January 2016

Measuring Students' Motivation: Validity Evidence for the MUSIC Model of Academic Motivation Inventory

Brett D. Jones
Virginia Tech, brettjones@vt.edu

Gary Skaggs
Virginia Tech, gskaggs@vt.edu

Measuring Students' Motivation: Validity Evidence for the MUSIC Model of Academic Motivation Inventory

Abstract

This study provides validity evidence for the MUSIC Model of Academic Motivation Inventory (MUSIC Inventory; Jones, 2012), which measures college students' beliefs related to the five components of the MUSIC Model of Motivation (MUSIC model; Jones, 2009). The MUSIC model is a conceptual framework for five categories of teaching strategies (i.e., eMpowerment, Usefulness, Success, Interest, and Caring) that were derived from research and theory as ones that are critical to students' motivation (Jones, 2009). Participants included 338 undergraduate students who provided questionnaire responses in reference to 221 different courses at a large public U.S. university. Our analyses included classical item analysis, confirmatory factor analysis, the calculation of Rasch measurement scales, and Pearson's correlation coefficients. Results support the validity of scores produced by the MUSIC Inventory for use with college students. This inventory could be useful to instructors and researchers interested in assessing the effects of instruction on students' motivational beliefs.

Cover Page Footnote

Brett D. Jones, Ph.D., is a Professor and leader of the Educational Psychology program in the School of Education at Virginia Tech. His research includes investigating how students' beliefs impact their motivation and examining methods instructors can use to design instructional environments that support students' motivation and learning (http://www.theMUSICmodel.com). Gary Skaggs, Ph.D., is an Associate Professor in the Educational Research and Evaluation program in the School of Education at Virginia Tech. His research focuses on psychometric issues related to large-scale assessments, such as educational achievement and certification/licensure tests. These issues include test equating and scaling, standard setting, and differential item functioning. This paper is a revised version of research presented by Jones and Skaggs at the International Conference on Motivation 2012 in Frankfurt, Germany.

Measuring Students' Motivation:
Validity Evidence for the MUSIC Model of Academic Motivation Inventory

Brett D. Jones[1] and Gary Skaggs[2]

Educational Psychology[1] and Educational Research and Evaluation[2], Virginia Tech, Blacksburg, VA 24061, USA

This study provides validity evidence for the MUSIC Model of Academic Motivation Inventory (MUSIC Inventory; Jones, 2012), which measures college students' beliefs related to the five components of the MUSIC Model of Motivation (MUSIC model; Jones, 2009). The MUSIC model is a conceptual framework for five categories of teaching strategies (i.e., eMpowerment, Usefulness, Success, Interest, and Caring) that were derived from research and theory as ones that are critical to students' motivation (Jones, 2009). Participants included 338 undergraduate students who provided questionnaire responses in reference to 221 different courses at a large public U.S. university. Our analyses included classical item analysis, confirmatory factor analysis, the calculation of Rasch measurement scales, and Pearson's correlation coefficients. Results support the validity of scores produced by the MUSIC Inventory for use with college students. This inventory could be useful to instructors and researchers interested in assessing the effects of instruction on students' motivational beliefs.

INTRODUCTION

Researchers have used various approaches to examine factors that affect students' motivation in learning environments. Early attempts were aimed at understanding students' traits and the ways in which students approached their academic work (Entwistle & Morton, 1989). Researchers then focused more on the influence of the learning environment on students' motivation and learning. These types of studies led Entwistle and Morton (1989) to conclude that, "Students' perceptions of their academic environment have a powerful influence on the quality of their learning" (p. 452). Since then, researchers have continued to explore how students' external environments interact with their internal needs, cognition, and affect (Ormrod & Jones, 2015), and many theories have been developed to explain these interactions (Schunk, Meece, & Pintrich, 2014).

Given the importance of students' perceptions of their learning environments, it is critical for instructors to consider these perceptions when designing or re-designing their courses. To determine students' perceptions of their learning environment in a course, instructors are faced with two primary questions: (1) What perceptions should be assessed? and (2) How should these perceptions be assessed? The research and theories related to student motivation can be difficult for instructors outside of the field of motivation to interpret because there are many theories of motivation from which to choose and the jargon used is often difficult to understand.

To address these issues, the MUSIC[SM] Model of Motivation (Jones, 2009; abbreviated in this article as the "MUSIC model") was designed to provide a conceptual framework to help instructors in any field understand how to apply current motivation research and theories to instruction. The MUSIC model consists of five categories of teaching strategies that Jones (2009) derived from research and theory as ones that are critical to students' engagement in academic settings: eMpowerment, Usefulness, Success, Interest, and Caring (MUSIC is an acronym). Instructors who address these MUSIC components are more likely to be successful at motivating their students to engage in learning (Jones, 2009, 2015). Thus, the MUSIC model provides instructors with some guidance in what student perceptions should be assessed.

The second question noted previously was, how should students' perceptions be assessed? Surveying students with a questionnaire can be an expeditious method to assess their perceptions. To

increase the validity of the scores produced from a questionnaire, responses to several questionnaire items can be averaged to form scales, which can comprise a larger instrument or inventory that can be used to measure one or more psychological constructs (e.g., interest). Instruments that measure motivation-related constructs usually include Likert-type items that are labeled at the endpoints (e.g., *strongly disagree* to *strongly agree*; Eccles & Wigfield, 1995; Vallerand, Pelletier, Blais, Briere, Senecal, & Vallieres, 1992; Williams & Deci, 1996). However, these instruments often measure constructs related to a specific motivation theory, which limits their use for instructors who want to examine a broader spectrum of motivation-related perceptions. Furthermore, because these instruments have often been designed by different researchers at different times for different purposes, the instruments can differ in many ways, including the tone of the items, the types of items (e.g., questions, statements), the number of response options (e.g., 1 to 5, 1 to 7), and the labels provided at the response options (e.g., *strongly agree*, *very interested*). These differences make it difficult to compare and interpret the scores obtained from the instruments. As a result, instructors and researchers who want to assess students' perceptions on all five MUSIC model components with these instruments have to modify the original measures, sometimes fairly significantly (e.g., Jones, 2010a).

The purpose of the present study was to provide validation evidence for the MUSIC[SM] Model of Academic Motivation Inventory (Jones, 2012; abbreviated in this article as the "MUSIC Inventory") that includes five scales to measure the five components of the MUSIC model. A strength of the MUSIC Inventory is that the scores produced from it can be used by instructors to improve their instruction because each MUSIC Inventory scale corresponds to one of the five key MUSIC components that can be used to increase students' motivation.

CONCEPTUAL FRAMEWORK

The MUSIC model is a conceptual model based on many different theories and focuses on how social contexts affect students' perceptions of key motivational beliefs (Jones, 2009, 2015). Five key principles of the model are that instructors need to ensure that students: (1) feel empowered by having the ability to make decisions about some aspects of their learning, (2) understand why what they are learning is useful for their short- or long-term goals,

(3) believe that they can succeed if they put forth the effort required, (4) are interested in the content and instructional activities, and (5) believe that the instructor and others in the learning environment care about their learning and about them as a person. Instructional recommendations are directly related to each of these five principles. Because instructors are typically not experts in the field of motivation, the terminology used for the MUSIC components emphasizes the importance of these five principles and the related instructional implications, rather than providing the names of the motivation jargon (e.g., autonomy, self-concept, utility value) that has specific meanings to motivation scholars. As a result, the MUSIC model principles can relate to more than one motivation construct, especially given that many constructs have similar or overlapping definitions (Schunk, 2000).

In measuring students' perceptions of the MUSIC components in a learning environment, it is necessary to relate each MUSIC component to a specific construct. Table 1 provides definitions for each of the five MUSIC model components assessed in the MUSIC Inventory and lists the constructs to which they most closely relate. In designing the MUSIC Inventory, Jones (2012) targeted one of the related constructs for each MUSIC component (shown in italics in Table 1): autonomy for empowerment, utility value for usefulness, expectancy for success, situational interest for interest, and caring for caring. Definitions of these constructs and the associated references are provided in the remainder of this section.

The empowerment component refers to the amount of control students perceive that they have over their learning environment in the course. Students who are empowered perceive that they can make choices and decisions, as opposed to feeling that they are being controlled by the instructor or other factors. For example, a student who feels empowered might say, "I did what I wanted to do during that activity." This student felt free to make her own decisions, and thus, felt empowered during the activity.

The empowerment component is consistent with the "autonomy" construct as defined by Deci and Ryan (1991) in the following description: "The need for autonomy (or self-determination) encompasses people's strivings to be agentic, to feel like the 'origin' (deCharms, 1968) of their actions, and to have a voice or input in determining their own behavior. It concerns the desire to experience an internal perceived locus of causality with regard to action—that is, to experience one's actions as emanating from the self" (p. 243).

It is important to note that the empowerment component and the autonomy construct are not synonymous with the "autonomy support" construct that may be familiar to some readers. The autonomy support construct includes aspects of the autonomy construct, but it is more broadly defined than autonomy. Definitions of autonomy support typically include most of the following five elements: providing meaningful rationales, acknowledging negative feelings, using non-controlling language, offering choices, and nurturing inner motivational resources (for a review, see Su & Reeve, 2011). For an example of how autonomy support is broader in scope than the autonomy construct, consider that nurturing inner motivational resources can include constructs such as interest, intrinsic motivation, competence, relatedness, sense of challenge, and intrinsic goals (Su & Reeve, 2011).

The usefulness component of the MUSIC model involves the extent to which students perceive that the coursework (e.g., assignments, activities, readings) is useful for their short- or long-term goals. One way in which a course could be perceived as being useful is if it provided knowledge or skills that were relevant to students' careers. The usefulness component is consistent with constructs such as utility value (Eccles et al., 1983; Eccles & Wigfield, 1995; Wigfield & Eccles, 2000) and instrumentality (De Volder & Lens, 1982; Kauffman & Husman, 2004). As explained by Wigfield and Eccles (2000), "Utility value or usefulness refers to how a task fits into an individual's future plans" (p. 72).

The success component of the MUSIC model is based on the idea that students need to perceive that they can succeed if they put forth the appropriate effort. Students who believe that they can be successful, expect to do well in future activities. The success component of the MUSIC model is consistent with constructs such as expectancy for success (Eccles et al., 1983; Eccles & Wigfield, 1995; Wigfield & Eccles, 2000), self-efficacy (Bandura, 1986), and competence motivation (Elliot & Dweck, 2005; White, 1959). For example, expectancy for success (Eccles et al., 1983) has been defined and measured as "children's beliefs about how well they will do on upcoming tasks, either in the immediate or longer term future" (Wigfield & Eccles, 2000, p. 70).

Consistent with current theories of interest (e.g., Hidi & Renninger, 2006), Jones (2009, 2010a) has discussed interest in the MUSIC model as consisting of two theoretically distinct constructs: sit-

MUSIC model components	The degree to which a student perceives that:	Related constructs[1]
eMpowerment	he or she has control of his or her learning environment in the course	•*autonomy*
Usefulness	the coursework is useful to his or her future	•*utility value* •instrumentality
Success	he or she can succeed at the coursework	•*expectancy for success* •self-efficacy •competence
Interest	the instructional methods and coursework are interesting	•*situational interest* •intrinsic motivation •intrinsic interest value •flow
Caring	the instructor cares about whether the student succeeds in the coursework and cares about the student's well-being	•*caring* •belongingness •relatedness •attachment

TABLE 1. Definitions of the MUSIC Model Components and Related Constructs

[1]Items in the MUSIC Inventory were designed to measure the constructs in italics, although the other constructs listed are closely related.

uational interest and individual interest. Situational interest refers to the immediate, short-term interest that arises spontaneously in a particular situation (Schraw, Flowerday, & Lehman, 2001), whereas individual interest refers to internally activated personal values about a topic that are more enduring. Although situational and individual interests can interact (Ainley & Ainley, 2011), students' perceptions of their situational and individual interest can be considered to be somewhat independent of one another. For example, a student could have a high individual interest in a topic and a low situational interest in an activity related to that topic. Or, the student could have a low individual interest in a topic and a high situational interest in an activity related to that topic. Because the intent of the MUSIC Inventory is to assess students' perceptions about a course and the topics within it, it includes a measure of situational interest and not a measure of individual interest. The rationale was that regardless of students' individual interest in a course or course topics, instructors should strive to design a course that triggers students' situational interest.

Students who are situationally interested pay attention to and are emotionally aroused by instructional methods and coursework. The situational interest component of the MUSIC model is consistent with constructs such as situational interest (Hidi & Renninger, 2006), intrinsic motivation (Deci, 1975), intrinsic interest value (Eccles & Wigfield, 1995), and flow (Csikszentmihalyi, 1990). As Hidi and Renninger (2006) explained: "Situational interest refers to focused attention and the affective reaction that is triggered in the moment by environmental stimuli, which may or may not last over time" (p. 113).

The caring component of the MUSIC model involves the extent to which a student believes that others in the learning environment (e.g., the instructor, other students) care about whether he or she succeeds in the coursework and cares about his or her well-being. When students feel cared for by the instructor, they perceive the instructor's caring through actions such as the instructor: listening to students, helping students with their academic needs, and being respectful of students. The caring component can be divided into academic caring (i.e., caring about students' success in the coursework) and personal caring (i.e., caring about students' well-being). However, because academic and personal caring have been shown to form one construct (Jones & Wilkins, 2013) and the educational implications for both are very similar (Jones, Watson, Rakes, & Akalin, 2012), both academic and personal caring are assessed with one scale in the MUSIC Inventory. The caring component is consistent with constructs such as caring (Noddings, 1984, 1992), belongingness (Baumeister & Leary, 1995), relatedness (Deci & Ryan, 2000), and attachment (Ainsworth, 1979; Bowlby, 1969). As Noddings (1992) explains: "A caring relation is in its most basic form, a connection or encounter between two human beings—a carer and a recipient of care, or cared-for" (p. 15). The carer (the one providing the caring) must be in a state of consciousness that is characterized by "open, nonselective receptivity to the cared-for" (p. 15); he must "really hear, see, or feel what the other tries to convey" (p. 16). Additionally, the cared-for must receive the caring and show that it has been received (p. 16).

PURPOSE
We designed this study to provide further validation evidence for the MUSIC Inventory. At the time of this study, the MUSIC Invento-

ry had not been tested across a variety of higher education courses in many different subject areas. Further, no studies had provided evidence for the convergent, divergent, and predictive validity of the MUSIC Inventory by correlating the MUSIC Inventory scales to other measures. The purpose of this study was to address these needs.

The MUSIC Inventory could be useful to instructors and researchers because one inventory that measures all five MUSIC model components does not currently exist. Consequently, instructors and researchers are forced to use instruments that vary in the tone and type of items, in the number of response options provided, and in the labels used for the response options. These differences make it difficult, if not impossible, to interpret the scores from any one instrument in relation to the scores from the instruments measuring the other MUSIC components. Furthermore, it is unwieldy for instructors to piece together various instruments designed by different researchers for different purposes and different populations. It is more practical and useful for instructors to have one inventory that measures all five of the MUSIC components.

METHODS
Participants and Procedure
Participants included 397 undergraduate students who were enrolled in an online course at a large public U.S. university. A little over half (56.4%) of the students were female and most of the students were White (77.3%), while the remaining students reported their race/ethnicity as Asian or Pacific Islander (15.6%), Black or African American (3.0%), Hispanic (2.0%), or other (2.0%). About half of the students reported their academic class standing as senior (49.9%), 22.4% reported it as junior, 20.9% reported it as sophomore, and 6.8% reported it as first year.

Near the end of the course, students completed a 121-item, online questionnaire that included the MUSIC Inventory items and the other motivation-related items described in the "Instruments" section below, as well as items about a course, an instructor, their personality, and their demographic information. Students received a small amount of course credit for completing the questionnaire, but their responses were not graded; rather, students received full credit if they completed the questionnaire items and were asked to answer them honestly. At the end of the questionnaire, they were asked whether or not they would allow their responses to be included in the present study.

At the beginning of the questionnaire, students provided their name and email address. Next, they were asked to do the following: "think of a course that you are enrolled in this semester OTHER THAN [THE CURRENT COURSE] and provide the information below. If you are not enrolled in any other courses, you may use [the current course]; otherwise, please think of another course besides [the current course]." The next three questions asked students to list the course name abbreviation, course number, and course title. Students were notified that they would need to refer to that course when answering the questions in the following sections. The final introduction question asked students to choose whether the course was primarily a face-to-face course, an online course, or a hybrid course that takes place face-to-face and online.

Instruments
In this section, we describe all of the instruments that we used in

TABLE 2. Example Items from the Instruments		
MUSIC model components	MUSIC Inventory example items	Comparison scale example items
eMpowerment/autonomy	I had the freedom to complete the coursework my own way.	My instructor listened to how I would like to do things. (Learning Climate Questionnaire)
Usefulness/ utility value	In general, the coursework was useful to me.	What I learned in this course was relevant to my life. (Utility Value scale)
Success/ expectancy for success	I was confident that I could succeed in the coursework.	During this course, I felt confident in my ability to learn the material. (Perceived Competence scale)
Interest/ situational interest	The coursework was interesting to me.	In general, how interested were you in learning the content material in this course? (Interest scale)
Caring	The instructor cared about how well I did in the course.	I believe that my instructor cared about how much I learned. (Classroom Life Instrument)

the study. Table 2 shows example items from the MUSIC Inventory and other instruments intended to measure similar constructs (hereafter referred to as the "comparison scales").

MUSIC Inventory. The MUSIC Inventory consists of 26 items that are organized into five scales that include five empowerment items, five usefulness items, four success items, six interest items, and six caring items (the complete inventory is provided in Jones, 2012). All items are rated on a 6-point Likert-type format scale ranging from *strongly disagree* to *strongly agree* with descriptors at each point.

Learning Climate Questionnaire (LCQ). We used five items modified from the 6-item, short version of the LCQ (Williams & Deci, 1996) that measures the degree to which students perceive the instructor as supporting their autonomy. Items were rated on a 7-point Likert-type format scale ranging from *strongly disagree* to *strongly agree*. Similar to Jones (2010a), we did not use one of the six items, shortened another item, and changed the wording in three of the items to make it past tense. In prior studies, the Cronbach's alpha values for this modified version have been good (in Jones, 2010a, α = .81, .85, .92, .91; in Jones et al., 2012, α = .93).

Utility Value Scale. We used items modified from the three-item utility value scale used by Hulleman, Durik, Schweigert, and Harackiewicz (2008) which measures the extent to which students perceived the course to be useful to their life and future. Items were rated on a 7-point Likert-type format scale ranging from *strongly disagree* to *strongly agree*. Similar to Jones (2010a), we changed the wording of the three items from "class" to "course" and changed the word "career" in one item to "my future" to make the question more general than the word "career." The Cronbach's alpha values for this modified version of the utility value scale have been good (in Jones, 2010a, α = .88, .86, .94, .93; in Jones et al., 2012, α = .95).

Perceived Competence Scale. We used a modified 4-item Perceived Competence Scale, which assesses students' feelings of competence in a course (Williams & Deci, 1996; Williams, Freedman, & Deci, 1998). Items were rated on a 7-point Likert-type format scale ranging from *very untrue* to *very true*. Similar to Jones (2010a), we changed the wording for one item slightly and this scale has led to good Cronbach alpha values (in Jones, 2010a, α = .88, .84, .92, .86; in Jones et al., 2012, α = .93).

Interest Scale. To measure situational interest, we used three items that are similar to those used by others to measure "interest" and "intrinsic interest value" (e.g., Eccles & Wigfield, 1995; Simpkins,

Davis-Kean, & Eccles, 2006). The scale assesses the degree to which students enjoy and are interested in the course. Items were rated on a 7-point Likert-type format scale ranging from *very uninterested/ strongly disliked* to *very interested/enjoyed a lot*. Researchers who have used this scale have reported good Cronbach alpha values (in Jones, 2010a, α = .88, .87, .90, .85; in Jones et al., 2012, α = .90).

Classroom Life Instrument (CLI). We used two, 4-item scales of the CLI (Johnson, Johnson, & Anderson, 1983) titled Teacher Academic Support and Teacher Personal Support. These scales measure students' perceptions of teacher support in academic and personal dimensions, respectively. Students provided their ratings on a 7-point Likert-type format scale ranging from *never to always*. We modified these scales similar to Jones (2010a) by changing the items from present to past tense and making a few wording changes to three of the items. The Cronbach alpha values have been found to be good with these modifications (for academic support, α = .84, .89, .94, .93 in Jones, 2010a, and α = .93 in Jones et al., 2012; for personal support, α = .83, .85, .92, .86 in Jones, 2010a, and α = .92 in Jones et al., 2012).

Effort/Importance Scale. We used a 4-item measure of effort used by Jones (2010a) that was based on the 5-item Effort/ Importance scale that is part of the Intrinsic Motivation Inventory (Plant & Ryan, 1985) and that assesses the amount of perceived effort that students put forth in a course. An example item, rated on a 7-point Likert-type format scale ranging from *very untrue* to *very true*, is: "I put a lot of effort into this course." Similar to Jones (2010a), we changed the items to focus on "this course" and we did not use one of the original five items because students in that study found it confusing and unclear (in Jones, 2010a, α = .84, .84, .86, .84).

Instructor and Course Ratings. To measure students' overall perceptions about the instructor and course ratings, we used one item for the instructor rating and another for the course rating. Both items had a 7-point Likert-type format scale with a descriptor at each point ranging from *terrible* to *excellent*. The exact wording of the items was: "My overall rating of the instructor for the course I listed is:" and "My overall rating of the course I listed is:" We selected these items because they are similar to the one-item scales completed by students on their official university course evaluation forms at the end of a course and they were also used by Jones (2010a).

RESULTS
First Validation Phase
The purpose of the first validation phase of this study was to assess

the statistical qualities of the MUSIC Inventory. Of the 397 students who responded, 68 indicated that they had completed the items in reference to the same course in which they were asked to complete the questionnaire. Because this subgroup scored significantly higher than the remaining sample, we included a random sample of only nine of these students. We chose to include nine students from this course because the highest frequency of responses from the other courses in the sample was 10 and we did not want to over- or under-sample the students from this course. Additionally, we were concerned that a systematic bias of responses could occur for students who referenced an online-only course. Because the responses for online courses were almost identical to those for face-to-face or hybrid courses, we did not find a systematic bias for online versus face-to-face courses. In sum, we conducted the study with 338 students who provided responses in reference to 221 different courses, of which 82.8% were reported to be primarily face-to-face courses, 10.1% were primarily online courses, and 7.1% were hybrid courses that took place face-to-face and online.

Classical item analysis showed that the items in the MUSIC Inventory maintained their statistical qualities and coefficient alphas were at least 0.90 for each scale. Next, we conducted a confirmatory factor analysis (CFA) of the five-factor model using MPlus statistical software. The purpose of the CFA was to determine whether the data fit our hypothesized model in which each of the five MUSIC Inventory scales comprised separate factors. We found that each item loaded on the assigned factor (e.g., the empowerment items loaded on the empowerment factor) and each loading was statistically significant with none of the items cross loading on any of the other factors (e.g., the empowerment items did not load on the usefulness, success, interest, or caring factors). The five-factor model fit the data well, as was evidenced by three fit indices: the Comparative Fit Index (CFI), the Standardized Root Mean Square Residual (SRMR), and the Root Mean Square Error of Approximation (RMSEA). The CFI index was 0.920, which represented a reasonable fit because values can range from 0 to 1 (Hu & Bentler, 1999). The SRMR also varies between 0 and 1, but values closer to 0 indicate better fit. SRMR was 0.055, which indicated a reasonable fit (Kline, 2005). The RMSEA, which also varies between 0 and 1 with values closer to 0 indicating better fit, was 0.085 and represented a borderline reasonable fit (Browne & Cudeck, 1993; Byrne, 2001; Kline, 2005). All of this evidence is taken into consideration as a whole because there is no one particular statistical test to determine whether the data fits a model or not.

To calculate scores for respondents, we calculated Rasch measurement scales for each of the five MUSIC scales. Rasch scaling is preferable to simply summing the item scores because it accounts for differences in overall scoring between items. That is, summed scores are probably ordinal at best, whereas Rasch scales are closer to equal-interval measurements. We used WINSTEPS (Linacre, 2006) to calibrate item parameters. Because all items had the same response choices, we investigated the application of a rating scale model (Andrich, 1978). For two scales, the rating scale model fit better than the partial credit model (Masters, 1982). For two other scales, the partial credit model fit better, and for the remaining scale, there was no difference in fit between the two models. As a result, for simplicity, we used the partial credit model for all scales. Individually, all items but one fit the model. For that one item, there was no obvious reason for the misfit and because this item otherwise displayed good response variance and discrimination, we included it.

The WINSTEPS program calculates a Rasch person measure, or scale score, for each possible summed score for each of the five scales. These Rasch measures typically span a scale of approximately -7.0 to +7.0. For reporting purposes, we applied a linear transformation to each scale Rasch measure to produce a scaling ranging from 0 to 100. The end result of these analyses was a table converting each summed scale score to its corresponding Rasch measure.

Descriptive statistics and coefficient alpha reliability estimates for each scale are shown in Table 3. For this sample, scores were relatively higher for the Usefulness, Success, and Caring scales and relatively lower for the Empowerment and Interest scales.

Second Validation Phase

The purpose of the second phase was to examine the construct and predictive validity of the MUSIC Inventory by comparing the scores obtained from the MUSIC Inventory to those from other instruments that measure (a) constructs similar to the MUSIC components and (b) constructs that the MUSIC components have been shown to predict. We compared the MUSIC Inventory scales to other instruments intended to measure similar constructs (the "comparison scales").

We calculated Pearson's correlation coefficients to examine the relationships between each of the MUSIC Inventory scales. Based on the results of prior studies that included similar constructs, we expected that the correlations between the MUSIC Inventory scales would be moderate because the MUSIC components have been found to be distinct, yet correlated constructs (Jones & Wilkins, 2013). As shown in Table 4, the correlation coefficients for the MUSIC Inventory scales ranged from .46 to .77 which indicated that the scales were correlated as expected. We calculated similar correlations for the comparison scales and the coefficients ranged from .45 to .82, as shown in parentheses in Table 4.

We also calculated Pearson's correlation coefficients between the five MUSIC Inventory Rasch scales and the five comparison scales. In addition, we correlated these 10 scales with students' reported effort, students' overall instructor rating, and students' overall course rating. The results in Table 5 show that the MUSIC Inventory scales are highly correlated with the comparison scales that measured similar constructs. That is, the bold values on the diagonal in Table 5 ranged from .82 to .88, with the exception of the correlation between the Empowerment scale and LCQ, which was .57. These strong correlations provide evidence for the convergent validity of the MUSIC Inventory scales (Campbell & Fiske, 1959).

The fact that the MUSIC Inventory scales were not as highly correlated with the scales that measured different constructs provides evidence for the discriminant validity of the MUSIC Inventory scales (see Table 5; Campbell & Fiske, 1959). That is, the values in bold on the

TABLE 3. Scale Mean Scores, Standard Deviations, and Reliability Coefficients

Scale	Raw item scores[a] M (SD)	Rasch scale scores[b] M (SD)	a
eMpowerment	4.26 (1.1)	61.7 (17.7)	0.91
Usefulness	4.78 (1.1)	70.8 (22.4)	0.96
Success	4.98 (0.9)	74.3 (20.7)	0.93
Interest	4.40 (1.2)	59.9 (21.0)	0.95
Caring	5.15 (0.9)	72.6 (20.9)	0.93

[a]Scale ranges from 1 to 6; [b]Scale ranges from 0 to 100

TABLE 4. Pearson's Correlation Coefficients for the MUSIC Inventory Scales and the Comparison Scales				
MUSIC Inventory scales (Comparison scales)	1	2	3	4
1. Empowerment (LCQ)				
2. Usefulness (Utility Value Scale)	.53 (.52)			
3. Success (Perceived Competence Scale)	.53 (.58)	.57 (.48)		
4. Interest (Interest Scale)	.67 (.60)	.77 (.76)	.65 (.65)	
5. Caring (CLI)	.46 (.82)	.54 (.45)	.62 (.49)	.60 (.57)

Note. Information for the comparison scales is shown in parentheses; p ≤ .001 for all coefficients.

diagonal of Table 5 are higher than the other values in the same row and column as that value. The only exception is that the correlation between empowerment and LCQ (r = .57) is lower than the values between LCQ and Interest (r = .63) and Caring (r = .78), an issue that we discuss further in the "Discussion and Conclusion" section.

To examine the predictive validity of the MUSIC Inventory scales, we calculated the correlation coefficients between the MUSIC Inventory scales and three outcomes that have been used in prior studies with the comparison scales: effort, instructor rating, and course rating. The results shown in Table 5 indicate that the MUSIC Inventory scales and the comparison scales are statistically correlated (p ≤ .001) with these three outcome variables.

DISCUSSION AND CONCLUSION
Validity Evidence
The purpose of the present study was to provide validation evidence for the MUSIC Inventory. The inventory was designed to be used as a measure of students' motivation in a college-level course and includes five scales that measure the five components of the MUSIC model (Jones, 2009). The results indicate that the scores produced by the MUSIC Inventory are valid for use with undergraduate students. Our sample included students' perceptions related to 221 different online, classroom based, and hybrid courses, which enhances the generalizability of the results.

Each MUSIC Inventory scale was moderately correlated with the other four MUSIC Inventory scales. This finding is consistent with the findings of Jones and Wilkins (2013) who documented through CFA that the MUSIC model components were well represented as a correlated five-factor model in which the components were correlated,

yet distinct factors. The fact that the scales used in the present study were correlated with existing instruments designed to measure similar constructs provided evidence for the convergent validity of the MUSIC Inventory scales. Discriminant validity evidence was provided by the fact that the MUSIC Inventory scales were not as highly correlated with the instruments that measured different constructs.

One correlation that did not fit the pattern of the others was that the empowerment scale did not correlate as highly with LCQ (r = .57) as LCQ correlated with Interest (r = .63) and Caring (r = .78). We believe that the explanation for the lower correlation between the empowerment scale and the LCQ is that the LCQ correlated highly with the CLI (r = .82). An examination of the LCQ items reveals that some of them overlap significantly with the caring component, such as: "My instructor tried to understand how I saw things" and "My instructor listened to how I would like to do things." This finding is not surprising given the theoretical difference between autonomy support and empowerment that we described previously. In short, empowerment refers specifically to meeting students' need for autonomy, whereas autonomy support also includes other factors, such as nurturing inner motivational resources (e.g., meeting students' needs for competence and relatedness; Su & Reeve, 2011). Thus, the empowerment scale provides a score that is a more "pure" representation of the autonomy construct (as defined by Deci & Ryan, 1991) than the autonomy support construct as measured by the LCQ. This finding does not diminish the validity of the LCQ scores; rather, it demonstrates that the empowerment scale in the MUSIC Inventory is a more valid representation of empowerment (as defined by Jones, 2009) and autonomy (as defined by Deci & Ryan, 1991) than the LCQ, which measures a broader spectrum of autonomy supportive factors.

The MUSIC Inventory scales and the comparison scales are correlated with effort, instructor rating, and course rating. Other researchers have found similar correlations between these constructs and outcomes (e.g., Filak & Sheldon, 2003), which provides evidence for the predictive validity of the MUSIC Inventory.

Implications
The MUSIC Inventory could be useful to instructors and researchers interested in assessing the effects of instruction on students' motivation-related beliefs. There are several advantages to using the MUSIC Inventory as compared to selecting existing scales to measure the five MUSIC model components. First, all five MUSIC Inventory scales are scored on the same 6-point Likert-type scale. Having the same number of scale points is more desirable than selecting scales with different numbers of points because: (a) students will likely be less

TABLE 5. Pearson's Correlation Coefficients for the Study Constructs								
	M	U	S	I	C	Effort	IR	CR
1. LCQ	**.57**	.53	.57	.63	.78	.32	.61	.54
2. Utility Value Scale	.47	**.88**	.46	.69	.45	.40	.38	.52
3. Perceived Competence Scale	.46	.56	**.84**	.63	.56	.24	.45	.55
4. Interest Scale	.54	.77	.65	**.82**	.61	.38	.60	.71
5. CLI	.43	.48	.52	.55	**.82**	.29	.64	.54
6. Effort Scale	.17	.43	.21	.38	.31	--	--	--
7. Instructor Rating (IR)	.41	.46	.47	.61	.69	.23	--	--
8. Course Rating (CR)	.55	.59	.60	.73	.60	.27	--	--

Note. The following abbreviations are used: M = eMpowerment, U = Usefulness, S = Success, I = Interest, C = Caring, Effort = Effort Scale, IR = Instructor Rating, and CR = Course Rating; p ≤ .001 for all coefficients; values on the diagonal are bolded for easier identification.

confused by the scaling, and thus, provide more reliable responses; (b) the items can be administered to students in random order without requiring students to respond on different rating scales on adjacent items; and (c) having a different number of points for some items could easily cue students that there are different scales and possibly lead to them discovering the pattern in items that relate to the same scale. Second, all of the MUSIC Inventory items have consistent labels at each of the six response options (i.e., *strongly disagree, disagree, somewhat disagree, somewhat agree, agree, strongly agree*). Labels allow students to better understand what each number represents. As Dillman, Smyth, and Christian (2009) explained: "fully labeled scales rate higher on reliability, validity, and respondent preference, and they are less susceptible to context effects (Krosnick & Fabringer, 1997)" (p. 143). Third, all five MUSIC Inventory scales have at least four items, which is important because longer instruments generally provide more reliable scores than shorter instruments with fewer than four items (Lord & Novick, 1968). Fourth, none of the five MUSIC Inventory scales have more than six items, which is important because it results in a shorter overall inventory; and therefore, it can be administered to students in a shorter amount of time than a longer inventory. Fifth, the empowerment scale is a more theoretically accurate measure of the autonomy construct as conceptualized by Jones (2009) in comparison to the LCQ that measures autonomy support more broadly. Sixth, the MUSIC Inventory can provide instructors with diagnostic information about a course very quickly because it can be administered rapidly. Seventh, the MUSIC Inventory provides a measurement option for instructors interested in assessing a broad variety of motivation perceptions not limited to one particular motivation theory. Eighth, because the MUSIC model was developed to summarize the major ideas in current academic motivation theory and research, it can teach instructors (or remind them if they already knew) about the variety of motivational concepts that are important to consider when designing instruction. Otherwise, in our experience with instructors, we have found that they generally consider motivation very narrowly or naively and might only measure one or two constructs such as self-efficacy or interest without considering empowerment, usefulness, and caring. In this way, the MUSIC Inventory could serve as a teaching tool for instructors. Ninth, and maybe most importantly, the scores obtained from the MUSIC Inventory can be linked directly to categories of instructional strategies that can be used to increase students' motivation (e.g., low empowerment scores indicate that the instructor could consider strategies related to increasing students' empowerment).

Taken together, these advantages of the MUSIC Inventory make it a practical, useful tool that can be used quickly and easily by instructors. These advantages are especially important for instructors who are unfamiliar with the variety of motivation constructs in the scholarly literature and do not want to search the many possible types of existing motivation measures. In this way, the MUSIC Inventory provides an easy "one-stop-shopping" experience for instructors who want to use a measure to assess students' motivational beliefs about their course. Open access versions of the MUSIC Inventory, instructor support materials, and scoring information are available (Jones 2012).

Instructors can use the MUSIC Inventory at any time in a course. Typically, the inventory is used in the middle of a course as a formative assessment or near the end of a course as a summative assessment (see Jones, 2012, for how to change the verbs in the items to past or present tense). However, the inventory can be administered as early as the first day of the course as long as students have enough information about the course to have formed perceptions about the course.

There is no one "correct" way to use the inventory scores. However, it is common to use the scores as a diagnostic tool to assess relative strengths and weaknesses of the instruction. For example, an instructor who receives lower scores on usefulness and interest than on empowerment, success, and caring, could reflect on the instruction to determine whether there was anything that could be done to increase the usefulness and interest scores in ways that met the objectives of the course. For MUSIC components that are rated lower, instructors can also collect more feedback from students using open-ended questionnaire items related to those components (see Jones, 2012, for possible items) or talk to students individually before or after class. In addition, instructors can conduct a class discussion to gather more feedback about students' perceptions of the MUSIC components.

Limitations and Future Research

For a more complete examination of the validity of the MUSIC Inventory than we were able to obtain from a sample at one university, these analyses could be replicated with other students in other settings. One future validity study could investigate differences between major subpopulation groups on the MUSIC Inventory scales. These analyses could include comparing overall group differences, but they could also examine differential item functioning and compare subgroups using factor analysis. In addition to gender and race/ethnicity, differences could be investigated among types of undergraduate majors and types of course structures.

The MUSIC Inventory presented in this study queried students about their perceptions of a particular course; however, it might be possible to obtain valid scores from the inventory by asking students to rate their perceptions about a particular activity by changing the word "course" and "coursework" in all of the items to "activity." Doing so would allow instructors to assess students' perceptions about particular activities within a course. Finally, another line of research could examine the extent to which instructors can be trained to improve their students' motivation, as evidenced by increased MUSIC Inventory scale scores.

REFERENCES

Ainley, M., & Ainley, J. (2011). Student engagement with science in early adolescence: The contribution of enjoyment to students' continuing interest in learning about science. *Contemporary Educational Psychology, 36,* 4-12.

Ainsworth, M. D. S. (1979). Infant-mother attachment. *The American Psychologist, 34*(10), 932-937.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561-573.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Baumeister, R., & Leary, M. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497-529.

Bowlby, J. (1969). *Attachment (vol. I).* New York: Basic.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 445–455). Newbury Park, CA: Sage.

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

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper & Row.

deCharms, R. (1968). *Personal causation: The internal affective determinants of behavior.* New York: Academic Press.

Deci, E. L. (1975). *Intrinsic motivation.* New York: Plenum Press.

Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality. In R. Dientsbier (Ed.), *Nebraska symposium on motivation (Vol. 38).* Lincoln: University of Nebraska Press.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268.

De Volder, M., & Lens, W. (1982). Academic achievement and future time perspective as a cognitive-motivational concept. *Journal of Personality and Social Psychology, 42*(3), 566–571.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method.* Hoboken, NJ: Wiley.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: Freeman.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*(3), 215-225.

Elliot, A. J., & Dweck, C. S. (2005). Competence and motivation: Competence as the core of achievement motivation. In A. J. Elliot, & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 3-12). New York: Guilford.

Entwistle, N., & Marton, F. (1989). Introduction. The psychology of student learning. *European Journal of Psychology of Education, 4*(4), 449-452.

Filak, V. F., & Sheldon, K. M. (2003). Student psychological need satisfaction and college teacher-course evaluations. *Educational Psychology, 23*(3), 235-247.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Hulleman, C. S., Durik, A. M., Schweigert, S. A., & Harackiewicz, J. M. (2008). Task values, achievement goals, and interest: An integrative analysis. *Journal of Educational Psychology, 100*(2), 398-416.

Johnson, D. W., Johnson, R. & Anderson, A. (1983). Social interdependence and classroom climate. *Journal of Psychology, 114*(1), 135-142.

Jones, B. D. (2009). Motivating students to engage in learning: The MUSIC Model of Academic Motivation. *International Journal of Teaching and Learning in Higher Education, 21*(2), 272-285.

Jones, B. D. (2010a). An examination of motivation model components in face-to-face and online instruction. *Electronic Journal of Research in Educational Psychology, 8*(3), 915-944.

Jones, B. D. (2012). *User guide for assessing the components of the MUSIC Model of Academic Motivation.* Retrieved from http://www.theMUSICmodel.com

Jones, B. D. (2015). *Motivating students by design: Practical strategies for professors.* Charleston, SC: CreateSpace.

Jones, B. D., Watson, J. M., Rakes, L., & Akalin, S. (2012). Factors that impact students' motivation in an online course: Using the MUSIC Model of Academic Motivation. *Journal of Teaching and Learning with Technology, 1*(1), 42-58

Jones, B. D., & Wilkins, J. L. M. (2013). Testing the MUSIC Model of Academic Motivation through confirmatory factor analysis. *Educational Psychology: An International Journal of Experimental Educational Psychology, 33*(4), 482-503. doi:10.1080/01443410.2013.785044

Kauffman, D. F., & Husman, J. (2004). Effects of time perspective on student motivation: Introduction to a special issue. *Educational Psychology Review, 16*(1), 1-7.

Kline, R. B. (2005). *Principles and practice of structural equation modeling.* New York: The Guilford Press.

Krosnick, J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. In L. Lyberg, P. Biemer, M. Collins, L. Decker, E. de Leeuw, C. Dippo, et al. (Eds.), *Survey measurement and process quality.* New York: Wiley-Interscience.

Linacre, J. M. (2006). *WINSTEPS: Rasch model computer program.* Chicago: winsteps.com.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47,* 149-174.

Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education.* Berkeley: University of California Press.

Noddings, N. (1992). *The challenge to care in schools: An alternative approach to education.* New York: Teachers College Press.

Ormrod, J. E., & Jones, B. D. (2015). *Essentials of educational psychology: Big ideas to guide effective teaching* (4th ed.). Boston: Pearson.

Plant, R. W., & Ryan, R. M. (1985). Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: An investigation of internally-controlling styles. *Journal of Personality, 53*(3), 435-449.

Schraw, G., Flowerday, T., & Lehman, S. (2001). Increasing situational interest in the classroom. *Educational Psychology Review, 13*(3), 211-224.

Schunk, D. H. (2000). Coming to terms with motivation constructs. *Contemporary Educational Psychology, 25,* 116-119.

Schunk, D. H., Meece, J. L., & Pintrich, P. R. (2014). *Motivation in education: Theory, research, and applications.* Upper Saddle River, NJ: Pearson.

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*(1), 70-83.

Su, Y., & Reeve, J. (2011). A meta-analysis of the effectiveness of intervention programs designed to support autonomy. *Educational Psychology Review, 23,* 159-188.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and a motivation in education. *Educational and Psychological Measurement, 52*(4), 1003-1017.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*(5), 297-333.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25,* 68-81.

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory.

Journal of Personality and Social Psychology, 70(4), 767-779.

Williams, G. C., Freedman, Z. R., & Deci, E. L. (1998). Supporting autonomy to motivate glucose control in patients with diabetes. *Diabetes Care, 21*(10), 1644-1651.