2)

# classes	Final Stage LL	AIC	BIC	ssBIC	Entropy	Vuong-Lo- Mendell-Rubin LRT (-2LL)*	Lo-Mendell-Rubin LRT (-2LL)*	Parametric Bootstrap LRT (-2LL)*+
2	-676.238	1398.48	1464.07	1391.33	.863	202.25, $p = .0002$	198.83, $p = .0002$	202.25, $p < .0001$
3	-651.123	1372.25	1472.07	1361.38	.834	50.23, $p = .062$	49.38, $p = .064$	50.23, $p < .0001$
4	-636.103	1366.21	1500.25	1351.61	.838	30.04, $p = .144$	29.53, $p = .150$	30.04, $p = .0100$
5	-623.973	1365.95	1534.22	1347.63	.846	24.26, $p = .311$	23.85, $p = .318$	24.26, $p = .0950$

Note. $N = 128$. All model results are based on 1000 random sets of starting values for initial stage optimization and 250 random sets of starting values for final stage optimization. The best likelihood value was replicated in all models.

* For model comparison tests, the baseline model is the model with one fewer classes

+ Based on 200 bootstrap draws

Another piece of information to consider is the entropy measure³, which provides an indication of the quality with which the model classifies individuals into classes. Mplus entropy values range between 0 and 1 with values greater than .8 considered to be “high” (Clark, 2010). All four models have entropy values greater than .83, with only small variations between models, indicating that all models function quite well in terms of classifying individuals.

The last piece of information to consider when determining the number of classes is the general interpretability and theoretical support. For example, if a 3-class model is marginally better than a 2-class model, but the 2-class model allows for a much cleaner interpretation of the classes, then one might consider choosing the 2-class model. For my data, the fit statistics and the model comparison tests were not in complete agreement with each other, so I examined the profile plots (see Appendix E). In a profile plot, each line color represents a different class. Normally, the lines for each class tend to cross over each other, indicating the groups of items to which each class tended to answer “yes”. However, Wang and Wang (2012) have said that parallel (non-crossing) lines in between the classes indicates classes of different *degree* as opposed to classes of qualitatively different *type*. For PEM, the 2-class and 3-class profile plots appeared to have essentially parallel profile lines. This suggests that for the 2-class model, the two classes represent *generally high overall PEM* and *generally low overall PEM*. The 3-class model could be interpreted as showing a third class in between, which could easily be labeled as *generally medium overall PEM*. Even the 4-class profile plot showed a similar pattern, with the addition of a fourth class that seemed to be characterized more specifically by individuals who have learned an instrument and taken formal music lessons.

³ Technically, the entropy value reported by Mplus is a relative entropy measure that adjusts for the number of classes and the sample size (see J. Wang & Wang, 2012, p. 294 for a more complete explanation)

After considering all of the information, I chose to retain the 3-class unconditional model for the following reasons: (1) the Bootstrap LRT clearly supported the advantage of a 3-class over the 2-class model; (2) the 3 class model provided a finer degree of classification beyond simple “low” and “high” PEM; (3) the 3-class model was more interpretable and had consistent meaning across classes compared to the 4-class model; (4) the 3-class model still had a high entropy value of .834; (5) the 3-class model still had average within-class classification probabilities above .90 for all three classes (meaning the probability of correct classification was greater than 90% for each of the three classes). After choosing the 3-class model, it was necessary to examine the conditional models, which would control for variables like gender and age.

Conditional models. Similar to the unconditional model analysis, a series of conditional models were also estimated with the inclusion of gender and age as covariates. Since 15 individuals had missing data for gender, age, or both, the sample size for this analysis was $N = 113$. Results from the latent class analyses for the conditional models are displayed in Table 4.3. Similar to the unconditional model analysis, there was clear statistical support for a 2-class model, but the fit statistics and model comparison tests were not all in agreement regarding the 3-class model. After considering all of the information, I decided to retain the 3-class model for the same reasons discussed above. The entropy of this model was .875, and the average latent class probability for most likely latent class membership was 1.00 for latent class 1, .949 for latent class 2, and .923 for latent class 3. Taken together, this indicates that the model does a very good job of separating individuals into PEM latent classes.

Table 4.3

Conditional Models (Age and Gender as Covariates) for Latent Class Analysis (Pilot Study 2)

# classes	Final Stage LL	AIC	BIC	ssBIC	Entropy	Vuong-Lo-Mendell- Rubin LRT (-2LL)*	Lo-Mendell-Rubin LRT (-2LL)*	Parametric Bootstrap LRT (-2LL)*+
2	-615.82	1281.65	1349.83	1270.82	0.904	181.099, $p < .0001$	178.404, $p < .0001$	181.099, $p < .0001$
3	-585.94	1249.87	1356.24	1232.98	0.875	59.775, $p = .101$	58.885, $p = .104$	59.775, $p < .0001$
4	-570.12	1246.23	1390.79	1223.28	0.901	31.637, $p = .211$	31.166, $p = .216$	31.637, $p = .0400$

Note. $N = 113$. Unless otherwise noted, all model results are based on 1000 random sets of starting values for initial stage optimization and 250 random sets of starting values for final stage optimization. The best likelihood value was replicated in all models.

* For model comparison tests, the baseline model is the model with one fewer classes.

+ Based on 200 bootstrap draws.

The best LL value for the 4-class model was not replicated, even after 5000 random starts and 1000 final stage optimizations. Results for the 4-class model may not be trustworthy due to local maxima.

After individuals were assigned to a class based on the highest posterior probabilities for both the unconditional 3-class and the conditional 3-class model, the extent to which the classification remained the same was examined. Only four individuals changed class after controlling for gender and age (one individual moved from class 3 to class 2, and three individuals moved from class 1 to class 2). The unconditional and conditional models were in 96.5% agreement on classification for individuals that had gender and age data (15 of the original 128 were missing gender and age data). Table 4.4 shows the proportions of individuals assigned to each of the three classes for both the unconditional and conditional 3-class models. After controlling for age and gender, class 1 became slightly more exclusive in that a lower percentage of individuals were assigned to that class. The fact that only 8.9% of the individuals were assigned to this class makes sense because this class represents individuals who, for every

single PEM item⁴, were more likely than individuals in the other two classes to have chosen “yes”. Individuals in class 1 are those that are most likely to have a very broad range of experience, including professional and semi-professional experience. The specific probabilities of answering “yes,” given membership in each of the classes, for all PEM items in the conditional 3-class model can be found in the profile plot Appendix E. Given the results of the LCA, I chose to name class 1 *Highly Experienced*, class 2 *Moderately Experienced*, and class 3 *Minimally Experienced*.

Table 4.4

Classification of Individuals Based on Their Most Likely PEM Class Membership

Unconditional 3-Class Model (No Covariates)

Latent Class	Number of Individuals Assigned to Class	Proportion of Individuals Assigned to Class
1 ("Highly Experienced")	14	.109
2 ("Moderately Experienced")	56	.438
3 ("Minimally Experienced")	58	.453

Note. $N = 128$

Conditional 3-Class Model (Age and Gender as Covariates)

Latent Class	Number of Individuals Assigned to Class	Proportion of Individuals Assigned to Class
1 ("Highly Experienced")	10	.089
2 ("Moderately Experienced")	56	.496
3 ("Minimally Experienced")	47	.416

Note. $N = 113$

⁴ Actually, the probability to choose “yes” to the singing experience item for class 1 was slightly lower than class 2 (.673 for class 1 compared to .716 for class 2). The difference in odds ratio for this parameter between class 1 and class 2 was not statistically significant ($p = .354$).

After an appropriate measurement model for PEM was established using LCA, it was necessary to consider how PEM related to MCSE. First, I discuss the measurement model for MCSE.

Musical creative self-efficacy. To examine the underlying structure of the data from the MCSE items, a series of exploratory factor analyses were conducted using both Mplus 7 and SPSS 22. While the total sample size was 132 individuals, 15 of these individuals did not complete any of the MCSE items, lowering the sample size to $N = 117$. Within this subset, there were many missing values, which lowered the dataset to $N = 99$ individuals with complete data. Mplus was utilized to conduct factor analyses using Full Information Maximum Likelihood, which makes use of whatever data are available for all 117 individuals. Geomin rotation, an oblique rotation method, was used with Mplus analyses. SPSS was used to conduct exploratory factor analyses with more traditional extraction methods (e.g., principal axis factoring) and oblique rotation methods (e.g., Direct Oblimin, Quartimax, Equimax). This approach was used in order to avoid severe capitalizations on chance that might result from following only one particular extraction and/or rotation method, given the somewhat small sample size for conducting a factor analysis.

Over the course of this analysis, several items were removed for various reasons. For example, items 4, 5, 14, and 15 seemed to cross load on several different factors (but not the same factors between them). Each of these items refers to a specific means of being creative that is ultimately dependent on one's performing experience and ability either on an instrument or with one's voice. In other words, these items were somewhat contaminated by other factors. Someone who plays an instrument but does not sing might have very different scores than someone who sings but does not play an instrument, even if they have the same level of MCSE.

This would be problematic because the observed score for such an item is not the result of a single MCSE factor.

Kaiser's criterion (eigenvalues > 1), parallel analysis (O'Connor, 2000), the Minimum Average Partial (MAP) test (Velicer et al., 2000), and a comparison of model fit statistics from Mplus were used to determine the number of factors to extract for each analysis. It should be noted that these different methods often did not agree. For example, for the final factor analysis that was completed, the revised version of the MAP test (Velicer et al., 2000) indicated 1 factor, parallel analysis indicated 1 factor, and Kaiser's criterion indicated 1 factor, but the single factor solution from Mplus indicated rather poor fit (RMSEA = .176, CFI = .821, SRMR = .062, $\chi^2(65) = 300.89, p < .001$).

The final three-factor solution included items 1-3 for the General MCSE factor, items 7-12 for the Component MCSE factor, and items 13, 17, 18, and 19 for the Improvise MCSE factor. A confirmatory factor analysis (CFA) was conducted using the same data for the purpose of examining reliability, particularly as it related to the possibility of correlated errors. Results from the CFA indicated several potential sets of correlated errors, although only the following were retained in the model:

- Item 9 with Item 10: this was deemed plausible because both items rely on terminology that may not be familiar to everyone, both of which deal with creating new music out of previously existing songs.
- Item 12 with Item 2: this was deemed plausible because both items use very similar wording in regards to creating "an entire song"
- Item 17 with Item 18: this was deemed plausible because both items have the exact same stem "improvise a melody (using either an instrument or your voice) to..."

- Item 7 with Item 17: this was deemed plausible because both items dealt with creating a “chord progression.”

A final CFA model was estimated including these four correlated errors, resulting in a fairly well-fitting model (SRMR = .043, CFI = .943, RMSEA = .105, $\chi^2(58) = 133.47, p < .001$). In addition, while Cronbach’s Alpha assumes that errors are uncorrelated, a Cronbach’s Alpha reliability estimate could still be estimated using the variance estimates from the CFA, thus adjusting the Cronbach’s Alpha estimate for the included correlated errors. Reliability estimates using both Cronbach’s Alpha and the adjusted Cronbach’s Alpha (and a 95% confidence interval) for each sub-scale are displayed in Table 4.5, all of which indicated very strong reliability of the MCSE scores.

It should be noted that I also considered the possibility of a bi-factor model that included the three factors discussed above, but with an additional general self-efficacy factor (not specific

Table 4.5

Reliability Estimates and Confidence Intervals for MCSE Sub-Scales (Pilot Study #2)

Scale	Cronbach's Alpha	Cronbach's Alpha Adjusted for Correlated Errors
MCSE (general)	.897 [.805, .950]	.914 [.886, .941]
MCSE (components)	.919 [.859, .959]	.903 [.874, .932]
MCSE (improvise)	.922 [.860, .961]	.908 [.878, .939]
MCSE (Overall 3-factor mean)	.919 [.846, .960]	x

Note. Cronbach's Alpha assumes errors are uncorrelated. The CFA model for the MCSE factors indicated several errors were significantly correlated. A new CFA model was estimated, including these correlated errors. A new Cronbach's Alpha was calculated using the correct variances from this correlated errors measurement model. There is no adjusted Cronbach's Alpha for MCSE Overall (3-factor mean) because this portion of the model did not have correlated errors.

to musical creative self-efficacy) on which all items load. I was unable to get this model to converge during estimation. This would be a model to consider further in future research.

Convergent validity. According to self-efficacy theory, the extent to which one has successful experiences will support one's self-efficacy for a given domain. While it was not possible to examine the participants' successful experiences in music, it was possible to examine their total experience in music. It is reasonable to expect that the more experience one has, the more likely one is to feel competent. We tend not to continue doing things at which we do not feel competent. As such, someone who has a large amount of experience in music is likely going to feel more competent at being creative with music than someone who has very little experience with music, simply because they have more skills and abilities upon which to draw. Convergent validity was examined from two perspectives, depending on whether PEM class was treated as an ordinal or a nominal variable. If PEM class is assumed to be at least an ordinal level variable, then the Spearman Rank Order Correlation (Spearman's rho) provides an analogous measure of association to the Pearson correlation. Both the Spearman and Pearson correlations (for comparison purposes) for PEM class and the three MCSE factors are displayed in Table 4.6. Regardless of which coefficient is used, there are clear positive, moderately strong, and statistically significant relationships between PEM class and all three factors of MCSE, thus providing mutual support for convergent validity. Note that the negative sign in the first column of the correlation matrices is an artifact of the fact that PEM class was coded with class 1 representing "highly experienced" and PEM class 3 representing "minimally experienced." Therefore, lower PEM values being associated with lower MCSE values shows up as a negative correlation coefficient. The actual relationship is positive though.

Table 4.6

Validity Coefficients for PEM Class and Musical Creative Self-Efficacy

Pearson Correlations				
	PEM Class	MCSE (general)	MCSE (components)	MCSE (improvise)
PEM Class	1			
MCSE (general)	-.524	1		
MCSE (components)	-.516	.834	1	
MCSE (improvise)	-.566	.748	.851	1

Spearman's Rank-Order Correlations				
	PEM Class	MCSE (general)	MCSE (components)	MCSE (improvise)
PEM Class	1			
MCSE (general)	-.478	1		
MCSE (components)	-.495	.831	1	
MCSE (improvise)	-.517	.775	.870	1

Note. $N = 105$ (listwise deletion). All correlations are significant at the 0.01 level (2-tailed). Correlation coefficients with pairwise deletion were comparable. The negative sign of the correlations between MCSE factors and PEM is an artifact of the fact that PEM class 1 (represented as a value of 1 in the data) actually represents a higher amount of PEM than PEM classes 2, which is in turn higher than PEM class3.

One way of examining the relationship between a nominal variable (indicating group/class membership) and a factor (indicating the level of some underlying construct) is by way of the *multiple-indicator multiple-cause* (MIMIC) model in the general structural equation modeling framework. In short, with the MIMIC model a series of dummy variables are created to represent group membership (or in this case PEM class membership) and then the factor of interest (in this case three different MCSE factors) is regressed on all dummy variables (Kline, 2011). The advantage to this approach over a series of one-way ANOVAs (or MANOVA) is that measurement error of each MCSE factor is directly accounted for in the model (i.e., each

MCSE is free of measurement error), thus allowing a better estimate of the effect. A significant effect of the dummy variables on the MCSE factors indicates group differences on the factor means. In this context, that would mean that being a member of a particular PEM class (which represents varying degrees of musical experience) is related to a different level of MCSE compared to members of all other PEM classes.

The MIMIC model for MCSE on PEM and results of the model estimation are displayed in Figure 4.2 and Table 4.7. Two models were estimated, one with the PEM class path coefficients freely estimated and one with those same coefficients fixed to zero, the former representing an overall null hypothesis that there is no PEM class effect for all three factors. This allowed a Likelihood Ratio Test (LRT) to be conducted between the two models. The significant LRT ($\chi^2(6) = 52.741, p < .001$) indicated that the model with the PEM class effects is significantly better fitting than the model with those effects constrained to zero. This is roughly analogous to the F-test with ANOVA, in that it functions like an omnibus test between the research model and a null model. As such, an examination of the specific PEM class effects was warranted. Model fit statistics and model comparison statistics are shown in Table 4.8.

The unstandardized path coefficients (see Figure 4.2) from each PEM class dummy variable represent the group difference in the factor mean between that specific PEM class and the reference PEM class, which in this case was PEM class 3 (“minimally experienced”). For example, the factor mean difference between PEM class 1 and PEM class 3 on MCSE General is .584. In considering the direction and magnitude of all unstandardized path coefficients, it appeared that individuals in PEM class 1 are associated with larger MCSE factor scores for all three MCSE factors than individuals in both PEM class 2 and PEM class 3. Remember, PEM class 1 represents “highly experienced” individuals, so this fits with the general theory that more

experience should be associated with greater self-efficacy. In addition, individuals in PEM class 2 (moderately experienced) are also associated with larger MCSE factor scores for all three MCSE factors than individuals in PEM class 3 (minimally experienced). A Wald Z-test for each path coefficient indicated that all of the aforementioned effects are significant at the .01 level.

Table 4.7

Logistic Regression MIMIC Model Results for MCSE Predicted by PEM Class

Covariate	Class 1 ("Highly Experienced") compared to Class 3 ("Minimally Experienced") in Logits	<i>p</i>	Class 1 ("Highly Experienced") compared to Class 3 ("Minimally Experienced") in Odds Ratio
Male	0.55 (1.975)	0.781	1.733
Age	-0.504 (0.382)	0.187	0.604
Education	0.182 (0.175)	0.297	1.2
Intercept	1.551 (3.325)	0.641	x

Covariate	Class 2 ("Moderately Experienced") compared to Class 3 ("Minimally Experienced") in Logits	<i>p</i>	Class 2 ("Moderately Experienced") compared to Class 3 ("Minimally Experienced") in Odds Ratio
Male	-0.962 (2.103)	0.647	0.382
Age	-0.429 (0.48)	0.372	0.651
Education	0.257 (0.203)	0.206	1.293
Intercept	2.127 (3.536)	0.548	x

Note. *N* = 173. Class 1 ("Highly Experienced"), Class 2 ("Moderately Experienced"), Class 3 ("Minimally Experienced"). Class 3 was the reference class. The *p*-value is for the Wald test of the logit estimate divided by the standard error.

Table 4.8

Fit Statistics and Model Comparison Statistics for MIMIC Models for MCSE on PEM

MIMIC Model for MCSE on PEM (treated as unordered, nominal classes)												
Model	χ^2	df	<i>p</i>	CFI	RMSEA	SRMR	Likelihood Ratio Test					
							$\Delta\chi^2$	Δ df	<i>p</i>	Δ CFI	Δ RMSEA	Δ SRMR
Constrained	226.73	84	<.0001	0.893	0.123	0.14	-	-	-	-	-	-
Unconstrained	173.99	78	<.0001	0.928	0.105	0.048	52.74	6	1.32E-09	0.035	-0.018	-0.092

MIMIC Model for MCSE on PEM (treated as ordered classes)												
Model	χ^2	df	<i>p</i>	CFI	RMSEA	SRMR	Likelihood Ratio Test					
							$\Delta\chi^2$	Δ df	<i>p</i>	Δ CFI	Δ RMSEA	Δ SRMR
Constrained	205.79	71	<.0001	0.899	0.13	0.164	-	-	-	-	-	-
Unconstrained	159.72	68	<.0001	0.931	0.11	0.046	46.06	3	5.50E-10	0.032	-0.020	-0.118

Single-Factor MIMIC Models						
Model	χ^2	df	<i>p</i>	CFI	RMSEA	SRMR
MCSE General	6.365	4	0.174	0.991	0.073	0.028
MCSE Components	62.489	18	<.0001	0.906	0.149	0.059
MCSE Improvise	17.721	7	0.013	0.97	0.121	0.028

Note. *N* = 112 for all models, except *N* = 105 for MCSE Improvise single-factor model and *N* = 111 for MCSE Components single-factor model.

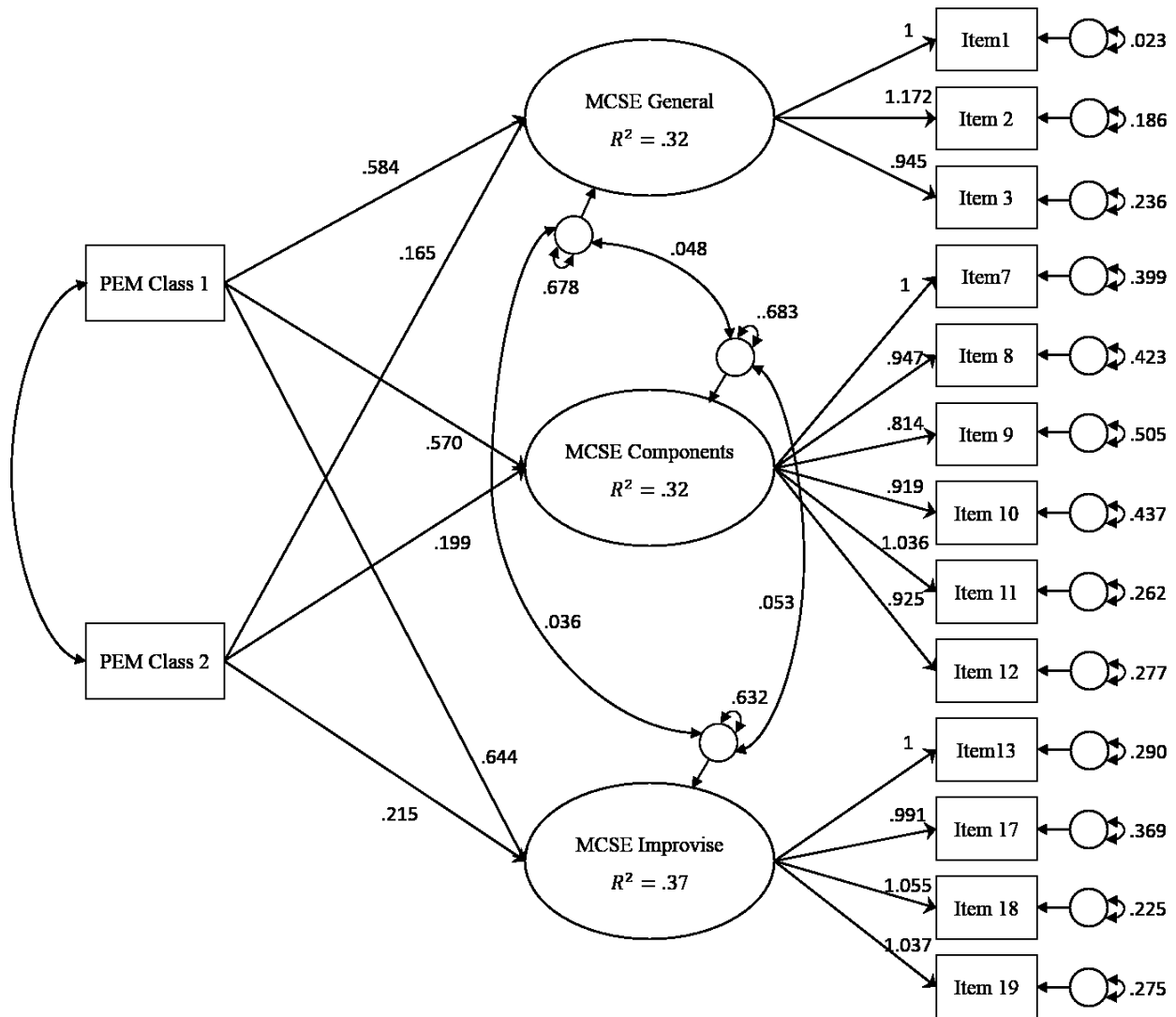


Figure 4.2. MIMIC model specification and results for testing PEM group/class differences on three MCSE factors. *Note.* MCSE = Musical Creative Self-Efficacy; PEM = Past Experience in Music. Residual/Error variances and disturbance variances are standardized, and can be interpreted as proportion of unexplained variance (i.e., $(1 - R^2)$). Covariances between disturbances are standardized, and can be interpreted as correlations. Path coefficients are unstandardized. $\chi^2(78) = 173.988$, $p < .001$, RMSEA = .105 [90% CI .084, .126], CFI = .928, SRMR = .048. The reference class for PEM was Class 3 (“Minimally Experienced”).

An overall effect size for PEM class on each factor (given there are more than two groups) can be calculated as $\hat{f} = \sqrt{\frac{R^2}{1-R^2}} = \sqrt{\frac{1-\varphi^2}{\varphi^2}}$ (Hancock, 2001). As such the effect sizes for PEM on the three factors are as follows:

$$\hat{f}_{MCSE\ General} = \sqrt{\frac{1 - .678}{.678}} = .689$$

$$\hat{f}_{MCSE\ Components} = \sqrt{\frac{1 - .683}{.683}} = .681$$

$$\hat{f}_{MCSE\ Improvise} = \sqrt{\frac{1 - .632}{.632}} = .763$$

The effect sizes provide an indication of the magnitude of the difference between PEM class means on each factor and is interpreted in the same way as the f effect size in univariate ANOVA (Hancock, 2001). Cohen (1988) recommended .1 to indicate a small effect, .25 to indicate a moderate effect, and .40 to indicate a large effect. The PEM class effect sizes, all greater than .68, represent very large factor mean differences between the PEM classes.

However, there are three issues to consider in interpreting these results. First, the MIMIC model assumes strict measurement invariance, which cannot be directly tested in the MIMIC model (Kline, 2011). It is possible to test this in a multiple group CFA framework, but this requires splitting the actual sample into three different groups, and estimating three different models. Given the sample size I had available ($N = 112$), this would have resulted in prohibitively low group sample sizes and far too many parameters to estimate with such small sample sizes for each group.

Second, the actual PEM class to which each individual was assigned is based on a probability, that is, individuals were assigned to whichever class they were most *likely* to be a member. As such, whether someone belongs to a specific PEM class is not as certain as, for example, someone's age, or gender, or primary language. The MIMIC model analysis described above does not take into account this degree of uncertainty regarding PEM class membership. That being said, recall that the 3-class conditional latent class model has an entropy value of .875 and probabilities of correct latent class assignment of 1.00, .949, and .923 for latent classes 1, 2, and 3, respectively, which support a strong degree of confidence in an individual having been assigned to the "correct" PEM latent class.

Third, this particular MIMIC model fits the data fairly well, but not as well as would be desired. In particular, the RMSEA of .105 is well above the recommended cutoff values of .06 (Hu & Bentler, 1998) or .08 (MacCallum, Browne, & Sugawara, 1996). RMSEA penalizes for more parameters in the model though, and given the three factors, 11 indicators, six direct effects from the dummy variables, three correlated disturbances, and four correlated errors, the RMSEA value may be picking up on this complexity. The CFI of .928 did not quite reach Hu and Bentler's (1998) recommended value of .95, and the chi-square test was significant, which is a formal test of model misfit. However, the SRMR value of .048 is below the recommended cutoff of .08 (Hu & Bentler, 1998), and is also sensitive to model misspecification. To compare, I estimated a model that included only the MCSE General factor (3 indicator items with no correlated errors) and the two PEM class dummy variables. The resulting model exhibited extremely good fit ($\chi^2(4) = 6.365, p = .174, RMSEA = .073, CFI = .991, and SRMR = .028$). On the other hand, the single-factor models for MCSE Components and MCSE Improvise did not exhibit very good fit (see Table 4.8). Taken together, this suggests that there is indeed some

model misspecification in the 3-factor MIMIC model, likely resulting from the MCSE Components and MCSE Improvise factors.

Given the apparent ordinal differences in PEM class (discussed above in the latent class analysis of PEM), it was also reasonable to consider a model in which PEM was treated as an ordinal variable instead of strictly nominal. For this model, each MCSE factor was regressed on a single PEM class variable (with values 1, 2, and 3 for the classes, with 1 actually representing the highest amount of experience). Similar to the approach discussed above, a freely estimated model was compared to a constrained model with the PEM class effect set at zero. As with the nominal models, the LRT was significant, warranting an examination of the PEM class effects. As can be seen from the model fit and comparison statistics displayed in Table 4.8, the fit the data slightly worse than the nominal model discussed above. Furthermore, since the models were not nested, the BIC, AIC, and sample-size adjusted BIC (ssBIC) were examined to make a more formal comparison of fit between the nominal and ordinal models. The AIC and ssBIC were both lower (indicating better fit) for the nominal model. For this reason, the specific parameter estimates for the ordinal model are not reported.

To recap, the validity coefficients in Table 4.6 combined with the MIMIC model results all imply a fairly clear relationship between PEM and MCSE. If these three PEM classes are viewed as at least ordinal in nature (representing more or less degree of experience), then the moderate to strong correlations between PEM and MCSE support the theory that previous experience will support self-efficacy. In addition, the increasingly large differences in factor means between PEM class 3 (“minimal experience”) and the other two PEM classes also fits within the general theory that PEM and MCSE should be positively related. In considering all of the results from the MCSE and PEM analyses, there appeared to be sufficient evidence to move

forward with the two measures, but keeping in mind there is still some clear misspecification in the measurement model for MCSE Components and MCSE Improvise. Further research on these instruments will need to explore this issue further and also the nature of the PEM classes in more detail (e.g., considering the possibility of additional classes, which are not related in an ordinal manner, as was originally proposed).

CHAPTER FIVE: RESULTS

In the following chapter I will report results from the primary MOOC study. First, I will describe the final dataset, including discussions of the various sample sizes and missing data. Second, I will describe the data for each variable that was measured. The final section will report results from the various statistical analyses.

Description of the Final Dataset

It should be noted up front that the final dataset for this study had many columns of missing data for many individual participants. As has been mentioned several times, the sample size and the extent to which participants actually participated became an issue fairly early in the study. The following is a description of the data that was actually collected.

Data from at least one questionnaire was collected from a total of 281 students. Of those 274 students, 2 students completed every single research item (including the questions from the Welcome to Canvas survey). What is clear from Figure 5.1 is that the variables for which there are larger amounts of data are also the variables that were measured via questionnaires at the very beginning of the course or within the first week.

As a result of having large amounts of missing data, the data available for different analyses varied quite a bit. I have made every effort to indicate the sample size for all results reported in this chapter so that it is clear exactly how many participants were included in the analysis. That being said, analyses that were conducted for inferential purposes primarily

utilized only the variables for which the sample size was appropriately large (in relation to the type of analysis).

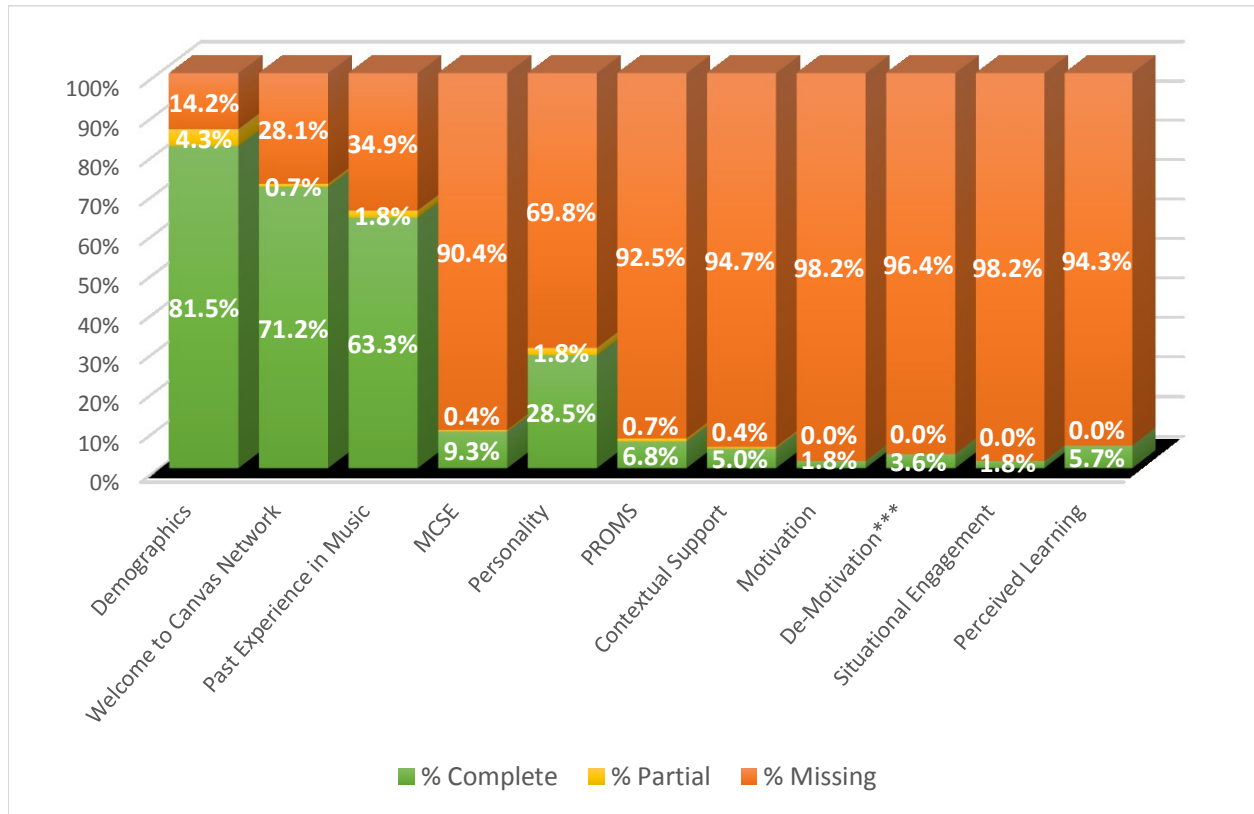


Figure 5.1. Complete and Missing Data by Measure

Note. Percentages based on $N = 281$ (the participant sample size), not the total course enrollment. This is because total enrollment numbers are not indicative of the number of people that actually attended and participated in the course.

Descriptive Statistics for Each Variable

This section is meant to provide a picture of the overall dataset by describing the data using descriptive statistics and the analyses used to determine the variable scores that would be used in further analyses (e.g., latent class analysis, factor analysis). Table 5.1 shows descriptive statistics for all measured variables, except the nominal variables.

Table 5.1

Descriptive Statistics for All Measured Variables

Variable	Mean [CI]	SD [CI]	Median	Skewness [CI]	Kurtosis [CI]	N
Gender (Male = 1)	0.443	x	x	x	x	237
Age (Ordinal)	3.606	1.465	3.00	0.248	-0.681	236
Education	3.682	1.828	4.00	0.015	-1.106	236
Active Participant	0.446 [.386, .505]	0.498 [.489, .501]	x	x	x	202
Expected Hours/Week	2.84	1.017	3.00	0.641	1.063	201
Past Experience in Music	1.85 [1.75, 1.96]	.70 [.64, .75]	2.00	0.215 [0.07, 0.37]	-0.94 [-1.25, -0.55]	173
Musical Creative Self-Efficacy	46.65 [36.42, 57.98]	28.71 [24.34, 31.72]	44.17	0.06 [-0.51, 0.61]	-0.95 [-1.63, 1.52]	27
Personality: Extraversion	2.72 [2.52, 2.89]	0.97 [0.87, 1.05]	2.75	0.06 [-0.24, 0.35]	-0.91 [-1.23, -0.41]	83
Personality: Agreeableness	4.00 [3.86, 4.13]	0.69 [0.56, 0.82]	4.00	-1.05 [-1.83, 0.11]	3.00 [-1.01, 5.13]	83
Personality: Conscientiousness	3.22 [3.05, 3.38]	0.86 [0.77, 0.94]	3.25	-0.14 [-0.49, 0.25]	-0.79 [-1.16, -0.32]	83
Personality: Neuroticism	2.79 [2.64, 2.97]	0.86 [0.75, 0.96]	2.75	0.12 [-0.22, 0.50]	-0.48 [-0.90, 0.14]	83
Personality: Openness	4.08 [3.95, 4.22]	0.67 [0.54, 0.80]	4.25	-1.07 [-1.73, -0.01]	2.14 [-0.59, 3.43]	83
Personality: Honesty-Humility	3.58 [3.4, 3.75]	0.83 [0.72, 0.94]	3.75	-0.60 [-1.11, 0.01]	0.09 [-0.83, 0.87]	83
Music Aptitude (PROMS): Melody	21.74 [19.58, 24.26]	4.85 [3.361 5.899]	22.00	0.73 [-0.19, 1.27]	1.01 [-0.95, 3.23]	19
Music Aptitude (PROMS): Time	25.00 [22.53, 27.50]	6.12 [4.826, 6.875]	26.00	0.52 [-0.71, 0.63]	-0.85 [-1.64, 1.85]	19
Music Aptitude (PROMS): Speed	27.53 [25.37, 29.58]	4.49 [3.47, 5.11]	27.00	-0.27 [-1.07, 0.75]	-1.06 [-1.26, 1.58]	19
Music Aptitude (PROMS): Beat	24.26 [22.11, 26.62]	4.27 [3.39, 4.80]	24.00	-0.14 [-0.79, 0.45]	-0.79 [-1.68, 1.29]	19

Variable	Mean [CI]	SD [CI]	Median	Skewness [CI]	Kurtosis [CI]	N
Contextual Support (Challenge)	3.90 [3.60, 4.23]	0.63 [0.44, 0.74]	4.00	0.31 [-0.48, 0.86]	-0.36 [-1.6, 2.88]	15
Contextual Support (Choice)	3.78 [3.38, 4.12]	0.72 [0.35, 0.95]	4.00	-1.84 [x, x]	3.76 [0.28, 3.63]	15
Contextual Support (Choice)	3.03 [2.67, 3.42]	0.93 [0.73, 1.04]	2.75	0.06 [-0.89, 1.11]	-1.07 [-1.92, 2.2]	15
Contextual Support (Curiosity)	4.12 [3.70, 4.47]	0.82 [0.49, 1.02]	4.25	-1.38 [x, x]	1.96 [-0.48, 2.84]	15
Contextual Support (Feedback)	3.20 [2.90, 3.55]	0.77 [0.45, 0.96]	3.00	0.91 [-0.15, 1.79]	0.79 [-1.65, 7.09]	15
Contextual Support (Interest)	3.95 [3.45, 4.38]	0.96 [0.58, 1.19]	4.00	-1.36 [x, x]	1.87 [-0.72, 3.65]	15
Motivation: Intrinsic Mot	4.53 [3.09, 5.73]	1.57 [0.72, 1.77]	4.67	-1.10 []*	2.53 []*	5
Motivation: Identified Reg	5.47 [4.24, 6.47]	1.5 [0.55, 1.73]	6.00	-1.37 []*	2.49 []*	5
Motivation: External Reg	3.67 [2.6, 4.73]	1.65 [1.07, 1.8]	3.67	0.37 [-0.94, 1.36]	-0.1 [-3.33, 5.00]	5
Motivation: Amotivation	2.20 [1.27, 3.20]	1.35 [1.28, 1.28]	1.33	0.56 []*	-3.25 []*	5
Situational Engagement	3.45 [3, 4.15]	0.891 [.18, 1.11]	3.25	1.929 [0.00, 1.98]	4.02 [-3.25, 4.57]	5
Perceived Learning: General	5.16 [4.30, 5.94]	1.75 [1.17, 2.12]	5.40	-0.96 []*	0.47 [-1.12, 1.05]	16

Note. 95% confidence intervals based on 1000 bootstrap re-samples and calculated using the bias corrected and accelerated (BCa) method.

* Bootstrap confidence interval could not be calculated.

Demographics. Of the 229 individuals for whom data on age, gender, and education was available, 126 (55%) were female. Figure 5.2 and Figure 5.3 display the distributions of age and education, respectively, for participants for whom such data were available. Nearly 60% of these participants had completed a bachelor’s degree (or equivalent 4-year degree) or higher, and nearly a third of the participants were between the age of 25 and 34. Additional descriptive statistics for all three demographic variables are shown in Table 5.1.

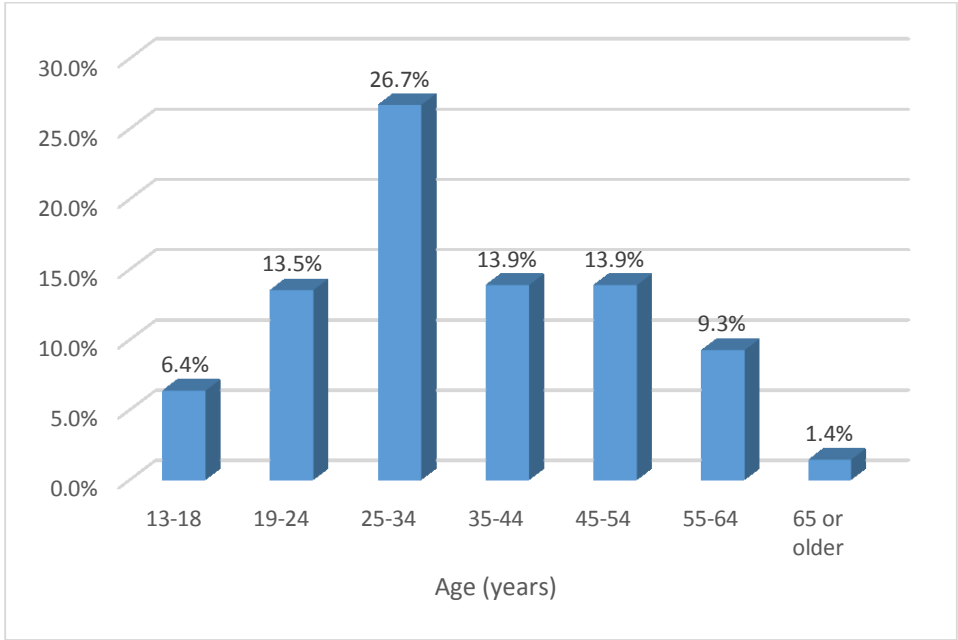


Figure 5.2: Distribution of Age of Participants

Note. $N = 229$

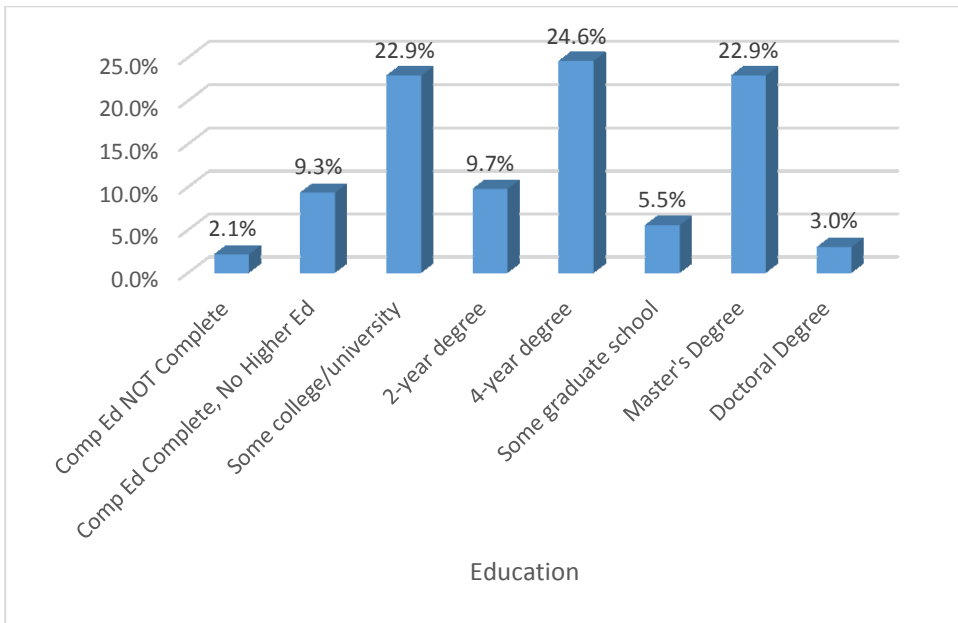


Figure 5.3: Distribution of Education of Participants

Note. $N = 229$

Past experience in music. As with Pilot Study #2, a series of latent class analyses were conducted to determine the underlying latent classes from the PEM variables. Since this approach to measuring PEM is still new and very exploratory, I decided to analyze the PEM data in the same exploratory manner that I did for Pilot Study #2. As with Pilot Study #2, I used only the dichotomous (yes/no) responses to the 11 PEM items as data for each LCA model. Results from the unconditional models are displayed in Table 5.2. As with Pilot Study #2, the fit statistics and model comparison tests were not all in agreement. The 2-class and 3-class models indicated a similar pattern of low to high degree of experience, but interpretability became difficult beyond three classes. I withheld a decision regarding which model to retain prior to examining the conditional models.

Table 5.2

Unconditional PEM Models for Latent Class Analysis

# classes	Unconditional Models					Vuong-Lo-Mendell-Rubin LRT (-2LL)*	Lo-Mendell-Rubin LRT (-2LL)*	Parametric Bootstrap LRT (-2LL)*+
	Final Stage LL	AIC	BIC	ssBIC	Entropy			
2	-1051.62	2149.24	2223.06	2150.21	.855	327.267, $p < .001$	322.114, $p < .001$	327.267, $p < .001$
3	-1023.80	2117.60	2229.94	2119.09	.754	55.633, $p = .114$	54.757, $p = .118$	55.633, $p < .001$
4	-1004.88	2103.76	2254.61	2105.75	.796	37.845, $p = .626$	37.249, $p = .630$	37.845, $p < .001$
5	-989.48	2096.96	2286.32	2099.46	.833	30.795, $p = .157$	30.311, $p = .159$	30.795, $p = .030$

Note. $N = 183$. All model results are based on 1000 random sets of starting values for initial stage optimization and 250 random sets of starting values for final stage optimization. The best likelihood value was replicated in all models.

* For model comparison tests, the baseline model is the model with one fewer classes

+ Based on 200 bootstrap draws

Similar to Pilot Study #2, I added age and gender as covariates, but since this sample had a much larger age range, and therefore a much greater degree of variability in education, I also

included education as a third covariate in the conditional models. Results from these conditional models are displayed in Table 5.3. The conditional models were very similar to the unconditional models in the lack of a clear decision indicated by the fit and model comparison statistics. I chose to retain the 3-class model for the following reasons: (1) the 3-class model was most interpretable, indicating three clear classes (see latent profile plots in Appendix F, particularly the 3-class conditional model); (2) the 3-class model also allowed for the creation of an ordinal categorical PEM variable to be used in further analyses, as opposed to a nominal variable, which fits better with the notion of having more or less overall musical experience; (3) given the exploratory nature of this measurement approach, I felt it more prudent to err on the side of being conservative, saving a more fine-grained analysis for a larger sample; and (4) the three-class model had a reasonably high entropy (.825, only slightly less than the entropy of the 2-class model), indicating the probability of correct classification was quite high; and (5) the three-class model matched the results and decisions from Pilot Study #2.

Table 5.3

Conditional (Age, Gender, and Education as Covariates) PEM Models for Latent Class Analysis

# classes	Log Likelihood	AIC	BIC	ssBIC	Entropy	Vuong-Lo-Mendell-Rubin LRT (-2LL)*	Lo-Mendell-Rubin LRT (-2LL)*	Parametric Bootstrap LRT (-2LL)*+
2	-992.743	2037.486	2119.472	2037.141	.864	323.301, $p < .001$	319.172, $p < .001$	323.301, $p < .001$
3	-961.288	2004.577	2133.862	2004.033	.822	62.909, $p = .503$	62.106, $p = .507$	62.909, $p < .001$
4	-935.415	1982.83	2159.414	1982.087	.815	51.747, $p = .622$	51.086, $p = .624$	51.747, $p < .001$
5	-918.446	1978.891	2202.775	1977.949	.863	33.938, $p = .021$	33.505, $p = .022$	33.938, $p = .105$

Note. $N = 173$. All model results are based on 1000 random sets of starting values for initial stage optimization and 250 random sets of starting values for final stage optimization. The best likelihood value was replicated in all models.

* For model comparison tests, the baseline model is the model with one fewer classes

+ Based on 200 bootstrap draws

As with Pilot Study #2, the classes were named and coded for further analyses in the following manner: Class 1 = *Highly Experienced*; Class 2 = *Moderately Experienced*; and Class 3 = *Minimally Experienced*. Shown in Table 5.4 are the proportions of individuals who were assigned to each class for both the unconditional and conditional 3-class models. Like Pilot Study #2, the number of individuals assigned to Class 1 (Highly Experienced) declined with the addition of the covariates.

Table 5.4

Classification of Individuals Based on Their Most Likely PEM Class Membership

Unconditional 3-Class Model (No Covariates)		
Latent Class	# Individuals Assigned to Class	Proportion Individuals Assigned to Class
1 ("Highly Experienced")	73	.399
2 ("Moderately Experienced")	53	.290
3 ("Minimally Experienced")	57	.311
TOTAL	183	1.000

Conditional 3-Class Model (Age, Gender, and Education as Covariates)		
Latent Class	# Individuals Assigned to Class	Proportion Individuals Assigned to Class
1 ("Highly Experienced")	57	.330
2 ("Moderately Experienced")	85	.491
3 ("Minimally Experienced")	31	.179
TOTAL	173	1.000

Musical creative self-efficacy. A total of 27 individuals completed the Module 2 Questionnaire, which contained the measure of Musical Creative Self-Efficacy (MSCE). Given the small sample size, it was not possible to calculate adjusted Cronbach's Alpha reliability estimates as in Pilot Study #2. Therefore, only Cronbach's Alpha (unadjusted) estimates (and

95% confidence intervals) are reported in Table 5.5. All estimates, including the lower bound of the 95% confidence interval are above the traditional cutoff of .8. Scores for each sub-scale were calculated by taking the mean of all items for the scale. The MCSE Overall score was then calculated by taking the mean of the three sub-scales. Descriptive statistics for all three sub-scales and the MCSE Overall scores are displayed in Table 5.6.

Table 5.5

Reliability Estimates for Musical Creative Self-Efficacy Sub-Scales and Overall Scale

Scale	Cronbach's Alpha	N
MCSE (general)	.897 [.805, .950]	27
MCSE (components)	.919 [.859, .959]	26
MCSE (improvise)	.922 [.860, .961]	27
MCSE (Overall 3-factor mean)	.919 [.846, .960]	27

Table 5.6

Descriptive Statistics for Musical Creative Self-Efficacy Sub-Scales and Overall Scale

Scale	Mean [CI]	SD [CI]	Median	Skewness [CI]	Kurtosis [CI]	N
MCSE (general)	56.11 [42.13, 68.96]	32.92 [27.32, 37.08]	56.67	-0.26 [-0.91, 0.34]	-1.14 [-1.74, 0.79]	27
MCSE (components)	44.80 [35.18, 54.75]	27.46 [23.57, 30.05]	45.00	-0.11 [-0.70, 0.40]	-1.22 [-1.80, 0.61]	27
MCSE (improvise)	39.03 [27.85, 51.62]	32.20 [26.02, 36.63]	32.50	0.66 [0.11, 1.21]	-0.67 [-1.61, 1.87]	27
MCSE (3-factor mean)	46.65 [36.42, 57.98]	28.71 [24.34, 31.72]	44.17	0.06 [-0.51, 0.61]	-0.95 [-1.63, 1.52]	27

It is worth noting that all MCSE scores did not indicate any substantial departures from a normal distribution and there were significant differences in means between sub-scales (see Figure 5.4). In particular, a series of paired sample t-tests (using Bonferroni corrected p -value of .0167) indicated the MCSE general mean was larger than both the MCSE Components mean, $t(26) = 3.282, p = .003$, and the MCSE Improvise mean, $t(26) = 4.010, p = .0005$, but the MCSE Components mean was not significantly greater than the MCSE Improvise mean, $t(26) = 1.516, p = .142$. Additionally, all three MCSE sub-scales had high, positive, and statistically significant positive intercorrelations (see Table 5.7), providing further support for their combination into an MCSE Overall score.

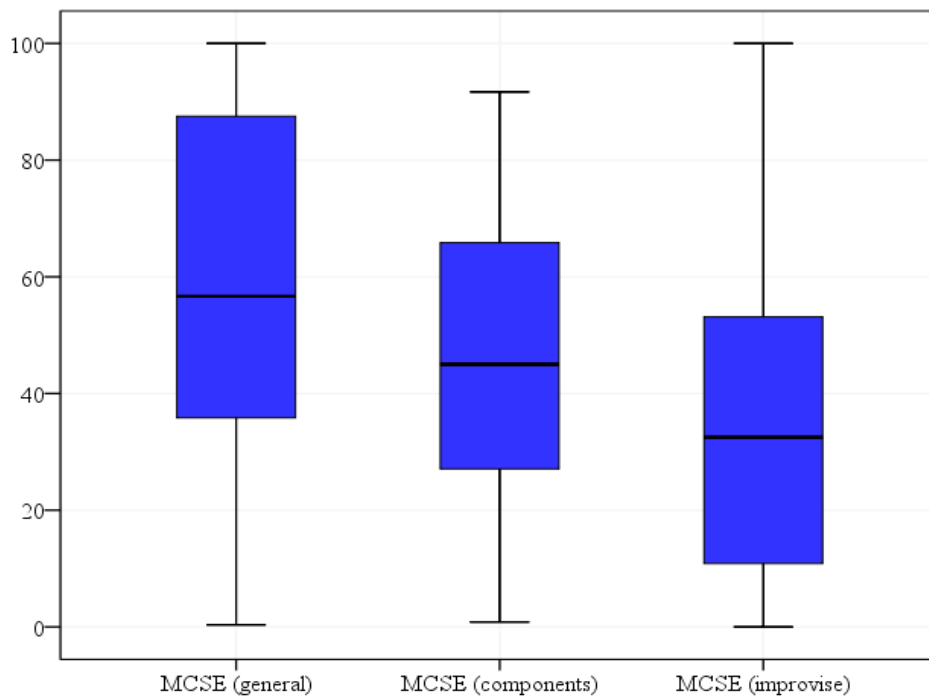


Figure 5.4. Boxplots for Scores on MCSE Sub-Scales

Table 5.7

Paired Sample T-Tests and Correlations for MCSE Sub-Scales

Pairwise Comparison	Correlation [95% CI]+	Paired Differences					
		Mean [95% CI]+	SD	SE	t	df	p
Pair 1 MCSE (general) - MCSE (components)	.84* [.67, .93]	11.31 [4.23, 18.39]	17.91	3.45	3.282	26	.003
Pair 2 MCSE (general) - MCSE (improvise)	.77* [.61, .89]	17.08 [8.32, 25.85]	22.15	4.26	4.008	26	.000
Pair 3 MCSE (components) - MCSE (improvise)	.79* [.61, .90]	5.77 [-2.05, 13.60]	19.79	3.81	1.516	26	.142

Note. $N = 27$.

* Correlations were statistically significant at $p < .001$.

+ 95% confidence intervals based on 1000 bootstrap re-samples and calculated using the bias corrected and accelerated (BCa) method

Music aptitude. The PROMS online survey tool was accessed 42 times, with 19 individuals completing the entire mini PROMS measure. Several individuals accessed the tool multiple times, although none of them actually completed the measure more than once. Descriptive statistics for each sub-scale of the PROMS and an Overall score (calculated by averaging the score from all four sub-scales) are displayed in Table 5.8. The correlation matrix (see Table 5.9) indicated somewhat low intercorrelations between the four sub-scales, which is also reflected in the less than ideal Cronbach's Alpha estimate of .783 [95% CI .564, .907]. That being said, the small sample size may have contributed to biased estimates of the correlations between sub-scales, as evidenced by the wide 95% confidence intervals for Cronbach's Alpha and for the correlation coefficients (not shown).

Table 5.8

Descriptive Statistics for Scores on the PROMS Sub-Scales

	Mean [95% CI]	SD [95% CI]	Median	Skewness [95% CI]	Kurtosis [95% CI]	N
Melody	21.74 [19.58, 24.26]	4.85 [3.361, 5.899]	22	0.73 [-0.19, 1.27]	1.01 [-0.95, 3.23]	19
Tuning	25.00 [22.53, 27.50]	6.12 [4.826, 6.875]	26	0.52 [-0.71, 0.63]	-0.85 [-1.64, 1.85]	19
Speed	27.53 [25.37, 29.58]	4.49 [3.469, 5.110]	27	-0.27 [-1.07, 0.75]	-1.06 [-1.26, 1.58]	19
Beat	24.26 [22.11, 26.62]	4.27 [3.391, 4.795]	24	-0.14 [-0.79, 0.45]	-0.79 [-1.68, 1.29]	19
Overall (4-factor mean)	24.10 [22.43, 25.75]	4.24 [3.05, 5.06]	25	-0.15 [-1.31, 0.82]	0.79 [-0.69, 2.32]	21

Note. 95% confidence intervals based on 1000 bootstrap re-samples and calculated using the bias corrected and accelerated (BCa) method

Table 5.9

Inter-Item Correlations for Scores on the PROMS Sub-Scales

	Melody	Tuning	Speed	Beat
Melody	1.00			
Tuning	.55	1.00		
Speed	.41	.56	1.00	
Beat	.61	.46	.28	1.00

Note. $N = 19$

Personality. A total of 80 individuals completed (5 more partially completed) the mini-IPIP6 measure in the Module 1 Questionnaire. After reverse scoring the necessary items, scores for each personality factor were calculated by taking the mean of the four items for each factor. Descriptive statistics for each factor are shown in Table 5.10. Unfortunately, all six personality factor scales did not exhibit high degrees of intercorrelation among items, as evidenced by the rather low estimates of reliability (see Table 5.11).

Table 5.10

Descriptive Statistics for Scale Scores for Each Personality Factor

Factor	Mean [CI]	SD [CI]	Median	Skewness [CI]	Kurtosis [CI]	<i>N</i>
Extraversion	2.72 [2.52, 2.89]	0.97 [0.87, 1.05]	2.75	0.06 [-0.24, 0.35]	-0.91 [-1.23, -0.41]	83
Agreeableness	4 [3.86, 4.13]	0.69 [0.56, 0.82]	4.00	-1.05 [-1.83, 0.11]	3 [-1.01, 5.13]	83
Conscientiousness	3.22 [3.05, 3.38]	0.86 [0.77, 0.94]	3.25	-0.14 [-0.49, 0.25]	-0.79 [-1.16, -0.32]	83
Neuroticism	2.79 [2.64, 2.97]	0.86 [0.75, 0.96]	2.75	0.12 [-0.22, 0.5]	-0.48 [-0.9, 0.14]	83
Openness	4.08 [3.95, 4.22]	0.67 [0.54, 0.8]	4.25	-1.07 [-1.73, -0.01]	2.14 [-0.59, 3.43]	83
Honest-Humility	3.58 [3.4, 3.75]	0.83 [0.72, 0.94]	3.75	-0.6 [-1.11, 0.01]	0.09 [-0.83, 0.87]	83

Note. 95% confidence intervals based on 1000 bootstrap re-samples and calculated using the bias corrected and accelerated (BCa) method

Table 5.11

Reliability Estimates for Scale Scores for Each Personality Factor

Factor	Cronbach's Alpha [95% CI]	<i>N</i>
Extraversion	.793 [.708, .858]	82
Agreeableness	.705 [.584, .797]	82
Conscientiousness	.729 [.619, .814]	82
Neuroticism	.645 [.500, .756]	82
Openness	.548 [.364, .689]	81
Honest-Humility	.680 [.550, .780]	81

Note. All scales included four items

Motivation. Only five individuals completed the motivation measure. Scores for each motivation factor were calculated by taking the mean of the three items for each factor. Descriptive statistics are reported in Table 5.12 for the sake of completeness, but due to the extremely low sample size the motivation variables were not included in any further analyses, with the exception of the exploratory analysis of the General Specified Model. That being said, it is worth noting that, even with such a small sample, the correlation matrix (see Table 5.13), using either Pearson correlations or Spearman’s rank order correlations, exhibited a very similar “simplex-like pattern” to that reported by Guay, Vallerand, and Blanchard (2000, p. 185).

Table 5.12

Descriptive Statistics for Motivation

Factor	Mean [CI]	SD [CI]	Median	Skewness [CI]	Kurtosis [CI]	<i>N</i>	Level of Measurement
Intrinsic Motivation	4.53 [3.09, 5.73]	1.57 [0.72, 1.77]	4.67	-1.10 []*	2.53 []*	5	Ordinal
Identified Regulation	5.47 [4.24, 6.47]	1.5 [0.55, 1.73]	6.00	-1.37 []*	2.49 []*	5	Ordinal
External Regulation	3.67 [2.60, 4.73]	1.65 [1.07, 1.80]	3.67	0.37 [-0.94, 1.36]	-0.10 [-3.33, 5.00]	5	Ordinal
Amotivation	2.20 [1.27, 3.20]	1.35 [1.28, 1.28]	1.33	0.56 []*	-3.25 []*	5	Ordinal

Note. 95% confidence intervals based on 1000 bootstrap re-samples and calculated using the bias corrected and accelerated (BCa) method.

* Bootstrap confidence interval could not be calculated

Table 5.13

Correlations Between Motivation Sub-Scales

Pearson Correlations				
	Intrinsic Motivation	Identified Regulation	External Regulation	Amotivation
Intrinsic Motivation	1.000			
Identified Regulation	.844	1		
External Regulation	-.299	-.090	1	
Amotivation	-.679	-.814	0.588	1

Spearman Rank-Order Correlations				
	Intrinsic Motivation	Identified Regulation	External Regulation	Amotivation
Intrinsic Motivation	1.000			
Identified Regulation	.763	1		
External Regulation	-.359	-.154	1	
Amotivation	-.703	-.973	0.264	1

Note. $N = 5$

Situational engagement. As with motivation, only five individuals completed the situational engagement items. Descriptive statistics are displayed below in Table 5.14. This variable was not included in any further analyses, except for the exploratory analysis of the General Specified Model.

Table 5.14

Descriptive Statistics for Situational Engagement

Factor	Mean [CI]	SD [CI]	Median	Skewness [CI]	Kurtosis [CI]	N
Situational Engagement	3.45 [3.00, 4.15]	0.891 [0.18, 1.11]	3.25	1.929 [0.00, 1.98]	4.02 [-3.25, 4.57]	5

Perceived learning. A total of 16 individuals completed the Perceived Learning items. The small sample size made factor analytic methods for examining the underlying structure of the Perceived Learning items not possible. Instead, I visually examined the correlation matrix (see Table 5.15). The items that were intended to measure a General factor of Perceived Learning all had high intercorrelations. The pattern of intercorrelations for items intended to measure the other more specific factors (affective, cognitive, and psychomotor) did not exhibit a clear pattern indicative of a well-specified latent factor measurement model. Based on the limited information available, I chose to utilize only the five General items. Scores for a General Perceived Learning factor were calculated by taking the mean of items 1-5. Descriptive statistics for the General Perceived Learning factor are shown in Table 5.16. Cronbach’s Alpha for the scores from this scale was .947 [95% CI .892, .979].

Table 5.15

Descriptive Statistics for Perceived Learning (General)

	Mean [CI]	SD [CI]	Median	Skewness [CI]	Kurtosis [CI]	N
Perceived Learning (General)	5.16 [4.30, 5.94]	1.75 [1.17, 2.12]	5.40	-0.96 []*	0.47 [-1.12, 1.05]	16

Project participation. The percentages of the 281 participants that submitted a project for each of the four creative projects are shown in Figure 5.5. Less than a third of the participants submitted a Top 10 Playlist, less than 6% submitted the Write Lyrics project, six

individuals submitted the Write & Analyze a Melody project, and only one individual actually submitted a completed song for the Write a Song project.

Table 5.16

Correlation Matrix for Perceived Learning Items

		Pearson Correlations														
Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	1															
2	.883**	1														
3	.804**	.776**	1													
4	.750**	.836**	.822**	1												
5	.696**	.820**	.814**	.811**	1											
6	.812**	.932**	.678**	.784**	.695**	1										
7	.512*	.712**	.383	.702**	.583*	.739**	1									
8	.036	.270	.310	.281	.377	.296	.446	1								
9	.424	.538*	.191	.526*	.409	.518*	.791**	.361	1							
10	.768**	.786**	.692**	.805**	.550*	.743**	.651**	.243	.466	1						
11	.599*	.661**	.669**	.526*	.706**	.590*	.353	.231	.180	.386	1					
12	.871**	.976**	.772**	.871**	.792**	.917**	.731**	.239	.544*	.837**	.718**	1				
13	.650**	.695**	.433	.610*	.534*	.792**	.829**	.294	.746**	.583*	.420	.698**	1			
14	.841**	.938**	.748**	.834**	.863**	.870**	.690**	.176	.460	.780**	.689**	.948**	.651**	1		
15	-.236	-.152	-.326	-.057	-.266	-.099	.415	.207	.451	-.019	-.072	-.064	.229	-.235	1	

Note. $N = 16$

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Blue = intercorrelations between General factor items

Red = intercorrelations between Affective factor items

Yellow = intercorrelations between Cognitive factor items

Green = intercorrelations between Psychomotor factor items

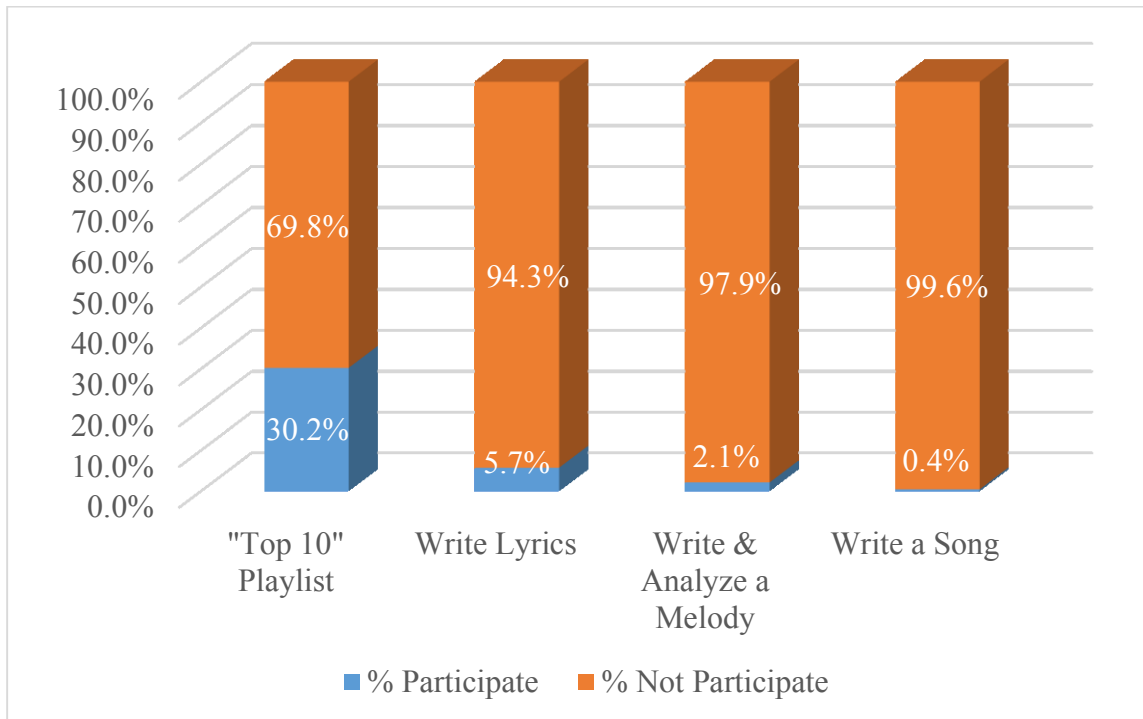


Figure 5.5. Percent Creative Music-Making Project Participation.
Note. $N = 281$

Primary reason. Percentages for each Primary Reason why individuals enroll in a MOOC are displayed in Figure 5.6. Nearly three quarters of the research participants indicated that their Primary Reason for taking an open online course was because they enjoy learning about topics that interest them. The next largest Primary Reason was “I like the format (online),” chosen by only 8.5% of participants. Most notable is that the two options related to gaining new skills (“I hope to gain skills for a new career” and “I hope to gain skills for a promotion at work”) combined to include less than 5% of the research participants. This of course does not mean that either of these reasons would not be a secondary reason for taking the course, given the question forced participants to choose only one reason, their primary reason.

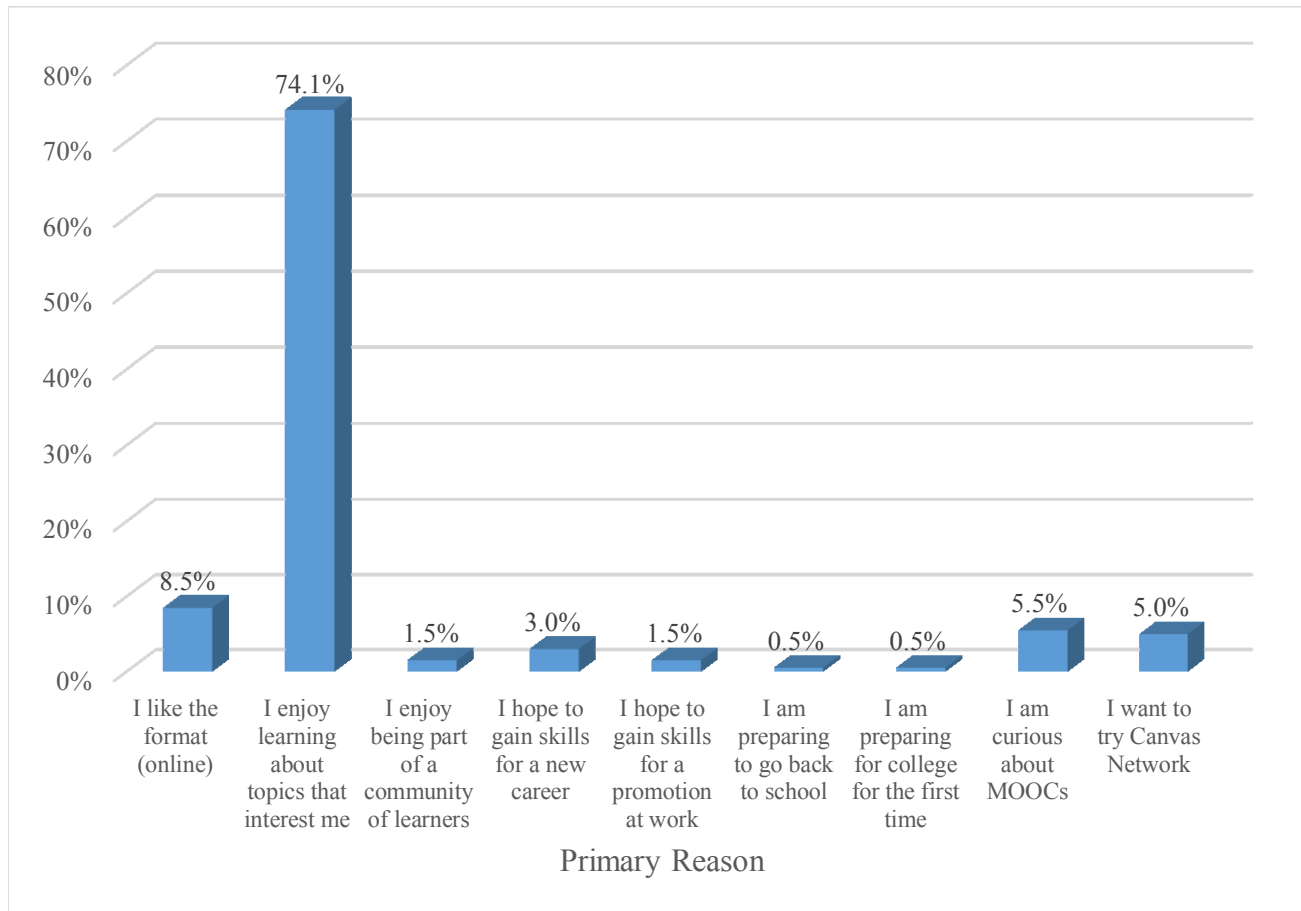


Figure 5.6. Percentages of each Primary Reason for Taking an Open Online Course
Note. $N = 201$.

Type of learner. Frequencies and proportions for each Type of Learner category are displayed in Table 5.17. In regards to project participation for different Types of Learners, Figure 5.7 shows what percentage of each type of learner participated in any project. Looking at this comparison from another perspective, Figure 5.8 displays what percentage of project participants (those that participated in any project) were from each Type of Learner group. Finally, Table 5.18 displays this information in cross tabulation format for Any Project Participation by Type of Learner.

Table 5.17

Frequencies and Proportions for Type of Learner

Type of Learner	Frequency	Proportion	95% CI for Proportion
"Observer"	8	.040	[.020, .059]
"Drop-In"	24	.119	[.079, .163]
"Passive Participant"	80	.396	[.337, .455]
"Active Participant"	90	.446	[.386, .505]
Total	202	1.000	

Note. 95% confidence intervals based on 1000 bootstrap re-samples and calculated using the bias corrected and accelerated (BCa) method

Table 5.18

Type of Learner by Any Project Participation Cross Tabulation

Type of Learner		Any Project Participation		Total
		No, Did not participate in ANY of four projects	Yes, Participated in at least one of four projects	
Observer	% within Type of Learner	87.5%	12.5%	100.0%
	% within Any Project Participation	4.8%	1.8%	4.0%
	% of Total	3.5%	.5%	4.0%
Drop-In	% within Type of Learner	87.5%	12.5%	100.0%
	% within Project Participation Any	14.5%	5.3%	11.9%
	% of Total	10.4%	1.5%	11.9%
Passive Participant	% within Type of Learner	82.5%	17.5%	100.0%
	% within Project Participation Any	45.5%	24.6%	39.6%
	% of Total	32.7%	6.9%	39.6%
Active Participant	% within Type of Learner	56.7%	43.3%	100.0%
	% within Project Participation Any	35.2%	68.4%	44.6%
	% of Total	25.2%	19.3%	44.6%
Total	Count	145	57	202
	% within Type of Learner	71.8%	28.2%	100.0%
	% within Project Participation Any	100.0%	100.0%	100.0%
	% of Total	71.8%	28.2%	100.0%

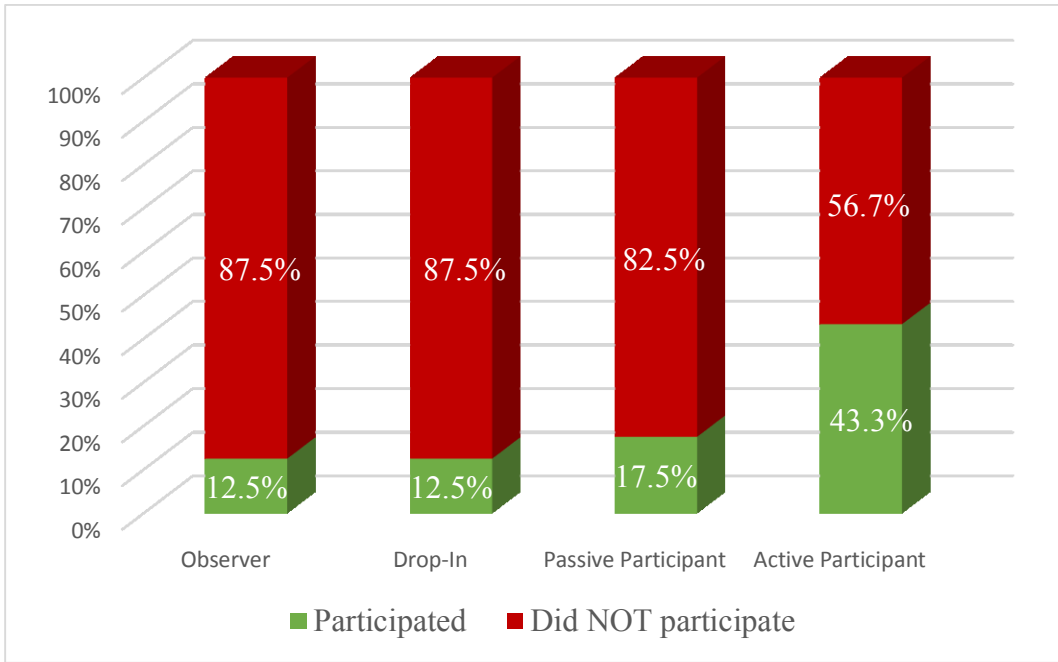


Figure 5.7. Any Project Participation by Type of Learner

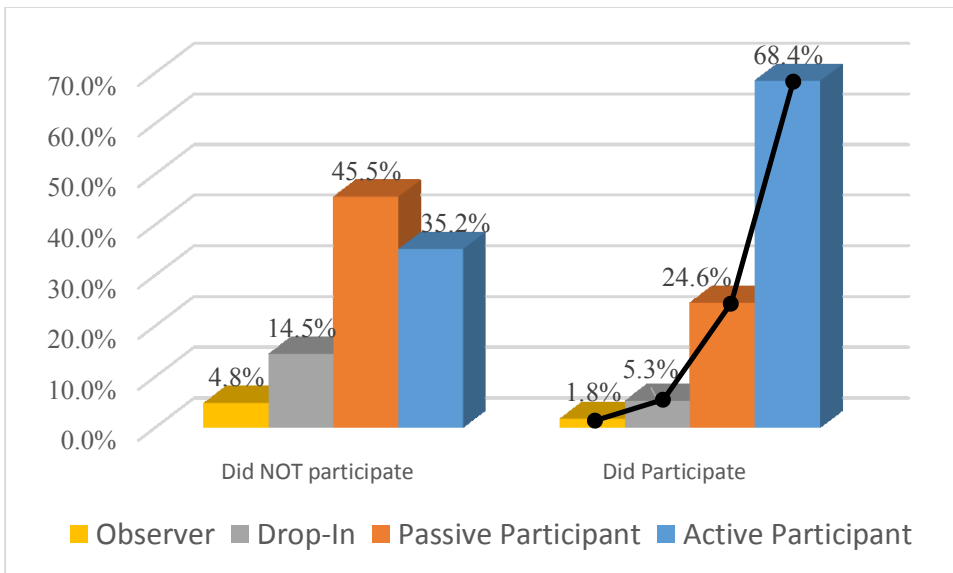


Figure 5.8. Proportions of Project Participants (and Non-Participants) by Type of Learner

Expected hours. Frequencies and proportions for each level of Expected Hours (the number of hours a student expected to spend on the course per week) are displayed in Table 5.19. A little less than three quarters of participants anticipated spending either between 1 and 2 hours or between 2 and 4 hours per week on the course.

Table 5.19

Frequencies and Proportions for Expected Hours/Week

Expected Hours/Week	Frequency	Proportion	95% CI
Less than 1 hour	14	.070	[.040, .100]
Between 1 and 2 hours	60	.299	[.239, .363]
Between 2 and 4 hours	86	.428	[.363, .489]
Between 4 and 6 hours	31	.154	[.114, .199]
Between 6 and 8 hours	5	.025	[.005*, .050]
More than 8 hours per week	5	.025	[.005*, .050]
Total	201	1.000	

Note. 95% confidence intervals based on 1000 bootstrap re-samples and calculated using the bias corrected and accelerated (BCa) method.

* These results could not be computed from jackknife samples, so this confidence interval is computed by the percentile method rather than the BCa method.

Type of learner and expected hours per week. Interestingly, over 44.6% of the participants, for whom data on these variables was available ($N = 202$), described themselves as an “active participant” at the beginning of the course (Week 1). Out of the individuals that described themselves as intending to be an “active participant,” 76.7% of them chose either the

category “2-4 hours” or categories with greater number of hours in regards to the number of hours they expected to spend on the course per week. By dichotomizing the four Type of Learner categories into “active participant” and NOT active participant” (includes passive participant, observer, and drop-in categories), a more complete picture can be seen in Table 5.20 regarding the hours per week individuals expected to spend on the course by whether they identified themselves as an “active participant.”

Table 5.20

Proportions of Individuals Expected Hours/Week Participation by Type of Learner (Active/Non-Active)

Type of Learner		< 1 hr.	1-2 hrs.	2-4 hrs.	4-6 hrs.	6-8 hrs.	> 8 hrs.	Total
	Count	11	42	45	8	2	3	111
NOT “Active Participant” (Passive Participant, Observer, Drop-in)	Proportion of NOT Active Participants	.099	.378	.405	.072	.018	.027	1.000
	Proportion of all participants	.055	.209	.224	.040	.010	.015	.552
	Count	3	18	41	23	3	2	90
“Active Participant”	Proportion of Active Participants	.033	.200	.456	.256	.033	.022	1.000
	Proportion of all participants	.015	.090	.204	.114	.015	.010	.448
	Total Proportion of all participants	0.070	0.299	0.428	0.154	0.025	0.025	1.000

Note. N = 201.

Logistic Regression Analyses

In order to examine the question of why individuals chose to participate (or not participate) in the creative music-making projects, a series of regression analyses were conducted, using the data that was available. There were two ways, out of many other possibilities, in which I conceptualized the dependent variable of project participation. First, I created a simple dichotomous variable indicating whether an individual participated in any project. With this approach, an individual who participated in all four projects was coded the same as an individual that participated in only one project. Second, I created a dichotomous variable to represent individuals who participated in projects beyond the Top 10 Playlist project. While this project involved some creative choices, it is difficult to construe the Top 10 Playlist project as on par with the other three creative music-making projects, mainly because actual music (or components of a song) were not created.

Any Project Participation and Project Participation Beyond Top 10 are dichotomous dependent variables, and as such, traditional linear regression is not appropriate. Instead, I chose to examine them within the framework of logistic regression. One important difference between traditional regression (with a continuous dependent variable) and logistic regression is that the model parameters are estimated using Maximum Likelihood estimation, which requires larger samples than Ordinary Least Squares regression. This limited the variables available to use as predictors of project participation to PEM items, demographics items, and items from the Welcome to Canvas Network survey (e.g., Expected Hours/Week Spent on the Course) because these were the only items for which I had a sufficiently large enough sample size.

Following a logistic regression model building procedure called “purposeful selection,” described by Hosmer, Lemeshow, and Sturdivant (2013), I began by running a series of

univariable logistic regression models. A likelihood ratio test (LRT) was used to determine the statistical significance of each individual predictor variable. The LRT compares the model with the predictor variable included to a null model (a model without the predictor included), and this difference in log likelihood values approximates a chi-square distribution. Results from step 1 (the univariable analyses) are displayed in Table 5.21. Hosmer and colleagues recommend retaining any variable that has a p -value less than .25 during this initial screening step. This very liberal p -value is recommended so that potentially important variables are not overlooked, given they may have an adjusting or interactive effect in combination with other variables.

Using this criterion, the variables retained for step two were Education, Type of Learner and Active Participant (these are essentially two different versions of the same variable), Primary Language, and PEM class (when treated as a categorical, not a continuous variable). Step two in the purposeful selection procedure is to fit a model that contains all the predictors retained from step one within one model. Because Type of Learner and Active Participant are based on the same measurement, I estimated two different models at this step, one with Active Participant included and one with Type of Learner included as a categorical predictor. As expected, the LRT (against the null model) was significant for both overall models ($\chi^2(5) = 18.684, p = .002$ with Active Participant included, and $\chi^2(7) = 19.400, p = .007$ for Type of Learner included). Results from both models are shown in Table 5.22. In the presence of the other predictors, the only variable that remained statistically significant was Type of Learner/Active Participant.

For the third step, Hosmer and colleagues (2013) recommend to fit a new model with the non-significant variables removed, and then compare the new smaller model to the larger model. In particular, one should look for changes in parameter estimates ($\Delta\beta$) that are greater than 20%. Such a large change would indicate that one or more of the variables have an important

“adjusting” role, although may not be statistically significant in their relationship to the dependent variable. Since all but one variable was removed, the larger model could be compared to the original univariable models for Type of Learner and Active Participant. A comparison of the change in β for Active Participant (4.33%) and change in each β for Type of Learner (Observer = 19.79%, Drop-In = 11.92%, and Passive Participant = 6.47%) indicated that none of the removed variables appeared to function in an “adjusting” role.

Table 5.21

Results from Univariable Logistic Regression (Any Project Participation)

Variable	-2LL	β	SE	Exp(β)	Likelihood Ratio Test			
					Chi-square	df	<i>p</i>	<i>N</i>
Male	278.448	0.017	0.293	1.018	0.011	1	.918	237
Age	277.696	-0.034	0.100	0.967	0.114	1	.736	236
Education	278.626	-0.137	0.080	0.872	11.389	7	.123	236
Primary Reason*					5.102	8	.747	201
Type of Learner	221.472	x	x	x	18.908	3	.000	202
Type of Learner (Observer) ^a	x	-1.678	1.090	0.187	x	x	x	x
Type of Learner (Drop-In) ^a	x	-1.678	0.653	0.187	x	x	x	x
Type of Learner (Passive Participant) ^a	x	-1.282	0.363	0.277	x	x	x	x
Active Participant	221.912	1.385	0.334	3.993	18.467	1	.000	202
Expected Hours (Continuous)	239.108	0.119	0.153	1.127	0.607	1	.436	201
Expected Hours (Categorical)*					6.674	5	.246	201
English	236.262	0.652	0.326	1.919	4.117	1	.042	202
Region*					12.303	11	.341	202
PEM Class (Continuous)	223.800	-0.201	0.230	0.818	0.769	1	.381	173
PEM Class (Categorical)	218.864	x	x	x	5.705	2	.058	173
PEM Class 1 ("Highly Experienced") ^b	x	0.734	0.534	2.083	x	x	x	x
PEM Class 2 ("Moderately Experienced") ^b	x	1.119	0.505	3.061	x	x	x	x

* Model is not identified, and therefore estimates are not trustworthy, so they are not reported.

a Reference group is "Active Participant"

b Reference group is PEM class 3 ("Minimally Experienced")

Table 5.22

Logistic Regression Results from Step 2 (Any Project Participation)

Model with Active Participant Included					
Variable	β	SE	Wald Z	p	Exp(β)
Education	-0.018	0.111	-0.162	0.871	0.982
Active Participant	1.325	0.389	3.406	0.001	3.761
English	0.378	0.399	0.948	0.343	1.460
PEM Class (Categorical)					
PEM Class 1 ("Highly Experienced")	0.607	0.602	1.008	0.313	1.835
PEM Class 2 ("Moderately Experienced")	0.881	0.556	1.585	0.113	2.414

Note. $N = 140$. $-2LL = 167.056$. $AIC = 179.057$, $BIC = 196.707$, $ssBIC = 177.723$, $R^2 = .166$ ($p = .019$)

Model with Type of Learner Included					
Variable	β	SE	Wald Z	p	Exp(β)
Education	-0.027	0.112	-0.238	0.812	0.974
Type of Learner					
Type of Learner (Observer)a	-1.346	1.214	-1.109	0.267	0.260
Type of Learner (Drop-In)a	-1.878	0.814	-2.306	0.021	0.153
Type of Learner (Passive Participant)a	-1.199	0.419	-2.860	0.004	0.301
English	0.342	0.405	0.846	0.398	1.408
PEM Class (Categorical)					
PEM Class 1 ("Highly Experienced")b	0.625	0.602	1.038	0.299	1.869
PEM Class 2 ("Moderately Experienced")b	0.912	0.560	1.630	0.103	2.490

Note. $N = 140$. $-2LL = 166.34$. $AIC = 182.341$, $BIC = 205.874$, $ssBIC = 180.563$, $R^2 = .177$ ($p = .018$)

At this point, two roughly equivalent models remained, one with Type of Learner and one with Active Participant as a predictor of Participation in Any Project. Since these predictors are essentially the same thing, I examined the fit statistics (BIC, AIC, and ssBIC) between the two models to determine which model to retain. The model with Active Participant had lower values

on all three indices, which indicates it is the better fitting model. This model is also more advantageous because it is more parsimonious, requiring only a distinguishing between those that identified themselves as “Active Participants” and those that identified themselves as something else. The odds ratio for Active Participant is 3.993, which means that those that identified themselves as being an “active participant” were about 4 times more likely to participate in any project than those that identified themselves as something else (e.g., observer, drop-in, or passive participant).

All that being said, the Nagelkerke R^2 value (which is really a pseudo- R^2 statistic) for the Active Participant predicts Participation in Any Project model was only .126. The R^2 value for the Type of Learner model was of similar magnitude, .132. Though statistically significant ($p = .02$), these values indicate there are many variables missing from the models.

An identical analysis to that describe above for Any Project Participation was also done for Participation Beyond Top 10 (participating in one of the other three projects beyond the Top 10 playlist project). The results were nearly identical, with Type of Learner/Active Participant emerging as the only significant predictor of Participation Beyond Top 10.

Type of Learner Group Differences

Given the apparent statistical importance of the Type of Learner variable in predicting project participation, as uncovered in the logistic regression analyses, I decided to pursue this variable a little further and search for differences between those that identified themselves as “active participants” and those that identified themselves as something else (passive participant, observer, drop-in). It is important to note that from this point forward, when I use the term “active participant” I am referring to those individuals that chose “active participant” from the four answer options of the Type of Learner question. I am not referring to those that actually

participated in any of the course projects. This is important because the “active participants” I am referring to are those that indicated an *intent* to participate, not those that *actually* participated (although certainly some “active participants” did in fact participate). I will continue to use this term inside of quotation marks so as to help remind the reader that it refers to what an individual has called himself/herself and not what he/she has actually done in the course.

Continuous variables. The results from a series of independent t-tests for all continuous variables are reported in Table 5.23. There was a significant difference in mean scores for all three MCSE factors between “Active Participants” and those that did not identify themselves as “Active Participants.” As can be seen in Figure 5.9, the differences were quite substantial, ranging from 37.57 for the Components factor, to 38.26 for the Improvise factor, to 49.93 for the General MCSE factor. The correlations between Active Participant and each MCSE factor were correspondingly large (.741 for general, .670 for components, and .570 for improvise), indicating a very strong relationship between MCSE and Active Participant.

Two other significant group differences were found for continuous variables: the Personality factor for Openness, $t(62) = 2.3, p = .025, r = .280$, and the Contextual Support factor for Challenge, $t(13) = 2.32, p = .037, r = .542$. It should also be noted that several other variables had substantial, although not statistically significant differences. Namely, the means for both Identified Regulation and External Regulation were notably higher for the Active Participants, with very high correlations of .588 and .553, respectively. The lack of statistical significance is likely a result of having only five participants that completed the motivation items. Active Participant was also highly correlated with Situational Engagement ($r = .461$) and both the Feedback factor ($r = .464$) and Competence factor ($r = .300$) of Contextual Support.

Again, the lack of statistical significance may have resulted from very low power associated with a very small sample size.

Table 5.23.

Results from Independent T-Tests for “Active Participant” Group Differences

	Variable	<i>N</i>	<i>t</i>	<i>df</i>	<i>p</i>	Mean Difference	<i>SE</i>	95% CI of Difference		<i>r</i>
								Lower	Upper	
MCSE	General	24	5.18	22	.000	49.93	9.643	69.93	29.93	.741
	Components	24	4.23	22	.000	37.57	8.887	56.00	19.14	.670
	Improvise	24	3.44	17.80	.003	38.26	11.133	61.67	14.85	.570
	PROMS Overall (4-factor mean)	10	0.23	8	.826	0.55	2.424	6.14	-5.04	.080
Personality	Extraversion	63	-0.28	61	.779	-0.07	0.239	0.41	-0.55	-.036
	Agreeableness	64	0.80	62	.424	0.13	0.155	0.44	-0.19	.102
	Conscientiousness	64	1.36	62	.180	0.29	0.213	0.71	-0.14	.170
	Neuroticism	64	-1.28	62	.205	-0.28	0.216	0.15	-0.71	-.161
	Openness	64	2.30	62	.025	0.35	0.152	0.65	0.05	.280
	Honesty-Humility	64	0.20	62	.844	0.04	0.200	0.44	-0.36	.025
Contextual Support	Challenge	15	2.32	13	.037	0.66	0.282	1.27	0.05	.542
	Choice	15	-0.69	13	.503	-0.26	0.382	0.56	-1.09	-.188
	Competence	15	1.13	13	.277	0.54	0.476	1.57	-0.49	.300
	Curiosity	15	-0.19	13	.851	-0.08	0.441	0.87	-1.04	-.053
	Feedback	15	1.80	7.84	.111	0.70	0.388	1.59	-0.20	.464
	Interest	15	0.18	13	.858	0.09	0.515	1.21	-1.02	.050
Motivation	Intrinsic	5	0.28	1.01	.825	0.61	2.170	27.84	-26.62	.213
	Identified Regulation	5	1.02	1.21	.468	1.61	1.576	14.98	-11.76	.588
	External Regulation	5	1.15	3	.333	1.67	1.449	6.28	-2.94	.553
	Amotivation	5	-0.36	3	.743	-0.50	1.389	3.92	-4.92	-.203
	Situational Engagement	5	0.90	3	.434	0.75	0.833	3.40	-1.90	.461
	Perceived Learning (General)	13	0.42	11	.681	0.38	0.894	2.35	-1.59	.126

Note. MCSE = Musical Creative Self-Efficacy. Variables that had a significant Levene's test used an adjusted *df* to account for unequal variances between groups. The correlation, *r*, between Active Participant and each variable is the point-biserial correlation.

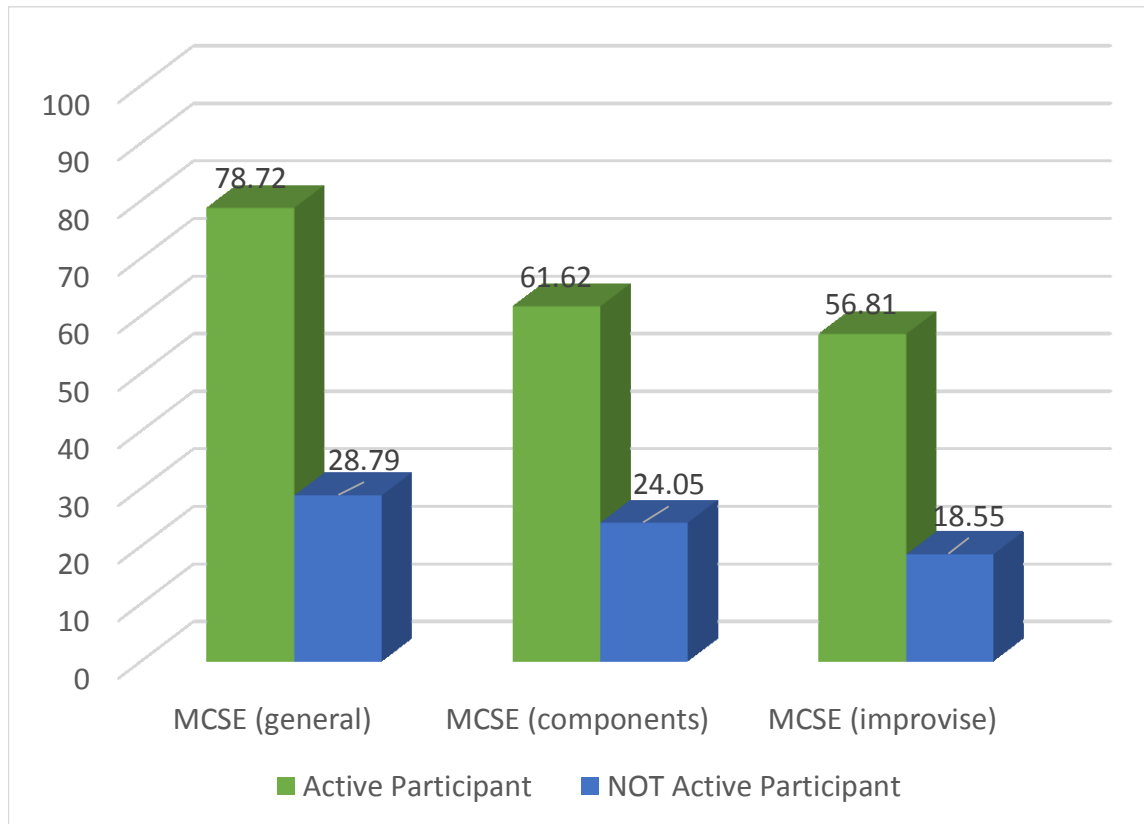


Figure 5.9. Differences in Scores for MCSE Factors Between "Active Participants" and Not "Active Participants"
Note. $N = 24$.

Categorical variables. For the analysis of group differences between “active participants” and not “active participants” I treated PEM and all demographic variables as categorical. Since Active Participant was also a categorical variable, I conducted a series of Chi-Square tests of independence. In addition, for variables that are not truly nominal, but represent ordered categories (PEM, age, education, and Hours per Week), Kendall’s tau was calculated as a measure of association, which treats both Active Participant and the respective comparison variable as ordinal variables.

The results of these analyses (see Table 5.24) indicated statistically significant differences between “Active Participants” and not “Active Participants” in regards to both the Number of Hours per week they expected to invest in the course and whether English was their primary language. When viewed as ordinal level variables, the relationship between Hours per Week and Active Participant is a weak and positive relationship (Kendall’s tau = .266).

Interestingly, the null hypothesis that the number of individuals within a particular PEM class were the same between those that identified themselves as “active participants” and those that did not identify themselves as “active participants” was not rejected, $\chi^2 (2) = 0.345, p = .841$. As such, there did not appear to be a relationship between PEM class and whether an individual planned to be an “active participant”.

Finally, although the Chi-Square test for Education was not statistically significant, the Kendall’s tau correlation of -.146 was statistically significant. This general trend is depicted in Figure 5.10, indicating that “active participants” tended to be less educated than non-“active participants.”

Table 5.24

Results from Chi-Square Tests for “Active Participant” Group Differences

Variable	N	χ^2	df	p	Kendall's tau
Male	200	0.074	1	.786	.019
Age (ordinal)	202	6.606	6	.359	-.007
Education	198	10.753	7	.150	-.146*
Primary Reason	201	10.617	8	.224	--
Hours/Week	201	40.040	5	.001	.266***
English	202	5.321	1	.021	--
Region	202	11.951	11	.367	--
PEM class	140	0.345	2	.841	-.030

Note. Kendall's tau is not reported for nominal variables. For ordinal variables, Kendall's tau treats Active Participant as an ordinal variable. * $p < .05$, ** $p < .01$, *** $p < .001$

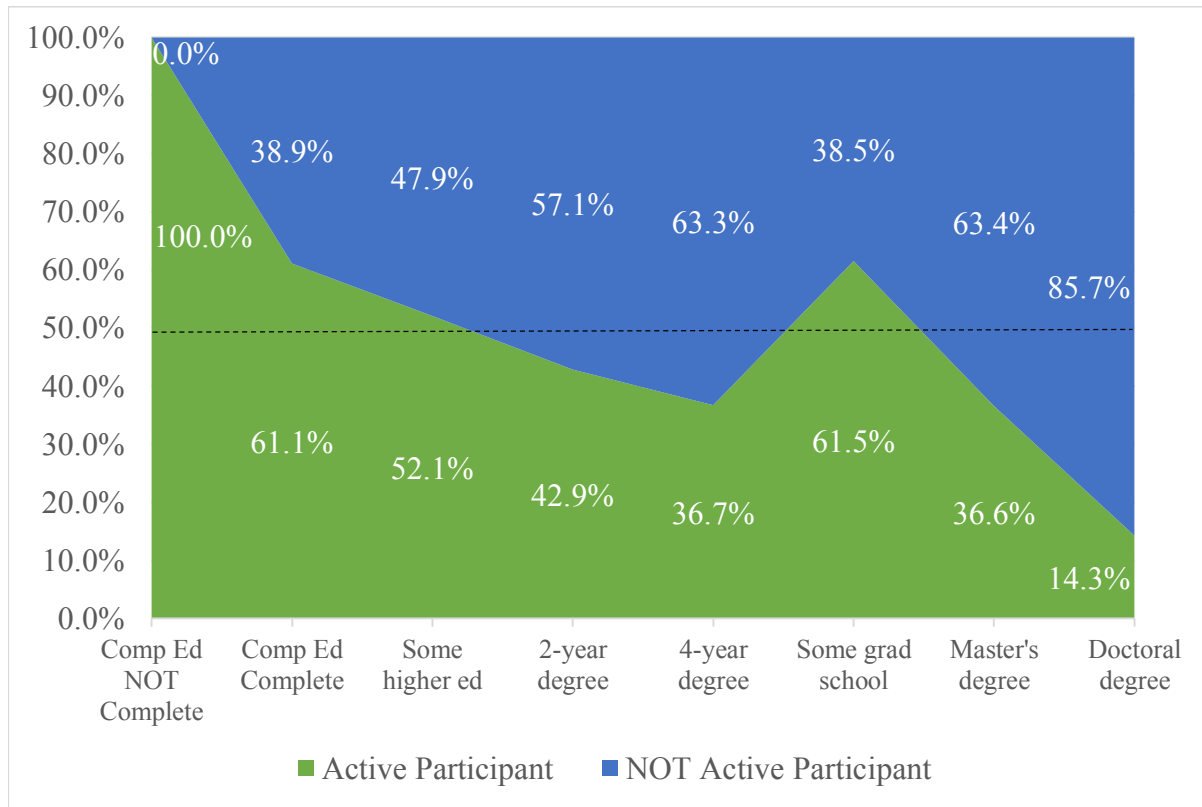


Figure 5.10. Percentages of "Active Participants" and Not "Active Participants" by Education *Note.* For example, of the participants that have a 2 year degree, 42.9% of them identified themselves as "active participants," while 63.3% of them identified themselves as one of the other four type of learners (passive participant, observer, or drop-in).

The General Specified Model Explored

After a targeted analysis of group differences in Active Participant, the final set of analyses widened the focus to all variable relationships. As was explained in the **Primary study data analysis overview** section of Chapter Three, Kendall's tau, a rank-order correlation coefficient, was calculated for all non-nominal variables. In effect, all variables in this analysis were assumed to be ordered categorical (ordinal) at a minimum. These intercorrelations can be found in Appendix G. Before presenting some important features of the correlation analysis, several points need to be made regarding statistical significance.

Statistical significance and missing data. It is important to note that the correlations shown in Appendix G were calculated using pairwise deletion for missing data, which means that the sample sizes are drastically different for each pairwise correlation, ranging from $N = 237$ to $N = 5$. This is particularly important if the goal was to identify only statistically significant correlations because a correlation of, for example, .40 may be significant with a sample size of 200, but not with a sample size of 14. This makes it difficult to interpret the implications of significance test results between two tests of drastically different sample sizes.

Furthermore, conducting significance tests on all pairwise correlations of 34 variables results in a total of 595 individual significance tests. Clearly, this results in a massive inflation of the nominal type I error rate of .05, so much so that the use of a correction procedure (e.g., Bonferroni), would result in such a huge loss of power that the testing becomes an utterly futile endeavor. For example, a Bonferroni correction for 595 tests would adjust the nominal alpha level of .05 to .000084. For these reasons, the use of statistical significance testing is not particularly useful for the present purpose.

A rough sketch. The primary purpose of this exploratory analysis is to gain a rough picture, more of a description of the data. As such, the results below are reported using more qualitative descriptors (e.g., large/small, high/low, positive/negative) instead of quantitative values. To assist in developing a rough sketch of the data overall, the correlation matrix in Appendix G is displayed as a heat map, with positive correlations represented by red and negative correlations represented by blue; the absolute value of the correlation is represented by change in hue, with darker hues representing higher absolute values. The advantage of this type of visualization is that patterns of correlations that are generally similar (in strength and direction) can be easily seen.

Before proceeding, it should also be made clear that this analysis is based solely on zero-order correlations. It will become clear that there is a fairly large degree of correlation between the various components of the General Specified Model with these data. Ideally, we would examine the relationships between these components after partialling out their shared variance (and also within a latent variable framework to account for measurement error). This was not possible for this study, which is another reason why I have chosen not to discuss the relationships in terms of quantitative value, but rather in terms of direction and qualitative magnitude. I will discuss each of the major components (e.g., motivation, personality) in turn.

Demographic variables. Gender has generally low correlations with all variables, with the exception of Identified Regulation, Amotivation, and Situational Engagement. These three specific correlations were based on only five participants. Age does not exhibit any clear patterns of relationships with other demographic variables (except Education), but does have a strong negative relationship with Music Aptitude, but again this is based on a small number of individuals ($N = 11$). Of particular interest is the pattern of relationships with Motivation showing an increasingly negative relationship with more autonomous forms of Motivation (e.g., Intrinsic Motivation), but an increasingly positive relationship with more controlled forms of Motivation (e.g., Amotivation). Age is also strongly negatively associated with Situational Engagement. Both Motivation and Engagement correlations are based on only five individuals.

The Number of Hours per week that an individual planned to participate in the course has a generally low positive correlation with MCSE, all Contextual Support factors, Situational Engagement, and all Project Participation variables. Interestingly, the strongest relationship for Hours per Week is with Perceived Learning ($N = 13$). The variable English (whether English is an individual's primary language) follows the same general pattern of relationships as Hours per

Week, with two notable exceptions. English has a strong *negative* relationship with Perceived Learning as well as weak negative relationships with all but one Contextual Support factor (Challenge).

Past experience in music. Recall that PEM classes were coded such that the higher numbered class represents less PEM, which means that a negative correlation actually represents a positive relationship with past experience. Based on the patterns in the correlation matrix, more experience is related to more MCSE, higher Music Aptitude, a higher score on the Openness factor of Personality, higher scores on all Contextual Support variables (except Choice), higher Identified Regulation, lower External Regulation and Amotivation, and higher Situational Engagement. Of these relationships, the only ones that are particularly strong are with the controlled motivations (ER and AM) and Music Aptitude.

Musical creative self-efficacy. What is immediately apparent when looking down the column and across the row of correlations for MCSE is the depth of color, indicating many moderate to strong relationships. All MCSE factors are highly positively interrelated. MCSE (all three factors) has a generally weak positive relationship with Extraversion, Agreeableness, and Openness, but generally very weak relationships with the other personality factors. MCSE is also positively related to all Contextual Support factors, both autonomous forms of Motivation, Situational Engagement, Perceived Learning, and all Project Participation variables, but very strongly negatively related to Amotivation.

Personality. In addition to the relationships already discussed, the most notable relationships for personality are a whole series of negative relationship between Conscientiousness and the following: all Contextual Support factors, the autonomous forms of Motivation (IM and IR), Situational Engagement, and Perceived Learning. This general pattern

is roughly paralleled by the Neuroticism factor, although with much weaker strength of relationships. Additionally, Agreeableness and Openness appear to have the strongest relationships (in either direction) between Contextual Support, Motivation, Situational Engagement, and Perceived Learning. Also, in regards to Contextual Support, we can see that the box for all Personality factors is generally similar in depth of color to Demographics, but certainly lighter in color compared to other components such as Motivation and MCSE.

Music aptitude (PROMS). Higher scores on the PROMS are very strongly associated with higher IR Motivation scores, lower controlled Motivation scores (ER and AM), higher Situational Engagement, and higher Perceived Learning. PROMS scores are also positively, but less strongly, associated with MCSE and all Contextual Support factors, except for Choice. There also appears to be strong negative correlations between PROMS scores and both Age and PEM class. Recall that the negative correlation between PROMS and PEM actually represents a positive relationship (higher PROMS score associated with more experience).

Contextual support. The box representing relationships between Contextual Support factors is light to moderately dark red, with no blue whatsoever, indicating all positive relationships between the components of Contextual Support. None of the relationships are so strong as to suggest conceptual redundancy, but the presence of a common antecedent variable or a shared general factor (as in a bi-factor model) could be further explored. For example, the box representing relationships between Contextual Support and MCSE has similar shades of red (no blue) as the box representing Contextual Support interrelationships. In the General Specified Model, both Contextual Support and MCSE are viewed as contextual level factors, but independent of each other. However, it is possible that there is a causal relationship between the two such that individuals who are confident in their creative music-making capabilities perceived

this course context as fitting well with what they are capable of doing precisely because this course was designed to support autonomy. Individuals who were less confident in their abilities might have found the course to be restrictive and unsupportive because it was less prescriptive; it is the idea of being paralyzed by too many choices or options.

It is also plausible to view MCSE and Contextual Support as mostly independent components whose correlations can be explained by a third component, a common cause. Music Aptitude may be a candidate worth considering in this scenario. For the participants that completed both the PROMS and the Contextual Support questionnaire ($N = 5$), their scores are moderately related to MCSE and moderately to strongly related to all Contextual Support factors, except for Choice. As noted above, there is also a fairly strong relationship between PROMS scores and PEM. The point here is that these relationships, if they are also present in other samples, need to be teased apart more. To do so via a structural equation modeling approach would require a much larger sample, as previously mentioned.

The relationships between Contextual Support and the situational factors in the General Specified Model (motivation, engagement) generally follow the pattern anticipated by the model. For example, the relationship between Contextual Support and Motivation is positive for autonomous types of Motivation (IM and IR), and either weakly positive or increasingly negative for less-autonomous (controlled) forms of Motivation (ER and AM). This follows the basic tenets of Self-Determination Theory, that the extent to which the environment supports the three basic psychological needs of autonomy, competence, and relatedness will determine the extent to which individuals are autonomously motivated. If the four types of Motivation measured in this study are viewed along a continuum from autonomous to controlled, SDT would predict the strongest relationships with Contextual Support to be with IM and AM, the former being

strongly positive and the latter being strongly negative. The general positive/negative pattern is clearly present in these data, although interestingly, the Identified Regulation factor, not the Intrinsic Motivation factor, has the strongest positive relationship with all Contextual Support factors, except for Competence.

Finally, there is a clear moderate to strong positive relationship between Contextual Support and both Situational Engagement and Perceived Learning. This is important because these relationships are at the heart of the General Specified Model. They are discussed in more detail below.

Motivation. Several relationships with Motivation and other components have already been presented above, so this section will focus on those relationships that have not been presented. To begin, the basic pattern of correlations between Motivation factors that is posited by SDT seems to be evident in the data, even if the Motivation data are from only five individuals. The two autonomous forms of Motivation are positively related, and the two controlled forms of Motivation (ER and AM) are positively related. Relationships between autonomous forms of Motivation and non-autonomous forms of Motivation are negative. This is precisely what SDT predicts. Furthermore, we see positive relationships with the autonomous forms of Motivation and both Situational Engagement and Perceived Learning, but negative relationships with the non-autonomous forms of Motivation. Again, this follows the premises of SDT.

One final result to note is that the pattern of correlations between the Motivation factors and Project Participation is different for each Project Participation variable. Whether an individual participated in any project (generally this meant they completed at least the Top 10 Playlist assignment) seems to be more related to a different form of Motivation than whether an

individual participated beyond the Top 10 playlist project (participated in at least one of the projects that required creation of lyrics, melodies, or a song). More specifically, it is possible that initial participation was based on Intrinsic Motivation, the desire to do something because it is enjoyable, but once the projects involved more work (something that may be viewed as less than pure enjoyment), the individuals that did continue to participate drew on more external sources of Motivation such as the feeling of obligation (External Regulation) or a belief that completing the project would be “good for me” (Identified Regulation). This would be an important area for further research, particularly in the MOOC context.

Situational engagement and perceived learning. Finally, the relationship between situational engagement and perceived learning was positive and quite strong. This lends some initial support to the notion that engagement fosters the perception of learning, although alternative explanations are certainly still possible. For example, Situational Engagement and Perceived Learning are both highly related to several other components, including MCSE, Music Aptitude, and Contextual Support, and several forms of Motivation, either of which could be potentially argued to be a common cause of both Situational Engagement and Perceived Learning. Partialling out the unique variances and examining different sorts of mediational models is possible with structural equation modeling, but would require a much larger sample, so these possible explanations are left for further research.

This concludes the presentation of results from the various analyses conducted on the data from this study. The following chapter considers the meaning and implications of the results presented in this chapter and also highlights many questions that have arisen from this study.

CHAPTER SIX: DISCUSSION

In considering the results from this research, there is much to be learned and much to be contemplated, questioned, and re-visited. I will begin with a discussion of the issues and results related to measurement of the constructs considered in this research. Next, I will discuss the general issue of creative music-making project participation in the course, and what it means in terms of future research in the MOOC environment. Finally, I will identify many more questions that have arisen from this research, both within and outside of the MOOC learning context.

Contributions and Issues with Measurement

A substantial amount of effort was put into selecting and/or developing appropriate, reliable, valid, and logistically feasible measures of the many constructs that I attempted to examine with this research. As such, there is as much to be said about the issues related to measurement in this line of research as there is to be said about the results for the primary research questions. I will discuss both the PEMI and the MCSES in turn.

Past experience in music. I view Past Experience in Music (PEM) as roughly analogous to the concept of a patient history in nursing and medicine. There are many different events and incidences that make up an individual's history, some of which may have some bearing to the current situation and some of which may not. To continue the patient history analogy, whether my father had heart disease is certainly a part of my patient history, as is whether I have been hospitalized for any particular reason, even though the two things might not be related in any easily noticeable way. In some cases, a particular event may have direct relevance (e.g., having

a history of high blood pressure may have an impact on my current heart health), but in other cases it may have an indirect, perhaps unmeasurable impact, on the current interest of a study (e.g., having a major lower body injury, which limits my ability to be active, which in turn causes me to gain weight, which in turn has an impact on my current heart health and susceptibility to heart disease). As such, one could sequentially look at individual direct effects of each individual historical event (e.g., a previous heart attack or a major lower body injury) on predicting certain outcomes (e.g., a future heart attack). But in following this approach, one may identify a particular variable that does not appear to have a direct measurable effect (e.g., major lower body injury) and unwittingly remove the variable from consideration. This is problematic when that variable works in some unmeasurable way in combination with other variables to produce the effect that appears to come from the one (or a few) “significant” variables.

Regarding the measurement of PEM, rather than focus on individual effects of specific previous experiences (e.g., having solo performance experience), I chose an approach that focuses less on identifying specific *items* that are more or less important than others, but rather focuses more on using the information that is available to identify *individuals* that are different than each other in substantive ways, based on the combination of responses to all items. The logic of Latent Class Analysis (LCA) is to identify groups of individuals based on the patterns of responses to all items included in the analysis. On the other hand, Exploratory Factor Analysis (EFA) treats all individuals as belonging to one homogeneous group, and then attempts to identify factors (combinations of items) based on the response patterns of all individuals. With EFA, the underlying factor(s) is continuous (i.e., individuals exist along a continuum on the factor; more or less of the factor), whereas with LCA the underlying factor is categorical (i.e., individuals belong to one category or another). LCA is common in the field of epidemiology

(the study of patterns and causes of diseases) because the goal is to identify patterns of symptoms for predicting (and defining) a particular disease, which is fundamentally about separating people into different groups (e.g., those with the disease, those that are at-risk, those that definitely do not have the disease and are not at-risk, etc.).

For PEM, my initial decision to use LCA instead of EFA was based on a belief that an individual's response patterns for the binary (yes/no) PEM items would be indicators of the individual's *type* of PEM, and less so of the individual's *amount* of PEM. This made intuitive sense when I considered, anecdotally, the experiences of my various friends, family, and colleagues. I knew that some of them had large amounts of experience in large ensembles and private lessons, but minimal experience in composing and improvising. Others had very minimal formal training, but had lots of experience with small ensembles and would create music and improvise on a regular basis. Even if these were the only two "types" of people there are in regards to PEM, to treat them as existing along some continuum of PEM seems counterintuitive. A simple sum of their yes/no responses would result in similar PEM scores, and therefore would not function to differentiate between these two very different types of experience. As such, I chose to use LCA to help identify the different *types* of experience instead of *different* amounts of experience.

However, even in choosing a latent class measurement model, it is interesting that the three classes that were emerged from the data still seemed to exist along some sort of continuum, coarse as it may be (e.g., highly experienced, moderately experienced, minimally experienced). Perhaps even more interesting, this 3-class model, with seemingly ordinal groups, was still retained even after controlling for gender, age, and education. There are several issues to consider with this result.

First, in choosing to retain the 3-class models, I made a conscious decision to be more conservative in my interpretation of the model fit and model comparison statistics in an attempt to avoid overfitting the data, particularly given the sample size. All statistical decisions have consequences. There was certainly statistical evidence that could have been used to support a decision to retain a model with 4 or more classes. This would have resulted in less interpretable classes/groups and the results became less stable (particularly for the 5-class models) due to the addition of more parameters to be estimated in the model without a sufficiently large sample size. In other words, the model might fit the data better, but it also increases the extent to which the model estimation capitalizes on chance, an issue of primary importance when sample sizes are only marginally large.

Second, while I was attempting to separate individuals into different types of PEM, the resulting 3-class model did seem to indicate some evidence for a continuum along which these three classes lie. This has important implications for how the class assignments are treated in terms of level of measurement. If they are truly unordered categories/groups, then class membership should be treated like a nominal variable (e.g., by dummy coding), but if they are ordered groups, then the actual level of the class is meaningful (e.g., class 2 represents “more” or “less” of something compared to class 1 or class 3). The decision about whether to treat class membership as ordered or nominal has important implications for how results would be interpreted.

Third, LCA utilizes a conditional independence assumption for the items with the latent class. This means that the items are mutually independent given membership in a particular class, or put another way, once the variance from the categorical latent variable is partitioned from the items, any remaining unexplained item variance (residual variance) is assumed to be

uncorrelated with other items. This may be too strict an assumption for the items in the PEMI. More research and theorizing needs to be done to investigate this assumption as it relates to the PEMI items. If it is found to be an unreasonable assumption, there are other models available to relax this assumption, including factor mixture models (FMM) (e.g., Lubke & Muthén, 2005).

Third, I chose a measurement model based solely on the binary (yes/no) PEM items. I had data for individuals' frequency, recency, and number of years of experience for each item to which they answered "yes." The issue with utilizing these data is deciding on a means to best combine it in a meaningful and statistically appropriate manner. For example, for any given type of experience (e.g., improvising experience), should the frequency value be added to the recency value or should the two be multiplied by each other? I considered at least five different ways in which these two values could be combined for each type of experience. The five different approaches I considered were (1) summing frequency and recency scores; (2) multiplying frequency and recency scores; (3) multiplying frequency and recency scores, and then taking the square root of that product; (4) taking the natural log of the frequency and recency score, and then multiplying those two log values together; and (5) taking the square root of the frequency and recency scores, and then multiplying those two square root values together. Each approach results in very different final scores and also very different distributions of scores (see Appendix H for graphical representations of these five different approaches to combining frequency and recency).

In the end, I decided that a thorough analysis of these different approaches, including the inherent assumptions and implications of each, was beyond the scope of this study. I include them here simply as documentation of ideas that need further deliberation and further research. Furthermore, a latent class analysis that utilizes continuous variables instead of categorical

variables (often called a latent profile analysis) is also much more computationally demanding. Given the small sample size and relatively little theory regarding the combination of frequency and recency items (let alone also combining “number of years” items), I decided to focus solely on the LCA measurement model using the dichotomous yes/no items. Needless to say, this approach to measuring PEM needs further research, both in terms of choosing appropriate measurement models and in terms of incorporating the additional frequency, recency, and number of years information.

Musical creative self-efficacy. The second area that I believe may contribute to measurement in musical creativity is the development and initial validation of a measure of Musical Creative Self-Efficacy (MCSE). While the measure exhibited marginally good fit as a measurement model in Pilot Study #2, there is much more research to be done on this measure. As with the latent class analysis for PEM, the exploratory factor analysis of the MCSE items did not result in unambiguous evidence regarding the number of factors. This could be indicative of many items that are not conceptually distinct enough from one factor to another (resulting in high cross loadings). It could also indicate a misspecified model. For example, items 1-3 were written to measure a general aspect of self-efficacy within musical creativity, and items 13, 17, 18, and 19 were clearly written to measure self-efficacy related to improvisation. However, items 7-12 were actually initially written to address two separate factors. Items 4-8, of which only 7 and 8 were retained, were intended to measure MCSE for specific components of music (e.g., writing melodies, creating accompaniments). Items 9-12 were intended to measure MCSE for activities that would fall under the general category of arranging, that is, re-working ideas originally created by someone else.

Another possibility that could not be explored further due to problems with model convergence is the bi-factor model. In a recent meta-analysis, Seltzer (2013) found evidence for a possible domain-general self-efficacy. A bi-factor model would account for the possibility of a general self-efficacy factor that impacts all items. Since the bi-factor model could not be examined with these data, future research will need to consider this possibility.

Finally, since not enough individuals completed the measure in the MOOC study, it was not possible to compare factor analytic results with Pilot Study #2. Nevertheless, some preliminary evidence of convergent validity was found in the significant, moderately large positive correlations between PEM and MCSE. The relationship was examined more closely using a MIMIC model. Results again supported the relationship between PEM and MCSE, with differences in PEM class being associated with fairly large factor mean differences in MCSE. Additional research on the MCSES will need to confirm the factor structure and investigate different forms of validity. For example, divergent validity could be established by comparing scores from the MCSES to scores from measure of musical performance self-efficacy. Predictive validity could be examined by relating MCSE scores to a variable that MCSE should theoretically predict, such as change in MCSE after master experiences (e.g., successfully writing a song for the first time).

Other considerations and future measurement research. So far I have addressed several contributions and issues related to the actual analyses I conducted. However, there are many other types of analysis that I potentially could have done, or if not with the current data, will need to consider with future research.

Measurement invariance and the PEMI. The latent class analysis on the PEM yes/no variables assumed measurement invariance, that is, the measurement model is the same for

individuals with different characteristics (e.g., age, education, gender, primary language, geographic region). However, it is feasible to argue that the items themselves may not function the same way for all people, which would be a lack of measurement invariance, and would represent bias in the measure. For example, it is possible that what people in East Asia interpret the term “formal training” to mean is different than what people in North America understand the term to mean. If this is the case, then the probability of answering “yes” to that item, given membership in the latent class, may be different across individuals in different regions. This results in a situation where the meaning and interpretation of the latent classes are not the same for individuals from different regions. This is particularly problematic in the MOOC context in which individuals enroll from all around the world, but an analogous problem could occur just as easily with school age students of different ages (e.g., younger students interpret the meaning of the items differently than older students). These sort of potential issues with measurement invariance need to be explored further if the PEMI is to be a useful research instrument.

Measurement invariance and the MCSES. Similar to the issues of measurement invariance in PEM latent classes, lack of measurement invariance in latent factors can also be an issue in MCSE factors. Again, the issue is potentially compounded in the MOOC context because individuals come from potentially very different cultures, and may interpret the items differently. Just as the instrument may be biased across geographic regions or languages, it is also possible that the instrument could be biased across PEM class, or more specifically, across different levels of musical education and/or musical expertise. For example, an individual who is a professional instrumentalist may hold him/herself to higher performance standards, so when asked if he sings, he may answer no because he cannot sing at the same level as his instrumental performance abilities, even though he likely has much experience with singing.

Issues with the latent class measurement model and the PEMI. In the current study, I utilized a latent class measurement model in order to produce a measurement (a numeric value) from the PEMI item responses that was based on an individual's most likely class membership. The measurement was the specific PEM class to which they were assigned. There are two issues to consider in regards to using this measurement model with this instrument. First, in studies with smaller sample sizes, a latent class analysis will be less appropriate because the maximum likelihood estimation procedure requires large sample sizes to produce good estimates. In this situation, the researcher would need some sort of non-latent approach to assigning individuals into PEM classes using the observed data. This would require some sort of pre-determined formula that assigns individuals based on specific response patterns. In the latent factor model, it is common for researchers to create composite variables for a factor by taking the sum or the average of all items presumed to measure a given factor. In this case the formula is very straightforward. This is not the case when working within a latent class framework.

Second, LCA is analogous to EFA in that both are exploratory procedures, and therefore a model that fits one dataset may not be the ideal model for another dataset. For example, I retained the 3-class model (conditioned on three demographic covariates), but data collected by another researcher might fit better with a four class or five class model, and will also vary depending on the inclusion of covariates and the manner in which those covariates are measured. This results in latent PEM classes that are not comparable between the two studies. In effect, the PEMI provides the initial input for the measurement, but the sample-specific characteristics are what determine the actual measurement. This is not what occurs when one uses a measurement instrument based on a continuous latent factor model. For example, when I used the mini-IPIP6 to measure personality factors, I created composite scores for each factor based on a priori

knowledge that certain items are intended to measure certain factors. The data did not dictate which items I chose to average together to create the score for each factor. The assumed factor structure is what dictated that the four extraversion items were the only items I would use to measure the extraversion factor.

The point here is that the use of LCA with the PEMI is useful in that it sorts individuals into the latent classes that are evident in the data from a specific sample. However, in order to create an analogous situation to instruments that assume a continuous latent factor measurement model, a confirmatory latent class analysis (CLCA) is necessary. In essence, the number of latent classes and other parameters such as the probabilities associated with each item for each class are specified in advance. This corresponds to CFA in which the paths from factors to items are specified in advance, and some paths are constrained to zero (e.g., with the mini-IPIP6, the agreeableness items do not load on the extraversion factor). Unfortunately, while CFA is fairly straightforward to conduct, CLCA is notably more complicated. That being said, it is certainly possible to create formulas (and syntax for different programs) that essentially create a specific latent class structure and then produce the measurements based on the pre-specified model and pre-specified parameters. This would require much more research.

Latent class analysis with type of learner. Finally, the Type of Learner variable can be conceived within a latent class framework as well. If we assume that there are certain characteristics that are indicators of whether an individual identifies himself as an active participant or not, then we can use latent class analysis⁵ to uncover what those characteristics are, and what variables function as indicators of those characteristics. If we already know an

⁵ I believe the Complier-Average Causal Effect (CACE) model(s) described by Muthén (2002) are one way to accomplish this. CACE modelling is a special case of the general latent variable modeling framework that incorporates both categorical latent classes and continuous latent factors.

individual's latent class membership in advance (e.g., "active participant" or not "active participant"), then this information (known as *training data* in the general latent variable modeling framework) can be incorporated into the model. Such an analysis could go a long way to uncovering more information about MOOC learners underlying motives, goals, and actual behavior in a MOOC learning context. I should also note that, as I understand it, this sort of analysis could be carried out in any context where the goal is to identify indicators of latent class membership, given a priori knowledge of class membership.

Moving beyond implicit measurement and structural modeling. The purpose of the discussion to this point has been to address several issues with the analysis and treatment of data from the present study, and also to suggest future avenues that could be explored from a measurement perspective. Beyond the specific instruments I have discussed here (PEMI and MCSES), I hope that this discussion has helped illuminate the need to be more explicit in specifying a measurement model for constructs, and not simply combine items together "willy-nilly" for the sake of creating a single variable that is easier to deal with in further analyses. In my experience thus far, it is all too common that researchers implicitly assume an underlying measurement model without carefully specifying the assumptions, let alone examining the appropriateness of the assumptions. For example, even the calculation of a Cronbach's Alpha as an estimate of reliability is bound up with many assumptions about the relationships of the items to the construct being measured (e.g., unidimensionality, essential tau-equivalence, uncorrelated errors)⁶. In other words, a particular measurement model is assumed when alpha is calculated,

⁶ Graham (2006) provided an excellent explanation of the essentially tau-equivalent model in regards to Cronbach's alpha.

an issue I have addressed in relation to the measurement of creative musical products using the Consensual Assessment Technique (Stefanic & Randles, 2014).

Similarly, when a researcher takes an average of several items intended to measure a single factor, by using the resulting average score as a measure of a construct, several assumptions are made about the underlying measurement model. In many cases, these assumptions are woefully unexplored, but religiously employed. Admittedly, I have done exactly this in this dissertation in calculating composite scores for many of the variables like Personality, Motivation, and Perceived Learning. The sample size for the items was prohibitively small to warrant a trustworthy examination the assumptions in a latent variable modelling framework (e.g., CFA in Mplus). However, I also intentionally limited my use of these variables in further analyses, precisely because of this. In any event, the point thus far is to highlight the need for more measurement research (i.e., research on the measurement qualities of research instruments), but also for music education researchers to make assumptions regarding the measurement model of a construct explicit, and to empirically evaluate those assumptions when possible with one's own data.

Having discussed several issues with measurement for some of the constructs I examined in this study, I turn now to the substantive results of the research.

Creative Music-Making Project Participation

Attempting to answer the question of why individuals chose to participate or not participate in the creative projects was difficult given the limited data. By looking at the data through an exploratory lens, I was able to uncover at least one important predictor: whether the student identified him/herself as an “active participant” at the beginning of the course. This result is important for at least two reasons.

Validity of the type of learner framework. First, this result provides at least some partial empirical validation of the framework around which this question from the Welcome to Canvas Network survey (“Which type of online learner best describes you?”) was developed (Hill, 2013). That is, people seemed to be fairly accurate (or perhaps realistic) in their expectations of themselves regarding participation in the course. Figure 5.7 illustrates that the proportion of people that identified themselves as “active participants” and participated in at least one project is clearly different than the proportion of “actual participants” in any of the other three types of learners. This is what the logistic regression results revealed.

An examination of the odds ratios, $\text{Exp}(\beta)$, in the logistic regression results in Table 5.21 indicates that observers and drop-ins had an odds of participating that was .187 times lower than active participants. Passive participants had an odds of participating that was .277 times lower than active participants. Put another way, observers and drop-ins were 5.34 (i.e., $1/.187$) times more likely to NOT participate compared to active participants, and passive participants were 3.61 (i.e., $1/.277$) times more likely to NOT participate in any projects compared to active participants.

Perhaps even clearer is the apparent trend visible when viewing the means of Any Project Participation across each Type of Learner. On the right side of Figure 5.8 there is a clear trend of increasing proportions of participation, which matches the general trend that would be expected, given the descriptions associated with each of the Type of Learner labels. For example, the “observer” description (see Additional Measures section in Chapter Three) clearly indicates that the student has no intention to participate, hence the term observer. The description for a “drop-in” implies some amount of participation, but that participation will be limited to a specific topic of interest within the course. As such, we would expect drop-ins to

participate more than observers. “Passive participants,” on the other hand, actually identify themselves as a “participant,” but are not interested in interaction with others. This category description is a bit troublesome because it includes the term “participant,” but the description indicates the student does not intend to engage “with other students or assignments.” This seems a bit contradictory. What the results suggest though is that passive participants were more likely than both observers and drop-ins to participate in any project. Finally, we would clearly expect “active participants” to participate more than any other Type of Learner, and this is indeed what was found to be the case.

What this implies is a possible ordinal relationship between the Types of Learners in regards to likelihood of participating in the creative projects. An interesting way to follow up this possibility would be to utilize latent class analyses (or maybe discriminant function analysis) to identify the variables that underlie whether a student identifies him/herself as a specific Type of Learner. In other words, if we assume there is an underlying categorical latent variable (Type of Learner) with four classes (corresponding to the four types of learners), then it may be possible to identify variables that are informative behavior indicators of this underlying latent variable. This is beyond the scope of this study, particularly given the data available. I also discussed this notion in a bit more detail in the previous section (see Latent class analysis with Type of Learner in the Contributions and Issues with Measurement section above).

Past experience in music and participation. The second reason this result (Type of Learner is the only significant predictor of Any Project Participation and Participation Beyond Top 10) is important is the fact that PEM class was not a significant predictor. This means that a simple explanation like, for example, “less musically experienced students are less likely to participate in the creative music-making projects,” is not tenable, at least given this particular 3-

class measurement model. So while it might make sense to argue that students did not participate because they just did not have enough musical experience, the data do not support this conclusion. Another equally simplistic explanation, “less musically experienced students are *more* likely to participate in the creative music-making projects,” is also not supported by these data. It could be argued that less musically experienced students would be *more* likely to participate because they are eager to learn, but again, this argument does not find support from the results of this study.

While an individual’s past experience in music did not predict whether they would participate in any of the projects, an alternative explanation is that individuals with more experience in music did not plan to participate in the projects in the first place. An indicator of their intent to participate would be how they identified themselves as learners (Type of Learner). This alternative explanation would be supported if an individual’s PEM class was related to an individual identifying him/herself as an active participant (i.e., indicating an intent to participate or not). The non-significant chi-square test of independence for PEM class and Active Participant (see Table 5.24), which is also evident in Figure 6.1, indicates such a relationship was not apparent with the participants in this course.

The continual lack of relationship found between PEM and other variables of interest is somewhat perplexing because PEM did not predict whether students participated in the creative music-making projects nor did it predict whether they intended to participate in the projects. What is clear is that a simple explanation is not possible, and perhaps that is to be expected, given the complexity of musical experience and the complexity of the many factors that may contribute to project participation that were not measured in this study. However, it is also possible that the 3-class model of PEM was just not sensitive enough to detect these

relationships, that the measurement was just too coarse. This would mean the ratio of signal to noise in the measure was too low, that there was too much error in the measurement. Such is the case when a measure is unreliable, the proportion of true score (signal) to total score (true score plus error/noise) is less than ideal. We know that correlations are attenuated as reliability gets lower so it is possible that the relationships exist, but they could not be detected with the measurement model I chose to use. Either way, the role of PEM in creativity-based learning requires much more research.

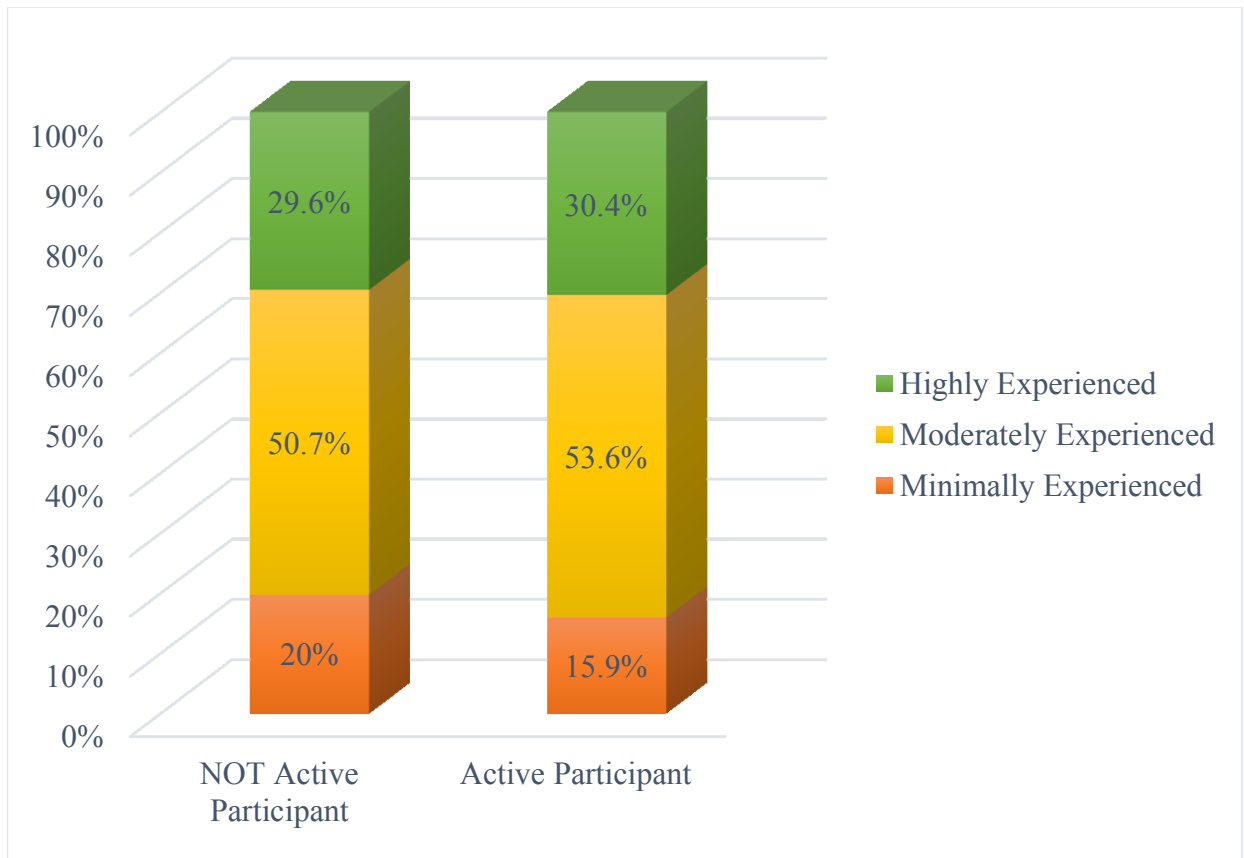


Figure 6.1. Proportions of "Active Participants" and Non-"Active Participants" by PEM Class

Group Differences for “Active Participants”

The exploratory analysis for group differences between two different types of learners, “active participants” and non-“active participants” (observers, drop-ins, and passive participants), revealed several importance differences. Before continuing, two points should be made. First, I remind the reader that the term “active participant” refers to how a student identified himself/herself at the beginning of the course in response to the Type of Learner question, and does not refer to an individual’s actual participation in the course. It is a label that refers to self-reported intent and not actual observed behavior. Second, due to the exploratory nature of this analysis, a specific theory for each potential group difference was not postulated a priori. In what follows I run the risk of committing what has been referred to as “HARKing” (Kerr, 1998), which is an acronym for hypothesizing after the results are known. Kerr defined HARKing as “presenting post hoc hypotheses in a research report as if they were, in fact, a priori hypotheses” (p. 197). Since my purpose was to search for possible group differences, and not confirm group differences I had suspected would be there in the first place, the discussion that follows should be regarded as an attempt at theory building, not an attempt at theory confirmation.

To begin, “active participants” had significantly higher MCSE scores for all three MCSE factors compared to non-“active participants.” The difference was also very large, as evidenced by effect size correlations well over .5 for all three factors. While the sample size was fairly small for this analysis ($N = 24$), the two groups were nearly equal in size (11 “active participants” and 13 non-“active participants”). Given the relationship between Active Participant and both MCSE and Project Participation, it is worth considering if MCSE would have been a good predictor of Project Participation. Recall that MCSE was not included in the

logistic regression analyses because logistic regression requires a large sample size, so this possibility could not be explored further for this study, but should be examined in future research.

“Active participants” also had statistically significantly higher scores on the Openness factor for Personality as well as the Challenge factor for Contextual Support. The items for the Openness factor have to do with an individual’s comfort with abstract ideas and imagination, so we might say that “active participants” were individuals who are a bit more comfortable using their imagination and dealing with abstract ideas. A course based heavily on creativity, which utilizes imagination, might be more appealing to those with a higher Openness component to their personality. It is also possible that the specific nature of this course had nothing to do with it. Instead, maybe individuals who are more “open” are simply more willing to participate in new things, regardless if they involve specific creative aspects or not. In other words, maybe people who are more “open” are just more likely to identify themselves as “active participants,” regardless of the course. To my knowledge, no research has looked specifically at personality characteristics of individuals who take MOOCs, so these potential explanations would require much more research.

That being said, the fact that Openness was the only factor that emerged as being different is also interesting because of the creative focus in the course and the importance of the Openness factor in previous creativity research (Feist, 1998, 2010; Prabhu et al., 2008). While I neither theorized nor anticipated this group difference being present in advance, I was certainly not surprised when the Openness factor seemed to have at least some level of importance, statistically speaking. Future research should consider the role of personality (especially the Openness factor) in a MOOC setting in general, in MOOC settings that specifically involve

creative activities (especially creative music-making), and in any educational setting that involves creative music-making.

The group difference on the Challenge factor of Contextual Support is a bit difficult to tease apart and explain. The items for the Challenge factor focus on the extent to which individuals perceived the course activities on the whole as being challenging. Unfortunately, the items for this factor could be interpreted several ways. For example, the first Challenge item, “I feel the course activities are challenging,” could be interpreted as meaning the activities are difficult, as in requiring a level of skill beyond one’s current abilities. In this case, “challenging” is an assessment of the difficulty of the activities, and could carry a negative connotation. This same item could also be interpreted as meaning the activities are provocative, stimulating, or even enticing, helping an individual to move beyond the current level of ability or knowledge. With this interpretation, the idea of being “challenged” has a positive connotation, and is perhaps something that is desired.

This latter interpretation is the intended interpretation because it matches well with the wording of the other three items. Within the Self-Determination Theory framework, especially as it is conceived within this particular measure of Contextual Support (Shroff & Vogel, 2009), Challenge is viewed as a lower-level indicator of the higher level construct of Competence. From this perspective, a higher score on the Challenge factor is potentially indicative of a greater sense of Competence. In light of the differences in MCSE discussed above, this finding is interesting because it provides another dimension to this group difference. “Active participants” began the course with much higher beliefs in their creative music-making capabilities (higher MCSE), and by the fifth week into the course (when the Contextual Support measure was administered), this group difference in competence shows up again, but this time in reference to

the course activities (as higher Challenge scores), which are designed to utilize one's creative music-making abilities.

The picture that emerges when considering the relationship between Active Participant and these three variables (Openness, Challenge, and MCSE) is far from clear at this point, especially when considering that Active Participant was also the only significant predictor of Project Participation. These results beg further research into the role that Competence plays (whether in the form of self-efficacy or perceptions of Challenge) in both how an individual identifies himself/herself as a Type of Learner at the beginning of a course, as well as how that individual actually participates within the course.

Following an exploratory approach, the analysis of group differences for the Active Participant variable was conducted only after the Active Participant variable emerged as the sole significant predictor of Project Participation. The logistic regression analyses that identified Active Participant as a significant predictor were based on only two of the many different possible ways in which Project Participation could be conceived and measured. In the following section I explore several other possibilities for exploring the Project Participation variable.

Other Avenues for Exploring Participation

My analysis of Project Participation focused on two binary dependent variables, Any Project Participation and Project Participation Beyond Top 10 (participation in either the Write Lyrics, Write a Melody, or Write a Song projects). Another possibility for analysis would be to treat Project Participation as count data and use the total number of projects a student participated in as the dependent variable. This would require a regression analysis that is appropriate for count data, such as the Poisson, Zero-Inflated, or negative binomial regression models, among others. One issue to consider with count data is how to deal with the exact sort

of data that I had, which had a large proportion of individuals with zero counts. One way to deal with this is to treat the count distribution as a mixture distribution, analogous to latent class analysis. Essentially, there are two classes of individuals, those that cannot have any value except zero (i.e., those that would not participate period) and then those that may have participated in anywhere between one to all four projects. I did a small amount of preliminary analysis with each of these models, focusing on the Active Participant variable as the sole predictor. I was not particularly surprised to find that it appears to still be an important predictor even when looking at total number of projects in which students participated. I did not report any full results from this analysis because it is still very preliminary.

Another potentially fruitful way to examine questions of participation and engagement in a course is from the perspective of survival analysis. Survival analysis is used in many fields, including education, to study things like student dropout rates in public education and higher education. Some researchers have also begun to apply it to MOOC research (e.g., Rosé et al., 2014). With survival analysis, the dependent variable is the amount of time that occurs before a particular event occurs (e.g., time until a student drops out). The goal is essentially to determine what variables can be used to predict the amount of time to the event (e.g., does self-efficacy at the beginning of a course predict the amount of time a student will remain in the course?). I should also mention that survival analysis would have applications in other areas of music education research, especially for researchers interested in examining attrition in music programs.

While this research study was a cross-sectional design, there is certainly much that could be learned from more longitudinal designs. Learning is clearly a process, and perhaps the best

way to study a process is longitudinally. Of course longitudinal designs have their own difficulties associated with them as well.

Finally, the qualitative researcher that is reading this may be burning up by this point because I have yet to mention the possibility of qualitative research to examine questions of engagement, participation, and learning experience in a creativity-based music learning context. In some ways, my attempt to uncover what may be important predictors of Project Participation was somewhat backwards. For example, a series of interviews could have been conducted to uncover broad trends and themes that emerged from the participants thoughts about their MOOC experience. These themes could have then been followed up with a more targeted quantitative approach that focuses on what appeared to be important from the perspective of the participants. There is certainly much that can be learned using qualitative methods and multiple-methods approaches.

What is probably clear at this point is that I am less comfortable making any strong or firm conclusions from this study, but rather I am keenly interested in the many possibilities and interesting questions that have arisen from this research. I turn now to some of the new questions that have emerged.

More Questions

In considering my experience of developing and teaching the course, “What is Music?” and in pondering the results of the research I conducted with the course, I have found myself asking more questions than providing answers. These questions seem to fall into one of two categories, questions about creativity-based learning and questions about MOOCs and music education. I address each below.

Creativity-based learning. When I designed this research study I sought to find some answers about the nature of music learning when it occurs via creative music-making. I chose to utilize the MOOC context for studying creativity-based learning because it offered a potential solution to several technical and logistical problems related to the type of large-scale, model-based research that was my goal. Students in my course chose, overwhelmingly, not to engage in the creative-music making projects that I believed would lead to musical learning. As a result, I found myself struggling to be able to make any claims about creativity-based learning. However, I think the very clear lack of participation in the creative projects is a very profound result, and a result that required little to no statistical analysis. The question that naturally follows from this glaring result is “why did they choose not to participate?” A very small part of the answer to that question, as discussed above, is that they simply had not planned on participating, as indicated by the different Type of Learner with which they identified. But this explains only a small amount of the variability in participation. It also does not explain why so many individuals who indicated that they planned to participate did not end up participating.

One explanation is that there is a misfit or a mismatch between the pedagogical model and the learners. It is possible that MOOC learners are simply not interested in completing projects, but are more interested in passive forms of learning (e.g., viewing videos or reading text content). Another explanation is that students simply did not have enough support (scaffolding) to allow them to participate in the creative projects. Perhaps more tutorials and demonstrations are necessary. Equally possible is that the creative projects required too much time to complete, or were too overwhelming. It is also possible that many people worked on and maybe even completed the projects, but were not comfortable sharing them with hundreds of strangers in their course. This implies the need to consider more sociological questions about the nature of

the power structures, social norms, and social interactions, as well as the role of traditional psychological constructs like self-esteem, social desirability, personality, and social risk, to name a few. The Community of Inquiry (CoI) framework, which looks at relationships between what CoI researchers refer to as social presence, cognitive presence, teaching presence, and learning presence (Garrison, 2011; Shea & Bidjerano, 2009, 2012), might be useful in examining these different explanations. The framework utilizes survey and content-analysis techniques to examine students' perceptions of their online learning environment, particularly as they relate to learning outcomes. A clearer understanding of students' perceptions of a creativity-based music learning context is necessary, including the relationship between those perceptions and the learners' specific learning goals and actual course engagement. But why are questions about MOOC participation important in the first place? Why should we care about MOOCs?

Why should we care about participation/engagement in a MOOC? I have to admit that as my study began to shift focus from modelling creativity-based learning and learning outcomes of that process, to focusing more on participation/non-participation in a creativity-based learning process, I initially fought this shift. I was disappointed that I could not collect the data that I wanted, and I felt a sense of defeat. Looking at participation/non-participation felt somehow less important or less meaningful. I have since come to reject that belief entirely. The meaning and importance of this research emerged when I began to take a step back from questions about the creativity-based learning pedagogy, and instead began to view this research within the much broader context of music education and musical engagement over a lifetime. If the overarching goal of all music education is, at a minimum, to encourage and facilitate lifelong musical engagement, then questions about engagement (including participation/non-participation and everything in between) are central to this fundamental goal.

The MOOC, as a relatively new phenomenon, and as one piece of a broader global philosophy of open learning and open knowledge, offers new opportunities for music education on the lifelong learning front. Beyond secondary school, there are relatively few opportunities for continuing one's music education (aside from choosing music as a career path or pursuing higher education degrees or coursework in music). Some have argued that the future of school music programs rests on our ability to create, maintain, and develop relationships between school music and community music programs (Myers, 2008). In other words, there is a gap that needs to be filled. MOOCs have begun filling in this space. At the time of this writing, using the MOOC search and aggregation website www.mooc-list.com, I found just under 30 music MOOCs that are currently available. Topics for these MOOCs included digital recording, songwriting, Western Art Music, history of rock, music theory, guitar, music business, and many others. The question I believe we need to consider as a field is whether we are filling in this MOOC space with what is already available elsewhere, or whether we are creating new opportunities. More specifically, are we leveraging the affordances of open, distance, online, networked, computer-mediated, and technology-based learning, given their availability in the MOOC context? Or are we simply transferring pedagogies from mediated environment (face-to-face) to another mediated environment (online)?

I have a sense that much of the non-music MOOC space has been filled in with moldy pedagogies and 20th century-style videotaped lectures. In essence, many universities and MOOC-providers have merely brought the passive lecture environment to global scale. That being said, I recognize that there is a place for and a value in the video recorded lecture. I myself have learned many things simply by searching for videos on YouTube or Vimeo, or even enrolling in some of these video lecture-based MOOCs. The video lecture-type MOOCs have

done the job of organizing the videos for me, presenting them in a coherent sequence and structure. There is certainly value in that. However, if lifelong musical engagement is the goal, *then engagement should be promoted in the venues that support lifelong learning*, which includes the MOOC. A passive music learning environment seems to support a passive musical life. The potential value of creativity-based music learning is a focus on active and creative engagement with music. But the potential remains to be seen. Indeed the results from this research provide a glaring example of the adage “you can lead a horse to water, but you can’t make it drink.” If there are potential pedagogies (or andragogies), approaches to teaching and learning, that may promote lifelong musical engagement, then the issue of “getting the horse to drink” is central to examining these potentials further.

Furthermore, it is also worth considering whether underlying patterns and issues in a MOOC context are unique to a MOOC context, or are simply different manifestations of the same broader societal issues. In the school music context (at any level), participation in course activities is generally mandatory, so the question of participation vs. non-participation (engage or not engage) becomes less crucial. However, the *extent* to which students engage may be a function of the same underlying factors. In academic contexts, the notion of Type of Learner can be envisioned within the frameworks of academic motivation and goal contents (Vansteenkiste et al., 2006; Wellborn, 1992). Students have different learning goals for different learning environments, and they have different types of motivation, which in turn relates to differences in outcomes like depth of learning, academic performance, and persistence (Vansteenkiste et al., 2004). Recall Rotgans and Schmidt’s definition of situational engagement: “the extent to which students’ are willing and able to take on the learning task at hand. This includes the amount of effort students are willing to invest in working on the task...and how long they persist” (p. 467).

The Type of Learner item can be viewed as a measure of the extent to which an individual *intends* to engage. The point here is that there is likely much that is known from research in motivation and engagement that could be applied to the MOOC context (particularly as it relates to course design), and in turn, there is likely much that can be learned about motivation and engagement in the MOOC context that can better inform course design, pedagogy, and curriculum in school music contexts.

Alexandra Lamont has written and spoken much about the idea of musical identity development and lifelong musical engagement. In attempting to clarify and dispel several myths regarding musical talent, motivation, opportunity, and continuity of music-making over a lifetime, Lamont (2011a) made the following claim:

...both opportunity and time are required to allow an individual to explore different kinds of activity before making a lasting commitment which will help give his or her life meaning. Contradicting most of the existing research which suggests that meaningful and positive musical experiences should occur early in life (e.g. Sloboda 1990), my new data from amateur adults seem to reflect the importance of later events (p. 382)...Providing favourable conditions for identity development throughout childhood and adulthood seems to be necessary to help as many as possible develop their own sense of musical identity and explore their passion. (p. 383)

I believe Lamont's point here is that the field of music education has tended to focus on developing young music-makers, but at the expense of developing lifelong musicians. I use the term musician not as "one who is musically skilled," but as "one who *identifies* him/herself as musical." Identity development is key. However, Lamont notes that the path to the musician identity is not linear, not necessarily continuous, and not necessarily a matter of simply having

opportunities. I would argue that the important thing for the field of music education is to not only ensure there are a multitude of opportunities, but more importantly, when a music student finds one of those opportunities, the educational, artistic, and aesthetic experience is one that promotes and facilitates meaningful and autonomous engagement. As such, understanding engagement is crucial, whether it be engagement in a MOOC, engagement in a second-grade general music class, engagement in a private piano lesson, or engagement in an undergraduate music elective. As I learned quite clearly with this research, if we cannot get students to meaningfully and autonomously engage in the musical *learning* we have planned for them, then we certainly should not expect that they will engage in music *making* over a lifetime.

What my data suggest is that, at least in some small part, if learners arrive at the learning context planning to engage, they are more likely to actually engage. One question to consider is “what conditions need to be present such that a learner arrives at the learning context already planning to meaningfully and autonomously engage?” I have repeatedly used the terms meaningful and autonomous in regards to engagement intentionally. Individuals that enroll in a MOOC are essentially completely autonomous in that environment. There is no initial investment (neither financial nor emotional) required to begin a MOOC and there is also often little external pressure to continue. Almost none of my participants said their primary reason for enrolling in the course was to gain work-related skills, to prepare for a new career, or to prepare for some other form of schooling. All of those reasons would represent some form of externally regulated motivation, which according to SDT is less autonomous, and more controlled.

On the contrary, the overwhelming majority of participants said they enrolled because they “*enjoy* learning about topics that interest [them].” The idea of doing something for enjoyment is a key indicator of intrinsic, autonomous motivation. Overall, students in my course

did not enroll because of external motives, and it is reasonable to assume that their activity in the course was generally internally motivated, at least much more so than most courses a student takes for some sort of grade or some sort of credit. So when operating autonomously, the students chose, on the whole, not to engage. Again, why?

One explanation, a very speculative one at that, has to do with what I said earlier, that “when a music student finds one of those opportunities [to learn music], the educational, artistic, and aesthetic experience is one that promotes and facilitates meaningful and autonomous engagement.” If, over one’s lifetime, the majority of one’s music educational experiences are not autonomy-supportive, then by definition, they are non-autonomous, what would be referred to as *controlled* in the SDT framework. We know from SDT research that autonomous motivation is related to greater persistence and deeper engagement, but controlled motivation is related to just the opposite. After ten to twelve years (during a student’s compulsory education) of non-autonomous, externally regulated music learning experiences, an individual becomes reliant on the learning context to provide the motivation to engage. It is no surprise then that when that individual wanders into a music learning context that is completely autonomous, the student does not engage because the context does not provide the external motivation to do so. From the perspective of lifelong musical engagement, it should also not be a surprise that once leaving the externally regulated music learning context, the student does not sustain musical engagement or music learning.

This situation is the general state of affairs for music education in the US. I cannot help but to think back on my experiences as a band director, especially working with a competitive marching band. On one hand, I know most of the students that chose to be in the band were intrinsically motivated to do so (i.e., they were enrolled in band primarily because they enjoyed

it), and as such their membership in the ensemble and their attendance at rehearsals was likely mostly autonomously motivated. However, when I consider the actual *musical* engagement of the individual student in the band, there was an incredibly tiny amount of music that came out of a student's instrument that was autonomously motivated. Decisions about what to play, how to play it, when to play it were entirely externally regulated. The motivation to make the specific musical sounds was entirely externally motivated at best, but was more likely completely unmotivated (i.e., the student is making the sounds someone else told him to make, but he is not even sure why he is doing so).

I have strayed a long way from MOOCs and creativity-based learning, and I do not intend for this to be an indictment of marching bands or any other ensemble, per se. Much of what I have just speculated is just that, speculation, and very little of it can actually be concluded from my research. But I titled this section "more questions" because what this research did uncover was many more questions, some of which are uncomfortable to have to consider. From the perspective of lifelong learning, if a student arrives at a MOOC expecting to actively participate (to fully engage), then we should wonder what about that individual's music education history has caused him/her to arrive at that music educational opportunity anticipating and intending to actively participate (to meaningfully and autonomously engage), as compared to arriving at that music educational opportunity and not expecting active participation to be the norm.

These are big questions, and I believe the MOOC context is uniquely situated to help address some of these questions, particularly because they involve global student populations, students with very diverse music education histories from a wide variety of musical cultures. The question of MOOCs as tools for research is the final question I consider.

What about MOOCs as a research tool? Aside from developing a model that could adequately represent the observed data to explain the relationships between some of the many variables involved in the creativity-based learning process, a secondary purpose of this research was to examine the feasibility of utilizing a MOOC for conducting substantive research in which the MOOC is not necessarily the subject of the research. Based on my experience attempting to conduct research within a MOOC, and after reviewing my journal notes, e-mails, and even various versions of documents that were components of the research, I have identified several issues that hurt feasibility, but I have also identified several opportunities that are unique to MOOCs.

Feasibility issues. First, as I just mentioned, MOOCs are unique because of the diversity of the students. This presents challenges in terms of both quantitative data collection (particularly for self-report measures), but also for qualitative data collection. For nearly half of the individuals in my course, English was not their primary language. For the actual teaching and facilitating of the course, there did not ever appear to be any major language barriers. However, having read the discussions and comments of many students throughout the course, I wondered how much nuance was “lost in translation” with some of the research items. This is likely the case whether the individual is responding to a researcher’s question by checking a box on a questionnaire or responding to a researcher’s question by talking about him/herself.

Second, I quickly realized the importance of verifying the global availability of any services or software that were necessary for participating in the course (or participating in the research). I had integrated the use of the music-streaming service Spotify into many of the discussions throughout the course. Prior to beginning the course, I completely overlooked the possibility that the service might not be available everywhere in the world. I quickly learned

within the first day of the course beginning that it was indeed not available worldwide. While this did not pose any issues for the research components of the course, the global availability of services and software is something a researcher would need to consider.

Third, because of the nature of my research, which involved collecting many different measurements, I decided to use the quiz/survey feature in Canvas to administer my research items so that merging the datasets and connecting individuals' responses across different questionnaires would be easier to manage. What I assumed would be a positive trade-off for data management was also a negative tradeoff in terms of question formatting and visual display for my questionnaires. A researcher would need to consider what sort of data collection means are available within the MOOCs learning management system to determine whether they are sufficient for the research purposes. This is important because, as I mentioned with point number two above, as soon as you move to software and services outside of the learning management system, there is a possibility that it will not be available for someone.

Fourth, I found it difficult to find a healthy balance between promoting my research within the course and becoming a nuisance to students who had no interest in the research. As the teacher of the MOOC, I felt an obligation to not allow the research component of the course to interfere with students' experience of the course. I did not directly seek out students' perceptions of having research included in the course, nor did I receive any comments or complaints from the students about the research component. The issue was always on my mind though, and probably limited the extent to which I solicited students' participation or repeatedly requested participants to complete research components. This might be less of an issue with research in which the research activities (i.e., how and what data are collected) are actually a component of the course. For example, if a researcher is interested in doing a content analysis of

the students' discussions, then students need not do anything beyond participating in the discussions, which are already a component of the course itself. This is different than asking students to complete questionnaires that are separate from the actual course activities, even if they are related to the course activities.

Unique opportunities. While I certainly experienced feasibility issues with my research, there were also some unique opportunities that I became aware of, but only after actually completing the research. To begin, the students in my course represented a truly *global* and *diverse* sample. To be clear, I am not suggesting it was a *representative* global population. Rather, it was a sample of students that represented a wide diversity of geographic locations, cultures, education levels, ages, and musical backgrounds, possibly more diverse than any other typical educational context. Even at institutions that have a large proportion of international students, the students within any given course will be relatively homogeneous. A globally diverse sample has its advantages and disadvantages. With large amounts of diversity, there are also greater numbers of differences between people on characteristics that are not of primary interest to the researcher. From a study design perspective, all of these differences represent potential confounds in the data, many of which are likely unmeasured. Of course this can be managed with traditional random assignment practices used in an experimental design, but random assignment is likely not an option in most MOOC contexts.

Conversely, this heterogeneity can also be a potential benefit. Any introductory statistics course will tell you that correlation is inherently a function of variability. A bivariate correlation, which is just covariance that has been standardized to the variance of both variables, will be greater with greater variance in the two variables of interest. As such, finding a very diverse sample is akin to adding statistical power to your study (increasing your ability to detect

statistically significant effects) because observed correlations (an effect size) will tend to be greater with greater variability, and power is a function of the alpha level (which is usually fixed), sample size, and effect size. It follows that whether heterogeneity in a sample is advantageous or problematic is dependent on the characteristics and design of a particular study.

A second unique opportunity concerning research in a MOOC is that it allows you to reach people that you most likely would not reach in more typical research contexts (e.g., universities, labs, schools, local communities). This is important at two levels. First, similar to the idea of diversity and heterogeneity discussed above, it allows for individuals to possibly be included in one's research that are possibly more unique than those encountered in typical research settings.

In addition, because of the uniqueness of the individuals that can be incorporated in one's research, it forces the researcher to confront issues of external and ecological validity head-on. Research that is conducted in the traditional settings discussed above can often be criticized for lacking any external or ecological validity. It is easy for critics to say things like "well you looked at a group of undergraduate students from a major American university. Is that really how things play out in the 'real-world'?" The way in which the individuals in a MOOC (each of whom are likely more unique than they are similar) react, respond, behave, think, and feel within the context of a research study is potentially a better representation of how "things play out" with people in the "real world." The glaring primary result from this study that, when not required to engage, people tended to not engage in the creative activities is the case in point for this argument. Such a clear and obvious result likely would not have been observed with a class of undergraduate students taking the same course at my university.

What about the model? This dissertation began as an attempt to further Webster's Model of Creative Thinking in Music, which included the development of the General Specified Model of Creativity-Based Learning and the subsequent attempt to examine this model in a real-world creativity-based learning context. The majority of the analysis has focused on the context itself (the MOOC), at the expense of an analysis of the model, mostly because of some extreme limitations of the data. At this point I wish to return to the model, in an effort to bring the dissertation full circle.

The exploratory analysis of the correlations between all components of the General Specified Model was not intended for inferential purposes or for making any generalizable truth claims, but rather was intended to raise questions and promote further investigation and more nuanced theorizing. As to the latter goal of more nuanced theorizing, the exploratory analysis can be illustrative of why such a goal is necessary. After even a casual glance at the correlation heat map shown in Appendix G, one can see how interrelated these components are, at least in terms of zero-order correlations. It was not possible with this study to partial out the shared variances between different components in order to look at their unique contributions. This is certainly necessary in future research, whether it be through structural equation modeling or other approaches. Nevertheless, what *is* clear from this study is how easy it would have been to claim the presence of a very clear picture (e.g., the clear relationship between Contextual Support and MCSE) when considering only a few components in isolation, when indeed the picture is much muddier when individual relationships are considered within the context of a broader theory.

Very generally speaking, the analysis of the small amount of data available for examining the whole model revealed support for much of the basic structure and primary theoretical

relationships articulated by the General Specified Model. At the core of this model are an individual's beliefs in their creative music-making abilities (MCSE), their reasons for engaging in creative music-making (Motivation), and the extent to which the environment is supportive of factors that promote autonomous motivation (Contextual Support). For the few individuals in this study for which data was available, each of these components appears to be highly interrelated, but also very much related to both their level of engagement and their perceptions of how much they learned from the course. Each of these components should be examined in further detail, particularly as they relate to each other in different contexts (both MOOC and non-MOOC contexts).

As I have previously mentioned, the General Specified Model should be seen as a starting point. I took one step—a possibly naively optimistic step and, at times, a somewhat ambiguous step, but a step nonetheless—toward advancing the study of musical creativity and musical learning and toward advancing this model, which is but one possible model of creativity-based learning. In this study it was not possible to take the next step with the model: to prune it, alter it, or reject it altogether. That being said, the study did provide enough evidence to warrant taking that next step. Designing and carrying out research on such a model requires huge amounts of time and resources. This study can be seen as a large-scale pilot study that provides both a rationale for further study of the model and some justification for the investment of those resources.

In the final chapter, the Epilogue, I detail what some of the future directions might look like for research on the General Specified Model of Creativity-Based Learning, for research on some of the research instruments developed for this dissertation, and for research on MOOCs in

music education. I close this chapter by considering something Webster (2013) said in a recent book chapter in regards to the importance of reflection in creative music-making:

Reflection. If I were to choose one word that represents the most important goal for creating teaching environments for enhancing compositional intelligence and on to creative products, it would be this one. As students play with sound and create more and more sophisticated products, I hope that teachers will encourage reflection... Learning to thoughtfully reflect on creative work is perhaps the one ultimate goal for learners, and for that matter, teachers. (p. 30)

To this I would add that thoughtful reflection is also a vitally important component of another creative activity: research. The process of theorizing, the generation of hypotheses, the designing of a study, the analysis of the data, the interpretation of results, and even the manner in which the results are presented are all activities that rely heavily on creative thought. After much reflection on the model, on the study I designed to investigate the model, and on the analysis and interpretation of the data, I have recognized many things I would do differently, many assumptions I had unknowingly made, and many weaknesses in my analyses and interpretations. I have attempted to be as transparent as possible by including many of these reflections in this document. If this research serves no other purpose, I hope that it at least serves as an instrument for reflection, a catalyst for further questions, and a tool that can be used to “interrogate intentionality” (Webster, 2013, p. 30) with future research. Reflecting on Webster’s (2002) Model of Creative Thinking in Music allowed me to interrogate my own intentions with my research, to situate my aims within the broader context and philosophy of music education. I hope this text may do the same for another.

EPILOGUE: FUTURE DIRECTIONS

I committed to the literary metaphor to begin the dissertation, so it is only fitting that I close with an epilogue to articulate some future directions for this research.

Try It Again

To begin, I intend to offer the course again in the future. I will take some time to consider everything I have learned from this experience so that I may adjust aspects of the course as necessary. I have considered splitting the course into two or three smaller portions, each 3-4 weeks in length. All three parts could be taught simultaneously, but repeated every other month or so. Each course could stand on its own and function as a complete course, without requiring prerequisite material. This would allow students to commit to a much smaller amount of time, and then take a break in between if changes in one's life occur that alter one's ability to participate. By repeating the course cycle every few months, students can choose the part they have not taken when they are ready to do so. These smaller 3-4 week courses would allow for a bit more detail to be covered in the content and also allow students more time to work on their creative projects, which could also be supported by additional demonstration and tutorial content.

From a research standpoint, this would create much smaller time periods of data collection. Instead of waiting for an entire 8 weeks, a new dataset would be generated for each 3-4 week course. Each data collection period would also then be associated with a particular creative project. This would allow for a much cleaner approach to dealing with the three levels of generality (global, contextual, and situational). The first week would involve data collection

for global level variables. During the second week, once students have acclimated to the course environment, measurement of contextual level variables could take place. Then finally, during the third and/or fourth week, measurements associated with a particular creative music-making task could be taken, allowing the measures to be associated with a much more specific learning situation.

In addition, because global level variables would be collected during the first week for every course, an individual who has already completed the global level measures would not need to complete them again. This might improve the chances of him/her completing other measures throughout the course by reducing the demand during each 3-4 week course period.

Advance the Research on the PEMI and MCSES

Now that the online surveys have been constructed, it will be relatively easy to seek out new samples to complete both instruments along with another instrument or two designed for assessing different types of validity. I hope to also continue to assess content coverage by consulting with additional experts and continuing to solicit feedback from individuals who complete the instrument. While the sample population for Pilot Study #2 was undergraduate students, I also wish to pursue other populations (e.g., high school students, younger children, and music majors/music professionals).

Institutional Collaboration

In addition to continuing the MOOC research, I also hope to establish some collaborative research projects across institutions. Once the course content is developed it would be fairly easy to offer the mini-courses as components of other existing courses at other institutions. This would allow the research to be conducted with smaller, more homogenous samples, which over time, could be combined using appropriate multilevel modeling techniques

to look at the between-course variability, something that would be particularly valuable in terms of studying the role of context.

Creativity-Based Learning Research

The idea that much can be learned through creative music-making activities, like composition, songwriting, arranging, making mashups, and improvisation, among other activities, is far from new, but also far from being thoroughly researched, particularly in regards to the what's, how's, and why's of learning that emerges from these activities. The General Specified Model of Creativity-Based Learning situates musical learning as the end goal of these activities. Whether future research utilizes this model as a starting point or not, many questions remain as to the efficacy, the feasibility, the applicability, the philosophical rationale, the ethics, the politics, the cultural biases, and the requisite teacher training related to a creativity-based learning model, all of which can and should be investigated. This is likely true of other teaching/learning models as well.

In describing potential music teacher identities that are appropriate for 21st century music education, Randles (2012) borrows a concept from Robert McKee, which is to “leave room for the actor” (p. 38). Randles asks the following: “If music students can be thought of as the actors in the teacher/writer’s screenplay, then does music education, as a profession, leave enough room for student creativity?” (p. 39). His point is that our profession has traditionally (particularly in large ensemble settings) identified its teachers as directors, not teachers (or writers, or producers). Students are both literally and figuratively the instruments to be used for the director’s design.

Creativity-based learning leaves “room for the actor,” the student, allowing the student’s voice to be heard and have an autonomous, meaningful role in the learning process. Research on

creativity-based learning needs to consider not just the *potentials* that result from opening up that space, leaving “room for the actor,” but also the *ramifications* and *implications* of opening that space. This study, and the model it is based upon, privileges psychological ramifications (e.g., self-efficacy, motivation, engagement), which are certainly important to understand in their own right. We must also consider ramifications and implications from an ethical perspective (e.g., is allowing the space a good/moral thing to do, and what ethical issues arise for the student, teacher, school, community when that space is opened?), political perspective (e.g., who is privileged and who is biased when this space is opened?), or sociological perspective (e.g., what are the effects of opening this space on students’ identity, particularly in regard to their “performances” of gender, race, or class?), to name but a few. Perhaps more importantly, each perspective will hopefully work in consideration and in collaboration with each other, promoting a broader understanding, and not simply a more heated debate.

Onward

These are but a few ideas of out of many other possibilities. Either way, this experience has opened my eyes to an entire world of new possibilities and a career’s worth of research. As a nod to the tradition that formed the foundation of my early musical development—a tradition towards which I have experienced a somewhat schizophrenic oscillation in attitude between aversion and fondness, respect and revile, confusion and clarity—I close by saying the movement is finished, but the work is incomplete. Please hold your applause.

REFERENCES

- Abbott, D. H. (2010). *Constructing a creative self-efficacy inventory: A mixed methods inquiry*. (Ph.D. in Music Education Dissertation), University of Nebraska - Lincoln. ProQuest Dissertations and Theses database. (Order No. 3402936)
- Amabile, T. M. (1983). *The social psychology of creativity*. New York, NY: Springer-Verlag.
- Amabile, T. M. (1996). *Creativity in context*. New York, NY: Westview.
- Ames, C. (1984). Achievement attributions and self-instructions under competitive and individualistic goal structures. *Journal of Educational Psychology*, 76(3), 478-487.
- Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument *Journal of School Psychology*, 44, 427-445.
- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11, 150-166.
- Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, 91, 340-345.
- Assor, A., Kaplan, H., Kanat-Maymon, Y., & Roth, G. (2005). Directly controlling teacher behaviors as predictors of poor motivation and engagement in girls and boys: The role of anger and anxiety *Learning and Instruction*, 15, 397-413.
- Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teaching behaviors predicting students' engagement in schoolwork. *British Journal of Educational Psychology* 27, 261-278.
- Auh, M.-S. (1999). *Enactive and reflective thinking during the compositional process by seventh-grade korean students*. Paper presented at the Proceedings of the International Music Education Research Symposium, Launceston, Tasmania.
- Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the consensual assessment technique to nonparallel creative products. *Creativity Research Journal*, 16(1), 113-117.
- Baldasaro, R. E., Shanahan, M. J., & Bauer, D. J. (2013). Psychometric properties of the Mini-IPIP in a large, nationally representative sample of young adults. *Journal of Personality Assessment*, 95, 74-84.
- Baltzer, S. W. (1990). *A factor analytic study of musical creativity in children in the primary grades*. (Doctoral Dissertation), Indiana University. ProQuest Dissertations and Theses database. (UMI #9029114)
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215. doi: 10.1037/0033-295X.84.2.191
- Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, 50, 248-287. doi: 10.1016/0749-5978(91)90022-L

- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. C. Urdan (Eds.), *Self-Efficacy Beliefs of Adolescents* (Vol. 5, pp. 307-337). Greenwich, CT: Information Age Publishing.
- Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, 38(1), 9-44.
- Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature *Genetic, Social, and General Psychology Monographs*, 132(4), 355-429.
- Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods*, 9(1), 3-29.
- Beghetto, R. A. (2006). Creative self-efficacy: Correlates in middle and secondary students. *Creativity Research Journal*, 18(4), 447-457.
- Beghetto, R. A., Kaufman, J. C., & Baxter, J. (2011). Answering the unexpected questions: Exploring the relationship between students' creative self-efficacy and teacher ratings of creativity. *Psychology of Aesthetics Creativity and the Arts*, 5(4), 342-349.
- Biasutti, M., & Frezza, L. (2009). Dimensions of music improvisation. *Creativity Research Journal*, 21(2-3), 232-242.
- Boden, M. A. (1991). *The creative mind: Myths and mechanisms* New York, NY: Basic Books.
- Boden, M. A. (2004). *The creative mind: Myths and mechanisms* (2nd ed.). New York, NY: Routledge.
- Brown, A. R., & Dillon, S. C. (2007). Networked improvisational musical environments: Learning through online collaborative music making. In J. Finney & P. Burnard (Eds.), *Music Education and Digital Technology* (pp. 95-106). New York, NY: Continuum.
- Brown, A. R., & Dillon, S. C. (2012). Meaningful engagement with music composition. In D. Collins (Ed.), *The Act of Musical Composition: Studies in the Creative Process* (pp. 79-109). Burlington, VT: Ashgate.
- Burnard, P. A. (2000a). Examining experiential differences between improvisation and composition in children's music-making. *British Journal of Music Education*, 17, 227-245.
- Burnard, P. A. (2000b). How children ascribe meaning to improvisation and composition: Rethinking pedagogy in music education. *Music education research*, 2(1), 7.
- Burnard, P. A. (2000c). Making a piece you don't play and forget: children composing and the role of context. *Australian Journal of Music Education*(1), 30-39.
- Burnard, P. A., & Younker, B. A. (2002). Mapping Pathways: fostering creativity in composition. *Music Education Research*, 4(2), 245-261. doi: 10.1080/1461380022000011948
- Burnard, P. A., & Younker, B. A. (2004). Problem-solving and creativity: Insights from students' individual composing pathways. *International Journal of Music Education*, 22(1), 59-76. doi: 10.1177/0255761404042375
- Burnard, P. A., & Younker, B. A. (2008). Investigating children's musical interactions within the activities systems of group composing and arranging: An application of Engeström's Activity Theory. *International Journal of Educational Research*, 47, 60-74. doi: 10.1016/j.ijer.2007.11.001

- Cadogan, J. W., & Lee, N. (2013). Improper use of endogenous formative variables. *Journal of Business Research*, 66, 233–241.
- Caspi, A., & Blau, I. (2011). Collaboration and psychological ownership: How does the tension between the two influence perceived learning? . *Social Psychology of Education*, 14, 283-298.
- Chamorro-Premuzic, T., & Furnham, A. (2005). *Personality and intellectual competence*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Chamorro-Premuzic, T., & Furnham, A. (2007). Personality and music: Can traits explain how people use music in everyday life? . *British Journal of Psychology*, 98, 175-185.
- Chamorro-Premuzic, T., Gomà-i-Freixanet, M., Furnham, A., & Muro, A. (2009). Personality, Self-Estimated Intelligence, and Uses of Music: A Spanish Replication and Extension Using Structural Equation Modeling. *Psychology of Aesthetics, Creativity, and the Arts*, 3(3), 149-155.
- Chavez-Eakle, R. A., Lara, M. C., & Cruz-Fuentes, C. (2006). Personality: A possible bridge between creativity and psychopathology? *Creativity Research Journal*, 18(1), 27-38.
- Chen, K. C., & Jang, S. J. (2010). Motivation in online learning: Testing a model of self-determination theory. *Computers in Human Behavior*, 26, 741–752.
- Chin, T. C., & Rickard, N. S. (2012). The music USE (MUSE) questionnaire: An instrument to measure engagement in music. *Music Perception*, 29(4), 429-446.
- Clark, S. L. (2010). *Mixture modeling with behavioral data*. (Doctoral Dissertation), University of California, Los Angeles, CA.
- Claxton, G. (2006). Thinking at the edge: Developing soft creativity. *Cambridge Journal of Education*, 36(3), 351-362.
- CMP, & MENC. (1965). Comprehensive musicianship (The Foundation for College Education in Music): A report of the seminar sponsored by the Contemporary Music Project at Northwestern University *Contemporary Music Project*. Ann Arbor, MI: University of Michigan.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed. ed.). Hillsdale, NJ: Lawrence Erlbaum Associates
- Collins, D. (2005). A synthesis process model of creative thinking in music composition. *Psychology of Music*, 33(2), 193.
- Collins, D. (2007). Real-time tracking of the creative music composition process. *Digital Creativity*, 18(4), 239-256. doi: 10.1080/14626260701743234
- Collins, D., & Dunn, M. (2011). Problem-solving strategies and processes in musical composition: Observations in real time. *Journal of Music, Technology & Education*, 4(1), 47-76.
- Connelly, B. S., Ones, D. S., & Chernyshenko, O. S. (2013). Introducing the special section on openness to experience: Review of openness taxonomies, measurement, and nomological net. *Journal of Personality Assessment*, 96(1), 1-16. doi: DOI: 10.1080/00223891.2013.830620
- Connelly, B. S., Ones, D. S., Davies, S. E., & Birkland, A. (2013). Opening up openness: A theoretical sort following critical incidents methodology and a meta-analytic investigation of the trait family measures. *Journal of Personality Assessment*, 96(1), 17-28. doi: DOI: 10.1080/00223891.2013.809355

- Corakli, E., & Batibay, D. (2012). The efficacy of a music education programme focused on creative thinking. *Procedia-Social and Behavioral Sciences*, 46, 3571-3576. doi: 10.1016/j.sbspro.2012.06.107
- Corbonneau, N., J., V. R., & Lafreniere, M. A. K. (2012). Toward a tripartite model of intrinsic motivation. *Journal of Personality*, 80(5), 1147-1178.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *NEO-PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81, 322-331.
- Covach, J. (2013). To MOOC or not to MOOC? *Music Theory Online*, 19(3). <http://mtosmt.org/issues/mt0.13.19.3/mt0.13.19.3.covach.php>
- Credé, M., Harms, P. D., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the big five personality traits. *Journal of Personality and Social Psychology*, 102(4), 874-888.
- Cropley, D. H., & Cropley, A. J. (2005). Engineering creativity: A systems concept of functional creativity. In J. C. Kaufman & J. Baer (Eds.), *Creativity across domains: Faces of the muse* (pp. 69-185). Hillsdale, NJ: Lawrence Erlbaum.
- Cropley, D. H., & Cropley, A. J. (2008). Elements of a universal aesthetic of creativity *Psychology of Aesthetics, Creativity, and the Arts*, 2(3), 155-161.
- Cropley, D. H., & Kaufman, J. C. (2012). Measuring functional creativity: Non-expert raters and the creative solutions diagnosis scale. *Journal of Creative Behavior*, 46(2), 119-137.
- Cropley, D. H., Kaufman, J. C., & Cropley, A. J. (2011). Measuring creativity for innovation management. *Journal of Technology Management and Innovation*, 6(3), 13-30.
- Csikszentmihalyi, M. (1988). Society, culture, and person: A systems view of creativity. In R. J. Sternberg (Ed.), *The Nature of Creativity: Contemporary Psychological Perspectives* (pp. 325-338). New York, NY: Cambridge University Press.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper & Row.
- Czikszentmihalyi, M., & Lefevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*, 56, 815-822.
- Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research*, 71(1), 1-27.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- Deci, E. L., & Ryan, R. M. (2012a). Motivation, personality, and development within embedded social contexts: An overview of self-determination theory. In R. M. Ryan (Ed.), *The Oxford Handbook of Human Motivation* (pp. 85-107). New York, NY US: Oxford University Press.
- Deci, E. L., & Ryan, R. M. (2012b). Self-determination theory. In P. A. M. Van Lange, A. W. Kruglanski & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology (Vol 1)* (pp. 416-436). Thousand Oaks, CA: Sage Publications Ltd.

- Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology, 73*(5), 642-650.
- Deshon, R. P., & Gillespie, J. Z. (2005). A motivated action theory account of goal orientation. *Journal of Applied Psychology, 90*(6), 1096-1127. doi: 10.1037/0021-9010.90.6.1096
- DeYoung, C. G. (2010). Intelligence and personality. In R. J. Sternberg (Ed.), *The Cambridge Handbook of Intelligence* (pp. 711-737). New York, NY: Cambridge University Press.
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*, 880-896.
- Dietrich, A. (2004). The cognitive neuroscience of creativity. *Psychonomic Bulletin & Review, 11*(6), 1011-1026.
- Dietrich, A., & Kanso, R. (2010). A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological Bulletin, 136*(5), 822-848.
- Dollinger, S. J., Urban, K. K., & James, T. A. (2004). Creativity and openness: Further validation of two creative product measures. *Creativity Research Journal, 16*(1), 35-47.
- Donin, N. (2009). Genetic criticism and cognitive anthropology: A reconstruction of Philippe Leroux's compositional process for *Voi(rex)*. In W. Kinderman & J. Jones (Eds.), *Genetic Criticism and the Creative Process: Essays from Music, Literature, and Theater* (pp. 192-215). Rochester, NY: University of Rochester Press.
- Donin, N., & Féron, F. X. (2012). Tracking the composer's cognition in the course of a creative process: Stefano Gervasoni and the beginning of *Gramigna*. *Musicae Scientiae, 16*(3), 262-285
- Donin, N., & Theureau, J. (2007). Theoretical and methodological issues related to long term creative cognition: The case of musical composition. *Cognition Technology & Work, 9*(4), 233-251.
- Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the big five across the life span: Evidence from two national samples. *Psychology and Aging, 23*(3), 558-556.
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*, 192-203.
- Downes, S. (2012). *Connectivism and connected knowledge: Essays on meaning and learning networks* (pp. 616). Retrieved from http://www.downes.ca/files/Connective_Knowledge-19May2012.pdf
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273.
- Eccles, J. S. (1983). Expectancies, values, and academic choice: Origins and changes. In J. Spence (Ed.), *Achievement and Achievement Motivation* (pp. 87-134). San Francisco, CA: W. H. Freeman and Company.
- Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity, 5*(3), 178-186.
- Esfandagheh, F. B., Harris, R., & Oreyzi, H. R. (2012). The impact of extraversion and pre-training self-efficacy on levels of training outcomes *Human Resource Development International, 15*(2), 175-191.

- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, 2(4), 290-309. doi: 10.1207/s15327957pspr0204_5
- Feist, G. J. (2010). The function of personality in creativity: The nature and nurture of the creative personality. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (pp. 113-130). New York, NY: Cambridge University Press.
- Folkestad, G., Hargreaves, D. J., & Lindström, B. (1998). Compositional strategies in computer-based music-making. *British journal of music education*, 15(1), 83-97.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research* 74(1), 59-109.
- Garrison, D. R. (2011). *E-learning in the 21st Century: A framework for research and practice*. New York, NY: Routledge.
- Getzels, J. W., & Jackson, P. W. (1958). The meaning of "Giftedness": An examination of an expanding concept. *The Phi Delta Kappan*, 40(2), 75-77.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt & F. Ostendorf (Eds.), *Personality Psychology in Europe*. Tilburg, The Netherlands: Tilburg University Press.
- Gong, Y. P., Huang, J. C., & Farh, J. L. (2009). Employee learning orientation, transformational leadership, and employee creativity: The mediating role of employee creative self-efficacy. *Academy of Management Journal*, 52(4), 765-778.
- Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement*, 66(6), 930-944.
- Greasley, A. E. (2008). *Engagement with music in everyday life: An in-depth study of adults' musical preferences and listening behaviours*. (Doctoral Dissertation), University of Keele. Available from Ethos Ethos database. (Thesis ID #496325)
- Greasley, A. E., & Lamont, A. (2011). Exploring engagement with music in everyday life using experience sampling methodology. *Musicae Scientiae*, 15(1), 45-71.
- Greene, B. A., & Miller, R. B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology*, 21(2), 181-192.
- Griffin, M., & McDermott, M. R. (1998). Exploring a tripartite relationship between rebelliousness, openness to experience, and creativity. *Social Behavior & Personality*, 26(4), 347-356.
- Gruber, H. E., & Wallace, D. B. (1999). The case study method and evolving systems approach for understanding unique creative people at work. In R. J. Sternberg (Ed.), *Handbook of Creativity* (pp. 93-115). New York, NY: Cambridge University Press.
- Grudnitski, A. M., & Krentler, K. A. (2011). Moving beyond satisfaction: Perceived learning as an assessment measure. *Journal of College Teaching And Learning*, 1(10), 7-16.
- Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*, 24(3), 175-213.
- Guilford, J. P. (1950). Creativity. *American Psychologist*, 5(9), 444-454.
- Guilford, J. P. (1968). *Intelligence, creativity and their educational implications*. San Diego, CA: Robert R. Knapp.

- Hancock, G. R. (2001). Effect size, power, and sample size determination for structured means modeling and mimic approaches to between-groups hypothesis testing of means on a single latent construct. *Psychometrika*, *66*(3), 373-388.
- Hardin, A. M., Chang, J. C., Fuller, M. A., & Torkzadeh, G. (2011). Formative measurement and academic research: In search of measurement theory. *Educational and Psychological Measurement*, *71*(2), 281-305.
- Hartnett, M., St. George, A., & Dron, J. (2011). Examining motivation in online distance learning environments: Complex, multifaceted, and situation-dependent. *The International Review of Research in Open and Distance Learning*, *12*(6), 20-38.
- Heilman, K. M., Nadeau, S. E., & Beversdorf, D. O. (2003). Creative innovation: Possible brain mechanisms. *NEUROCASE*, *9*(5), 369-379.
- Hennessey, B. A. (2004). *Developing creativity in gifted children: The central importance of motivation and classroom climate*. Storrs, CT: NRC/GT, University of Connecticut.
- Hennessey, B. A. (2010). The creativity-motivation connection. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (pp. 342-365). New York, NY: Cambridge University Press.
- Hickey, M. (2001). An application of Amabile's consensual assessment technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education*, *49*(3), 234-244. doi: 10.2307/3345709
- Hickey, M., & Webster, P. R. (2001). Creative thinking in music. *Music Educators Journal*, *88*(1), 19-23. doi: 10.2307/3399772
- Hill, P. (2013). Emerging Student Patterns in MOOCs: A (Revised) Graphical View. Retrieved from <http://mfeldstein.com/emerging-student-patterns-in-moocs-a-revised-graphical-view/>
- Hirche, K. L. (2011). *The potential impact of generative systems on learning experiences*. (Master of Arts), Queensland University of Technology. QUT ePrints database. (52760)
- Hosmer, D. W. J., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Hoboken, NJ: Wiley.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424-453.
- Jackson, J. J., Hill, P. L., & Roberts, B. W. (2012). Misconceptions of traits continue to persist: A response to Bandura. *Journal of Management*, *38*(3), 745-752.
- Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology*, *102*(3), 588-600.
- Jimerson, S. R., Campos, E., & Greif, J. L. (2003). Toward and understanding of definitions and measures of school engagement and related terms *California School Psychologist*, *8*, 7-27.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory - Versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.
- Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology*, *87*(4), 797-807.

- Judge, T. A., Jackson, C. L., Shaw, J. C., Scott, B. A., & Rich, B. L. (2007). Self-efficacy and work-related performance: The integral role of individual differences. *Journal of Applied Psychology, 92*(1), 107-127.
- Karwowski, M. (2011). It doesn't hurt to ask...But sometimes it hurts to believe: Polish students' creative self-efficacy and its predictors. *Psychology of Aesthetics, Creativity, and the Arts, 5*(2), 154-164. doi: 10.1037/a0021427
- Karwowski, M. (2012). Did curiosity kill the cat? Relationship between trait curiosity, creative self-efficacy and creative personal identity. *Europe's Journal of Psychology, 8*(4), 547-558. doi: 10.5964/ejop.v8i4.513
- Kaufman, A. S. (2009). *Creativity 101*. New York, NY: Springer.
- Kaufman, J. C., Baer, J., Cole, J. C., & Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the consensual assessment technique. *Creativity Research Journal, 20*(2), 171-178.
- Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four C model of creativity. *Review of General Psychology, 13*(1), 1-12.
- Kaufman, J. C., Evans, M. L., & Baer, J. (2010). The American Idol Effect: Are students good judges of their creativity across domains? *Empirical Studies of the Arts, 28*(1), 3-17.
- Kenny, B. J., & Gellrich, M. (2002). Improvisation. In R. Parncutt & G. E. McPherson (Eds.), (pp. 117-134). New York, NY: Oxford University Press.
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review, 2*(3), 196-217.
- Kim, K. H. (2005). Can only intelligent people be creative? A meta-analysis. *Journal of Secondary Gifted Education, 16*(2-3), 57-66.
- Kim, K. H., Cramond, B., & VanTassel-Baska, J. (2010). The relationship between creativity and intelligence. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (pp. 395-412). New York, NY: Cambridge University Press.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: The Guilford Press.
- Koestner, R., Ryan, R. M., Bernieri, F., & Holt, K. (1984). Setting limits on children's behavior: The differential effects of controlling vs. informational styles of intrinsic motivation and creativity. *Journal of Personality, 52*(3), 233-248.
- Koutsoupidou, T., & Hargreaves, D. J. (2009). An experimental study of the effects of improvisation on the development of children's creative thinking in music. *Psychology of Music, 37*(3), 251-278.
- Kozbelt, A., Beghetto, R. A., & Runco, M. A. (2010). Theories of creativity. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (pp. 20-47). New York, NY: Cambridge University Press.
- Kratus, J. (1989). A time analysis of the compositional processes used by children ages 7 to 11. *Journal of Research in Music Education, 37*(1), 5-20.
- Kratus, J. (1990a). Characterization of the compositional strategies used by children to compose a melody. *Canadian Journal of Research in Music Education, 33*, 95-103.
- Kratus, J. (1990b). Structuring the music curriculum for creative learning. *Music Educators Journal, 76*(9), 33-37.
- Kratus, J. (1994). Relationships among children's music audiation and their compositional processes and products. *Journal of Research in Music Education, 42*(2), 115-130.

- Kratus, J. (2001). Effect of available tonality and pitch options on children's compositional processes and products. *Journal of Research in Music Education*, 49(4), 294-306.
- Lamont, A. (2011a). The beat goes on: Music education, identity and lifelong learning. *Music Education Research*, 13(4), 369-388.
- Lamont, A. (2011b). University students' strong experiences of music: Pleasure, engagement, and meaning. *Musicae Scientiae*, 15(2), 229-249.
- Laverdière, O., Morin, A. J. S., & St-Hilaire, F. (2013). Factor structure and measurement invariance of a short measure of the Big Five personality traits *Personality and Individual Differences*, 55(7), 739–743.
- Law, L. N. C., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the Profile of Music Perception Skills. *PLOS ONE*, 7(12). doi: 10.1371/journal.pone.0052508
- Laycock, R. P. (1992). *The relationship of musical experience, musical aptitude, self-concept, age, and academic achievement to the musical problem-solving abilities of high*. (Doctoral Dissertation), Case Western Reserve University. (UMI #9236315)
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research*, 39, 329-358.
- Lee, K., & Ashton, M. C. (2008). The HEXACO personality factors in the indigenous personality lexicons of English and 11 other languages. *Journal of Personality*, 76, 1001-1053.
- Lee, N., & Cadogan, J. W. (2013). Problems with formative and higher-order reflective variables. *Journal of Business Research*, 66, 242-247.
- Legutki, A. (2010). *Self-determined music participation: The role of psychological needs satisfaction, intrinsic motivation, and self-regulation in the high school band experience*. (Doctoral Dissertation), University of Illinois at Urbana/Champaign. Retrieved from Illinois Digital Environment for Access to Learning and Scholarship (IDEALS) <http://hdl.handle.net/2142/16850>
- Lehmann, A. C., Sloboda, J. A., & Woody, R. H. (2007). *Psychology for musicians: Understanding and acquiring the skills*. New York, NY: Oxford University Press.
- Lemons, G. (2010). Bar drinks, rugas, and gay pride parades: Is creative behavior a function of creative self-efficacy? *Creativity Research Journal*, 22(2), 151-161. doi: 10.1080/10400419.2010.481502
- Levesque, C. S., Williams, G. C., D., E., Pickering, M. A., Bodenhamer, B., & Finley, P. J. (2007). Validating the theoretical structure of the treatment self-regulation questionnaire (TSRQ) across three different health behaviors. *Health Education Research*, 21, 691-702.
- Lo, C. C. (2010). How student satisfaction factors affect perceived learning. *Journal of the Scholarship of Teaching and Learning*, 10(1), 47-54.
- Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, 10(1), 21-39.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. W. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130-149.
- Mark, M. L., & Gary, C. L. (2007). *A history of American music education* (3rd ed.). New York, NY: Rowman & Littlefield Education.

- Martin, L. (2012). The musical self-efficacy beliefs of middle school band students: An investigation of sources, meanings, and relationships with attributions for success and failure. *Bulletin of the Council for Research in Music Education, 191*, 45-60.
- McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport, 60*(1), 48-58.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research *Intelligence, 37*, 1-10.
- McPherson, G. (2006). Self-efficacy and music performance. *Psychology of Music, 34*(3), 322-336.
- Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review, 3*, 220-332.
- Milojev, P., Osborne, D., Greaves, L. M., Barlow, F. K., & Sibley, C. G. (2013). The Mini-IPIP6: Tiny yet highly stable markers of Big Six personality. *Journal of Research in Personality, 47*, 936-944.
- Moore, B., & Stanley, T. (2010). How do critical thinking skills enhance student achievement? In B. Moore & T. Stanley (Eds.), *Critical Thinking and Formative Assessments: Increasing the Rigor in Your Classroom* (pp. 14-22). Larchmont, NY: Eye on Education, Inc.
- Morris, C. E. (2010). *A comparison of the use of composition as a teaching tool in music classrooms of the United States and United Kingdom*. (Master of Music Master's Thesis), University of Miami, Coral Gables, FL. Open Access Theses database.
- Mumford, M. D., Medeiros, K. E., & Partlow, P. J. (2012). Creative thinking: Processes, strategies, and knowledge. *Journal of Creative Behavior, 46*(1), 30-47.
- Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika, 29*(1), 81-117.
- Myers, D. E. (2008). Freeing music education from schooling: Toward a lifespan perspective on music learning and teaching. *International Journal of Community Music, 1*(1), 49-61.
- Nilsson, B., & Folkestad, G. (2005). Children's practice of computer-based composition. *Music Education Research, 7*(1), 21-37. doi: 10.1080/14613800500042042
- North, A. C., Hargreaves, D. J., & Hargreaves, J. J. (2004). Uses of music in everyday life. *Music Perception, 22*(1), 41-77.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535-569.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test *Behavior Research Methods, Instruments, & Computers, 32*(3), 396-402.
- Orman, E. K. (2002). Comparison of the National Standards for Music Education and elementary music specialists' use of class time. *Journal of Research in Music Education, 50*(2), 155-164.
- Pelletier, L. G., Seguin-Levesque, C., & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers' motivation and teaching behaviors *Journal of Educational Psychology, 94*, 186-196.

- Pike, G. R., Smart, J. C., & Ethington, C. A. (2012). The mediating effects of student engagement on the relationships between academic disciplines and learning outcomes: An extension of Holland's theory. *Research in Higher Education*, 53(5), 550-575.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322-338. doi: 10.1037/a0014996
- Prabhu, V., Sutton, C., & Sauser, W. (2008). Creativity and certain personality traits: Understanding the mediating effect of intrinsic motivation. *Creativity Research Journal*, 20(1), 53-66.
- Priest, T. (2006a). Self-evaluation, creativity, and musical achievement. *Psychology of Music*, 34(1), 47-61.
- Priest, T. (2006b). The reliability of three groups of judges' assessments of creativity under three conditions. *Bulletin of the Council for Research in Music Education*(167), 47-60.
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *JOURNAL OF RESEARCH IN PERSONALITY*, 41(1), 203-212.
- Randles, C. A. (2012). Music teacher as writer and producer. *Journal of Aesthetic Education*, 46(3), 36-52.
- Randles, C. A., & Stringham, D. (Eds.). (2013). *Musicianship: Composing in band and orchestra*. Chicago, IL: GIA Publications, Inc.
- Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, 44(3), 159-175.
- Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, 98(1), 209-218.
- Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion*, 28, 147-169.
- Reimer, B. (2003). *A philosophy of music education: Advancing the vision*. Upper Saddle River, NJ: Prentice Hall.
- Rhodes, M. (1961). An analysis of creativity. *The Phi Delta Kappan*, 42(7), 305-310. doi: 10.2307/20342603
- Rigdon, E. E. (2014). Comment on "Improper use of endogenous formative variables". *Journal of Business Research*, 67(1), 2800-2802.
- Ritchie, L., & Williamon, A. (2011a). Measuring distinct types of musical self-efficacy. *Psychology of Music*, 39(3), 328.
- Ritchie, L., & Williamon, A. (2011b). Primary school children's self-efficacy for music learning. *Journal of Research in Music Education*, 59(2), 146-161.
- Ritchie, L., & Williamon, A. (2012). Self-efficacy as a predictor of musical performance quality. *Psychology of Aesthetics, Creativity, and the Arts*, 6(4), 334-340. doi: 10.1037/a0029619
- Roca, J. C., & Gagné, M. (2008). Understanding e-learning continuance intention in the workplace: A self-determination theory perspective. *Computers in Human Behavior* 24(4), 1585-1604.
- Rosé, C. P., Carlson, R., Yang, D., Wen, M., Resnick, L., Goldman, P., & Sherer, J. (2014). *Social factors that contribute to attrition in MOOCs*. Paper presented at the The First ACM conference on Learning@ scale conference, Atlanta, GA.
- Rotgans, J. I., & Schmidt, H. G. (2011). Cognitive engagement in the problem-based learning classroom. *Advances in Health Sciences Education*, 16(4), 465-479.

- Rovai, A. P., Wighting, M. J., Baker, J. D., & Grooms, L. D. (2009). Development of an instrument to measure perceived cognitive, affective, and psychomotor learning in traditional and virtual classroom higher education settings. *Internet and Higher Education, 12*(1), 7-13.
- Rowley, J., & O'Dea, J. (2009). *How do students perceive the enhancement of their own learning? A comparison of two education faculties' experiences in building an online learning community for bachelor of music education and bachelor of education students*. Paper presented at the Teacher education crossing borders: Cultures, contexts, communities and curriculum: The annual conference of the Australian Teacher Education Association (ATEA), Albury, Australia.
<http://atea.edu.au/ConfPapers/2009/Refereed/Rowley.pdf>
- Runco, M. A. (2003). Education for creative potential. *Scandinavian Journal of Education, 47*(3), 317-324.
- Runco, M. A. (2007a). A hierarchical framework for the study of creativity. *New Horizons in Education, 55*(3), 1-9.
- Runco, M. A. (2007b). *Creativity: Theories and themes: Research, development, and practice*. New York, NY: Academic Press.
- Runco, M. A. (2008). Creativity and education. *New Horizons in Education, 56*(1), 96-104.
- Runco, M. A. (2010). Divergent thinking, creativity, and ideation. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge Handbook of Creativity* (pp. 413-446). New York, NY: Cambridge University Press.
- Runco, M. A. (in press). *The new science of creativity*. Philadelphia, PA: Taylor & Francis.
- Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal, 24*(1), 92-96.
- Rupayana, D. D. (2010). *Developing SAENS: Development and validation of a student academic engagement scale (SAENS)*. (Doctor of Philosophy Dissertation), Kansas State University. Available from ProQuest (UMI #3408149)
- Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57*, 749-761.
- Saucier, G. (2002). Orthogonal markers for orthogonal factors: The case of the Big Five. *Journal of Research in Personality, 36*, 1-31.
- Sawyer, R. K. (2011). The cognitive neuroscience of creativity: A critical review. *Creativity Research Journal, 23*(2), 137-154.
- Sawyer, R. K. (2012). *Explaining creativity: The science of human innovation* (2nd ed.). New York, NY: Oxford University Press.
- Schmidt, C. P. (2007). Intrinsic-mastery motivation in instrumental music: Extension of a higher order construct. *Bulletin of the Council for Research in Music Education, 173*, 7-23.
- Schmidt, C. P., & Sinor, J. (1986). An investigation of the relationships among music audiation, musical creativity, and cognitive style. *Journal of Research in Music Education, 34*(3), 160-172.
- Schunk, D. H., & Pajares, F. (2005). Competence perceptions and academic functioning. *Handbook of Competence and Motivation, 85-104*.

- Seddon, F. A., & O'Neill, S. A. (2003). Creative thinking processes in adolescent computer-based composition: An analysis of strategies adopted and the influence of instrumental music training. *Music Education Research*, 5(2), 125-137.
- Seltzer, B. K. (2013). *The nomological network of self-efficacy and psychometric properties of its general and specific measures*. (Doctor of Philosophy Dissertation), University of Minnesota. Retrieved from <http://purl.umn.edu/148904> Available from University of Minnesota Digital Conservancy database.
- Shea, P., & Bidjerano, T. (2009). Community of inquiry as a theoretical framework to foster “epistemic engagement” and “cognitive presence” in online education. *Computers & Education*, 52(3), 543-553.
- Shea, P., & Bidjerano, T. (2012). Learning presence as a moderator in the community of inquiry model. *Computers & Education*, 59(2), 316-326.
- Shea, P., Hayes, S., Smith, S. U., Vickers, J., Bidjerano, T., Pickett, A., . . . Jian, S. (2012). Learning presence: Additional research on a new conceptual element within the Community of Inquiry (CoI) framework. *Internet and Higher Education*.
- Shroff, R. H., & Vogel, D. R. (2009). Assessing the factors deemed to support individual student intrinsic motivation in technology supported online and face-to-face discussions. *Journal of Information Technology Education*, 8, 59-85.
- Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2(1). http://www.itdl.org/Journal/Jan_05/article01.htm
- Silvia, P. J. (2008). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual Differences*, 44, 1012-1021. doi: 10.1016/j.paid.2007.10.027
- Simonton, D. K. (1990). History, chemistry, psychology, and genius: An intellectual autobiography of historiometry. In M. A. Runco & R. S. Albert (Eds.), *Theories of Creativity* (pp. 92-115): Sage Publications, Inc.
- Simonton, D. K. (2012). Taking the U.S. Patent Office criteria seriously: A quantitative three-criterion creativity definition and its implications. *Creativity Research Journal*, 24(2-3), 97-106.
- Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions *Journal of Applied Psychology*, 93(2), 280-295.
- Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education*, 9(2), 169-191.
- Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85(4), 571-581.
- Sloboda, J. A. (Ed.). (2000). *Generative processes in music: The psychology of performance, improvisation, and composition* (reprinted paperback ed.). New York, NY: Oxford University Press.
- Sloboda, J. A., O'Neill, S. A., & Ivaldi, A. (2001). Functions of music in everyday life: An exploratory study using the Experience Sampling Method. *Musicae Scientiae*, 5(1), 9-32.

- Standage, M., Duda, J. L., & Ntoumanis, N. (2006). Students' motivational processes and their relationship to teacher ratings in school physical education: A self-determination theory approach. *Research Quarterly for Exercise and Sport*, 77(1), 100-110.
- Stefanic, N. M. (2011). *Developing notational understanding through iconic representation of rhythm: A descriptive study*. (Master of Music Education Unpublished Master's Thesis), University of Illinois.
- Stefanic, N. M. (in press). Structural equation modeling and multilevel modeling in music education: Advancing quantitative research data analysis. In C. A. Randles (Ed.), *Music Education: Navigating the Future*. New York, NY: Routledge.
- Stefanic, N. M., & Randles, C. (2014). Examining the reliability of scores from the consensual assessment technique in the measurement of individual and small group creativity. *Music Education Research*. doi: 10.1080/14613808.2014.909398
- Stringham, D. A. (2010). *Improvisation and composition in a high school instrumental music curriculum* (Doctoral dissertation), University of Rochester, Rochester, NY. ProQuest database. (AAT 3445843)
- Swanwick, K., & Tillman, J. (1986). The sequence of musical development: A study of children's composition. *British Journal of Music Education*, 3(3), 305-340.
- Tierney, P., & Farmer, S. M. (2002). Creative self-efficacy: Its potential antecedents and relationship to creative performance. *The Academy of Management Journal*(6), 1137. doi: 10.2307/3069429
- Tierney, P., & Farmer, S. M. (2011). Creative self-efficacy development and creative performance over time. *Journal of Applied Psychology*, 96(2), 277-293. doi: 10.1037/a0020952
- Torrance, E. P. (1966). *The Torrance Tests of Creative Thinking-Norms-Technical manual research edition-Verbal tests, forms A and B-Figural tests, forms A and B*. Princeton, NJ: Personnel Press.
- Torrance, E. P. (1998). *The Torrance tests of creative thinking norms—technical manual figural (streamlined) forms A & B*. Bensenville, IL: Scholastic Testing Service, Inc.
- Troum, J. F. (2010). *Empowered for practice: The relationship among perceived autonomy support, competence, and task persistence of undergraduate applied music students*. (Doctor of Philosophy Dissertation), University of South Florida. Retrieved from <http://scholarcommons.usf.edu/etd/1795> Available from University of South Florida Scholar Commons Graduate School Theses and Dissertations database.
- Tsai, Y. M., Kunter, M., Lüdtke, O., Trautwein, U., & Ryan, R. M. (2008). What makes lessons interesting? The role of situational and individual factors in three school subjects. *Journal of Educational Psychology*, 100(2), 460-472.
- Vallerand, R. J. (1997). *Toward a hierarchical model of intrinsic and extrinsic motivation*. San Diego, CA: Academic Press.
- Vallerand, R. J., & Bissonnette, R. (1992). Intrinsic, extrinsic, and amotivational styles as predictors of behavior: A prospective study. *Journal of Personality*, 60(3), 599-620.
- Vallerand, R. J., Blais, M. R., Briere, N. M., & Pelletier, L. G. (1989). Construction et validation de l'Echelle de Motivation en Education. *Canadian Journal of Behavioral Sciences*, 321, 323-349.

- Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senécal, C. B., & Vallières, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement, 52*, 1003-1017.
- Van Nuland, H. J. C., Taris, T. W., Boekaerts, M., & Martens, R. L. (2012). Testing the hierarchical SDT model: The case of performance-oriented classrooms. *European Journal of Psychology of Education, 27*(4), 467-482.
- Vandewalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement, 57*(6), 995-1015.
- Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic versus extrinsic goal contents in self-determination theory: Another look at the quality of academic motivation. *Educational Psychologist, 41*(1), 19-31.
- Vansteenkiste, M., Simons, J., Lens, W., & Sheldon, K. M. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology, 87*(2), 246-260.
- Vansteenkiste, M., Timmermans, T., Lens, W., Soenens, B., & Van den Broeck, A. (2008). Does extrinsic goal framing enhance extrinsic goal-oriented individuals' learning and performance? An experimental test of the match perspective versus self-determination theory. *Journal of Educational Psychology, 100*(2), 387-397.
- Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and Solutions in Human Assessment* (pp. 41-71). Boston, MA: Kluwer.
- Vold, J. N. (1986). *A study of musical problem solving behavior in kindergarten children and a comparison with other aspects of creative behavior*. Unpublished Dissertation. University of Alabama.
- Waldron, J. L. (2009). Exploring a virtual music 'community of practice': Informal music learning on the Internet. *Journal of music, technology, and education, 2*(2-3), 97.
- Waldron, J. L. (2011a). Conceptual frameworks, theoretical models and the role of YouTube: Investigating informal music learning and teaching in online music community. *Collected Work: International perspectives on music, technology & education. (AN: 2011-10551)*, 4(2-3), 189.
- Waldron, J. L. (2011b). Locating narratives in postmodern spaces: A cyber ethnographic field study of informal music learning in online community. *Action, Criticism, and Theory for Music Education, 10*(2), 32-60.
- Walker, C. O., Greene, B. A., & Mansell, R. A. (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learning and Individual Differences, 16*(1), 1-12.
- Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children*. New York, NY: Holt, Rinehart & Winston.
- Wallas, G. (1926). *The art of thought*. London: J. Cape.
- Wang, C. (1985). *Measures of creativity in sound and music*. Unpublished Testing Instrument. Retrieved from <http://www.uky.edu/~cecilia/MCSM/mcsm.htm>
- Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. Chichester, West Sussex, UK: Wiley.

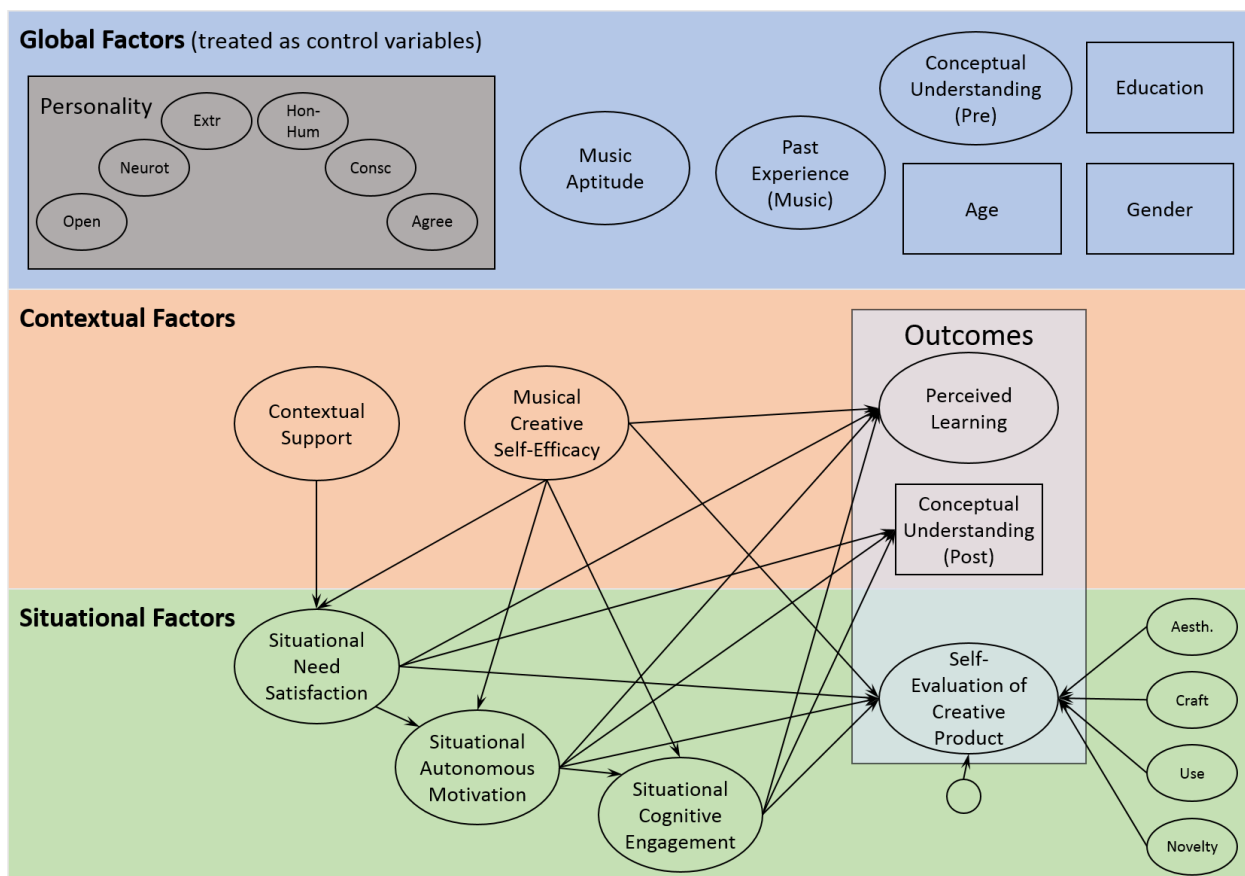
- Webster, P. R. (1979). Relationship between creative behavior in music and selected variables as measured in high school students. *Journal of Research in Music Education*, 27(4), 227-242.
- Webster, P. R. (1987a). Conceptual bases for creative thinking in music. In J. Peery, I. Peery & T. Draper (Eds.), *Music and Child Development* (pp. 158-174). New York, NY: Springer-Verlag.
- Webster, P. R. (1987b). Refinement of a measure of creative thinking in music. In C. Madsen & C. Prickett (Eds.), *Applications of Research in Music Behavior* (pp. 257-271). Tuscaloosa, AL: The University of Alabama Press.
- Webster, P. R. (1990). Creativity as creative thinking. *Music Educators Journal*, 76(9), 22.
- Webster, P. R. (2002). Creative thinking in music: Advancing a model. In T. Sullivan & L. Willingham (Eds.), *Creativity and Music Education (Research to Practice)* (pp. 16-33). Edmonton, AB: Canadian Music Educators' Association.
- Webster, P. R. (2013). Music composition intelligence and creative thinking in music. *Composing our Future: Preparing Music Educators to Teach Composition* doi: 10.1093/acprof:oso/9780199832286.001.0001
- Wellborn, J. G. (1992). *Engaged and disaffected action: The conceptualization and measurement of motivation in the academic domain*. (Doctor of Philosophy Dissertation), University of Rochester. Available from ProQuest ProQuest database. (UMI #9218563)
- Werner, P. D., Swope, A. J., & Heide, F. J. (2006). The music experience questionnaire: Development and correlates *Journal of Psychology*, 140(4), 329-345.
- Wiggins, G. A., Pearce, M. T., & Müllensiefen, D. (2009). Computational modelling of music cognition and musical creativity. In R. T. Dean (Ed.), *Oxford Handbook of Computer Music* (pp. 383-420). New York, NY: Oxford University Press.
- Wiggins, J. (1994). Children's strategies for solving compositional problems with peers. *Journal of Research in Music Education*, 42(3), 232-252.
- Wiggins, J. (2005). Fostering revision and extension in student composing. *Music Educators Journal*, 91(3), 35-42.
- Wiggins, J. (2009). *Teaching for musical understanding* (2nd ed.). Rochester, MI: Center for Applied Research in Musical Understanding, Oakland University.
- Wiggins, J., & Medvinsky, M. (2013). Scaffolding student composers. *Composing our Future: Preparing Music Educators to Teach Composition*. doi: 10.1093/acprof:oso/9780199832286.003.0006
- Williams, D. A. (2007). What are music educators doing and how well are we doing it? *Music Educators Journal*, 94(1), 18-23. doi: 10.1177/002743210709400105
- Willing, T. (2009). *Kindergarten students' tonal pattern creativity and developmental music aptitude*. (Master's Thesis), University of South Carolina. ProQuest database. (UMI #1483687)
- Willis, G. B. (1999). *Cognitive interviewing: A "how to" guide*. Paper presented at the Meeting of the American Statistical Association, Baltimore, MD.
- Woo, S. E., Chernyshenko, O. S., Longley, A., Zhang, Z. X., Chiu, C. Y., & Stark, S. E. (2013). Openness to experience: Its lower level structure, measurement, and cross-cultural equivalence. *Journal of Personality Assessment*, 96(1), 29-45. doi: DOI: 10.1080/00223891.2013.806328

- Xie, K., Debacker, T. K., & Ferguson, C. (2006). Extending the traditional classroom through online discussion: The role of student motivation *Journal of Educational Computing Research* 34(1), 67-89.
- Younker, B. A. (2000a). Composing with voice: Students' strategies and processes and the influence of text on the composed music. *The Canadian Music Educator*, 41(3 & 4), 91-99.
- Younker, B. A. (2000b). Thought processes and strategies of students engaged in music composition. *Research Studies in Music Education*, 14(1), 24-39.
- Younker, B. A., & Smith, W. H., Jr. (1996). Comparing and modeling musical thought processes of expert and novice composers. *Bulletin of the Council for Research in Music Education*, 128, 25-36.
- Zelenak, M. (2011). *Self-efficacy in music performance: Measuring the sources among secondary school music students*. (Doctoral Dissertation), University of South Florida. ProQuest database. (UMI #3466262)

APPENDIX A:

THE GENERAL SPECIFIED MODEL OF CREATIVITY-BASED LEARNING

This model represents the path model only (no measurement model), and in the ideal specification of the model. Several constructs shown here were not actually measured in this study.

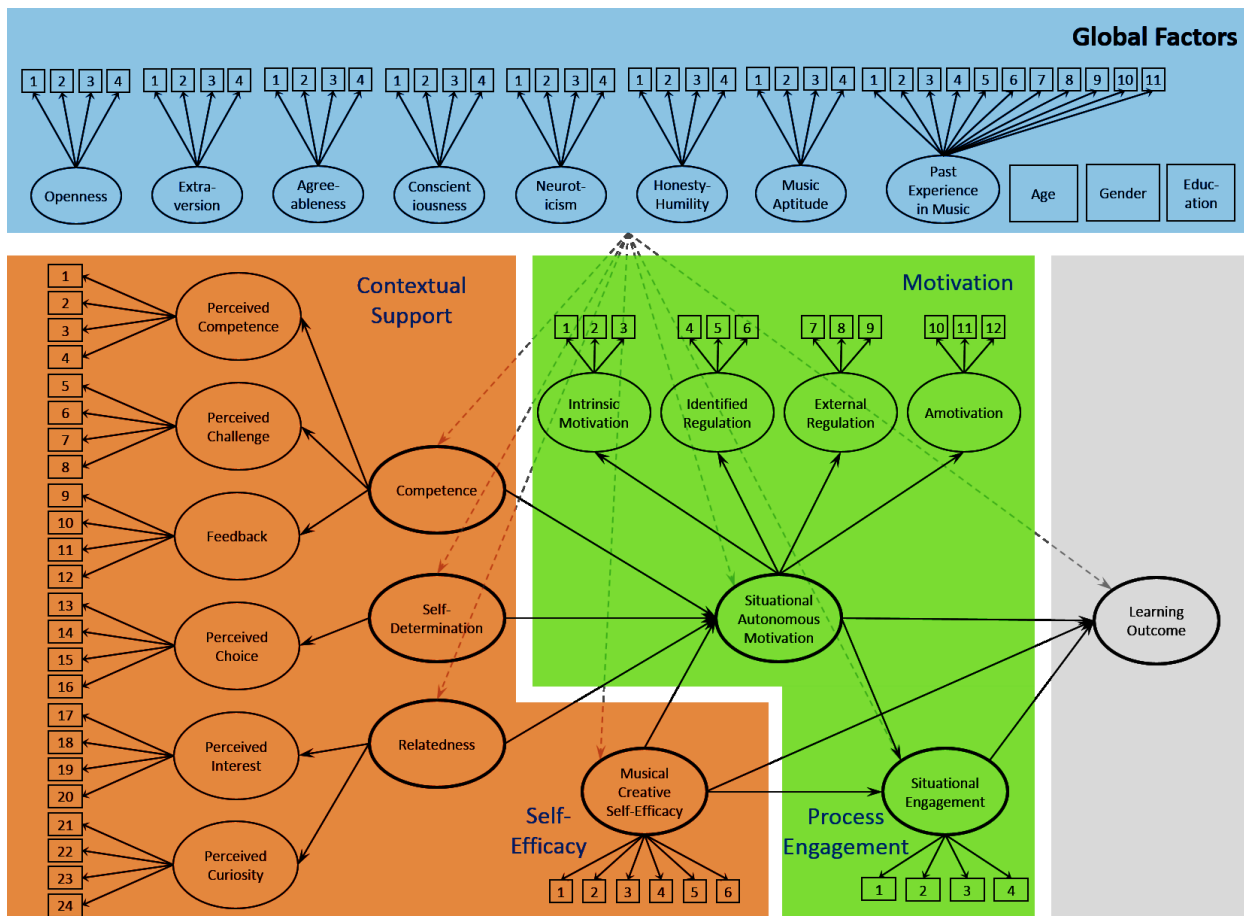


APPENDIX B:

THE GENERAL SPECIFICIED MODEL OF CREATIVITY-BASED LEARNING

(FULL STRUCTURAL MODEL)

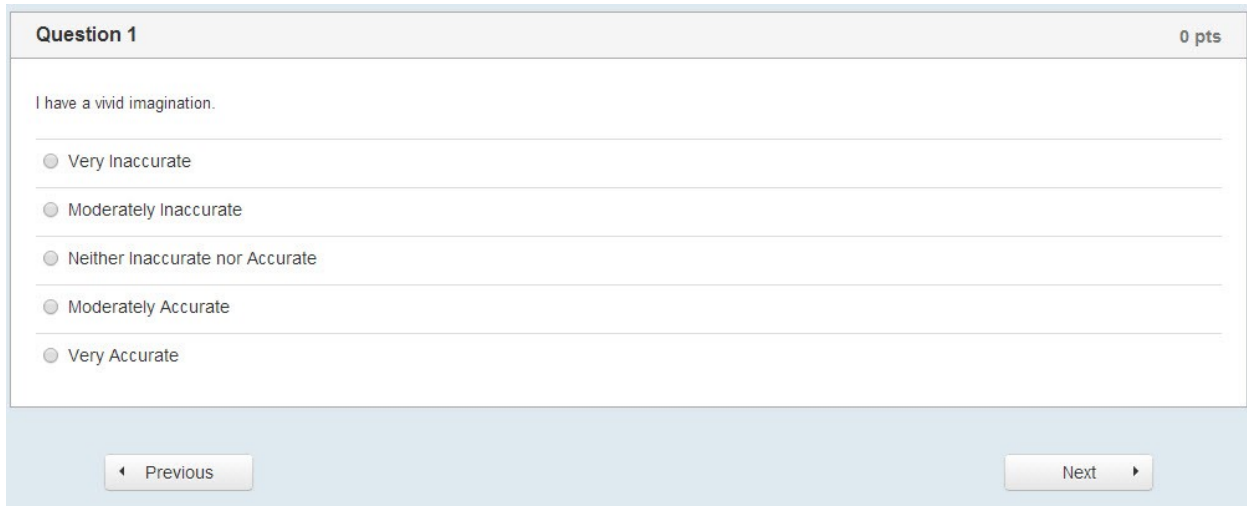
This is the full structural model (path model and measurement model). Had enough data been collected with a large enough sample this is the initial model that would have been estimated. To avoid too much visual confusion in the depiction of all paths, each dotted line represents a path from every single global factor to the factor to which that arrow points. Also, the measurement model for “Learning Outcome” is not depicted because it would be different for each specific learning outcome.



APPENDIX C: MEASUREMENT ITEMS

Items for each measured variable are presented below. While the items listed in this appendix are the actual text shown to participants, what is shown here is not how the items actually *looked* in the questionnaire. Most questionnaires were administered within Canvas, so response scales were either in the form of a multiple choice or drop-down menu. Two examples of what the questionnaires looked like are presented first, followed by items for each measurement instrument.

Examples of Questionnaire Appearance in Canvas



Question 1 0 pts

I have a vivid imagination.

Very Inaccurate

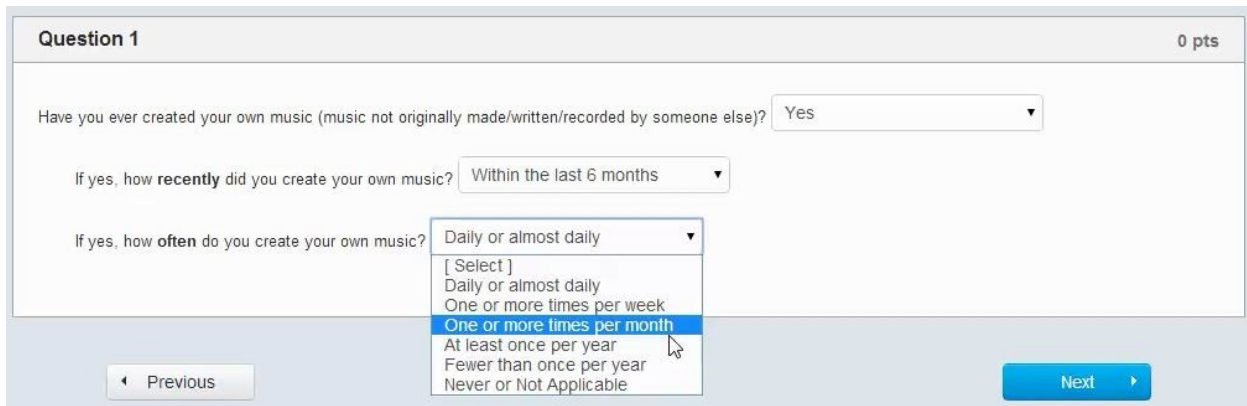
Moderately Inaccurate

Neither Inaccurate nor Accurate

Moderately Accurate

Very Accurate

◀ Previous Next ▶



Question 1 0 pts

Have you ever created your own music (music not originally made/written/recorded by someone else)? Yes

If yes, how **recently** did you create your own music? Within the last 6 months

If yes, how **often** do you create your own music?

Daily or almost daily

[Select]

Daily or almost daily

One or more times per week

One or more times per month

At least once per year

Fewer than once per year

Never or Not Applicable

◀ Previous Next ▶

Welcome to Canvas Network Survey

Item	Response Options	Code	Measurement Level
What is your primary reason for taking an open online course?			
	I like the format (online)	1	Unordered Categorical
	I enjoy learning about topics that interest me	2	
	I enjoy being part of a community of learners	3	
	I hope to gain skills for a new career	4	
	I hope to gain skills for a promotion at work	5	
	I am preparing to go back to school	6	
	I am preparing for college for the first time	7	
	I am curious about MOOCs	8	
	I want to try Canvas Network	9	
Not everyone has the same participation and learning goals. We welcome the diversity. Which type of online learner best describes you?			
	An observer. I just want to check the course out. Count on me to “surf” the content, discussions, and videos but don’t count on me to take any form of assessment.	1	Unordered Categorical
	A drop-in. I am looking to learn more about a specific topic within the course. Once I find it and learn it I will consider myself done with the course.	2	
	A passive participant. I plan on completing the course but on my own schedule and without having to engage with other students or assignments.	3	
	An active participant. Bring it on. If it’s in the course, I plan on doing it.	4	
How many hours a week are you planning to spend on this course?			
	Less than 1 hour	1	Ordered Categorical
	Between 1 and 2 hours	2	
	Between 2 and 4 hours	3	
	Between 4 and 6 hours	4	
	Between 6 and 8 hours	5	
	More than 8 hours per week		
What is your highest level of education?			
	High School or College Preparatory School	1	Ordered Categorical
	Some college, but have not finished a degree	2	
	Completed 2-year college degree	3	
	Completed 4-year college degree	4	
	Some graduate school	5	
	Master's Degree (or equivalent)	6	
	Ph.D., J.D., or M.D. (or equivalent)	7	
	None of these	.	

Is English your primary spoken language?		
Yes	1	Binary
No	0	
Where do you live?		
North America	1	Unordered Categorical
Central America	2	
South America	3	
Caribbean	4	
West Europe	5	
East Europe or Former Soviet Union	6	
Africa	7	
Middle East	8	
South Asia	9	
East Asia	10	
Southeast Asia	11	
Australia & South Pacific	12	
What is your gender?		
Male	1	Binary
Female	0	
How old are you?		
13-18	1	Ordered Categorical
19-24	2	
25-34	3	
35-44	4	
45-54	5	
55-64	6	
65 or older	7	

Pre-Course Questionnaire

Items 1-11 are the Past Experience in Music Inventory (PEMI). Items 12-14 are demographic items.

Instructions:

The first set of questions are about your previous experience with music and music-related activities.

If you answer "No" to any question, be sure to choose "Not Applicable" for the remaining questions on the page.

Please answer honestly and to the best of your knowledge.

Again, your answers will remain confidential.

Item	Response Options	Coding
1. Have you ever created your own music (music not originally made/written/recorded by someone else)?		
Yes	No	
1a.If yes, how recently did you create your own music?		
	Within the last week	6
	Within the last month	5
	Within the last 6 months	4
	Within the last year	3
	Within the last 5 years	2
	Longer than 5 years ago	1
	Never	0
1b.If yes, how often do you create your own music?		
	Daily or almost daily	5
	One or more times per week	4
	One or more times per month	3
	At least once per year	2
	Fewer than once per year	1
	Never	0
2. Have you ever improvised music?		
(Improvising is making music "on the spot" or "on the fly")		
Yes	No	
2a.If yes, how recently did you improvise music?		
	Within the last week	6
	Within the last month	5
	Within the last 6 months	4
	Within the last year	3
	Within the last 5 years	2
	Longer than 5 years ago	1
	Never	0
2b.If yes, how often do you improvise music?		
	Daily or almost daily	5
	One or more times per week	4
	One or more times per month	3
	At least once per year	2
	Fewer than once per year	1
	Never	0

3. Have you ever learned to play an instrument?

Yes No

3b.If yes, how recently did you play that instrument?

Within the last week	6
Within the last month	5
Within the last 6 months	4
Within the last year	3
Within the last 5 years	2
Longer than 5 years ago	1
Never	0

3c.If yes, how often do you play that instrument?

Daily or almost daily	5
One or more times per week	4
One or more times per month	3
At least once per year	2
Fewer than once per year	1
Never	0

4. Do you sing? (You do NOT have to sing professionally to answer yes)

Yes No

4a.If yes, how recently have you sung?

Within the last week	6
Within the last month	5
Within the last 6 months	4
Within the last year	3
Within the last 5 years	2
Longer than 5 years ago	1
Never	0

4b.If yes, how often do you sing?

Daily or almost daily	5
One or more times per week	4
One or more times per month	3
At least once per year	2
Fewer than once per year	1
Never	0

5. Have you ever recorded someone playing music, whether for professional recording purposes or as a hobby (this could include recording yourself)?

Yes No

5a.If yes, how recently did you play that instrument?

Within the last week	6
Within the last month	5
Within the last 6 months	4
Within the last year	3
Within the last 5 years	2
Longer than 5 years ago	1
Never	0

5b.If yes, how often do you play that instrument?

Daily or almost daily	5
One or more times per week	4

One or more times per month	3
At least once per year	2
Fewer than once per year	1
Never	0

6. Have you ever performed with a large music ensemble (greater than 5 people)?
Examples include (but are not limited to): a choir, large a cappella group, jazz big band, concert band, orchestra, large church music ensemble, steel drum band

Yes No

6a.If Yes, how many years of experience do you have?

_____ years

7. Have you ever performed with a small music ensemble (2-5 people)?Examples include (but are not limited to): in a small chamber ensemble (2-5 people), as a piano accompanist with a solo singer/instrumentalist (or vice versa), in a vocal/instrumental quartet, in a jazz combo, in a rock band

Yes No

7a.If Yes, how many years of experience do you have?

_____ years

8. Have you ever performed music, either live or for a recording, as a soloist (by yourself)?
Examples include (but are NOT limited to): as a singer/songwriter playing guitar and singing, as a DJ, as a solo pianist, as a solo singer with a backing track, as a performing street musician

Yes No

8a.If Yes, how many years of experience do you have?

_____ years

9. Do you currently work or have you worked in the past in a professional music career?
Examples include (but are not limited to): professional performing musician (solo or part of a group), recording or studio musician/artist, music critic, music producer, music recording engineer.

Yes No

9a.If Yes, how many years of experience do you have?

_____ years

10.Although you may or may not have worked in a professional music career, do you currently or have you in the past worked (for pay) as a part time, semi-professional or amateur musician?
Examples include (but are not limited to): performing musician (solo or part of a group), recording or studio musician/artist, music critic, music producer, music recording engineer.

Yes No

10a. If yes, for how many years did you/have you worked in this semi-professional/part-time/amateur role?

_____ years

11. Have you ever taken formal music lessons?

Formal music lessons would be a lesson that you take with an individual teacher separate from any group general music classes you currently have or have had in primary, middle, or secondary school.

Yes No

11a. Are you currently taking lessons

Yes No

11b. How recently was your last lesson?

Within the last week	6
Within the last month	5
Within the last 6 months	4
Within the last year	3
Within the last 5 years	2
Longer than 5 years ago	1
Never or Not Applicable	0

11c. How often do you have formal music lessons?

More than once per week	6
Weekly	5
About every two weeks	4
Monthly	3
Once every few months	2
Only a few times per year	1
Never or Not Applicable	0

11d. How many years of music lessons have you had?

_____ years

12. What is your age?

_____ years

13. Are you male or female?

Male

Female

Do not wish to answer

14. Have you completed your compulsory education for your country?

For example, in the U.S., compulsory education ends after the 12th grade (the end of high school)

Yes No

14a. If yes, what is the highest amount of education/degree you have completed?

Some College/University	1
Vocational/Technical School (2 year)	2
Associate's Degree (2 year)	3
Bachelor's degree	4
Master's degree	5
Doctoral Degree (Ph.D., Ed.D., etc.)	6
Professional Degree (MD, JD, etc.)	6
Other	Missing

Musical Creative Self-Efficacy

The items below were used in both the pilot study and the primary study. However, after analysis of the scale, the final three-factor solution used only the following items: items 1-3 for the General MCSE factor, items 7-12 for the Component MCSE factor, and items 13, 17, 18, and 19 for the Improve MCSE factor.

Confidence

(Enter a number between 0 and 100)

1. I am _____% confident that I can... **create new music.**
2. I am _____% confident that I can... **create an entire new simple song.**
3. I am _____% confident that I can... **create an entire music album or collection of songs.**

4. I am _____% confident that I can... **create a simple melody with an instrument.**
5. I am _____% confident that I can... **create a simple melody with my voice (with or without lyrics).**
6. I am _____% confident that I can... **create lyrics for a song.**
7. I am _____% confident that I can... **create a chord progression.**
8. I am _____% confident that I can... **create a "beat" or background accompaniment for a song.**

9. I am _____% confident that I can... **create a mashup from two or more other songs.**
10. I am _____% confident that I can... **create a new version of a song by remixing it.**
11. I am _____% confident that I can... **create an entirely new version of a song's melody by singing/playing it in a new or different way.**
12. I am _____% confident that I can... **make a new arrangement or entirely new version of an entire song.**

13. I am _____% confident that I can... **create new music spontaneously "on-the-spot" or "in the moment".**
14. I am _____% confident that I can... **improvise on an instrument.**
15. I am _____% confident that I can... **improvise with my voice (singing).**
16. I am _____% confident that I can... **improvise lyrics.**
17. I am _____% confident that I can... **improvise a melody (using either an instrument or your voice) to match a chord progression.**
18. I am _____% confident that I can... **improvise a melody (using either an instrument or your voice) to match a certain style/genre.**
19. I am _____% confident that I can... **carry on a musical "conversation" by singing/playing my instrument along with another musician.**

20. I am _____% confident that I can... **create music that I would think is pleasing or enjoyable to listen to.**
21. I am _____% confident that I can... **create music that others would think is pleasing or enjoyable to listen to.**
22. I am _____% confident that I can... **create music of similar quality to music I hear on the radio (or from other mass media sources).**
23. I am _____% confident that I can... **create music of better quality than the music I hear on the radio (or from other mass media sources).**
24. I am _____% confident that I can... **create music that would gain international popularity and high praise from music critics.**

Personality

Items from the mini-IPIP6 (Milojev et al., 2013). Response scale:

Very Inaccurate

Moderately Inaccurate

Neither Inaccurate nor Accurate

Moderately Accurate

Very Accurate

Instructions: On the following pages, there are phrases describing people's behaviors. Please use the rating scale below each phrase to describe how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Please read each statement carefully, and then choose an answer.

Personality Factor	Item	(R) indicates reverse-coded
Extraversion		
E1	I am the life of the party	
E2	I don't talk a lot	(R)
E3	I keep in the background	(R)
E4	I talk to a lot of different people at parties	
Agreeableness		
A1	I sympathize with others' feelings	
A2	I am not interested in other people's problems	(R)
A3	I feel others' emotions	
A4	I am not really interested in others	(R)
Conscientiousness		
C1	I get chores done right away	
C2	I like order	
C3	I make a mess of things	(R)
C4	I often forget to put things back in their proper place	(R)
Neuroticism		
N1	I have frequent mood swings	
N2	I am relaxed most of the time	(R)
N3	I get upset easily	
N4	I seldom feel blue	(R)
Openness to experience		
O1	I have a vivid imagination	
O2	I have difficulty understanding abstract ideas	(R)
O3	I do not have a good imagination	(R)
O4	I am not interested in abstract ideas	(R)
Honesty-Humility		
H1	I would like to be seen driving around in a very expensive car	(R)
H2	I would get a lot of pleasure from owning expensive luxury goods	(R)
H3	I feel entitled to more of everything	(R)
H4	I deserve more things in life	(R)

Contextual Support

Items modified from Schroff and Vogel's (2009) Intrinsic Motivation Inventory
The response scale was a 5-point Likert scale (Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree)

Instruction: For the following questions, you will be asked about your experience in the course so far, as it relates to the course activities.

The term "Course activities" includes discussions, projects, quizzes, and any other interactions you may have had as part of this course.

Please answer honestly. Your responses will remain confidential.

Constructs	Items
Perceived Competence	
COM1	I feel I am competent in my performance in the course activities.
COM2	I feel that my engagement in the course activities gives me competence.
COM3	I feel I am skilled in the course activities.
COM4	I feel I am capable in the course activities.
Perceived Challenge	
CHA1	I feel the course activities are challenging.
CHA2	I participate in the course activities because they are challenging.
CHA3	I like being challenged in the course activities.
CHA4	I like exerting effort in the course activities.
Feedback	
FEE1	The discussions (or the instructor) provide(s) positive feedback.
FEE2	I receive positive responses in the discussions or from the instructor.
FEE3	The comments I receive in the discussions (or from the instructor) are encouraging.
FEE4	I receive compliments in the course discussions (or from my instructor).
Perceived Choice	
CHO1	I believe I have some choice in the course activities.
CHO2	I feel like it is my own choice as to how much I participate in the course activities.
CHO3	I contribute in the course activities because I want to.
CHO4	I could make alternative selections in the course activities.
Perceived Interest	
INT1	I would say the course activities are very interesting.
INT2	I enjoy the course activities.
INT3	I feel that the course activities hold my attention.
INT4	I feel the course activities are fun to do.

Perceived Curiosity

- CUR1 I feel the course activities encouraged me to explore a variety of different issues.
- CUR2 I feel the course activities arouse my curiosity about the topics being addressed.
- CUR3 The course activities encourage me to discover issues that I may not have otherwise considered.
- CUR4 The course activities encourage me to look into issues that I may not have otherwise thought of.

Situational Motivation

Items were modified from the Situational Motivation Scale (SIMS) (Guay et al., 2000). Items were presented in random order. Items in red were created as modifications from the original SIMS, but for the sake of brevity, they were not included in the questionnaire administered to participants.

Participants responded using the following response scale:

- 1: corresponds not at all
- 2: corresponds a very little
- 3: corresponds a little
- 4: corresponds moderately
- 5: corresponds enough
- 6: corresponds a lot
- 7: corresponds exactly

Intrinsic Motivation

Because I thought that the projects/discussions were interesting.

Because I thought that doing the projects or being involved in the discussions was enjoyable.

Because the projects/discussions were fun.

Because I felt good when doing the projects/discussions.

Identified Regulation

Because I did it for my own good.

Because I thought that doing the projects/discussions was good for me.

Because I believed that the projects/discussions were important for me.

I did the projects/discussions by personal decision.

External Regulation

Because I was supposed to.

Because it was something that I had to do.

Because I felt that I had to do them.

Because I didn't have any choice.

Amotivation

There may have been good reasons to do the projects/discussions, but personally I didn't see any.

I did the projects/discussions, but I am not sure if it was worth it.

I don't know; I don't see what the projects/discussions brought me.

I did one or more of the projects/discussions, but I am not sure it was a good thing to pursue them.

Situational Cognitive Engagement

Items from Rotgans and Schmidt (2011)

Participants responded using the following response scale:

1 (not true at all for me)

2 (not true for me)

3 (neutral)

4 (true for me)

5 (very true for me)

Items

I was engaged with the topic at hand

I put in a lot of effort

I wish we could still continue with the work for a while

I was so involved that I forgot everything around me

Perceived Learning

The Perceived Learning items were written based on research in perceived learning and other measures reported in the literature. Due to the limited sample size, the factor structure could not be examined. Much more work is needed in developing this research instrument. See the discussion of this instrument in the section entitled *Descriptive Statistics for Each Variable* in Chapter Five.

Response Scale:

1	2	3	4	5	6	7
Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree

Instructions: The following questions are about your perceptions of your learning as a result of taking this course. Indicate the degree to which you agree or disagree with the following statements:

“As a result of taking this course...”

Items by (proposed) Perceived Learning Factor	R indicates reverse-coded
---	---------------------------

General Learning

I clearly understand the course content.	
I know more things.	
I do NOT expect to remember the content and issues covered in this course.	R
I learned.	
I learned almost nothing new.	R

Affective Learning

I understand the musical side of myself better	
I value music more.	
I am NOT more curious about music.	R
I am more interested in making music.	

Cognitive Learning

I can speak more intelligently about music.	
I do NOT feel that I am a more sophisticated musical thinker.	R
I can better analyze music.	

Psychomotor Learning

I have expanded my abilities to perform/play/sing music.	
I can perceive more or finer detail in the music I listen to.	
I do NOT feel more musically self-reliant (better able to make music on my own)	R

APPENDIX D:

EQUATIONS FOR COMBINING FREQUENCY AND RECENCY VARIABLES

Recency = time to present (in months)

$$\text{Frequency} = \frac{\text{count (in number of times)}}{\text{interval of time (in months)}}$$

$$\text{Combination of Recency and Frequency} = \sqrt{\text{Recency} \times \text{Frequency}}$$

$$= \sqrt{\frac{\text{count (in number of times)}}{\text{interval of time (in months)}} \times \text{time to present (in months)}}$$

$$= \sqrt{\frac{\text{number of times} \times \text{number of months}}{\text{number of months}}}$$

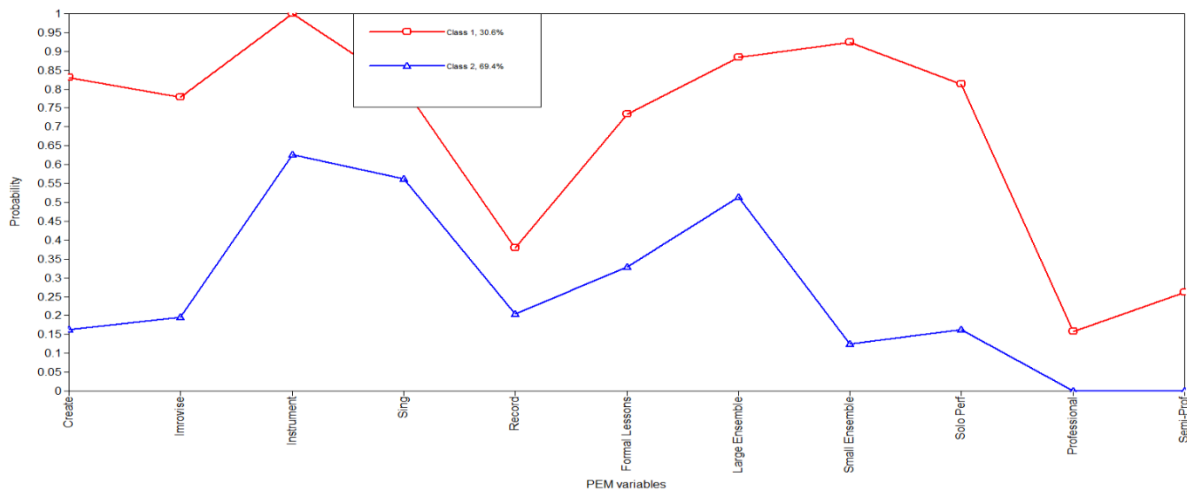
$$= \sqrt{\text{number of times}}$$

APPENDIX E:

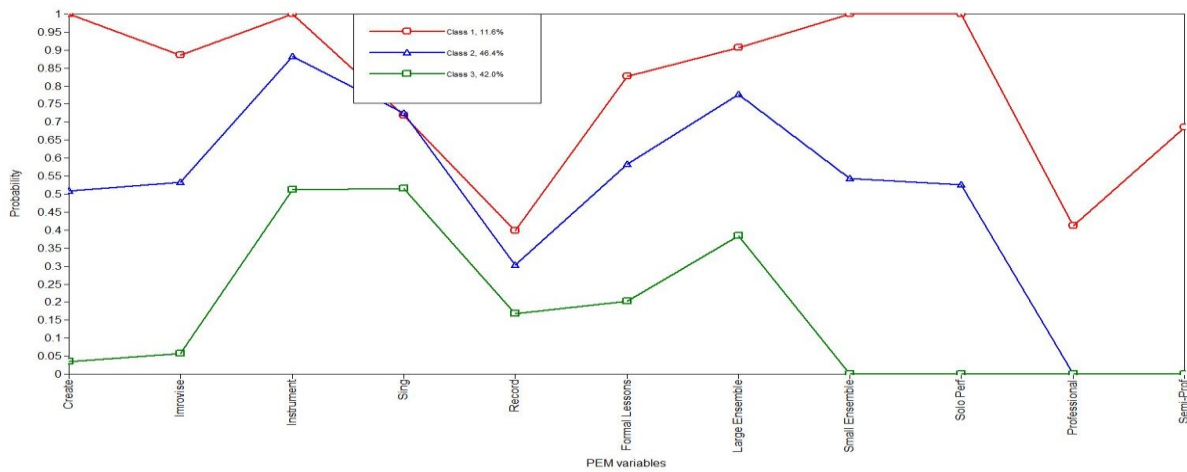
PROFILE PLOTS FROM PEM LATENT CLASS ANALYSIS (PILOT STUDY #2)

Note: For all profile plots below, the x-axis represents the 11 different PEM variables and the y-axis represents the probability that individuals answer “yes” for that item, given membership in the particular class.

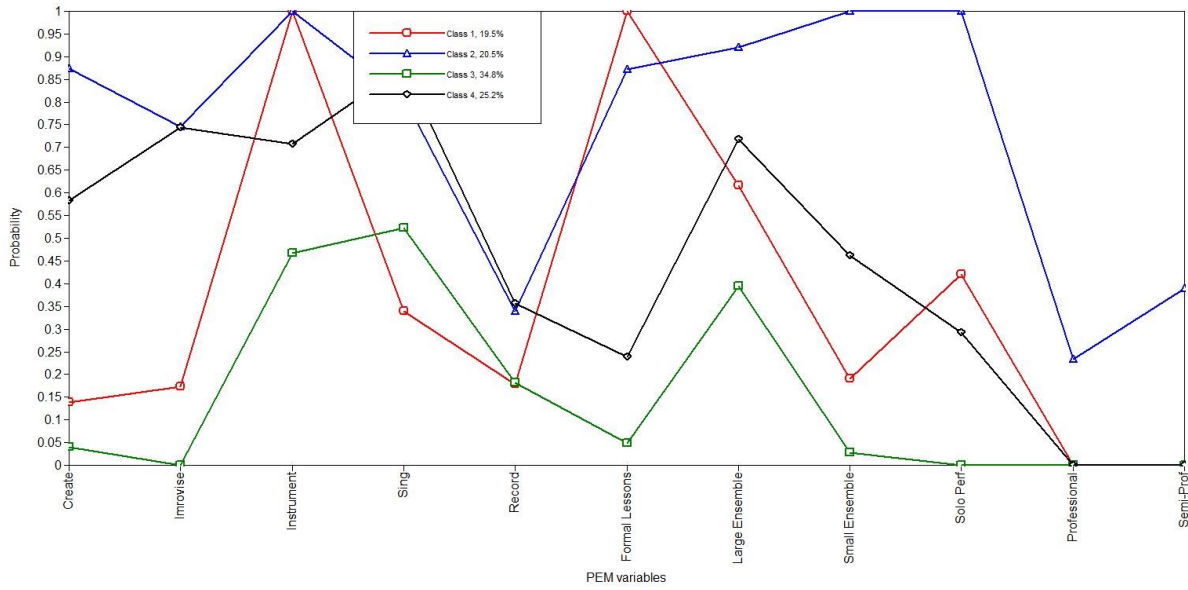
2-Class Unconditional Model



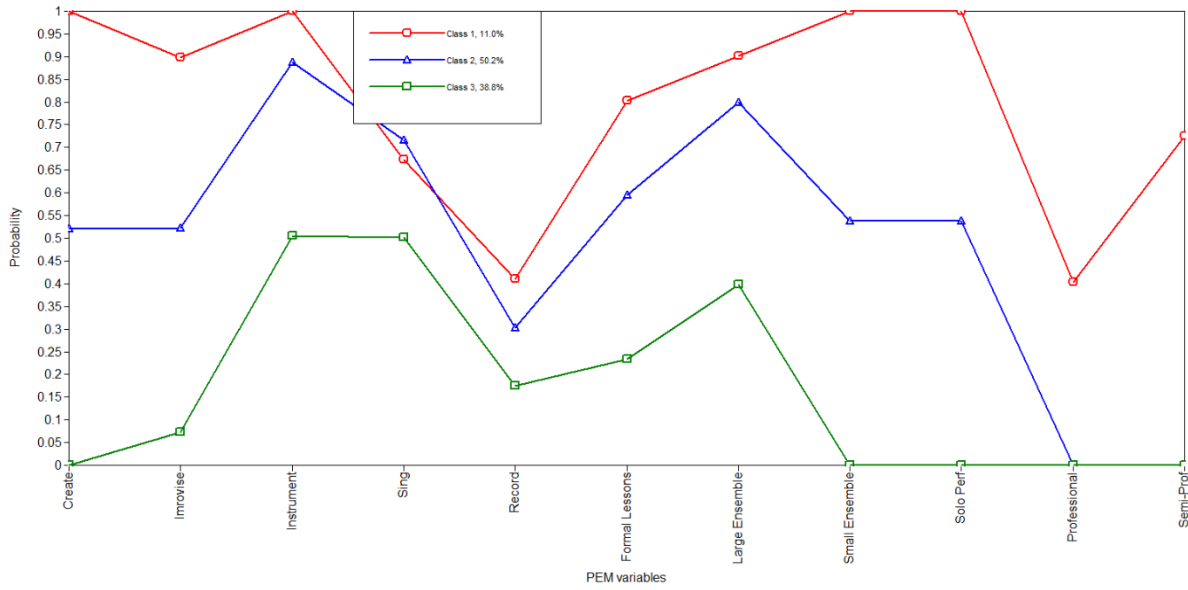
3-Class Unconditional Model



4-Class Unconditional Model



3-Class Conditional Model (Age and Gender as Covariates)

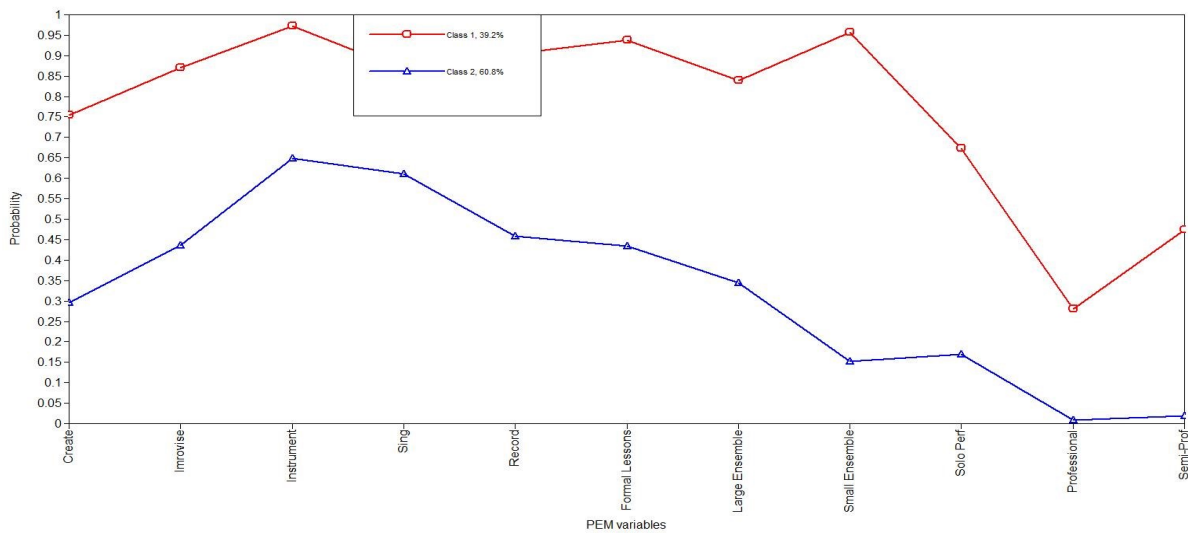


APPENDIX F:

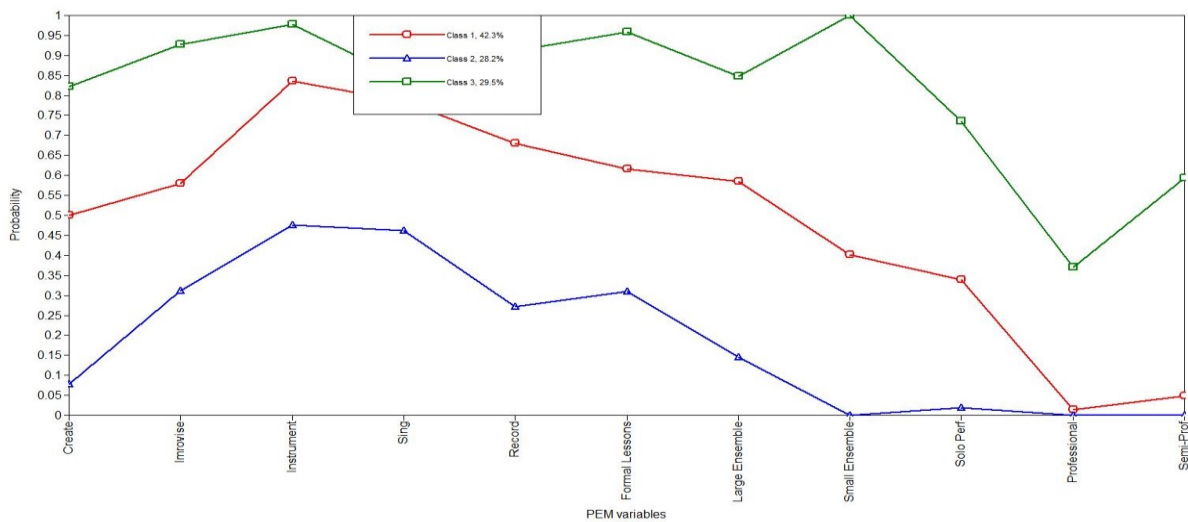
PROFILE PLOTS FOR PEM LATENT CLASS ANALYSIS (MOOC STUDY DATA)

Note: For all profile plots below, the x-axis represents the 11 different PEM variables and the y-axis represents the probability that individuals answer “yes” for that item, given membership in the particular class.

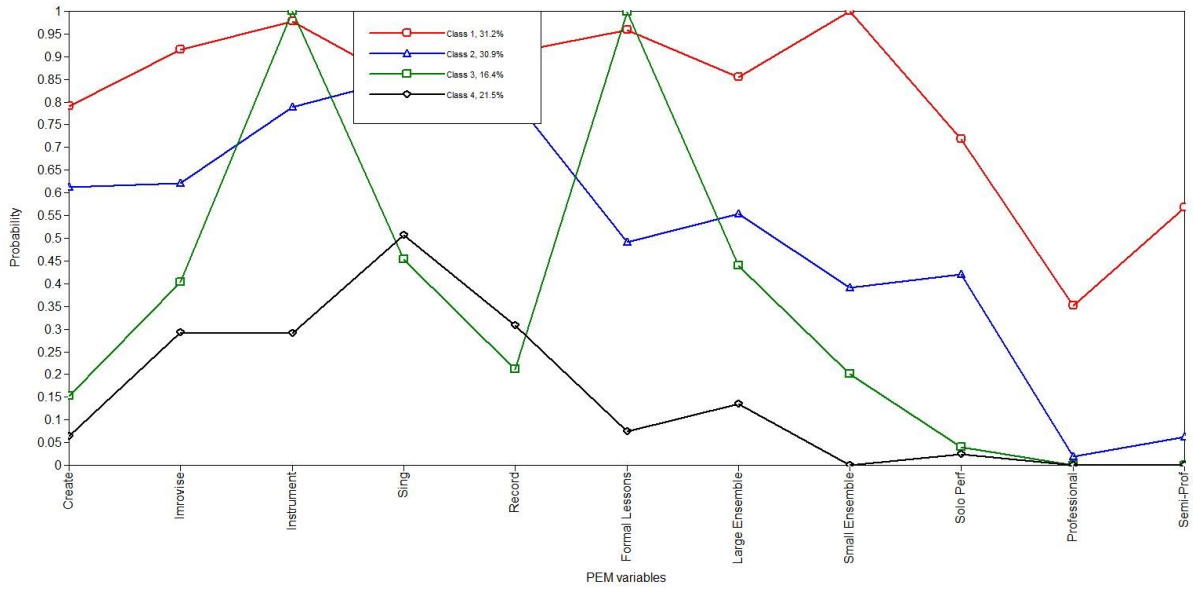
2-Class Unconditional Model



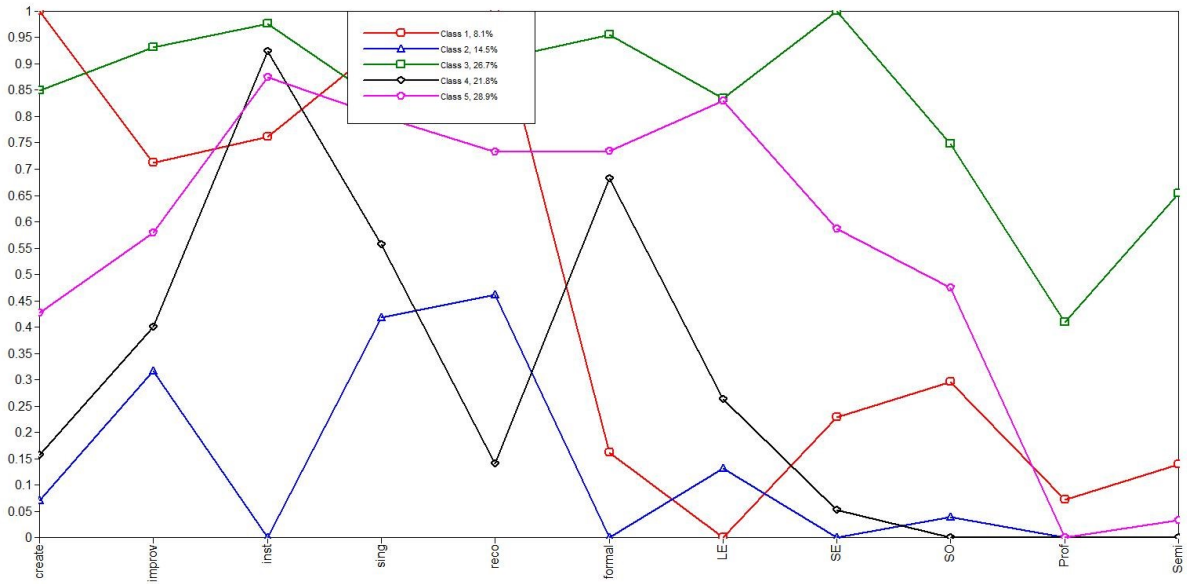
3-Class Unconditional Model



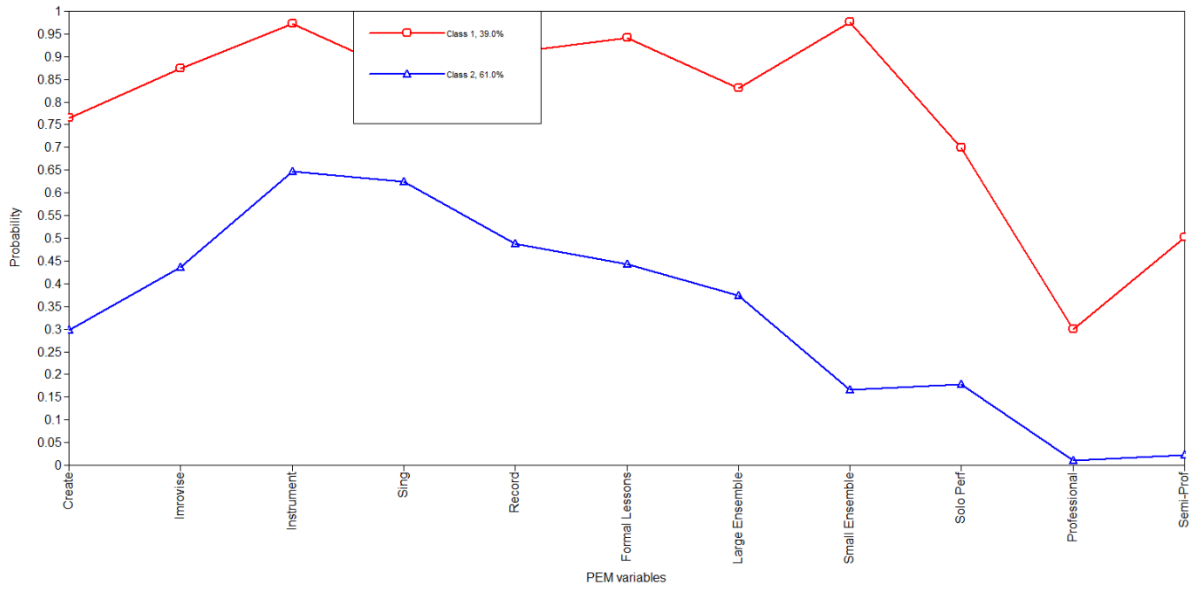
4-Class Unconditional Model



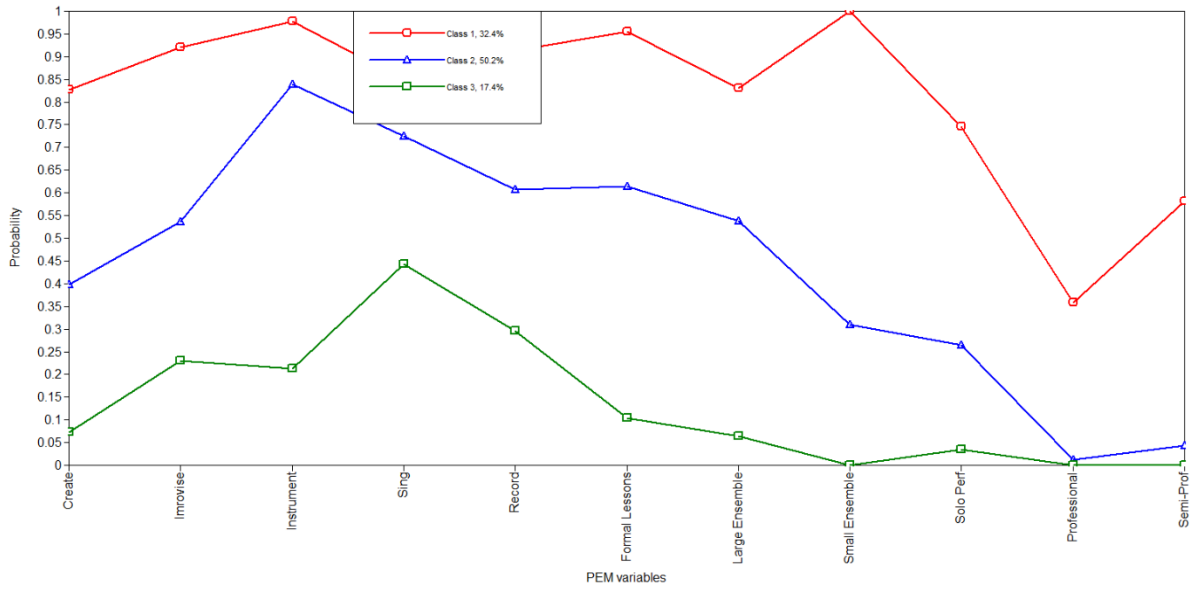
5-Class Unconditional Model



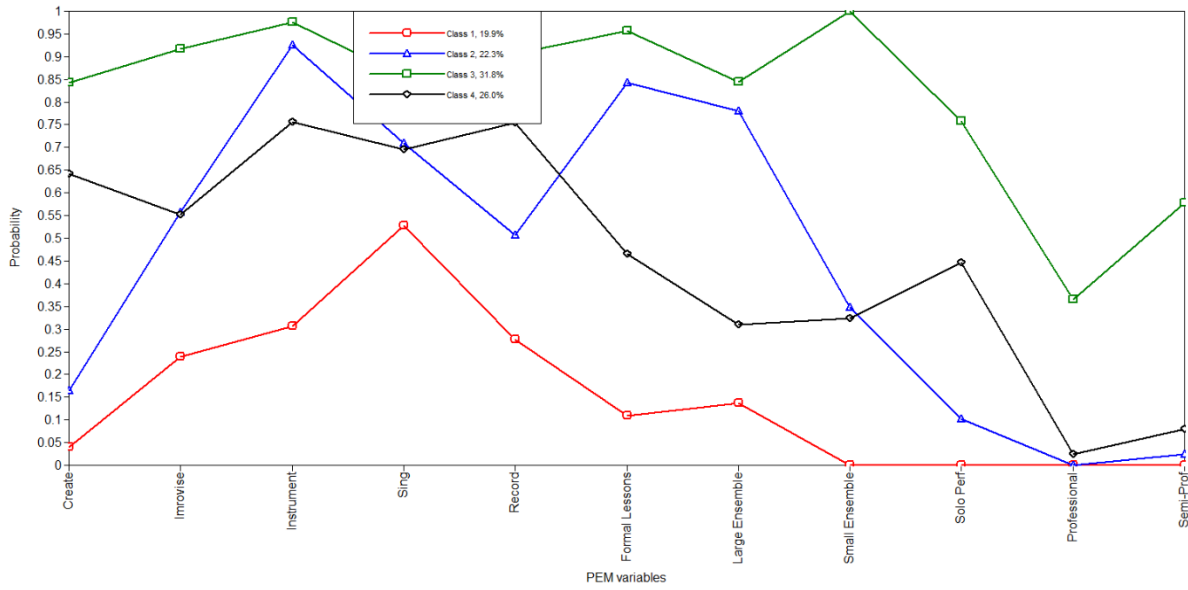
2-Class Conditional Model (Age, Gender, and Education as Covariates)



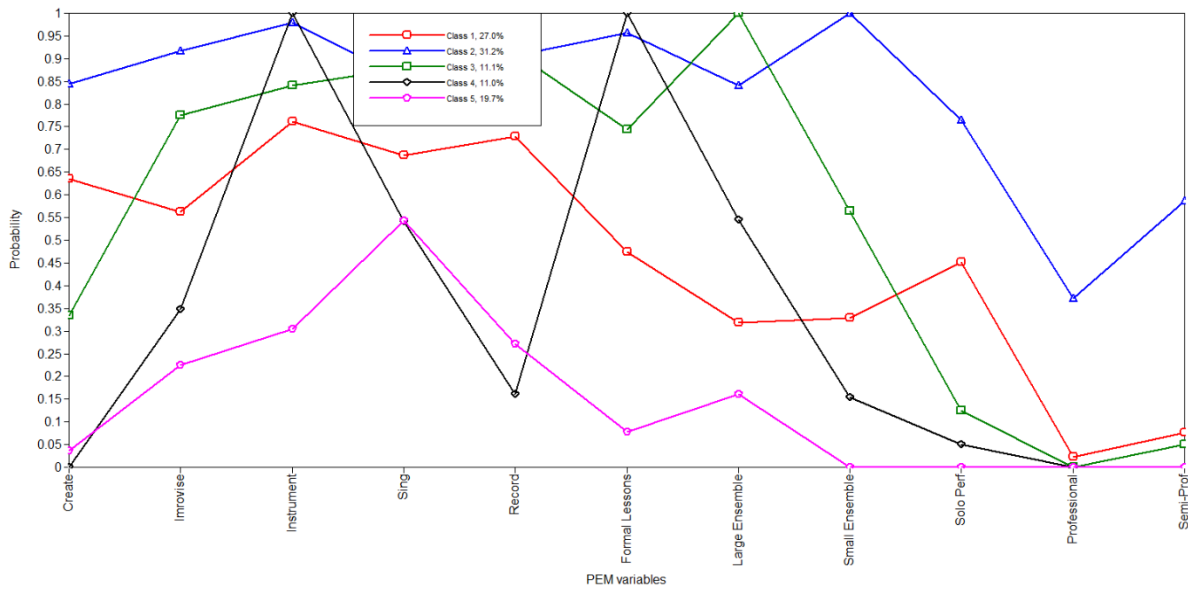
3-Class Conditional Model (Age, Gender, and Education as Covariates)



4-Class Conditional Model (Age, Gender, and Education as Covariates)



5-Class Conditional Model (Age, Gender, and Education as Covariates)



APPENDIX G: INTERCORRELATIONS (KENDALL'S TAU)

BETWEEN ALL NON-NOMINAL VARIABLES

Darker hues indicate higher absolute values of the correlation coefficient. Red indicates correlations with a positive sign. Blue indicates correlations with a negative sign.

	Variable	Demographics					MCSE				Personality					Contextual Support			Motivation				Project Participation													
		Male	Age	Educ	Actv Part	Hrs /WK	Engl	PEM	Gen	Compe	Improv	PROMS	Extrv	Agree	Consc	Neurot	Open	Hum-Hum	Chall	Choice	Compe	Cur	Fidbk	Int	IM	IR	ER	AM	Engage	PL	Any	BeyTop10	TotalNum			
Demographics	Male	1.000																																		
	Age (Ordinal)	0.16	1.000																																	
	Education	0.12	0.32	1.000																																
	Active Participant	0.19	-0.07	-0.146	1.000																															
	HoursWeek	0.128	-0.069	-0.092	0.266	1.000																														
	English	-0.242	0.069	-0.094	0.162	0.021	1.000																													
PEM class		-0.145	0.143	-0.001	-0.030	0.029	0.113	1.000																												
MCSE	General	0.164	0.019	-0.112	0.225	0.258	0.246	0.300	1.000																											
	Components	0.301	-0.008	-0.075	0.579	0.159	0.089	-0.160	0.676	1.000																										
	Improve	0.264	0.053	-0.019	0.479	0.252	0.140	-0.271	0.649	0.656	1.000																									
	PROMS Overall (4-factor mean)		-0.197	0.630	0.167	0.030	0.025	-0.268	-0.523	0.586	0.429	0.333	1.000																							
Personality	Extraversion	0.115	-0.038	0.080	-0.030	-0.120	-0.017	-0.032	0.367	0.299	0.293	0.056	1.000																							
	Agreeableness	-0.147	0.036	-0.018	0.069	0.038	0.031	0.061	0.293	0.112	0.279	-0.059	0.052	1.000																						
	Conscientiousness	0.140	0.226	0.100	0.124	0.091	0.044	0.014	-0.127	-0.147	-0.069	0.184	-0.029	-0.061	1.000																					
	Neuroticism	-0.148	-0.153	-0.119	-0.140	0.004	-0.171	-0.008	-0.156	-0.122	0.008	-0.047	-0.157	-0.012	-0.077	1.000																				
	Openness	0.162	-0.010	0.083	0.235	0.132	0.091	-0.267	0.285	0.269	0.303	0.210	0.057	0.096	0.024	-0.179	1.000																			
	Honesty-Humility	0.034	0.152	0.163	-0.001	0.015	0.077	-0.050	0.076	-0.057	0.177	0.070	0.063	0.176	-0.015	-0.234	0.076	1.000																		
Contextual Support	Challenge	0.069	0.124	-0.131	0.471	0.376	0.220	-0.095	0.512	0.524	0.419	0.600	0.072	0.315	-0.338	0.084	0.525	-0.147	1.000																	
	Choice	0.185	0.489	0.320	-0.128	0.125	-0.075	0.388	0.107	0.131	0.155	0.000	0.223	0.087	-0.282	-0.124	0.064	-0.088	0.365	1.000																
	Competence	0.356	-0.011	0.129	0.247	0.156	-0.261	-0.094	0.244	0.338	0.268	0.738	-0.060	0.145	-0.199	-0.120	0.348	0.317	0.330	0.197	1.000															
	Curiosity	0.302	0.168	0.259	0.137	0.048	-0.145	-0.152	0.393	0.463	0.370	0.316	0.071	0.262	-0.184	-0.071	0.331	0.157	0.372	0.339	0.632	1.000														
	Feedback	0.028	-0.012	0.013	0.366	0.123	-0.089	-0.237	0.334	0.346	0.310	0.527	-0.157	0.376	-0.213	0.012	0.338	0.245	0.372	0.157	0.681	0.627	1.000													
	Interest	0.165	0.067	0.079	0.165	0.240	-0.261	-0.058	0.372	0.431	0.466	0.316	0.129	0.331	-0.232	0.129	0.366	0.168	0.436	0.372	0.632	0.651	0.692	1.000												
Motivation	Intrinsic	-0.158	0.016	0.222	0.000	-0.167	-0.008	0.000	0.913	0.913	0.833	N.D.	0.927	0.222	-0.316	-0.105	0.589	-0.105	0.252	0.252	0.845	0.818	0.667	0.778	1.000											
	Identified Regulation	0.880	0.880	0.111	0.408	0.167	0.136	0.236	0.880	0.673	0.333	0.500	0.105	0.667	0.738	-0.105	0.589	-0.316	0.258	0.258	0.504	0.825	0.800	0.667	0.800	1.000										
	External Regulation	0.000	0.258	0.000	0.516	-0.316	0.000	0.307	0.000	0.333	0.000	0.000	-0.200	-0.105	-0.290	0.400	-0.447	0.000	0.120	0.120	-0.359	0.224	-0.105	-0.105	-0.316	-0.105	1.000									
	Amotivation	0.722	0.722	-0.256	-0.289	-0.177	-0.289	0.516	0.713	-0.548	-0.548	0.000	-0.224	0.825	0.671	0.000	-0.750	0.447	0.802	0.802	-0.401	0.750	0.843	0.825	0.589	0.843	0.224	1.000								
Situational Engagement	-0.680	-0.680	0.111	0.408	0.167	0.136	-0.236	0.680	0.673	0.333	0.680	0.105	0.667	-0.738	-0.105	0.589	-0.316	0.258	0.258	0.504	0.825	0.800	0.667	0.800	-0.105	0.800	1.000									
Perceived Learning (General)	0.160	-0.338	-0.144	0.061	0.021	-0.561	0.088	0.725	0.725	0.322	0.816	0.072	-0.084	-0.195	0.071	-0.110	-0.177	0.435	0.465	0.827	0.591	0.409	0.597	0.882	0.504	-0.359	-0.401	0.504	1.000							
Project Participation	Project Participation Any	0.004	-0.023	-0.092	0.301	0.078	0.142	-0.052	0.248	0.095	0.140	-0.070	-0.100	0.078	0.191	-0.107	0.278	0.096	0.122	-0.167	0.443	0.241	0.475	0.382	0.667	0.667	0.000	-0.530	0.667	0.144	1.000					
	Project Participation Beyond Top 10	-0.023	-0.049	-0.118	0.211	0.149	0.154	0.012	0.410	0.360	0.274	0.140	-0.139	0.008	0.123	-0.054	0.057	0.029	0.513	-0.014	0.178	0.192	0.239	0.123	0.000	0.408	0.516	-0.289	0.408	0.275	0.382	1.000				
	Total Num Projects Participated	0.000	-0.026	-0.101	0.304	0.095	0.150	-0.040	0.306	0.250	0.223	-0.039	-0.122	0.066	0.178	-0.093	0.232	0.078	0.600	0.057	0.363	0.264	0.475	0.330	0.333	0.667	0.316	-0.589	0.667	0.220	0.492	1.000				

Demographics

Variable		Male	Age	Educ	Actv Part	Hrs /Wk	Engl	Regn
Demographics	Male	1.000						
	Age (Ordinal)	.016	1.000					
	Education	.112	.321	1.000				
	Active Participant	.019	-.007	-.146	1.000			
	Hours/Week	.128	-.069	-.092	.266	1.000		
	English	-.242	.069	-.094	.162	.021	1.000	
	Region	.185	-.072	.114	-.115	.018	-.626	1.000
PEM class	-.145	.143	-.001	-.030	.029	.113	-.018	
MCSE	General	.164	.019	-.112	.625	.258	.246	-.222
	Components	.301	-.008	-.075	.579	.159	.089	-.071
	Improvise	.264	.053	-.019	.479	.252	.140	-.062
PROMS Overall (4-factor mean)	-.197	-.610	.167	.030	.025	.268	-.285	
Personality	Extraversion	.115	-.038	.080	-.030	-.120	-.017	.011
	Agreeableness	-.147	.036	-.018	.069	.038	.031	.071
	Conscientiousness	.146	.226	.100	.124	.091	.104	-.152
	Neuroticism	-.148	-.153	-.119	-.140	.004	-.171	.174
	Openness	.162	-.010	.083	.235	.132	.091	-.244
	Honesty-Humility	.034	.152	.163	-.001	.015	.077	.004
Contextual Support	Challenge	.069	.124	-.131	.471	.376	.220	-.253
	Choice	.185	.489	.320	-.128	.125	-.075	.300
	Competence	.356	-.011	.129	.247	.156	-.261	.013
	Curiosity	.302	.168	.259	.137	.048	-.145	.025
	Feedback	.028	-.012	.013	.366	.123	-.089	.013
	Interest	.165	.067	.079	.165	.240	-.261	.163
Motiv- ation	Intrinsic	-.136	-.816	.222	.000	-.167	-.408	.680
	Identified Regulation	-.680	-.680	.111	.408	.167	.136	.408
	External Regulation	.000	.258	-.949	.516	-.316	.000	.258
	Amotivation	.722	.722	-.236	-.289	-.177	-.289	-.289
Situational Engagement	-.680	-.680	.111	.408	.167	.136	.408	
Perceived Learning (General)	.160	-.338	-.144	.061	.621	-.561	.583	
Project Particip- ation	Project Participation Any	.004	-.023	-.092	.301	.078	.142	-.134
	Project Participation Beyond Top 10	-.023	-.049	-.118	.211	.149	.154	-.123
	Total Num Projects Participated	.000	-.026	-.101	.304	.095	.150	-.139

MCSE

Variable		Compo-				
		PEM	Gen	nent	Imprv	PROMS
PEM class		1.000				
MCSE	General	-.300	1.000			
	Components	-.160	.676	1.000		
	Improvise	-.271	.649	.656	1.000	
PROMS Overall (4-factor mean)		-.523	.586	.429	.333	1.000
Personality	Extraversion	-.032	.367	.290	.293	.056
	Agreeableness	-.061	.293	.112	.279	-.059
	Conscientiousness	.014	-.127	-.147	-.069	-.184
	Neuroticism	-.008	-.156	-.122	.008	-.047
	Openness	-.267	.285	.269	.303	.210
	Honesty-Humility	-.050	.076	-.057	.177	.070
Contextual Support	Challenge	-.095	.512	.524	.419	.800
	Choice	.388	.107	.131	.155	.000
	Competence	-.094	.244	.338	.268	.738
	Curiosity	-.152	.393	.463	.370	.316
	Feedback	-.237	.334	.346	.310	.527
	Interest	-.058	.372	.431	.466	.316
Motiv- ation	Intrinsic	.000	.913	.913	.183	
	Identified Regulation	-.236	1.000	.667	.333	1.000
	External Regulation	.707	.000	.333	.000	-1.000
	Amotivation	.516	-.913	-.548	-.548	-1.000
Situational Engagement		-.236	1.000	.667	.333	1.000
Perceived Learning (General)		.088	.725	.725	.322	.816
Project Particip- ation	Project Participation Any	-.052	.248	.095	.140	-.070
	Project Participation Beyond Top 10	.012	.410	.360	.274	.140
	Total Num Projects Participated	-.040	.306	.250	.223	-.039

		Personality					
Variable		Extrv	Agree	Consc	Neurot	Open	Hon-Hum
Personality	Extraversion	1.000					
	Agreeableness	.052	1.000				
	Conscientiousness	-.029	-.061	1.000			
	Neuroticism	-.157	-.012	-.077	1.000		
	Openness	.057	.096	.024	-.179	1.000	
	Honesty-Humility	.063	.176	-.015	-.234	.076	1.000
Contextual Support	Challenge	.072	.315	-.338	.084	.525	-.147
	Choice	.223	.087	-.282	-.124	.064	-.088
	Competence	-.060	.145	-.199	-.120	.348	.317
	Curiosity	.071	.262	-.184	-.071	.331	.157
	Feedback	-.157	.376	-.213	.012	.338	.245
	Interest	.129	.331	-.232	.129	.366	.168
Motiv-ation	Intrinsic	.527	.222	-.316	-.105	.589	-.105
	Identified Regulation	.105	.667	-.738	-.105	.589	-.316
	External Regulation	-.200	-.105	-.200	.400	-.447	.000
	Amotivation	-.224	-.825	.671	.000	-.750	.447
Project Particip-ation	Situational Engagement	.105	.667	-.738	-.105	.589	-.316
	Perceived Learning (General)	.072	-.084	-.195	.071	-.110	-.177
	Project Participation Any	-.100	.078	.191	-.107	.278	.096
Project Particip-ation	Project Participation Beyond Top 10	-.139	.008	.123	-.054	.057	.029
	Total Num Projects Participated	-.122	.066	.178	-.093	.232	.078

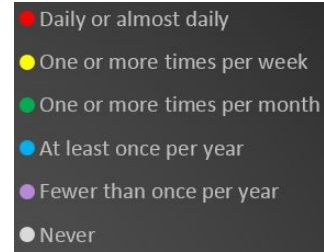
		Contextual Support					Motivation				
Variable		Chall	Choice	Compe- tence	Cur	Fdbk	Int	IM	IR	ER	AM
Contextual Support	Challenge	1.000									
	Choice	.365	1.000								
	Competence	.330	.197	1.000							
	Curiosity	.372	.339	.632	1.000						
	Feedback	.372	.157	.681	.627	1.000					
	Interest	.436	.372	.632	.653	.692	1.000				
Motiv- ation	Intrinsic	.252	.252	.882	.943	.667	.778	1.000			
	Identified Regulation	.756	.756	.504	.825	1.000	.889	.667	1.000		
	External Regulation	.120	.120	-.359	-.224	-.105	-.105	-.316	-.105	1.000	
	Amotivation	-.802	-.802	-.401	-.750	-.943	-.825	-.589	-.943	.224	1.000
	Situational Engagement	.756	.756	.504	.825	1.000	.889	.667	1.000	-.105	-.943
	Perceived Learning (General)	.435	.465	.827	.591	.409	.597	.882	.504	-.359	-.401
Project Particip- ation	Project Participation Any	.122	-.167	.443	.241	.475	.382	.667	.667	.000	-.530
	Project Participation Beyond Top 10	.513	-.014	.178	.192	.239	.123	.000	.408	.516	-.289
	Total Num Projects Participated	.600	.057	.363	.264	.475	.330	.333	.667	.316	-.589

		Project Participation				
Variable		Engage	PL	Any	BeyTop 10	Total Num
Situational Engagement		1.000				
Perceived Learning (General)		.504	1.000			
Project Particip- ation	Project Participation Any	.667	.144	1.000		
	Project Participation Beyond Top 10	.408	.275	.382	1.000	
	Total Num Projects Participated	.667	.220	.967	.492	1.000

APPENDIX H:

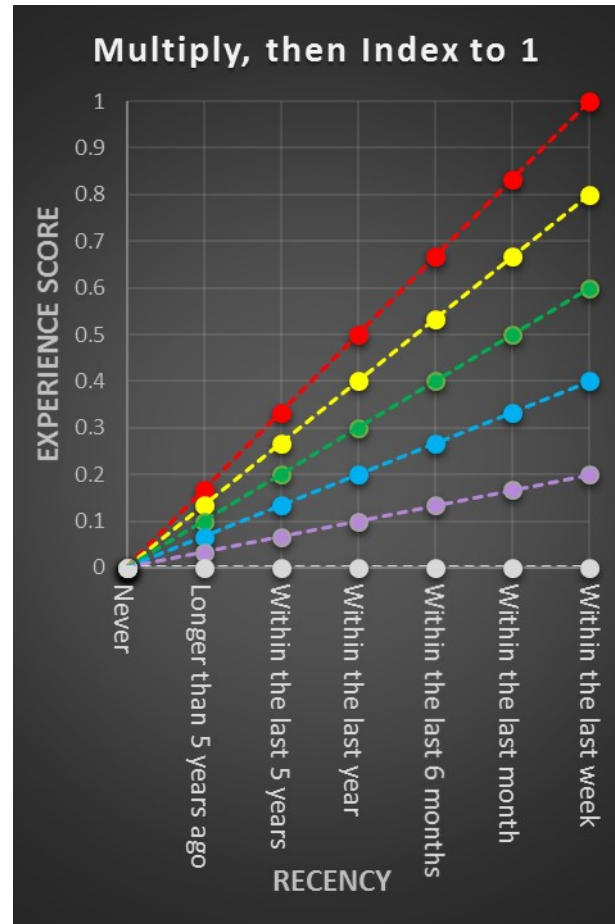
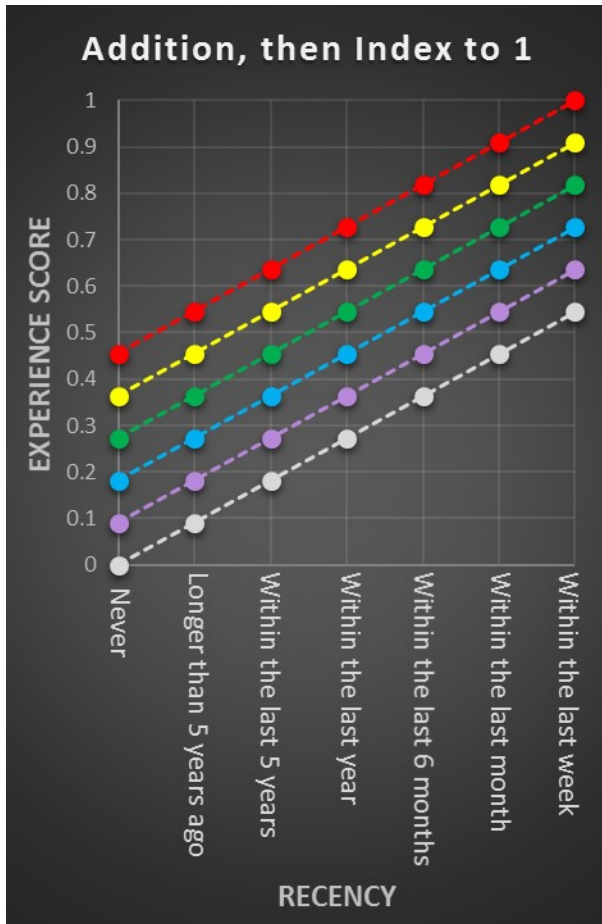
PLOTS FOR POSSIBLE PEM FREQUENCY AND RECENCY COMPARISONS

For all graphs in this appendix, use the following legend:



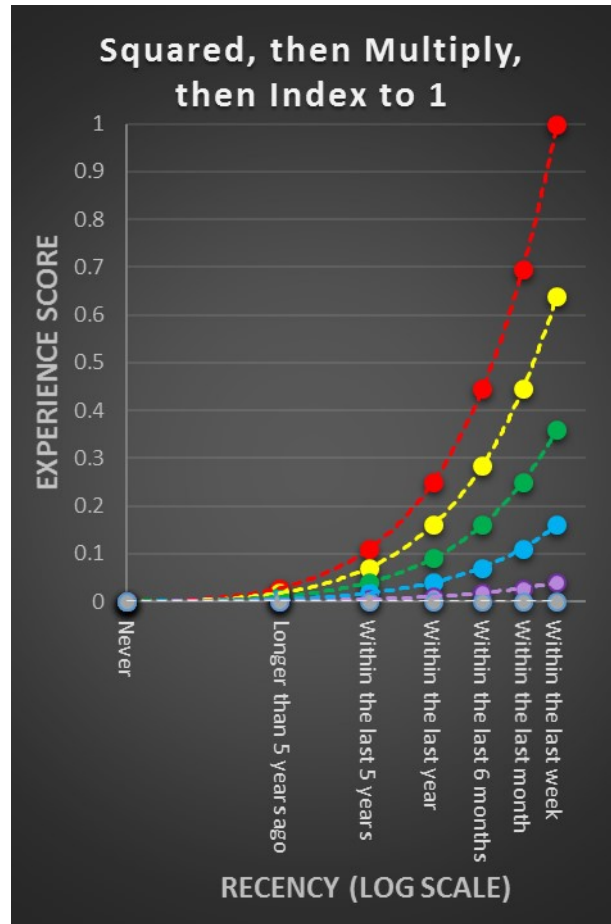
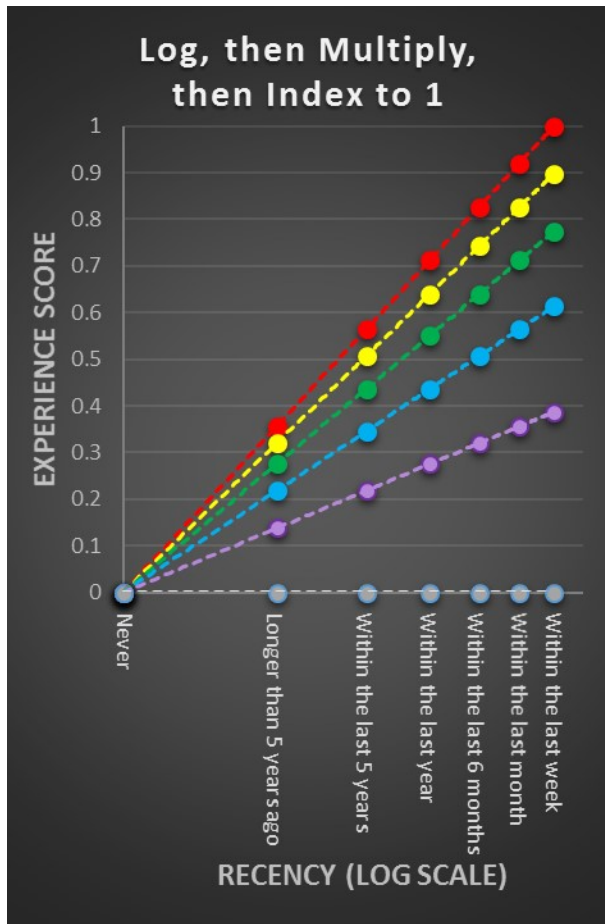
Option #1: $\frac{(\text{Recency} + \text{Frequency})}{\text{Total Possible}}$

Option #2: $\frac{(\text{Recency} \times \text{Frequency})}{\text{Total Possible}}$

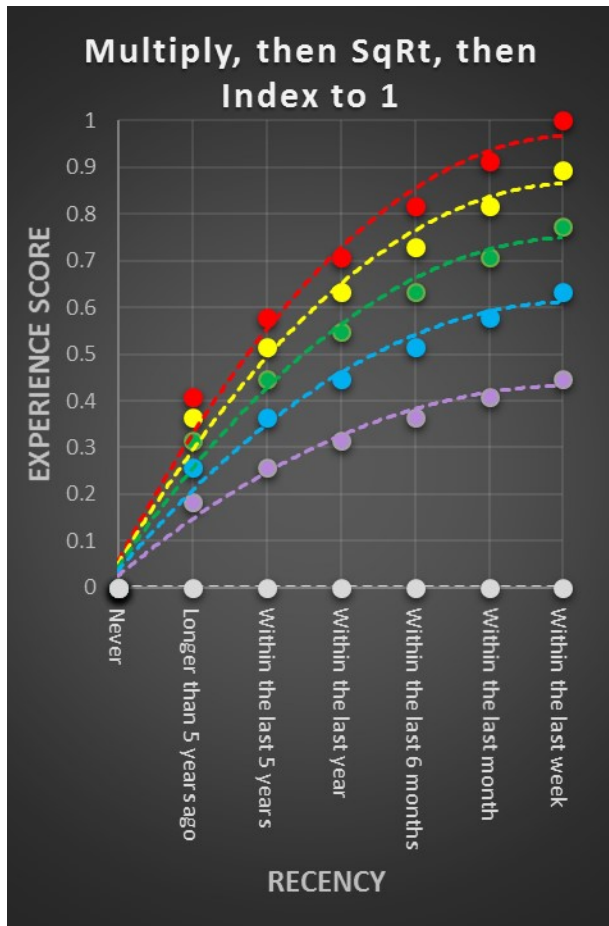


Option #3: $\frac{[\text{Log}(\text{Frequency}) \times \text{Log}(\text{Recency})]}{\text{Total Possible}}$

Option #4: $\frac{(\text{Recency}^2 \times \text{Frequency}^2)}{\text{Total Possible}}$



Option #5: $\frac{\sqrt{(\text{Recency} \times \text{Frequency})}}{\text{Total Possible}}$



APPENDIX I: IRB APPROVAL LETTER



RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-7091

January 21, 2014

Nicholas Stefanic
School of Music
Tampa, FL 33612

RE: Expedited Approval for Initial Review
IRB#: Pro00015900
Title: Creativity as Pedagogy: Modeling the Process and Learning Outcomes in a Massive Open Online Course

Study Approval Period: 1/20/2014 to 1/20/2015

Dear Mr. Stefanic:

On 1/20/2014, the Institutional Review Board (IRB) reviewed and **APPROVED** the above application and all documents outlined below.

Approved Item(s):

Protocol Document(s):

[Study Protocol-Creativity as Pedagogy-Stefanic-ver1 1-10-14.docx](#)

Consent/Assent Document(s)*:

[15900 Online ICF ver1 1-10-14.docx](#) (**granted a waiver)

*Please use only the official IRB stamped informed consent/assent document(s) found under the "Attachments" tab. Please note, these consent/assent document(s) are only valid during the approval period indicated at the top of the form(s). **Waivers are not stamped.

It was the determination of the IRB that your study qualified for expedited review which includes activities that (1) present no more than minimal risk to human subjects, and (2) involve only procedures listed in one or more of the categories outlined below. The IRB may review research through the expedited review procedure authorized by 45CFR46.110 and 21 CFR 56.110. The research proposed in this study is categorized under the following expedited review category:

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Your study qualifies for a waiver of the requirements for the documentation of informed consent as outlined in the federal regulations at 45CFR46.117(c) which states that an IRB may waive the requirement for the investigator to obtain a signed consent form for some or all subjects.

As the principal investigator of this study, it is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB. Any changes to the approved research must be submitted to the IRB for review and approval by an amendment.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

A handwritten signature in cursive script that reads "John A. Schinka, Ph.D.".

John Schinka, Ph.D., Chairperson
USF Institutional Review Board