UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MEASURING ONLINE STUDENT ENGAGEMENT IN HIGHER EDUCATION: SCALE DEVELOPMENT, VALIDATION, AND PSYCHOMETRIC PROPERTIES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

MURAT TURK
Norman, Oklahoma
2022

MEASURING ONLINE STUDENT ENGAGEMENT IN HIGHER EDUCATION: SCALE DEVELOPMENT, VALIDATION, AND PSYCHOMETRIC PROPERTIES

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Xun Ge, Chair

Dr. Amy C. Bradshaw

Dr. Benjamin C. Heddy

Dr. Doo Hun Lim

Dr. H. Michael Crowson

Dedication

I dedicate this dissertation to

my beloved daughter *Begum Turk*, who has always supported me emotionally and psychologically during this difficult and challenging journey with her unconditional love and presence. This dissertation would not have been possible without you, my *Poncikita*.

Acknowledgements

First of all, I would like to thank all my committee members for their support and guidance throughout my program and this dissertation study. I would like to thank my academic advisor *Dr. Xun Ge* for her invaluable feedback, constant guidance, and support during this dissertation study; *Dr. H. Michael Crowson* for his statistical guidance and mentoring throughout my doctoral program and during this dissertation study; *Dr. Benjamin C. Heddy* for his continuous support and encouragement during my most difficult times during the program; *Dr. Amy C. Bradshaw* for her support and guidance during my program and throughout this dissertation study; and *Dr. Doo Hun Lim* for his inspiring feedback and guidance during this dissertation study.

I would also like to thank *Dr. Maeghan N. Hennessey* for her guidance and invaluable feedback during the instrument development process. Similarly, I would like to thank all the experts who reviewed the survey items and provided their feedback during the instrument development process in this dissertation study.

I would also like to thank the Turkish Fulbright Commission and Fulbright Program for giving me this wonderful opportunity to get my doctoral degree in the United States.

I would like to extend my special thanks to my dear father *Muzaffer Turk* and my dear mother *Sema Turk* for supporting me with their unconditional love and care whenever I needed it during the turbulent times.

Finally, I would like to express my gratitude to my dear wife and colleague *Dr. Sinem Toraman Turk* for her unconditional and constant support, care, and understanding during this challenging journey in my life.

Table of Contents

List of Tables

List of Figures

Abstract

Student engagement in online learning environments is of particular importance to successful learning experiences due to the unique features of online learning environments (e.g., physical distance, sense of isolation from peers and instructors). Reliable and valid assessment of student engagement is vitally important for making evidence-based decisions for online learning environments. Addressing this need and the limitations of the existing measures, this study presents a tool to assess student engagement in online learning environments, the Online Engagement in Higher Education (OEHE), which has been validated through a series of confirmatory factor analyses, as well as additional validity analyses, and internal consistency reliability analysis. Using data from 235 undergraduate and graduate students, who took at least one online course during the time of data collection in Fall 2021, a hypothesized three-factor model of online student engagement based on Fredricks et al.'s (2004) engagement framework focusing on three core dimensions of behavioral, emotional, and cognitive engagement was adequately confirmed in the context of CFA. The OEHE was also shown to have reasonable evidence of convergent validity and criterion validity, and strong evidence of discriminant validity through Pearson correlations. The OEHE subscales and the final validated instrument with 20 items also were found to have adequate or very high internal consistency reliability. Implications for the OEHE instrument development and validation and recommendations for future studies are discussed to provide insights for online engagement researchers, practitioners, and other stakeholders in online education.

Keywords: behavioral engagement, cognitive engagement, emotional engagement, instrument development, online student engagement.

Chapter 1: Introduction

Background of the Study

Online courses and online learning have steadily become more popular and widespread in higher education for the last couple of decades (Allen et al., 2016; Ferrer et al., 2020; Hsu et al., 2019; Martin et al., 2019; Martin, Stamper, & Flowers, 2020). The number of college students enrolled in higher education institutions who took at least one online distance course in Fall 2015 was over six million, up from just 1,602,970 in Fall 2002 (Allen & Seaman, 2013, 2017). The terms *online education* and *online learning* are used interchangeably today. Online education can be simply defined as any kind or form of learning and instruction taking place with the use of the Internet and web-based technologies (Picciano, 2019). Online courses delivered to students at a distance through technology-mediated web-based learning environments can provide a wide range of advantages and benefits for students such as convenience, accessibility, and flexibility (Borup et al., 2011, 2012; Caskurlu et al., 2021; Picciano, 2019).

Despite such affordances of online education, however, challenges remain. Lack of face-to-face and real-time interactions with peers and instructors in online learning environments leads to feelings of isolation and disconnectedness over time (Dixson, 2015; Hoi & Le Hang, 2021; Sherblom, 2010). This aspect and other inherent challenges of online education may negatively influence students' online learning experiences including student engagement (Kucuk & Richardson, 2019). Although the number of students enrolling in online programs or taking online courses as part of their higher education is steadily increasing (Henrie et al., 2015; Hoi & Le Hang, 2021) and more and more higher education institutions have been offering online courses (Caskurlu et al., 2021), lack of student engagement remains an important concern in online education (Bolliger & Martin, 2018; Martin et al., 2021). Accurately identifying students' level of engagement in online learning environments is thus an important issue for consideration by all stakeholders in online education to ensure students are optimally involved and engaged in their online learning experiences (Henrie et al., 2015; Hoi & Le Hang, 2021; Martin et al., 2021). Accurate assessment of online student engagement requires a clear and consistent understanding of what student engagement is and how it manifests in online learning environments (Ferrer et al., 2020; Hoi & Le Hang, 2021). Based on Fredricks et al.'s (2004) three-dimensional conceptualization of engagement, student engagement is defined in this study as the extent to which students *behaviorally*, *emotionally*, and *cognitively* engage with their online learning process.

Student engagement is one of the most significant conditions that need to be fulfilled and maintained adequately to achieve desirable learning outcomes in online learning environments (Dixson, 2015; Kucuk & Richardson, 2019; Meyer, 2014; Oncu & Cakir, 2011; Revere & Kovach, 2011). Similar to student engagement in traditional face-to-face classrooms of K-12 and higher education (Fredricks et al., 2004; Kuh, 2016; Redmond et al., 2018), student engagement in online learning environments has been associated with several educational outcomes, such as student satisfaction, academic performance, and persistence (Bolliger & Martin, 2018; Meyer, 2014), and student retention (Vayre & Vonthron, 2017).

Higher attrition rates in online programs compared to their traditional face-to-face counterparts in higher education connect online student engagement with the significant issue of student retention (Boston & Ice, 2011; Hoi & Le Hang, 2021; Meyer, 2014). With fewer opportunities to interact with the course instructor and peers and lack of physical and social contact and cues, student engagement is considered much more critical to student retention and student success in online courses compared to face-to-face courses (Bigatel & Williams, 2015;

Meyer, 2014). This is particularly so for non-traditional or part-time students with various life commitments (e.g., parenting, full-time employment) who prefer online education due to its accessibility and flexibility (Meyer, 2014; Redmond et al., 2018; Stone et al., 2019). If those non-traditional online students can remain actively engaged and enrolled in their online courses and programs, they can successfully complete their higher education and get their college or more advanced degrees through online education (Bigatel & Williams, 2015; Meyer, 2014; Stone et al, 2019).

In addition to non-traditional students, student engagement may be particularly important, if not vital, to underrepresented or marginalized student populations (e.g., underrepresented ethnic minorities) who may confront more serious challenges and obstacles (e.g., systemic barriers, microaggression, racial bias) to remaining actively engaged and enrolled in their online courses and completing their higher education programs successfully (Salvo et al., 2019). For example, African American male students are particularly considered to be among at-risk students in terms of attrition in online education, with lack of technology skills and technical support and accessibility problems in online courses cited as some major reasons for this (Salvo et al., 2019). In light of Chen et al.'s (2010) research indicating marginalized or underrepresented student populations are more likely to prefer online courses, keeping students engaged in their online courses and enrolled in their programs is particularly important so that they can successfully complete online courses and programs in higher education.

Student Engagement and Quality of Online Education

With the ever-increasing demand for and popularity of online higher education, quality issues associated with online learning experiences and online teaching practices emerge because online learning has unique requirements and characteristics different from traditional education

(Dixson, 2010; Kozan & Richardson, 2014; Kucuk & Richardson, 2019; Redmond et al., 2018). Higher education institutions have been offering more and more online courses, programs, and degrees (Caskurlu et al., 2021) and they are expanding their efforts to assure the quality of their online courses and programs, viewing online education as a significant component of their long-term strategies in this rapidly growing and highly competitive sector of higher education (Kozan & Richardson, 2014; Kucuk & Richardson, 2019).

Keeping online students actively involved, on task, and engaged with their online learning experiences (e.g., interactions with peers) is one of the significant factors that influence the quality of online education and online student success (Dixson, 2015; Kucuk & Richardson, 2019; Revere & Kovach, 2011). In other words, student engagement is seen as one of the important factors influencing students' online learning experiences (Robinson & Hullinger, 2008). Similarly, engagement in learning is often regarded as one of the most important standards of quality of overall university education (Coates, 2005; Redmond et al., 2018). As such, accurately measuring and thoroughly understanding the extent to which online students are behaviorally, emotionally, and cognitively engaged in their learning process is of paramount importance not only to online learning researchers but also to online educators, instructional designers, and higher education institutions because results of engagement assessments can be used as empirical evidence to make informed decisions regarding the quality of online education (Dixson, 2015; Meyer, 2014). Accordingly, the stakeholders of online education can design and implement evidence-based interventions to improve student engagement as a pliable state sensitive to contextual variations and changes (Fredricks, Wang et al., 2016; Manwaring et al., 2017), to increase the quality of learning experiences, and to enhance the retention rate, which in turn leads to successful online student graduations (Meyer, 2014). Further, accurate and useful assessments are much needed by all stakeholders in higher education, both traditional and online, to ensure accountability and to demonstrate empirical evidence of the effectiveness of courses and programs (Robinson & Hullinger, 2008; Sriram, 2014).

Problem Statement

Although student engagement is a key factor in online learning and is associated with a variety of positive student outcomes including student satisfaction and academic performance (Bolliger & Martin, 2018; Dixson, 2015; Kucuk & Richardson, 2019), research specifically examining student engagement in online learning environments is still new and therefore quite limited (Dixson, 2015; Kucuk & Richardson, 2019; Paulsen & McCormick, 2020; Vayre & Vonthron, 2017). Researchers have extensively studied this elusive and complex construct in traditional classrooms, using different conceptualizations and measures with mostly multidimensional operationalizations of the construct (Appleton et al., 2008; Deng et al., 2020; Fredricks & McColskey, 2012; Hoi & Le Hang, 2021; Kuh, 2003, 2009, 2016). One of the most common conceptualizations and operationalizations used by engagement researchers is Fredricks et al.'s (2004) tripartite conceptualization as behavioral engagement (i.e., involvement and participation in learning tasks and activities), emotional engagement (i.e., positive and negative reactions to learning experiences), and cognitive engagement (i.e., use of cognitive and metacognitive strategies and psychological investment in learning) (Henrie et al., 2015). In considering this multidimensional construct in online learning environments, three issues warranting empirical examination have been identified.

The first major issue is that there is not sufficient empirical work and systematic scale development and validation evidence showing whether and how Fredricks et al.'s (2004) three-dimensional conception of student engagement applies to online learning to explain how student engagement works in online learning environments. Although extensively studied in traditional face-to-face learning environments (e.g., Fredricks et al., 2005; Fredricks & McColskey, 2012; Reeve, 2013; Reeve & Tseng, 2011), only a few studies have used this tripartite view of engagement to measure online student engagement, with multiple indicators of each sub-dimension being applied to online learning environments of higher education (e.g., Kucuk & Richardson, 2019; Park & Yun, 2018). The issue of limitation is that researchers who have measured these three dimensions of student engagement in online learning environments have attempted to create a scale for their specific research purposes without systematic scale development and validation work. For example, Sun and Rueda (2012) adapted certain items from Fredricks et al.'s (2005) validated engagement scale and aimed to measure three types of engagement of university students in an online learning context via a new scale. Although they did some exploratory factor analytic work, they provided no confirmatory evidence for the construct validity of their scale. Similarly, Park and Yun (2018) adapted items from Sun and Rueda's (2012) measure to assess the three types of engagement of undergraduate and graduate students in the online learning context without doing any systematic instrument validation work to confirm the construct validity of their engagement scale. Most recently, Kucuk and Richardson (2019) modified Reeve's (2013) 17-item engagement scale, which included the three dimensions of Fredricks et al.'s (2004) engagement as well as agentic engagement as a fourth dimension (see Reeve, 2013). They reported modifying the wording of each item to fit them to the online learning context and stated that they did not need to validate the scale again on the grounds that the meaning of the items remained the same (Kucuk & Richardson, 2019). Although Reeve (2013) had provided reliability and validity evidence for the original scale, there was no validity evidence specifically for the online version of the scale used by Kucuk and

Richardson (2019), who reported no factorial or construct validity work special to their online student engagement measure. Such uses of modified or ad hoc scales to measure online engagement are quite common in the existing literature (e.g., Ferrer et al., 2020).

Given that online education is different from traditional education in terms of the characteristics of online learning and online learners (Kucuk & Richardson, 2019) and that engaging in online learning environments has its own challenges and complexities (Ferrer et al., 2020), more systematic scale development and validation studies need to be conducted to test the tenability of Fredricks et al.'s (2004) three-dimensional conception of student engagement in online learning environments. It would be wrong to simply assume that "engagement theories established in relation to the traditional on-campus classroom translate to the online learning environment" (Ferrer et al., 2020, p.319). There have been a few scale development and validation studies testing the tenability of behavioral, emotional, and cognitive dimensions of engagement within online learning environments. However, they were intended to develop and validate scales to measure engagement in special online learning contexts with idiosyncratic features such as Massive Open Online Courses (MOOCs) rather than regular university courses (e.g., Deng et al., 2020) and they are also conflated with the issues of either construct irrelevance or construct underrepresentation or both (e.g., Deng et al., 2020; Hoi & Le Hang, 2021).

The second major issue is construct irrelevance (Spurgeon, 2017). Some authors developed scales purported to measure online student engagement, but their scale items actually tap motivational elements or constructs (e.g., intrinsic motivation, goal orientation) rather than engagement (e.g., Schumacher, 2018). The significant distinction that needs to be drawn between motivation and engagement is often obfuscated in measures and studies of engagement, which is often stated as a serious problem in the relevant literature (Appleton et al., 2008; Fredricks &

McColskey, 2012). While motivation is basically one's underlying energies or reasons to perform or avoid a particular behavior or action, engagement is one's active and actual involvement in that behavior or action (Appleton et al., 2008; Fredricks & McColskey, 2012). For example, an online student may be motivated to participate in an academic learning task but not actually get engaged in the task for some other reasons and so their engagement may not actually occur at all. That is, motivation as one's energy or volition to do something may exist but it may not be sufficient for one's actual engagement in that activity or action (Appleton et al., 2008).

The final issue identified in the literature is construct underrepresentation (Spurgeon, 2017). Most of the studies examining student engagement in the online learning context and in higher education settings are primarily dealing with student engagement as a unidimensional construct (Vayre & Vonthron, 2017), more specifically as a behavioral phenomenon (Lee et al., 2019; Zhoc et al., 2019) identified by behavioral indicators such as how often college students perform certain behaviors (e.g., participating in asynchronous discussions) during their courses (e.g., Chen et al., 2010) or the number of logins (e.g., Lee et al., 2019). However, there is a significant body of theoretical and empirical literature pointing to the multidimensional nature of student engagement rather than considering it a unidimensional construct (Reeve et al., 2020). In addition, some online engagement scales developed and validated via scale development and validation studies were also identified as measuring behavioral, emotional, and cognitive dimensions of engagement with limited representation of each dimension (e.g., Deng et al., 2020; Hoi & Le Hang, 2021). Thus, in this study, online student engagement is conceptualized and operationalized as a three-dimensional construct and with items intentionally designed and developed to thoroughly tap and adequately represent each core dimension of engagement (Fredricks et al., 2004).

In support of the importance and relevance of all these issues identified to address in this dissertation study, it has been suggested by several scholars in the engagement literature that there is still a pressing need for more systematic and meticulous approaches and endeavors regarding the measurement of engagement since the existing measures have serious theoretical, conceptual, and operational inconsistencies and limitations as well as lacking empirical evidence for their scale validity (Fredricks & McColskey, 2012; Fredricks, Wang et al., 2016; Hoi & Le Hang, 2021; Veiga et al., 2014).

In summary, there is a need to address these three major issues identified regarding the measurement of student engagement in online learning environments. More thorough, reliable, and valid assessments of student engagement may become possible in online learning environments of higher education with the development of a new scale by using Fredricks et al.'s (2004) tripartite conception of engagement (addressing the first issue), by tapping the experience-based construct of engagement itself rather than the energy or drive-based construct of motivation or any other construct other than engagement (addressing the second issue), and by adequately representing all three core dimensions of engagement (addressing the third issue).

Purpose of the Study

Despite the theoretical and empirical justifications for using Fredricks et al.'s (2004) tripartite conception for measuring student engagement in online learning environments of higher education (e.g., Kucuk & Richardson, 2019; Park & Yun, 2018), there is still little research attempt to develop a theoretically well-defined and psychometrically strong scale for measuring student engagement in online learning environments by using this tripartite conception (e.g., Sun & Rueda, 2012). Although Sun and Rueda's (2012) study represents an initial attempt to measure online student engagement in a higher education context by developing a new scale using Fredricks et al.'s (2004) tripartite conception, their scale is still limited especially in terms of fully capturing online engagement due to the limited number of items they used for measuring each dimension of student engagement. The practice of merging certain cognitive engagement items with behavioral and emotional engagement items after their exploratory factor analysis and their use of only exploratory factor analysis for construct validation without doing any confirmatory analysis also raise serious concerns and questions about the construct validity of their scale. Thus, the problem associated with a lack of fully and distinctly capturing each dimension of online student engagement also called for this study to develop a new scale that can reliably and validly measure each unique dimension of student engagement with a sufficient number of items for each dimension.

Therefore, the purpose of this study was to develop and validate a new self-report survey scale of student engagement in online learning environments of higher education, titled as Online Engagement in Higher Education (OEHE). The OEHE scale was intended to have three subscales: online behavioral engagement, online emotional engagement, and online cognitive engagement, each of which has specific items intended to operationalize and measure behavioral (i.e., participation and involvement in online learning tasks and activities), emotional (i.e., experience of emotional reactions toward online learning experiences), and cognitive (i.e., use of cognitive and metacognitive strategies) dimensions of student engagement in online learning environments of higher education respectively. The definitions of the three distinct dimensions of online student engagement in this study are as follows:

Online behavioral engagement in this study refers to the extent to which students of higher education (i.e., undergraduate and graduate) exhibit specific behaviors of participation and involvement in their online learning process within online courses of higher education, corresponding to *participation* and *involvement* in the relevant literature of behavioral engagement (Finn, 1989; Fredricks et al., 2004; Fredricks & McColskey, 2012). These specific online student behaviors include but are not limited to reading discussion posts, responding to discussion posts/prompts, studying online learning materials, and submitting assignments online, all of which indicate online students' behavioral engagement in the academic learning process within online courses of higher education.

Online emotional engagement in this study refers to the extent to which online students of higher education experience certain positive emotions such as enjoyment and curiosity regarding their learning experiences in online courses.

Online cognitive engagement refers to the extent to which online students of higher education expend and regulate conscious mental effort to use cognitive and metacognitive strategies during online learning, such as trying to make connections between online course concepts, putting forth effort to understand diverse perspectives during online discussions, and reflecting on their own understanding.

Behavioral, emotional, and cognitive dimensions were purposefully selected and incorporated into the measurement of online student engagement in this study because they have been shown to be the three core dimensions or major components of student engagement across studies (Fredricks & McColskey, 2012; Reeve, 2013). Despite great variations in terms, definitions, and conceptualizations of the construct of engagement, Fredricks et al. (2004) in a seminal review of the engagement literature came to the conclusion that "engagement can be thought of as a 'meta' construct" (p.60) and put forward *behavior*, *emotion*, and *cognition* as the three major facets of engagement. Within this tripartite conceptualization, each dimension of engagement is investigated individually but in a dynamic relation to or simultaneous interaction with one another, which makes the meta or multidimensional conceptualization of engagement more capable of reflecting what individuals actually experience in real-life situations which are much more complex relationships occurring among behaviors, emotions, and cognitions (Fredricks et al., 2004). Therefore, the OEHE scale with its three subscales used Fredricks et al.'s (2004) multidimensional conception of student engagement in its conceptualization and operationalization of online student engagement in this study.

Research Questions

This study focused on developing and validating a self-report survey scale intended to measure behavioral engagement, emotional engagement, and cognitive engagement of students in online higher education settings as the three core dimensions of student engagement based on Fredricks et al.'s (2004, 2005) tripartite conception of engagement. For this purpose, the current study aimed to assess the reliability and validity evidence associated with the OEHE instrument by conducting factor analyses to confirm its a priori hypothesized three-factor structure and by assessing internal consistency reliability of the items in the overall scale and within each subscale. In addition to factorial and structural validity through confirmatory analytic work, construct validity evidence for the scale and the subscales was also evaluated through testing for convergent, discriminant, and criterion validity evidence that was collectively used to evaluate the appropriateness of the use of the OEHE to measure online student engagement in higher education. In line with the research purpose, the following research questions guided this instrument development and validation study:

RQ1: Can a three-factor hypothesized model of online student engagement based on theory and prior research be confirmed in a validation sample of online students of higher education?

RQ2: Assuming the hypothesized three-factor model of online student engagement is at least an acceptable fitting model, does convergent, discriminant, and criterion validity evidence support the use of the scale factors as indicators of online student engagement in higher education contexts?

RQ3: Do the confirmed OEHE scale factors exhibit evidence of internal consistency reliability? If yes, to what extent?

The first research question addresses whether a three-factor model of student engagement that is a priori hypothesized on the basis of Fredricks et al.'s (2004) conception of engagement can be validated in online learning environments of higher education with a sample of undergraduate and graduate students taking online courses as part of their education. The second research question addresses the construct validity evidence to be obtained through convergent, discriminant, and criterion validity evidence supporting the use of the OEHE factors to represent online student engagement in higher education. Finally, the third research question addresses the internal consistency reliability of the OEHE scales and of the OEHE instrument as a whole.

Significance of the Study

Online education continues to grow at a rapid rate in higher education and student engagement in online learning environments remains a significant issue and an important factor influencing online learning experiences (Hoi & Le Hang, 2021; Kucuk & Richardson, 2019; Meyer, 2014; Robinson & Hullinger, 2008). Accurate assessments and evaluations of online student engagement are thus necessary for all stakeholders of online education to make appropriate and accurate evidence-based decisions about instructional design and practices, and to introduce possible interventions in online courses to improve student engagement and so ensure successful online learning experiences (Czerkawski & Lyman, 2016; Dixson, 2015; Kozan & Richardson, 2014; Kucuk & Richardson, 2019).

The findings of this instrument development and validation study have significant implications regarding student engagement in online learning environments. First, this study contributes to the literature with the development of an accurate and useful assessment tool of online student engagement in higher education settings. The OEHE instrument can be used to assess students' behavioral, emotional, and cognitive engagement levels in online learning environments to make evidence-based decisions for learning and instruction (e.g., course design, development of tasks and activities). Second, this study contributes to extending our theoretical understanding of the complex construct of engagement in online learning environments. Third, researchers in the field may benefit from this study as the study offers a robust measure of engagement that they can use to examine relationships between engagement and other constructs of their interest in online learning environments. Lastly, the findings of this study help community of practice to highlight the issues of diversity, inclusiveness, and accessibility in online education, which may be used to identify weaknesses in online course design and facilitation and to improve students' online learning experience through improving the quality and effectiveness of online courses.

Chapter 2: Literature Review

In this chapter, a review of the literature on engagement in both traditional face-to-face and online learning environments is presented, primarily focusing on the major definitions and conceptualizations of the construct and subconstructs for the specific purposes of the current study. Then, an overview of operationalizations and measurements of each dimension of student engagement together with methodological limitations associated with them is presented. For the specific purpose of the study, this section focuses on self-report survey measures used for each subdimension of engagement. Finally, two major issues of limitation are discussed regarding the operationalization and measurement of online student engagement through a critical review of existing measures of online student engagement, highlighting their theoretical and/or methodological limitations that are in turn used by the current study to justify the need for the development and validation of a new instrument in this study.

Student Engagement: An Overview

Student engagement has been studied for decades both as an important educational outcome sensitive to different factors in learning environments (Reeve et al., 2020) and as an important variable empirically associated with other significant outcomes, including better learning, academic progress and achievement (Appleton et al., 2006; Azevedo, 2015; Finn & Zimmer, 2012; Fredricks, 2015; Fredricks, Filsecker, & Lawson et al., 2016; Glanville & Wildhagen, 2007; Henrie et al., 2015; Klem & Connell, 2004; Reeve et al., 2004; Reeve & Tseng, 2011; Reeve et al., 2020), student behaviors in the classroom and at school (Azevedo, 2015; Fredricks, 2015; Klem & Connell, 2004; Wang & Fredricks, 2014), student persistence and student satisfaction (Bolliger & Martin, 2018; Meyer, 2014), school completion (Fredricks, Filsecker, & Lawson et al., 2016; Reeve et al., 2004), and withdrawal from school and dropout decisions (Archambault et al., 2009; Fredricks, 2015; Glanville & Wildhagen, 2007; Henrie et al., 2015; Wang & Fredricks, 2014).

Student engagement has been extensively studied especially in K-12 classrooms and its major theoretical models and conceptualizations mostly emerged from research conducted in K-12 education contexts (Appleton et al., 2008; Fredricks et al., 2004; Fredricks & McColskey, 2012). Theoretical and empirical research interest in the construct of engagement and also disengagement started in the 1980s with a view of examining and gaining insights into why students would feel bored, isolated or disconnected from others in their classrooms and schools, gradually become at-risk students, and eventually decide to drop out (Finn & Zimmer, 2012). Student engagement became a popular construct of educational research especially among researchers and authors who were interested in understanding such issues as absenteeism, withdrawal or disengagement from school, alienation from school, poor student participation in school-related activities, negative student behaviors including delinquency, and school dropouts in US secondary schools (Finn, 1989; Natriello, 1984).

Engagement was also used as a practical model to develop effective school interventions to address the problem of high school dropouts and to actualize school reforms (Reschly & Christenson, 2006). Since then, engagement has been extensively investigated in educational research settings by various researchers to understand its relations to other significant educational variables and learning outcomes (Fredricks, 2015). Especially in the past decade, there has been an increasing amount of interest and attention in student engagement as a research construct and as an educational phenomenon among researchers, educators, education institutions, and similar other stakeholders (Bond et al., 2020). In line with this growing interest in the topic, student engagement has become an important phenomenon of research interest within the contexts of traditional and online higher education as well (Groccia, 2018; Kucuk & Richardson, 2019; Kuh, 2016; Martin, Sun, & Westine, 2020; Vayre & Vonthron, 2017; Xu et al., 2020).

In higher education, student engagement has come to be regarded as a significant factor influencing the quality of education (Coates, 2005; Groccia, 2018; Xu et al., 2020). Similar to the positive effects of student engagement empirically shown in traditional face-to-face classrooms (Fredricks, 2015), student engagement in online learning environments has been associated with several positive outcomes, including student satisfaction, academic performance, and student persistence (Bolliger & Martin, 2018; Kucuk & Richardson, 2019; Meyer, 2014), student learning and academic success (Dixson, 2015; Vayre & Vonthron, 2017) and student retention (Vayre & Vonthron, 2017). However, compared to the substantial body of empirical literature on student engagement in traditional learning environments, there is less research examining student engagement in online learning environments (Vayre & Vonthron, 2017). In addition, although there is a wealth of online learning literature with studies examining online engagement through such constructs as online presences, collaboration, and interaction (Bolliger & Martin, 2021), which were reviewed and categorized as sub-themes of online engagement (e.g., Martin, Sun, & Westine, 2020), studies examining engagement in online learning environments through the lens of three core dimensions of behavioral, emotional, and cognitive engagement (Fredricks et al., 2004) are still needed, compared to research on these dimensions in traditional classrooms and brick-and-mortar schools.

Sun and Rueda (2012), for example, examined the hypothesized relationships among interest, computer self-efficacy, self-regulation skills and online graduate student engagement that they conceptualized on the basis of Fredricks et al.'s (2004, 2005) behavioral, emotional, and cognitive engagement. They found that certain specific self-regulation skills of online students significantly predicted all three types of online student engagement. Park and Yun (2018) investigated the impact of eight specific motivational regulation strategies (e.g., mastery self-talk, environmental control, enhancement of personal significance) employed by undergraduate and graduate online students on behavioral, emotional, and cognitive engagement (Fredricks et al., 2004, 2005). They found that each type of online student engagement was influenced and predicted by certain motivational regulation strategies when students' academic level was controlled for. Kucuk and Richardson (2019) examined the structural equation relationships among online presences (i.e., teaching, social, and cognitive presences), online student engagement based on Fredricks et al.'s (2004, 2005) conceptualization of engagement in addition to agentic engagement (Reeve, 2013), and online student satisfaction in a fully online graduate program. Behavioral, emotional, and cognitive engagement all had significant effects on online student satisfaction (Kucuk & Richardson, 2019).

To sum up, the existing studies indicate that student engagement with its three core dimensions can come into play in relation to significant variables in online learning environments and can yield positive outcomes. Therefore, the key construct of student engagement needs to be incorporated into empirical examinations of online learning environments either as a predictor or outcome variable to get a thorough understanding of online learning environments. However, such empirical examinations need reliable and valid measures to precisely assess this significant construct. In light of the inconsistencies and variations in the conceptualizations, operationalizations, and measurements of engagement in the online learning literature (Hoi & Le Hang, 2021), this study aims to provide a reliable and valid assessment tool to make robust examinations of student engagement possible in online learning environments of higher education.

Student Engagement: Myriad Terms, Definitions, and Conceptualizations

With the increasing popularity of and scholarly research attention to the construct of engagement, variations and inconsistencies in terms, definitions, and conceptualizations of engagement as well as diverse operationalizations and measurements of engagement have emerged in the relevant literature (Ainley, 2012; Appleton et al., 2008; Deng et al., 2020; Fredricks, Filsecker, & Lawson, 2016; Fredricks & McColskey, 2012; Hoi & Le Hang, 2021). For instance, engagement as a construct has been referred to by various terms in the relevant literature, including but not limited to *student involvement* (e.g., Astin, 1984; Finn, 1989), *student engagement* (e.g., Christenson et al., 2012; Fredricks, 2015; Marks, 2000), *school engagement* (e.g., Fredricks et al., 2004; Furlong et al., 2003; Jimerson et al., 2003; Wang & Fredricks, 2014), *academic engagement* (e.g., Fredricks, 2015), and *learner engagement* (Deng et al., 2020). Among all these different terms and names, *student engagement* and *school engagement* are most widely used two terms in the relevant literature (Fredricks & McColskey, 2012). In this study, the terms *engagement* and *student engagement* are used interchangeably to refer to students' behavioral, emotional, and cognitive involvement in their own academic learning process in online courses of higher education.

The diversity of the engagement terms is also seen in myriad definitions and conceptualizations of engagement (Henrie et al., 2015; Wang et al., 2016) used by researchers who theoretically and empirically studied engagement in both K-12 and higher education settings (e.g., Astin, 1984; Coates, 2007; Kuh, 2003; Fredricks et al., 2004; Richardson & Newby, 2006). For instance, some researchers have defined and conceptualized engagement by referring to the amount of physical and psychological energy, time, attention, interest, personal or psychological investment, and effort students devote to their academic learning process as well as their involvement in non-academic activities at their institutions (e.g., Astin, 1984; Coates, 2007; Kuh, 2003; Marks, 2000; Newmann et al., 1992; Richardson & Newby, 2006; Sun & Rueda, 2012). Others have described engagement as school engagement and identification with school by using students' behavioral and affective characteristics, such as student participation in learning activities and student participation in school-level activities, their sense of belonging to school, positive emotions about schooling, and valuing school-related outcomes and education in general, as the subdimensions of engagement (e.g., Finn, 1989; Skinner & Belmont, 1993; Voelkl, 1997). As Gibbs (2014) stated, the term *student engagement* has become such a buzzword in higher education used in diverse ways and for many different purposes, and what people actually mean by this term often remains vague and obfuscated with so many other terms or constructs (Groccia, 2018).

A careful review of the seminal literature of student engagement also indicates that affective/psychological and cognitive dimensions have been less frequently incorporated into the earlier definitions and conceptualizations of engagement than its behavioral dimension, which is included in almost all definitions of engagement (Appleton et al., 2008; Jimerson et al., 2003). Moreover, despite the consensus on the multifaceted composition of student engagement (Ainley, 2012; Deng et al., 2020; Fredricks, 2015; Hoi & Le Hang, 2021), there is still no scholarly consensus on the exact number of the subdimensions of engagement "which range from two (i.e., behavior and emotion) to four (i.e., academic, behavioral, cognitive, and psychological/affective)" (Fredricks, 2015, p. 31). Agentic engagement (Reeve, 2013) and social engagement (Hoi & Le Hang, 2021) are other dimensions of student engagement found in the relevant literature. There is also no consistency in the naming of these different dimensions of engagement (Appleton et al., 2008; Fredricks, 2015). For example, while Fredricks et al. (2004) used the terms *behavioral*, *emotional*, and *cognitive* to describe the three subdimensions of engagement, Wang and Holcombe (2010) used the terms *use of self-regulation strategies*, *sense of identification with school*, and *school participation* to actually refer to cognitive engagement, emotional/affective engagement, and behavioral engagement respectively.

Conceptualizations of Student Engagement in Online Learning Environments

There is still no agreed conceptualization of online student engagement in the relevant literature since different researchers use different terms or constructs to refer to and examine online student engagement (Bolliger & Martin, 2021; Dixson et al., 2017; Martin, Sun & Westine, 2020). One camp of researchers studying engagement in online courses defined it by highlighting the role of student effort in engagement. For example, Kucuk and Richardson (2019) defined student engagement in their study as “active involvement in course activities with continuous efforts to attain desired learning outcomes” (p. 199). Park and Yun (2018) defined engagement as “the quality of students’ efforts to achieve designed outcomes” (p. 45). Similarly, Sun and Rueda (2012) defined engagement in academic learning environments as “the quality of effort students make to perform well and achieve desired outcomes” (p. 193). With a similar emphasis on conscious student effort and investment in learning, Bigatel and Edel-Malizia (2018) defined student engagement as “the time and physical energy that students expend on activities in their academic experience” (p. 59).

All these engagement definitions used by online engagement researchers emphasize the importance of student effort in achieving certain learning outcomes. Extending the scope of student engagement beyond student efforts and investment to be successful, Dixson (2015) defined online student engagement as “students putting time, energy, thought, effort, and, to some extent, feelings into their learning” (p. 4). According to this definition, when online students are engaged in online learning, they not only display certain behavioral indicators of engagement such as allocating sufficient amounts of time, energy, and efforts, but they also are cognitively and emotionally involved in the online learning process (Dixson, 2015). This shift from mere behavioral indicators of student engagement towards more cognitive or emotional dimensions was emphasized in earlier definitions of student engagement in online learning environments. For example, Richardson and Newby (2006) defined the cognitive dimension of student engagement in online learning as “the integration and utilization of students’ motivations and strategies in the course of their learning” (p. 25), highlighting what motivates online learners to study and learn, and what learning strategies or approaches they employ to achieve their different motivations.

Online engagement has also been described by another camp of researchers who followed an interaction and sense of community framework in online learning. For example, Lear et al. (2010) defined an engaged online learner as a function of online learner interactions with other learners, with their instructors, and with the content and a concomitant sense of community in an online learning environment. Following a similar approach, Martin and Bolliger (2018) also defined online student engagement as emerging from online students interacting with their peers, with their instructors, and with the content.

Other online engagement researchers conceptualized engagement as a three-component construct in line with Fredricks et al.’s (2004) tripartite conception, but with different concepts. For instance, Vayre and Vonthron (2017) conceptualized online learner engagement, based on Brault-Labbe and Dube’s (2010) theoretical framework, as being made up of perseverance, enthusiasm, and reconciliation, which corresponds to behavioral engagement, emotional engagement, and cognitive engagement in online learning environments respectively (Vayre & Vonthron, 2017). Others have extended the three core dimensions of student engagement (i.e., behavioral, emotional, cognitive) in online learning environments. For example, social engagement has been added as a fourth dimension to the behavioral, emotional, and cognitive dimensions in order to conceptualize and assess online student engagement (e.g., Deng et al., 2020; Hoi & Le Hang, 2021).

In the online learning research literature, engagement has been referred to and investigated through different constructs in online learning including interaction, community building, presences, communication, and collaboration (Bolliger & Martin, 2021; Martin, Sun, & Westine, 2020), which clearly indicates the messiness and inconsistency of the conceptualizations of the construct of student engagement in online learning environments. Furthermore, although student engagement is generally conceptualized as a multifaceted construct by consensus in the relevant literature (Hoi & Le Hang, 2021), there are still online learning studies treating the construct of engagement as a one-dimensional single overarching construct (e.g., Imlawi et al., 2015; Murillo-Zamorano et al., 2019).

Despite all these variations in conceptualizations and subsequent operationalizations of engagement, since Fredricks et al.'s (2004) seminal review of the engagement literature concluded that "engagement can be thought of as a 'meta' construct" (p.60), their tripartite conceptualization of student engagement with subdimensions conceptualized as *behavioral engagement*, *emotional engagement*, and *cognitive engagement* has become the common dimensions of engagement (Henrie et al., 2015). Each of these dimensions is investigated individually but in a dynamic interaction with one another, which makes this meta-conceptualization of engagement more capable of reflecting what individuals actually experience in real life situations, which are much more complex relationships occurring among behaviors, emotions, and cognitions (Fredricks et al., 2004). Each dimension of engagement is also viewed as varying in intensity or quality (Fredricks et al., 2004). For example, cognitive engagement may mean simply memorizing a list of items and reproducing it when asked, or deeply understanding the items by using certain learning strategies and self-regulation skills (Fredricks et al., 2004). Moreover, engagement in this conceptualization can be specific to certain contexts or situations and can be temporary or it can sustain over longer periods of time and so can be more permanent (Fredricks et al., 2004).

Behavioral Engagement

The relevant literature of engagement defines behavioral engagement in three main ways (Fredricks et al., 2004). The first group of definitions conceptualizes behavioral engagement as positive and compliant student behavior (Fredricks et al., 2004). Student behavior, such as breaking school and classroom rules, paying no attention to the teacher and the classroom work, and misbehaving in the classroom or school, is considered non-compliant student behavior (Finn & Rock, 1997).

In the second group of definitions, behavioral engagement is conceptualized as students' active involvement and participation in instructional activities and academic learning tasks through proactive behaviors of asking questions, having learning-related conversations with their teachers and the class, doing more homework than they are expected to do, making effort, paying attention to classroom instruction, and showing persistence when faced with challenge or difficulty (Finn & Rock, 1997; Fredricks et al., 2004; Sinatra et al., 2015). This conceptualization of behavioral engagement as positive and active student involvement with and participation in classroom learning activities is commonly used as a core dimension of student engagement in most definitions and conceptualizations in the relevant literature (Skinner et al., 2009). In the third and final group of definitions of behavioral engagement, behavioral engagement is viewed as student participation in school-level academic or non-academic activities including sports and extracurricular activities at school (Finn & Rock, 1997; Fredricks et al., 2004; Sinatra et al., 2015).

Emotional Engagement

Within the tripartite conceptualization of engagement, emotional engagement is usually defined as “students’ affective reactions in the classroom, including interest, boredom, happiness, sadness, and anxiety” (Fredricks et al., 2004, p. 63). Nevertheless, myriad definitions and conceptualizations of emotional engagement exist in the relevant literature. Finn’s (1989) participation-identification model with its *identification with school* component characterized by students’ sense of belongingness and connectedness with their school and their valuing education or school-related outcomes has been used as one major conceptualization of emotional engagement in the relevant literature (Fredricks et al., 2004). This conceptualization of emotional engagement using belongingness and value overlaps with an extensive body of theory and research on students’ sense of belonging, relatedness, and value perceptions (Appleton et al., 2008; Fredricks et al., 2004).

The basic human need to belong to others in general and students’ sense of belonging and connectedness to school in particular have been considerably studied for decades by researchers across different academic contexts and cultural settings in relation to different outcomes or variables and it has a large theoretical and empirical knowledge base of its own (e.g., Baumeister & Leary, 1995; Bonny et al., 2000; Gillen-O’Neel & Fuligni, 2013; Goodenow & Grady, 1993).

Value as the second element of this conceptualization has also been extensively theorized and empirically studied for decades as another major area of educational psychology and motivation research, especially under the Expectancy-Value theory of achievement motivation (Wigfield & Eccles, 2000).

Another major conceptualization of emotional engagement involves students' affective responses to or affective perceptions of their teachers, peers, academic work, and their schools (i.e., positive and negative emotions) (Fredricks et al., 2004). This conceptualization of emotional engagement overlaps with the extensive literature of achievement emotions that have been empirically shown to have influence on several significant learning outcomes in the relevant literature of motivation and emotions (Pekrun, 2006; Pekrun et al., 2009).

Cognitive Engagement

Similar to diverse definitions and conceptualizations of behavioral engagement and emotional engagement, cognitive engagement also is "a loosely defined construct" (Cleary & Zimmerman, 2012, p. 238) and has been conceptualized from different theoretical perspectives in the engagement literature (Ainley, 2012; Fredricks et al., 2004). Accurately defining cognitive engagement, which is much less observable than behavioral and emotional engagement, is considered to be much more challenging (Sinatra et al., 2015). The use of the concept of cognitive engagement dates back to the 1980s, when Corno and Mandinach (1983) in their seminal manuscript described cognitive engagement as one form of student engagement.

Corno and Mandinach (1983) conceptualized cognitive engagement as active student engagement during which students interpret their personal and social experiences within their classrooms regarding their learning experiences (e.g., making causal attributions, having expectancies for success, making peer comparisons) and accordingly regulate and expend a certain amount and quality of effort (e.g., learning strategies) to carry out and achieve learning tasks in their classrooms. In fact, Corno and Mandinach (1983) conceptualized self-regulated learning as a specific form or type of cognitive engagement, incorporating such cognitive interpretations as success expectations, goals, or attributions into self-regulated learning. They specifically stated that "self-regulated learning is the highest form of cognitive engagement" (Corno & Mandinach, 1983, p. 90). This self-regulated learning (SRL) theoretical perspective accounts for one major camp of definitions and conceptualizations of cognitive engagement in the engagement literature (Fredricks et al., 2004). According to the SRL-based conceptualization, cognitively engaged students are those self-regulated learners who are "metacognitively, motivationally, and behaviorally active participants in their own learning" (Zimmerman, 1990, p. 4). That is, self-regulated students utilizing their metacognitive skills and strategies approach the learning task with certain goals and make a plan before diving into it, then start using certain strategies to deal with the task and regulate their attention and monitor their progress while doing the task, then evaluate their performance through self-reflection after completing the task, and then make necessary revisions to improve their learning experience in the future (Cleary & Zimmerman, 2012).

Cognitively engaged students are also known for their use of different cognitive strategies, such as rehearsal (e.g., repeating to memorize), elaboration (e.g., paraphrasing, summarizing), and organization strategies (e.g., preparing outlines, drawing mind maps), that may vary depending on the specific conditions or specific goals while working on the learning material to achieve their specific goals whatever they are (Pintrich & De Groot, 1990; Weinstein et al., 2011; Weinstein & Mayer, 1986). Deep versus shallow learning strategies used by students while studying the learning material at hand are another important component of cognitive engagement whereby "deep cognitive engagement involves elaboration processes, while shallow involves more rote memorization and other strategies that engage the new information in more superficial ways" (Xie et al., 2019, p. 184). Researchers have commonly used this self-regulated learning and cognitive-metacognitive strategy use conceptualization of cognitive engagement (e.g., Wolters, 2004).

Another major conceptualization of cognitive engagement is defining and conceptualizing it as *psychological investment* in learning, a construct heavily drawing from the motivation literature (Fredricks et al., 2004). A similar term or concept of *personal investment* in the motivation literature was originally proposed and used by Maehr (1984) to explain how human motivation works as personal investment that individuals prefer to make in terms of pursuing certain goals and performing certain actions or behaviors as opposed to others. Similar to Maehr's (1984) concept of personal investment, some researchers have widely defined and conceptualized cognitive engagement as psychological investment (Sinatra et al., 2015). According to this conceptualization of cognitive engagement, students who make cognitive effort to truly understand, learn, and master the knowledge and skills under study, volitionally do more than what they are required to do, can be flexible in problem solving, and prefer to be challenged by new things to master are considered to be psychologically invested in their own learning (Sinatra et al., 2015). Conceptualizing cognitive engagement from such a psychological investment perspective involves the use of several major constructs extensively studied in the motivation literature for decades, such as mastery goal orientations (Ames & Archer, 1988; Elliott & Dweck, 1988) and intrinsic motivation (Ryan & Deci, 2002), which often leads to great confusions over the crucial distinctions between engagement itself and its antecedents (Finn & Zimmer, 2012).

Student Engagement: Measures and Methodological Challenges

Different methods (e.g., surveys, observations, teacher ratings, experience sampling, interviews) have been utilized to measure student engagement (Fredricks & McColskey, 2012). Of all these different ways of collecting empirical data about student engagement, surveys are most commonly used by engagement researchers to measure student engagement (Almutairi & White, 2018; Bond et al., 2020; Fredricks & McColskey, 2012; Hoi & Le Hang, 2021). The consensus is that self-report survey measures should be the best method especially for the internal, less observable dimensions of student engagement (i.e., cognitive and emotional engagement) (Henrie et al., 2015; Hoi & Le Hang, 2021) because researchers would otherwise need to infer those dimensions of engagement from observable student behaviors rated by teachers or observers (Appleton et al., 2006; Fredricks, 2015). Fredricks and McColskey (2012) described eleven self-report measures of student engagement, presenting a detailed comparative analysis of the measures with their subscales corresponding to behavioral, emotional, and cognitive dimensions of engagement (see Fredricks & McColskey, 2012 for instruments). A careful examination indicates that "student engagement can be reliably measured through self-report methods" (Fredricks & McColskey, 2012, p. 777).

Measuring Behavioral Engagement. Researchers using self-report survey scales have operationalized behavioral engagement by focusing on various indicators such as attendance, compliance with classroom and school rules, time spent doing homework and studying, student effort, persistence, attention and concentration, initiative or self-directed learning behavior, and classroom participation exhibited by students during the process of academic learning as well as participation in school-level academic and non-academic activities (Fredricks & McColskey, 2012; Sinatra et al., 2015).

Different scales have been developed to date and used for the purpose of measuring behavioral engagement in the research literature (Fredricks et al., 2005) (e.g., The Rochester Assessment Package for Schools, the RAPS, Wellborn & Connell, 1987; the Rochester Assessment of Intellectual and Social Engagement, the RAISE, Miserandino, 1996; the Teacher Rating Scale of School Adjustment, the TRSSA, Birch & Ladd, 1997; and the National Educational Longitudinal Study, the NELS, as a nationwide representative dataset, Glanville & Wildhagen, 2007).

Fredricks et al. (2005) constructed and validated a self-report survey measure of engagement based on the previous measures of student engagement (e.g., Finn et al., 1995; Wellborn & Connell, 1987) and also used teacher ratings of students' behavioral engagement as well as the other dimensions of student engagement. To operationalize and measure behavioral engagement in these self-report survey measures and teacher ratings, Fredricks et al. (2005) used such indicators as effort, completion of schoolwork, attention, and compliance with school rules, and assessed self-reported behavioral engagement. For example, they used items "I follow the rules at school," "I pay attention in class," and "I complete my work on time." Similar other measures were also developed and used by researchers in the relevant literature (e.g., Reeve, 2013). Several other self-report survey measures have also been used by other researchers to operationalize and assess students' behavioral engagement under different subscales using different indicators of behavioral engagement (e.g., High School Survey of Student Engagement, Motivation and Engagement Scale) (see Fredricks & McColskey, 2012 for scales).

Limitations. The first limitation emanates from researchers merging varying levels of behavioral indicators in their operationalization and measurement of behavioral engagement, ranging from compliance with school rules to participation in extracurricular activities to self-directed autonomous behavior and participation in classroom learning activities. For example, in their seminal scale of student engagement adapted and used by many researchers to date, Fredricks et al. (2005) used both school-level positive conduct (e.g., "I follow the rules at school") and classroom-level student behavior (e.g., "I pay attention in class") as the indicators of behavioral engagement. Most recently, Fletcher et al. (2020) used indicators of both academic student behavior (e.g., "Studying for tests or quizzes") and extra-curricular non-academic student behavior (e.g., "I go to school to participate in athletics") to measure behavioral engagement. When researchers combine such differential indicators, their findings might be quite difficult, if not impossible, to interpret or compare across studies of behavioral engagement since their reported behavioral engagement would mean something different across studies due to such different operationalizations and measures (Fredricks & McColskey, 2012). The second limitation is concerned with the extraction of items from large databases (e.g., NELS) without having a strong conceptual or theoretical foundation, which further adds to the lack of clarity in definitions and conceptualizations of student engagement and its different subdimensions (Appleton et al., 2006; Henrie et al., 2015).

Measuring Emotional Engagement and Limitations. Emotional engagement has also been operationalized and measured in different ways (Fredricks et al., 2004). Some researchers have operationalized it with students' positive and negative emotions during learning experiences used as indicators of their emotional engagement in academic work and learning (e.g., enjoyment, boredom), while others have assessed students' sense of belonging to school, relationships with significant others (i.e., teachers, peers), identification with school, perceptions of valuing school and educational outcomes, and overall student attitudes about schooling and academic learning (Finn, 1993; Fredricks et al., 2004; Fredricks & McColskey, 2012; Sinatra et al., 2015).

Measures. Self-report survey measures have also been employed widely by engagement researchers to operationalize and measure students' emotional engagement (e.g., the RAPS, Connell & Wellborn, 1991; Wellborn & Connell, 1987; the RAISE, Miserandino, 1996). Furrer and Skinner (2003) used both teacher ratings and student self-reports of their emotional engagement by operationalizing it with such emotional indicators as positive and negative emotions during classroom learning experiences in the classroom (e.g., interest, frustration). In their tripartite conceptualization of student engagement, Fredricks et al. (2005) used both student self-reports and teacher ratings to measure students' emotional engagement. Their operationalization of emotional engagement included indicators of both students' overall attitudes about school (e.g., "I like being at school") and their positive/negative emotions about schoolwork (e.g., "I feel excited by my work at school"). Similar to Fredricks et al.'s (2005) use of individual emotional reactions to operationalize and measure students' emotional engagement, several other leading engagement researchers in the literature have also used such emotions or emotional reactions (e.g., enjoyment, interest, curiosity, happiness, enthusiasm) as the indicators of emotional engagement rated by students' self-reports and/or teacher ratings (e.g., Reeve, 2013; Reeve & Tseng, 2011; Skinner et al., 2009). Students' emotional engagement was also measured in the form of students' sense of belonging, identification with their schools and teachers, peer relationships, and also utility value perceptions (e.g., the NELS, Finn, 1993)

Limitations. The first limitation, which is actually a common limitation across engagement scales, is that most of the items are worded too broadly (e.g., "I like being at school" and "Teachers care about me") rather than specifying emotional engagement regarding specific activities, tasks, or situations (Fredricks & McColskey, 2012). Another major limitation is that indicators associated with different constructs other than emotional engagement are all merged under the scale purported to measure emotional engagement. For example, students' sense of belonging with teachers, peer relationships, identification with school, and utility value are all distinct constructs (e.g., basic psychological need for relatedness in Self-Determination Theory) in the motivation literature. However, some engagement researchers have combined and used such distinct constructs as indicators of emotional engagement (e.g., Ding et al., 2017; Xu et al., 2020). When the construct of emotional engagement is operationalized with such motivational indicators, serious questions emerge regarding construct validity. The final limitation is concerned with the practice of operationalizing and assessing emotional engagement together with another dimension of engagement under a single scale and variable (e.g., Marks, 2000). Such combined scales run the risks of damaging the construct validity of the scales and blurring the identification of the unique impact of each distinct dimension of student engagement on positive student outcomes (Fredricks et al., 2004).

Measuring Cognitive Engagement and Limitations. Self-report survey measures are also considered and have been commonly used as the most appropriate method of measuring cognitive engagement as an internal state (Henrie et al., 2015; Hoi & Le Hang, 2021).

Measures. Seminal studies on student engagement have commonly used self-report measures to operationalize and measure cognitive engagement (e.g., Connell & Wellborn, 1991). The assessment of cognitive engagement operationalized as one's self-investment and independence comprising continued effort and persistence in the learning process has been commonly employed by other researchers in the relevant literature (e.g., Miller et al., 1996). Fredricks et al. (2005) also used self-report survey items to operationalize and measure children's cognitive engagement in school and schoolwork via such items as "I study at home even when I don't have a test," "If I don't know what a word means when I am reading, I do something to figure it out," and "If I don't understand what I read, I go back and read it over again," all of which indicate the extent to which students expend effort, show persistence, and display psychological investment in their own learning.

Cognitive engagement was also operationalized and measured with indicators of students' use of cognitive and metacognitive self-regulated learning strategies. This operationalization of cognitive engagement actually dates back to the earlier work on study methods (Entwistle & Entwistle, 1970), self-reported learning strategies (Kardash & Amlund, 1991), learning processes (Schmeck et al., 1977), metacognition (Schraw & Dennison, 1994), cognitive and metacognitive learning strategies (Pintrich & De Groot, 1990), and self-regulated learning (Zimmerman & Martinez-Pons, 1990). Several engagement researchers operationalized and measured cognitive engagement with self-report survey scales tapping into cognitive and metacognitive strategies (e.g., Greene & Miller; 1996; Miller et al. 1996). Fredricks et al. (2005) also measured cognitive engagement via such indicators as "When I read a book, I ask myself questions to make sure I understand what it is about" and "I check my schoolwork for mistakes." Adapting certain items from Fredricks et al.'s (2005) scale, Sun and Rueda (2012) similarly used such items, utilizing online students' metacognitive self-regulatory strategy use as an indicator of online college students' cognitive engagement.

Limitations. The first limitation is the use of motivation constructs in the operationalizations and measures of cognitive engagement (e.g., Appleton et al., 2006). The crucial distinction between motivation and engagement is commonly obfuscated this way across engagement scales (Finn & Zimmer, 2012). The other significant limitation is the use of indicators of the other two dimensions to operationalize and measure cognitive engagement (e.g., Fletcher et al., 2020; Parnes et al., 2020).

A Critical Review of Issues Associated with the Existing Measures of Online Engagement

Student engagement has been operationalized and measured in different ways in online learning environments as well (Henrie et al., 2015). In addition to survey measures (e.g., Dixson, 2015; Hoi & Le Hang, 2021; Park & Yun, 2019; Sun & Rueda, 2012), online behavioral engagement has been measured through data analytics information such as frequency of logins, number of discussion posts and responses, or number of assignments submitted (Henrie et al., 2015; Hoi & Le Hang, 2021). Emotional engagement, apart from being measured through surveys, has been measured through students' explicit and observable use of positive emotion expressions during their online interactions (e.g., Wang, 2010). Cognitive engagement, as an internal state just like emotional engagement, has also been commonly measured through survey items in online education research (e.g., Deng et al., 2020; Hoi & Le Hang, 2021; Kucuk & Richardson, 2019; Sun & Rueda, 2012).

Two major issues of limitation have been identified in the existing measures of online engagement accessed through database searches and snowball referencing: (a) construct irrelevance (i.e., not measuring the intended construct of interest or measuring beyond the construct) and (b) construct underrepresentation (i.e., limited or broad measurements of engagement).

Construct Irrelevance. The first major issue of limitation regarding the existing measures of student engagement in online learning environments is associated with measuring other constructs rather than engagement. Some of the existing measures of online engagement purported to assess engagement were actually found to be tapping motivational elements such as intrinsic motivation or goal orientation (e.g., Schumacher, 2018). The crucial distinction between indicators of engagement and facilitators of engagement (Finn & Zimmer, 2012) seems to have been seriously obfuscated in such measures of online engagement, which raises serious questions and concerns over construct validity. Similar issues of construct irrelevance and construct validity were also identified in other existing measures used to assess engagement in online learning environments (e.g., Bigatel & Williams, 2015; Buelow et al., 2018; Dixson, 2015; Lee et al., 2019; Young & Bruce, 2011). Such confusion over the intended construct or subconstructs of interest adds to the conceptual and operational inconsistencies and ambiguities in the existing literature.

Construct Underrepresentation. The other important limitation identified in the existing measures of online student engagement is concerned with construct underrepresentation. Some existing measures of online engagement were identified as missing a core dimension of engagement such as emotional engagement (e.g., Anderson (2017). Missing the emotional dimension is a limited measure of the multifaceted construct of engagement, which in turn results in a limited representation of the construct. Therefore, it raises concerns over construct validity (DeVellis, 2012). The limited representations of the construct or subconstructs of engagement, including a unidimensional focus (e.g., Chen et al., 2010), limited number of dimensions (e.g., Manwaring et al., 2017), measuring engagement as a single overarching construct (e.g., Imlawi et al., 2015), measuring engagement at a broad institutional level (e.g., Robinson & Hullinger, 2008), and limited representation of each core dimension of engagement (e.g., Deng et al., 2020; Hoi & Le Hang, 2021; Sun & Rueda, 2012) were also identified in several other existing measures of online engagement.

To conclude, what is evident from the extensive review of the existing literature of student engagement and the critical review of online engagement measures is that student engagement and its dimensions have been conceptualized, operationalized, and measured in diverse ways (Henrie et al., 2015), thus leading to inconsistencies in the assessment online student engagement (Hoi & Le Hang, 2021). In addition, the existing measures of online student engagement, either modified as ad hoc scales or developed and validated through scale development studies, have several limitations regarding construct irrelevance and construct underrepresentation.

Chapter 3: Methods

Overview

The purpose of this study is to develop and validate a new self-report survey scale of student engagement in online courses of higher education titled *Online Engagement in Higher Education (OEHE)* by analyzing and evaluating its reliability and validity evidence. Based on an extensive review of the relevant literature and a critical analysis of the existing measures of online student engagement, a significant gap was identified in the relevant literature regarding the insufficient scale development and validation work testing the tenability of Fredricks et al.'s (2004) three core dimensions of engagement to online learning environments. The issues of construct irrelevance and construct underrepresentation were also identified in the existing measures of online student engagement. In response, a multi-stage scale development process was initiated in which a pool of items was created and reviewed by experts, pilot tested via cognitive interviews with actual respondents, and necessary revisions and modifications were conducted. The final set of sixty-six items that emerged as a result of this developmental process was submitted for large-group data collection since the aim of this dissertation study was to test the tripartite conceptualization of engagement and the developed items in online learning environments. After data collection, a series of item screening and evaluations, CFA analyses, validity coefficients, and reliability coefficients were conducted to test for the reliability and validity of the OEHE scales.

Research Design

Because the goal of this study was to develop a new self-report survey measure of online student engagement that can accurately and appropriately measure student engagement in online learning environments of higher education and to assess and evaluate its psychometric properties (i.e., reliability, validity), multiple steps of a scale development process were followed (DeVellis, 2012; Holmes, 2018). Initial item development, initial dissertation committee reviews, expert reviews, and pilot cognitive tests were conducted before the testing of the items in the validation sample constitute the multi-stage design of this study. For the large-group administration of the survey items, a cross-sectional survey research design was used in this study (Creswell, 2014).

Summary of Item Development and Validation Process

Since the reviewed measures of online student engagement were found to be limited because of the issues of measuring other constructs rather than engagement (i.e., construct irrelevance) and measuring engagement in a manner that was too limited or too broad (i.e., construct underrepresentation), the primary purpose of the initial item construction process was to ensure, as much as possible, that the constructed items would validly and sufficiently represent all three core dimensions of online student engagement in terms of their content and language. To achieve this purpose, construct definitions and conceptualization of each dimension of online student engagement with their indicators were used as a guide to the initial item construction process (see Table 3.1). While the initial pool of draft items was created, the scales of Fredricks et al. (2005), Reeve and Tseng (2011), Reeve (2013), Sun and Rueda (2012), and Young and Bruce (2011) were also consulted. Pekrun et al. (2002, 2005, 2011) was used for the operationalization of positive academic achievement emotions (i.e., enjoyment, pride, hope) and Pekrun and Linnenbrink-Garcia (2012) and Litman and Spielberger (2003) were consulted for the operationalization of epistemic emotion of curiosity. Redundant items were initially constructed since redundancy at the initial stages of scale development is considered to be “the foundation of internal-consistency reliability which, in turn, is the foundation of validity” (DeVellis, 2012, p. 85). Based on the feedback comments of the dissertation committee members and in light of the relevant literature and existing measures of engagement, iterative revisions, modifications, and additions were made before selecting and submitting a set of candidate items as the first draft of the instrument to the expert review process. Based on two rounds of expert reviews, iterative item revisions, additions, and deletions were made, leading to a second draft to be submitted to pilot cognitive interviews. As a result of the entire pilot cognitive testing process, a third and final draft of the survey emerged and was submitted to the actual large-group validation study to test the OEHE items in a validation sample (see Figure 3.1).

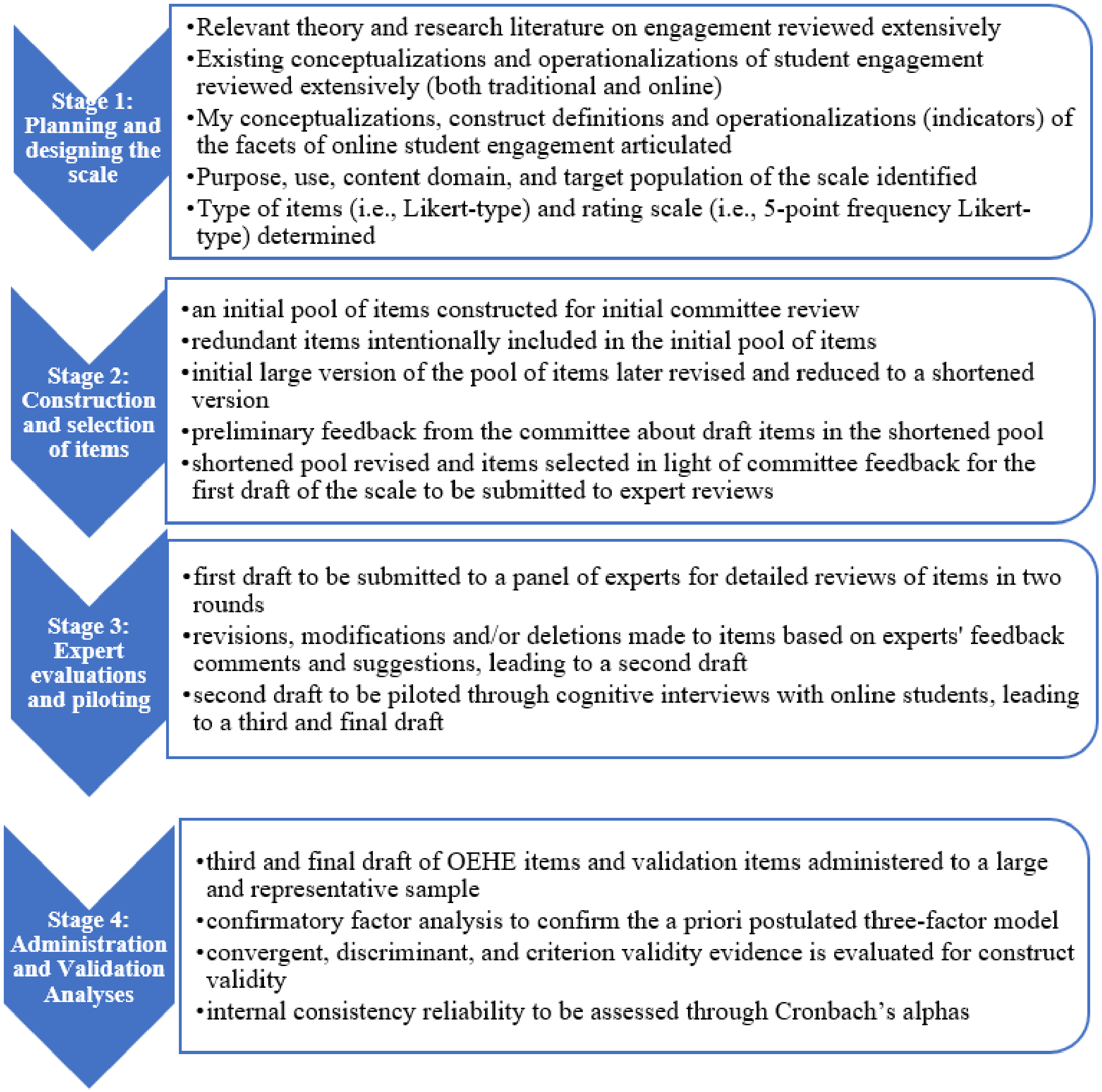
Table 3.1

Construct Definitions and Indicators of Online Student Engagement in the OEHE

Subdimension	Construct Definition	Indicators (Examples)
Online Behavioral Engagement	The extent to which students of higher education (i.e., undergraduate and graduate) exhibit specific behaviors of participation and involvement in their online learning process within online courses of higher education	• reading posts • responding to posts • sharing resources (e.g., articles, website links, videos) with others • studying text materials (e.g., pdfs, slides) • studying lecture videos • completing online tasks, activities, and assignments • submitting assignments online
Online Emotional Engagement	The extent to which online students of higher education experience certain positive emotions regarding their learning experiences in online courses	• enjoyment • pride • hope • curiosity
Online Cognitive Engagement	The extent to which online students of higher education expend and regulate conscious mental effort to use cognitive and metacognitive strategies during online learning	• making logical connections • reflecting on ideas • deriving logical conclusions • trying to fully understand • creating examples • formulating questions • making inferences • relating new information to prior knowledge • planning • monitoring • self-evaluating

Figure 3.1

Four Stages and Their Corresponding Steps to Be Followed during the Scale Development



Five response options were used in a Likert-type frequency response scale (1=Never, 5=Always) for two reasons. First, there is no consensus among researchers on the ideal number of points to be used in response scales (Leung, 2011) and second, "researchers often use five or seven response options, balancing fine-gradation, subtlety, and psychometric quality" (Furr, 2011, p.18). The optimal number of items and the final length of the OEHE instrument was determined in light of psychometric properties (i.e., internal consistency reliability, factorial validity) (DeVellis, 2012; Furr, 2011).

Construct Definition and Operationalization of Online Behavioral Engagement

Due to the course-level student participation and involvement intended to be measured, online behavioral engagement is defined in this study as the dimension which refers to *the extent to which students of higher education (i.e., undergraduate and graduate) exhibit specific behaviors of participation and involvement in their online learning process within online courses of higher education* (e.g., participating in asynchronous online discussions, participating in online activities, submitting assignments by due dates). This construct definition corresponds to conceptualization of behavioral engagement as student participation and involvement in academic learning (Finn, 1989; Fredricks et al., 2004; Fredricks & McColskey, 2012). Next, in light of the engagement literature suggesting common indicators of behavioral engagement (e.g., homework completion, studying, participation/involvement, asking for help, see Bond et al., 2020; Henrie et al., 2015) and in alignment with the relevant literature indicating common online student participation behaviors and engagement actions, such as participation in online discussions (e.g., discussion posting, reading others' discussion posts, responding to discussion posts) (Chen et al., 2010; Dennen, 2008; Goggins & Xing, 2016; Ramos & Yudko, 2008; Warren, 2018), knowledge sharing behaviors in online discussions (Kumi & Sabherwal, 2019; Yilmaz & Karaoglan Yilmaz, 2019), exchanging or sharing resources/contents with others (Park et al., 2016; Shackelford & Maxwell, 2012), watching lecture videos (Ozan & Ozarslan, 2016), note-taking behaviors of online students (van de Sande et al., 2017; Watkins et al., 2015), completing online tasks and activities, submitting assignments online (Cerezo et al., 2016; Kokoc et al., 2021; Park et al., 2016), and getting and using feedback from the instructor

(Watson et al., 2017), certain indicators intended to adequately and accurately represent the content domain or universe of online behavioral engagement were developed. To address the issues of construct irrelevance (CI) and construct underrepresentation (CU), all the constructed items were kept within the relevant domain of online student behaviors but nothing else and as many draft items as possible were created to adequately represent the relevant content domain of online behavioral engagement. Fredricks et al. (2005), Reeve and Tseng (2011), Reeve (2013), and Sun and Rueda (2012), in which student engagement was similarly operationalized and measured in line with Fredricks et al.'s (2004) multi-dimensional operationalization of engagement were also consulted. Table 3.2 presents the construct definition and the corresponding indicators of online behavioral engagement.

Table 3.2

Construct Definition and Indicators of Online Behavioral Engagement (OBE)

Subdimension	Construct Definition	Indicators (Examples)
Online Behavioral Engagement	The extent to which students of higher education (i.e., undergraduate and graduate) exhibit specific behaviors of participation and involvement in their online learning process within online courses of higher education	• posting • reading posts • responding to posts • sharing additional resources with peers • studying course materials • completing assignments • submitting assignments • read feedback • incorporate feedback into future work

Construct Definition and Operationalization of Online Emotional Engagement

Engagement is when an individual is actually doing something or actively experiencing some phenomenon or a certain state rather than being a reason or motivation behind that action or experience (Fredricks & McColskey, 2012). According to this definition, emotional engagement in this study is defined as *the extent to which online students of higher education experience certain positive emotions regarding their learning experiences in online courses*. This conceptualization of online emotional engagement is in line with the achievement emotions literature stating that "emotions are critically important for students' engagement with academic tasks" (Pekrun & Linnenbrink-Garcia, 2012, p. 278). As defined by Pekrun et al. (2011) in the control-value theory, achievement emotions refer to those emotions experienced by students in relation to their achievement activities or outcomes in academic achievement settings. According to the control-value theory, both activity-related emotions such as a student's *enjoyment* of a particular learning activity in progress and outcome-related emotions such as prospective outcome emotion of *hope* about possible success in the future and retrospective outcome emotion of *pride* about past success are all achievement emotions experienced in academic learning settings (Pekrun et al., 2011).

The specific achievement emotions used in the current study to conceptualize and operationalize online emotional engagement include *enjoyment*, *pride,* and *hope*. In this study, *enjoyment* refers to an online student's sense of joy and liking experienced in relation to engaging in a particular activity or behavior in an online course (e.g., enjoying reading peers' discussion posts, enjoying sharing perspectives during online discussions). *Pride* refers in this study to an online student's sense of feeling good, happy, and/or pleased about themselves in relation to a certain successful or pleasant outcome (e.g., feeling proud about getting positive feedback, feeling proud about being able to answer peers' questions during online discussions). *Hope* refers in this study to an online student's sense of being positive and remaining optimistic about their successful learning and good performance in the online course.

These three achievement emotions were incorporated into the OEHE subscale of emotional engagement on the empirical grounds that these positive-activating emotions are among the achievement emotions commonly experienced by university students in academic settings and that they represent the major achievement emotions as indicated within the control-value theory (Pekrun et al., 2011; Pekrun & Linnenbrink-Garcia, 2012). *Curiosity* is also used as an indicator of emotional engagement because curiosity is given as an epistemic emotion triggered by cognitive incongruities that are also commonly experienced by students in academic settings (Pekrun & Linnenbrink-Garcia, 2012). In line with its epistemic nature indicated by Pekrun and Linnenbrink-Garcia (2012) and in line with Litman and Spielberger's (2003) conceptualization of epistemic curiosity, *curiosity* in this study is defined as an epistemic emotion that refers to an online student's desire or wish to learn and know about what others think or believe about a course concept or issue, to learn about the course concepts, and their desire to explore, find out, and acquire more about the course concepts being taught in the online course. Negative emotions (e.g., anxiety, shame, anger, guilt) were not used in the current study as indicators of online emotional engagement due to their more complex nature and their ambiguous effects on students' engagement (Daniels & Stupnisky, 2012; Pekrun et al., 2009, 2011; Pekrun & Linnenbrink-Garcia, 2012). To address the issues of CI and CU, all the OEE items were intentionally kept within the relevant domain of online student emotions but nothing else and as many draft items as possible were created to adequately represent the relevant domain of online emotional engagement. While determining the indicators, Fredricks et al. (2005), Reeve and Tseng (2011), Reeve (2013), and Sun and Rueda (2012) were also consulted. Table 3.3 presents the construct definition and the corresponding indicators of online emotional engagement.

Table 3.3

Construct Definition and Indicators of Online Emotional Engagement

Subdimension	Construct Definition	Indicators
Online Emotional Engagement	The extent to which online students of higher education experience certain positive emotions regarding their learning experiences in online courses	• enjoyment • pride • hope • curiosity

Construct Definition and Operationalization of Online Cognitive Engagement

Based on a critical review of the literature of cognitive engagement (Fredricks et al., 2004; Fredricks & McColskey, 2012), online cognitive engagement in this study is defined as *the extent to which online students of higher education expend and regulate conscious mental effort to use cognitive and metacognitive strategies during online learning*. This conceptualization of online cognitive engagement as expending conscious mental effort to use cognitive and metacognitive strategies versus psychological investment (motivation) conceptualization rests on the argument that students' making conscious mental effort for the use of cognitive and metacognitive strategies is conceptually and theoretically closer to the action or experience-based nature of the construct of engagement (Fredricks & McColskey, 2012).

Next, based on this construct definition, in light of the relevant literature of cognitive engagement (Henrie et al., 2015) and of cognitive and metacognitive strategies (e.g., Greene & Miller, 1996; Wolters, 2004) as well as the relevant literature of cognitive engagement indicators in online learning environments such as reflection on the discussion board (e.g., Dennen, 2008), certain indicators intended to adequately and accurately represent the content domain or universe of online cognitive engagement were created. Specific indicators used to operationalize online cognitive engagement include but are not limited to making logical connections between ideas, creating one's examples, making conscious mental effort to better understand, elaborating on ideas, reflection, and thinking and asking about one's own learning, which are some common cognitive and metacognitive learning strategies used by students in academic settings (Greene & Miller, 1996; Wolters, 2004). To address the issues of CI and CU, all the OCE items were kept within the relevant domain of online students' use of cognitive and metacognitive learning strategies but nothing else and as many draft items as possible were created to adequately represent the relevant domain of online cognitive engagement. While determining these indicators, the indicators of cognitive engagement used by Fredricks et al. (2005), Miller et al. (1996), Reeve and Tseng (2011), Reeve (2013), and Sun and Rueda (2012) were also examined. Table 3.4 presents the construct definition and the corresponding indicators of online cognitive engagement.

Table 3.4

Construct Definition and Indicators of Online Cognitive Engagement

Subdimension	Construct Definition	Indicators (Examples)
Online Cognitive Engagement	The extent to which online students of higher education expend and regulate conscious mental effort to use cognitive and metacognitive strategies during online learning	• making logical connections • reflecting on ideas • creating examples • formulating questions • making inferences • relating prior knowledge to new knowledge • planning • monitoring • self-evaluating

Expert Review and Cognitive Pilot Testing

A shortened version of the initial pool was submitted to the dissertation committee members for their reviews of and selections from among the draft items. Based on the feedback, comments, and suggestions, several iterative item revisions, additions, and deletions were made. Next, a set of candidate items from this reviewed initial pool was selected, coming up with the first draft of the OEHE instrument for expert reviews. Invitation emails were sent to more than ten experts in the field of online education, instructional design and technology, and educational psychology. Expert reviews provide a good deal of useful information about the face validity, content validity, and language quality of the survey items and questions (DeVellis, 2012; Groves et al., 2009). The expert reviews were conducted in two rounds (6 experts in round 1 and 5 experts in round 2). The purpose of the first round was to focus on the qualitative comments and suggestions of the experts and the primary focus of the second round was to focus on experts' item ratings and orderings to land on the best items. A Likert-type frequency scale of 1=Never to 5=Always was adopted for the OEHE based on the experts' feedback and suggestions as well. As a result of the expert review process, several item revisions and modifications were applied and 18 OBE, 21 OEE, and 22 OCE items were selected as the best items to be submitted as the second survey draft to pilot cognitive testing (see Appendix A).

Cognitive interviewing as a pilot testing of newly developed scales is one of the most commonly utilized methods to reveal how actual respondents would understand and interpret the survey items and how they would decide to choose a particular response (Fowler, 2014; Ryan et al., 2012). Cognitive interviews are usually done with a very small number of participants, and there is no agreed sample size (Fowler, 2014). After getting the IRB approval of the study (see Appendix B), four cognitive interviews were conducted with four potential respondents: two undergraduate-level students and two graduate-level students who had taken and/or were taking at least one online course as part of their higher education degree or program at the time of the interview. During the cognitive interviews video-recorded after getting the signed consents approved by the IRB, the Qualtrics online survey form was shared with the interviewees (see Appendix C) and they synchronously went through the online survey instructions and items while they were asked to respond to five standardized questions adapted from Holmes (2018) about the OEHE survey instructions, candidate items, and their response choices (see Table 3.5 for cognitive interview questions). The cognitive interview data were open-coded, based on a *Success* and *Revision Needed* coding scheme. Table 3.5 presents the coding scheme adapted from Holmes (2018). In light of the cognitive interview data, some issues were identified, and certain decisions were made about the issues. Accordingly, multiple iterative minor and major revisions were applied to some of the OEHE items, and some new items were also added when deemed necessary to thoroughly respond to the issues (see Appendix D for a summary of item issues and responses). Appendix E presents the final list of items used for the large-group data collection.

Table 3.5

Coding Scheme for Open-Coding Cognitive Interview Data

Cognitive Interview Question (CIQ)	Success	Revision Needed
#1. Do you think the survey instructions are clear and easy to understand?	Instructions clearly and easily understood	Instructions not clearly and easily understood
#2. What do you think this item is asking you about? Please explain.	Item was understood in the intended way	Item was not understood in the intended way
#3. Why did you choose the response you chose over the other response options? Please explain.	Response options were understood and used in the intended way	Response options were not understood and used in the intended way
#4. Do you think the item needs to be re-worded for clarity? If yes, how would you re-word it?	Item does not need to be re-worded	Item needs re-wording
#5 How relevant do you think this item is to you as an online student? Please explain.	Relevant	Not Relevant

Participants and Sampling Design

Potential participants were recruited from a population of undergraduate and graduate students currently enrolled at a higher education institution in the United States and currently taking at least one online course during an academic semester at the time of data collection. Snowball convenience sampling was used as the main sampling strategy to recruit potential participants through key individuals and organizations as gatekeepers (Creswell, 2014; Gall et al., 2007) (e.g., online program coordinators, online course instructors) who have access to potential participants (i.e., undergraduate and graduate students taking online courses). Within this survey design, purposeful criterion-based sampling (Creswell, 2014) was also used by only recruiting those individuals: (a) who were currently enrolled as undergraduate or graduate students at a higher education institution in the United States and (b) who were currently taking at least one online course during Fall 2021 academic semester. Due to the nature of snowball convenience sampling, the research invitation emails, and the survey link could have been shared with an unknown number of students, rendering the computation of a response rate impossible. After cleaning the raw data of 813 recorded responses on Qualtrics, screening the data for minimum sampling criteria and missing data (see Appendix F), and removing outliers, a final sample of 235 participants provided complete and minimum-criteria-fulfilling OEHE data used for the confirmatory factor analyses and reliability analyses of the OEHE items, and 234 participants provided data for other validity analyses in this study. This sample size was deemed adequate for the CFA analyses because a sample size of $200 \leq \mathit{235} \leq 400$ would help us ensure that the sample mean can differentiate among the online undergraduate and graduate students at a 95% confidence interval (Asempapa, 2016; Creswell, 2005). The specific context for this study was online courses which contained undergraduate and graduate students as the participants drawn from the target population through snowball convenience sampling.

Participants' Demographic Characteristics. The participants (*N*=235) were all university students currently enrolled at a higher education institution at the time of data collection in the United States (inclusion criterion 1). Most of them were female (80.8%) and White (73.1%). Seventy-nine (33.6%) were undergraduate and 66.4% were graduate-level students (*n*=156) currently enrolled in various degree programs ranging from education to biology at 51 different higher education institutions in the United States, 86% (*n*=44) of which were public and 14% (*n*=7) were private institutions. The snowball sampling strategy made it unfeasible to describe all those specific courses, academic majors, and universities indicated by the participants. All of the participants had taken at least one online course before, with 57% (*n*=134) having taken more than 5 online courses. All the participants were currently taking, at the time of data collection, at least one online course as part of their higher education degree or program (inclusion criterion 2). Appendix G presents detailed information about participants' characteristics.

Participants' Online Learning Characteristics. The participants indicated one specific online course as the online course to think of while responding to the survey items and there were a wide range of online courses reported by the participants, ranging from adult education to human sexuality. The convenience sample recruited through snowball sampling made it impossible to describe any course specifics including the instructor and program characteristics. The majority of the participants (70.2 %) reported learning fully asynchronous and mostly asynchronous online (70.2%). All participants (*N*=235) reported that in their online course there were regular (e.g., weekly, biweekly) online discussions happening on the discussion board/in the discussion forum (inclusion criterion 3).

Participants' Inclusiveness, Bias, and Accessibility Characteristics. A total of 234 participants responded to the inclusiveness, bias, and accessibility questions. Of the 234 participants, 36.8% (*n*=86) reported that they were a first-generation college student in their family. Appendix H presents a summary of the participants' perceptions regarding inclusiveness and bias in their online courses. A large majority of the participants (88.5 %, *n*=207) reported that they had not experienced any accessibility issues while taking online courses, while 27 participants (11.5%) reported that they had experienced some accessibility issues (see Appendix H).

Procedure

After the IRB approval of the study, through online program coordinators and department heads, online course instructors teaching online within their programs/departments during Fall 2021 semester at my institution were accessed via an IRB-approved invitation email. Further communication with the instructors wishing to encourage their students to participate in this study was conducted about the details and then they started sharing the survey with their students. The university's mass email system was also used to share the online survey with all undergraduate and graduate students at the institution. Online program coordinators and course instructors at other U.S. institutions both in Oklahoma and other U.S. states were also invited and sent the recruitment materials as well. In doing so, the aim was to obtain a multi-course, multi-disciplinary, and multi-institutional sample of online undergraduate and graduate students. The online survey was administered around the middle of Fall 2021 academic semester so that the participants had spent a considerable amount of time in their online courses and developed their online course practices and experiences.

The participants first responded to a set of online course context (OCC) questions, next the OEHE items developed in this study. Then, they took another measure of student engagement, a measure of epistemic beliefs, and a measure of task value perceptions. The participants were kindly asked to think of a specific online course they were currently taking at the time of data collection while responding to all these measures in order to help them contextualize their ratings and responses. The participants finally responded to a set of demographic questions as well as a couple of inclusiveness and accessibility questions.

Measures

The Online Engagement in Higher Education (OEHE) Scale

The sixty-six OEHE items (see Appendix E) were rated by the respondents on a 5-point Likert-type frequency scale of 1=Never, 2=Seldom, 3=Sometimes, 4=Often, and 5=Always. The survey was delivered via Qualtrics, the institution-hosted online questionnaire system. There were 20 items for online behavioral engagement (OBE), 27 items for online emotional engagement (OEE), and 19 items for online cognitive engagement (OCE). Items within each subscale asked the participants about the extent to which they experienced and exhibited certain indicators of behavioral, emotional, and cognitive engagement in their online courses. Higher scores on each subscale of the OEHE scale represent a higher level of engagement on that particular dimension.

Classroom Engagement Scale

For convergent validity evidence, a modified version of the classroom engagement scale developed by Reeve (2013) designed to assess college students' experiences of engagement in traditional college courses was used. Convergent validity can be assessed by examining the relationship between the scores on a new measure of a particular construct or attribute and the scores on an existing and validated measure of the same construct or attribute, looking for reasonably high correlations between the two (Campbell & Fiske, 1959; Chin & Yao, 2014; Warner, 2008).

For the purpose of this study, only behavioral, emotional, and cognitive engagement subscales in Reeve's (2013) Study 2 were included. There was a total of 12 items rated on a 5-point Likert-type scale (1=Strongly Disagree, 2 = Disagree, 3=Neither Agree nor Disagree, 4=Agree, and 5=Strongly Agree), with four items for behavioral engagement ($\alpha = .87$), four items for emotional engagement ($\alpha = .91$) and four items for cognitive engagement ($\alpha = .87$), all indicating acceptable levels of internal consistency reliability. The construct validity and factorial validity evidence for the overall engagement scale was documented by Reeve (2013). Each of the items on each subscale of engagement was slightly adapted to fit each item to the online learning context of the current study (for the full scale see Appendix I).

Reeve's (2013) three dimensions of student engagement were used to correlate each of them with each corresponding dimension of the OEHE instrument for convergent validity purposes because similar to the conceptualization of engagement in the OEHE instrument, Reeve's (2013) engagement scale was also based on Fredricks et al.'s. (2004) three core dimensions of engagement (along with agentic engagement), which made it a theoretically appropriate measure to use for convergent validity testing in the current study because measures of the same construct (i.e., engagement) are expected to converge (Campbell & Fiske, 1959; Furr, 2011).

Epistemic Beliefs Inventory

For discriminant validity evidence, a seven-item subset of the 32-item version of the Epistemic Beliefs Inventory (EBI) was used (DeBacker et al., 2008; Schraw et al., 2002). Discriminant validity can be assessed by examining the relationship between the scores on the new measure and the scores on an existing measure of those constructs or attributes that the new instrument is not supposed to measure, looking for very weak or no correlations at all between them (Warner, 2008). Epistemic beliefs basically refer to individuals' beliefs about what knowledge is, what it is made up of, and how certain knowledge is and beliefs about how we come to know something (DeBacker et al., 2008). The epistemic beliefs measure was used to correlate it with the OEHE instrument for discriminant validity purposes because the underlying construct that the epistemic beliefs items are intended to tap (e.g., fixed ability mindset) is a different enough construct that should theoretically not be related to the construct of student engagement. Due to the theoretical or conceptual dissimilarity, weak or negligible correlations were expected between the OEHE instrument and students' epistemic beliefs about ability. The subset of fixed ability mindset was also chosen for the current study because the Fixed (Innate Ability) subscale was shown to have the most desirable psychometric properties when compared to the other subscales of the EBI (DeBacker et al., 2008). The reliabilities for the Fixed subscale of the EBI were reported by DeBacker et al. (2008) to be .67 (Sample 1) and .68 (Sample 2) (for items see Appendix J) Participants ranked all the items on a 5-point Likert-type scale from 1=Strongly Disagree to 5=Strongly Agree. Higher scores on this scale represent a higher level of epistemic beliefs about fixed ability.

Interpreting Convergent and Discriminant Validity Evidence. When interpreting convergent and discriminant validity evidence, researchers primarily evaluate the magnitudes of the correlation coefficients to see whether the magnitudes are strong enough to obtain sufficient convergent validity evidence and whether the magnitudes are weak enough to obtain sufficient discriminant validity evidence (Furr, 2011). Researchers also evaluate the statistical significance of the validity coefficients while still considering the potential impact of our sample size on the significance of the validity coefficients (Furr, 2011). Researchers usually look for the statistical significance of moderate to strong validity coefficients when interpreting convergent validity evidence, and they usually look for non-significance of weaker validity coefficients while interpreting discriminant validity evidence, although it should be noted that statistical significance is very sensitive to large sample sizes (Furr, 2011).

In this study, discriminant validity correlation coefficients were expected to be considerably lower in strength than those of convergent validity (Hubley, 2014). Discriminant validity coefficients were interpreted in comparison to convergent validity coefficients, which is the recommended approach to the interpretation of convergent and discriminant validity evidence (Hubley, 2014). Translating this principle into the context of this study, relatively higher and statistically significant correlations were expected between the subscales of the OEHE instrument and those of Reeve's (2013) engagement scale as two measures of the same construct of engagement for convergent validity evidence. Relatively lower or negligible and non-significant correlations were expected between the subscales of the OEHE instrument and the epistemic beliefs subscale of fixed ability as two measures of different constructs for discriminant validity evidence.

Task Value

For criterion-related validity evidence, a modified version of the six-item task value subscale of the 81-item Motivated Strategies for Learning Questionnaire Survey (MSLQ) designed to assess college students' motivational orientations and learning strategies (Pintrich et al., 1991, 1993) was used. In order for a new scale to have criterion-related validity, researchers look for some empirical relationship with a measure of a criterion variable (DeVellis, 2017). In other words, for criterion-related validity evidence, as DeVellis (2017) points out, "an item or scale is required only to have an empirical association with some criterion" (p.92). In addition, criterion-related validity as a validity term is preferable since it does not suggest any time relationship between two measures (DeVellis, 2017).

The MSLQ is one of the popular measures consistently used in educational psychology research for decades (Duncan & McKeachie, 2005; Hilpert et al., 2013; Jackson, 2018). The items on the task value subscale are intended to assess how important, interesting, valuable, and useful a student finds a particular course in terms of the course material, content area, and subject matter (Pintrich et al., 1991, 1993). The MSLQ uses a seven-point Likert scale from 1=Not at all true of me to 7= Very true of me. Reliability and validity evidence of the MSLQ instrument was documented by prior research (Credé & Phillips, 2011; Pintrich et al., 1993). Pintrich et al. (1993) originally reported the Cronbach's alpha of .90 for the subscale of six task value items. Each of the six items on this subscale was slightly modified to fit the items to the online learning context (see Appendix K).

Task value is one of those underlying motivational factors influencing student engagement in the academic learning process (Olivier et al., 2020). It basically refers to students' perceptions regarding how interesting, enjoyable, important, and useful they find a particular academic task or the course content as well as the amount of time and effort needed to master (i.e., cost) (Wigfield & Eccles, 2000). Such theoretical relationships between motivation constructs, such as task value, and engagement are also pointed out in the relevant engagement literature (Finn & Zimmer, 2012). Prior research has also empirically supported these theoretical relationships between the two constructs by reporting significant positive correlations between task value and student engagement (e.g., Olivier et al., 2020). Moderate, significant, and positive correlations were expected between the OEHE scores and task value scores as evidence of criterion-related validity. Criterion validity was assessed in this study in addition to convergent and discriminant validity because a new educational measure would be of little use and practical value if it were not related to any other important educational or motivational criteria (e.g., task value) and if it did not provide any further insights into informed decisions about individuals being measured (Messick, 1989, 1995a, 1995b).

Online Course Context, Demographic, and Equity/Inclusiveness Questions

In the first section of the online survey, participants responded to a couple of context questions about their specific online learning environments and experiences. In the last section of the online survey, participants also responded to a couple of demographic and educational background questions. They also responded to a couple of inclusiveness, bias, and accessibility questions regarding their online learning experiences. The data regarding inclusiveness, bias, and accessibility perceptions were collected to evaluate the extent to which the sample of the study included and was representative of individuals from diverse populations and their social contexts (Cintron & Hagan, 2021).

Data Analysis

After the large-group online survey was closed, the raw data set was downloaded from Qualtrics as an SPSS file and started cleaning the raw data set using IBM SPSS Statistics 27.

Data Cleaning, Minimum Criteria Screening, and Missing Data

The raw data set before cleaning had a total of 813 recorded responses on Qualtrics. As a result of data cleaning and minimum-criteria screening, 364 minimum-criteria fulfilling cases remained (see Appendix F). These 364 cases were screened for their missing values and missing data patterns on the OEHE items, the measure under investigation in this study. Thirty respondents having missing data on the OEHE items were listwise deleted as well, leading to 334 minimum-criteria-fulfilling cases with complete data on the OEHE items. The Little's MCAR test was conducted on the sixty-six OEHE indicator variables to identify the pattern of missingness and the Little's test was found to be non-significant, Little's MCAR Test Chi-square = 69.294, df = 67, p = .400, indicating that the missing data on the OEHE variables were missing completely at random (MCAR), which further makes the use of listwise deletion a reasonable choice (Schafer & Graham, 2002).

Preliminary Analyses of Data

After reverse-coding the three negatively worded OEHE items OBE18, OEE21, and OCE16, all of the OEHE variables were screened for normality and potential outlying cases. The 66 OEHE items were screening for both univariate and multivariate outliers (Tabachnick & Fidell, 2013), using the conventional criterion of the standardized z-scores of each case on each observed variable greater than the 3.29 cut-off in absolute value (Tabachnick & Fidell, 2013). The extreme values table and box plot of each indicator variable were also visually screened for potential outliers. The OEHE items were also screened for any potential multivariate outliers using the Mahalanobis distances (p < .001). The very conservative probability value of p < .001 to identify a case as a multivariate outlier is considered appropriate to be used with the Mahalanobis distance (Tabachnick & Fidell, 2013). As a result of this outlier screening process, 87 seriously outlying cases were removed, leading to a data set of 247 cases. Except for six online behavioral engagement (OBE) items, which had extreme skewness and kurtosis values, the skewness and kurtosis values of most of the other OEHE indicator variables were found to fall in between -2 and +2 interval (Lomax & Hahs-Vaughn, 2012). The distribution of most of the variables visually inspected via their histogram was also found to be approximately normally distributed.

Item Screening and Evaluation Before CFA. As a result of individual item evaluations (e.g., interitem correlations, internal consistency reliability statistics), 31 best items (9 OBE, 12 OEE, 10 OCE) were selected to be used in testing the hypothesized three-factor model in the context of CFA. The purpose of this individual item screening and evaluation stage was to obtain subscales involving a set of items that were positively and strongly intercorrelated so that desirable model fit indices could be obtained during the CFA (Russell, 2002). It should be noted that screening the OEHE items to remove potentially problematic items at this preliminary analysis stage is important to obtain better model fit indices during the CFA analyses because the number of variables under analysis has been shown mathematically to negatively influence several model fit indices in the context of SEM (Kenny & McCoach, 2003). Linearity and multicollinearity assumptions were also reasonably met by the data.

Confirmatory Factor Analysis (CFA)

The sample size of the initial CFA data set after the removal of the outliers was 247, which was considered to be reasonably adequate in light of all recommendations of sample size for CFA procedures (Russell, 2002; Worthington & Whittaker, 2006). Confirmatory factor analysis (CFA) as a special structural equation modeling (SEM) technique and a member of the SEM family is used to test theory-based hypotheses postulated a priori regarding the number of factors and the loadings of certain items or indicator variables on certain underlying factors or latent constructs (Jackson et al., 2009; Kahn, 2006; Russell, 2002). The rationale for using CFA without EFA is that the OEHE items were already based on an a priori theoretical conceptualization of student engagement with a priori determined three dimensions, namely behavioral, emotional, and cognitive (Fredricks et al., 2004) and the aim of this study was to test this a priori conceptualization and the corresponding items. Researchers can appropriately use CFA when they already have a strong theoretical and/or empirical knowledge base regarding the underlying latent variable or factor structure (Byrne, 2005; DiStefano & Hess, 2005; Kahn, 2006).

The initial CFA model of the OEHE structure hypothesized that (a) the sample participants' responses to the OEHE items could be explained by three underlying factors (i.e., OBE, OEE, OCE), (b) each OEHE item would load strongly on its corresponding factor for which it was intended and would not load on the other two factors, (c) the three factors were correlated with each other although distinct factors, and (d) the error terms associated with the OEHE items were uncorrelated (Byrne, 2005, 2016). In line with the purpose of CFA, it was specified which items would load on which dimension of online student engagement (see Appendix E)

The Maximum Likelihood (ML) estimation approach was used since it is one of the most commonly utilized estimation approaches in CFA (Byrne, 2005; Jackson et al., 2009; Kahn, 2006; Kline, 2016; Russell, 2002; Schermelleh-Engel et al., 2003; Tabachnick & Fidell, 2013). ML estimation is an effective estimation method that can provide accurate results as long as the observed variables are multivariate normally distributed, models are specified correctly, and the sample size is large enough (Schermelleh-Engel et al., 2003). Its main limitation or challenge is the fundamental assumption of multivariate normality. The good news, however, is that "ML seems to be quite robust against the violation of the normality assumption" (Schermelleh-Engel et al., 2003, p. 26). Bootstrapping as a resampling strategy is also offered as a strategy to deal with multivariate nonnormal data when using ML estimation method in the context of SEM (Byrne, 2010, 2016).

In the evaluation of the initial CFA model with 31 variables (N=247), indices of univariate skewness and univariate kurtosis were used to screen for univariate normality first as a prerequisite for multivariate normality (Byrne, 2016). Given that CFA procedures in the context of SEM are based on covariance matrices, kurtosis, and more specifically multivariate kurtosis which can be quite detrimental to the robustness of the CFA findings, is particularly more important in CFA analyses (Byrne, 2010). Multivariate kurtosis basically refers to "the situation where the multivariate distribution of the observed variables has both tails and peaks that differ from those characteristic of a multivariate normal distribution" (Byrne, 2010, p. 103). There is no clear agreement on the size of univariate kurtosis indices to make conclusions of extreme kurtosis about the data (Byrne, 2010; Kline, 2016). Given that univariate normality is not sufficient for multivariate normality and that the univariate normal distribution of the observed variables does not necessarily guarantee the multivariate normal distribution of the variables, multivariate normality assumption was assessed by inspecting the multivariate kurtosis value and most importantly its critical ratio value, "which in essence represents Mardia's (1970, 1974) normalized estimate of multivariate kurtosis, although it is not explicitly labeled as such" (Byrne, 2010, p. 104). According to Bentler (2005), as cited in Byrne (2010), values in excess of 5.00 indicate nonnormally distributed data. In order to identify potential multivariate outliers that contribute to multivariate non-normality in the data (Kline, 2016), the squared Mahalanobis distance (D^2), which basically indicates "the distance in standard deviation units between a set of scores for one case and the sample means for all variables (centroids)" (Byrne, 2010, p. 106) was used. A multivariate outlying case typically has a D^2 value that is quite different from all the other D^2 values in the data set (Byrne, 2016). Using a conservative significance criterion of p=.001 for detecting potential multivariate outliers (Kline, 2016; Tabachnick & Fidell, 2013) and checking the p1 column in the Amos output, serious multivariate outliers were checked for in the dataset (Byrne, 2010; Kline, 2016; Pituch & Stevens 2016; Tabachnick & Fidell 2013). Based on the Amos output of observations farthest from the centroid (Mahalanobis distance), twelve potential multivariate outlier cases were further removed from the CFA dataset, resulting in the CFA dataset of 235 cases used for the subsequent model tests. Bootstrapping procedure was used to address multivariate nonnormality in the data since it "allows the researcher to assess the stability of parameter estimates and thereby report their values with a greater degree of accuracy" (Byrne, 2010, p. 332), although it still has its own limitations (see Byrne, 2010).

In SPSS Amos 27 Graphics, a first-order CFA model was specified to test and confirm the hypothesized three-factor model of online student engagement and the data were entered in this model. SEM and CFA researchers are strongly recommended to utilize multiple model fit indices or criteria while assessing and evaluating the fit of their hypothesized models so that their models can be assessed from a wide range of perspectives and angles (Byrne, 2016). While evaluating model fit, both the goodness of fit of the individual parameter estimates (local fit) and the goodness of fit of the hypothesized model as a whole (global fit) are evaluated (Byrne, 2005, 2016). Several goodness-of-fit statistics were used to evaluate the fit of the CFA models in this study (e.g., chi-square, GFI, TLI, CFI, SRMR, RMSEA). Despite the use of these fit indices to evaluate the overall model fit, all those model fit indices alone cannot guarantee that a particular model, even if it is statistically found to be a good fitting model using the model fit indices, is theoretically plausible and practically useful (Byrne, 2016). It is still incumbent on the researcher to judge the extent to which the model is not only statistically adequate but also theoretically plausible and practically useful (Byrne, 2016). When the SEM/CFA researcher finds a poor fitting model, they may continue their analyses in an exploratory fashion because they start looking for those mis-specified parameters leading to the poor fit of the model (Byrne, 2005). To improve model fit, standardized residual covariances and modification indices are also recommended to be considered for model modifications (Byrne, 2016). However, as noted by Byrne (2005, 2016), while deciding whether to add new parameters into the model based on modification indices, the new parameters suggested by the MIs should make both logical and theoretical sense.

Validity and Reliability Coefficients

In addition to the factorial validity evidence provided by the CFA analyses, Pearson correlations were used to test the OEHE items for convergent, discriminant, and criterion-related validity. Moderate to strong, positive and significant correlations were sought after for convergent validity evidence, weak to non-existent and non-significant correlations were expected for discriminant validity evidence, and moderate to strong, positive, and significant correlations were desired for criterion-related validity evidence. For reliability evidence, Cronbach’s alphas were used to test the OEHE items for internal consistency reliability.

Chapter 4: Findings

The purpose of this instrument development study was to develop and validate a new self-report survey scale of student engagement in online courses of higher education titled *Online Engagement in Higher Education* (OEHE) by analyzing and evaluating its reliability and validity. This chapter presents the empirical findings regarding confirmatory factor analysis, convergent, discriminant, and criterion validity, and internal consistency reliability of the OEHE instrument. For factorial validation, a series of confirmatory factor analyses (CFAs) was carried out to test and confirm the a priori hypothesized measurement model of the OEHE. Pearson correlations between the OEHE subscales and other validation measures were run to seek empirical evidence for convergent, discriminant, and criterion validity as other forms of construct validation. Finally, Cronbach's alpha statistics were obtained to assess and evaluate the internal consistency reliability of the OEHE instrument. The findings indicated that the hypothesized three-factor structure of engagement based on Fredricks et al.'s (2004) three core dimensions of student engagement was adequately confirmed in the context of online learning. Pearson correlations also provided reasonable evidence for convergent and criterion-related validity and strong evidence for discriminant validity. Finally, the Cronbach's alpha findings provided adequate to very good internal consistency reliability for the subscales of the OEHE and for the OEHE instrument as a whole.

Confirmatory Factor Analysis (CFA)

The first research question *RQ1: Can a three-factor hypothesized model of online student engagement based on theory and prior research be confirmed in a validation sample?* was answered satisfactorily as the hypothesized three-factor structure of student engagement in online learning environments was adequately confirmed through a series of CFA. The measurement model to be tested in this study postulated a priori that online student engagement is a three-factor structure involving online behavioral engagement (OBE), online emotional engagement (OEE), and online cognitive engagement (OCE). In the initial CFA model drawn up in the Amos graphics (see Appendix L) after the cleaning of the poor items at the initial item screening process, there was a total of 31 observed indicator variables, with each representing a particular facet of online student engagement only. In this CFA model, nine items were hypothesized to load only on the OBE factor; twelve items were hypothesized to load only on the OEE factor, and ten items were hypothesized to load only on the OCE factor. In this hypothesized model to be tested through CFA procedures, OBE, OEE, and OCE as three distinct factors of online student engagement were assumed, based on theory and research, to be intercorrelated. On the other hand, error terms associated with each indicator variable were assumed to be uncorrelated in this initially specified model. This specified CFA model hypothesized a priori that the responses to the OEHE instrument could be explained by three factors, namely OBE, OEE, and OCE; that each indicator variable would have a nonzero loading on the target factor that it was theoretically expected to load on and would have a zero loading on all the other factors; the three factors were correlated with one another, as consistent with theory and prior research (Fredricks et al., 2004; Kucuk & Richardson, 2019); and error variances associated with each indicator variable were uncorrelated with one another (Byrne, 2016).

Next, the CFA data set (N = 247) was entered into this initial model with 31 items and maximum likelihood (ML) estimation method was used as the default option in Amos. Assessment of normality was also requested since ML estimation assumes multivariate normality and this assumption has to be analyzed before undertaking any further CFA analyses or making any interpretations (Byrne, 2016). When the initial specified CFA model for the hypothesized three-factor structure of online student engagement was estimated, it produced a chi-square value of 1309.613 with 431 degrees of freedom and the probability level of .000 (p <.001), indicating a poor model fit. In terms of assessment of normality, multivariate normality is the major assumption to be met in SEM-based CFA applications using maximum likelihood (ML) as the estimation method (Arbuckle, 2020; Byrne, 2010, 2016). In addition to the preliminary analyses where the data were screened for univariate normality through an inspection of skewness and kurtosis indices of the indicator variables and their box plots and histograms as well as eliminating univariate and multivariate outliers, the indices of univariate skewness and univariate kurtosis were screened in the Amos output as well to examine the CFA data for univariate normality first as a prerequisite for multivariate normality before evaluating and interpreting any model fit parameters or indices (Byrne, 2016). According to Kline (2016), skewness values greater than 3 indicate more severely skewed data, while skewness values falling in between -2 and +2 are considered to be reasonably consistent with univariate normality (Lomax & Hahs-Vaughn, 2012). Taking both rules of thumb into consideration, none of the OEHE items were found to have a skew index in excess of 2 in absolute value, with all the 31 items having skewness values between -2 and +2, indicating reasonable consistency with univariate normality and no severe skewness, although the twenty-nine indicator variables were negatively skewed with negative skew values ranging from -.241 to -1.651 and the two indicator variables were positively skewed with values of .164 and .229.

Given that CFA procedures in the context of SEM are based on covariance matrices, kurtosis, and more specifically multivariate kurtosis which can be quite detrimental to the robustness of the CFA findings, is particularly more important in CFA analyses (Byrne, 2010. 2016). First, the univariate kurtosis values given for each of the OEHE items were screened in the assessment of normality output produced by Amos. There is no clear agreement on the size of univariate kurtosis indices to make conclusions of extreme kurtosis about the data (Byrne, 2010, 2016; Kline, 2016). Lomax and Hahs-Vaughn (2012) suggest that univariate kurtosis values between -2 and +2 are consistent with univariate normality, although Byrne (2010) recommends interpreting the univariate kurtosis values "equal to or greater than 7 to be indicative of early departure from normality" (p. 103). None of the OEHE items were found to be severely kurtotic, with the positive kurtosis values ranging from .009 to 2.143 and the negative kurtosis values ranging from -.048 to -1.191. Overall, the univariate skewness and kurtosis values did not indicate any major issues with univariate non-normality on the thirty-one OEHE variables. The index of multivariate kurtosis was found to be 144.009, which clearly indicated departure from multivariate normality. The critical ratio was z=25.018, which also indicated that the multivariate kurtosis significantly departed from multivariate normal distribution.

In terms of the detection of multivariate outliers, the squared Mahalanobis distance (D^2) indices were inspected for the cases in the AMOS output. Using a conservative significance criterion of p=.001 for detecting potential multivariate outliers (Kline, 2016; Tabachnick & Fidell, 2013) and checking the p1 column in the Amos output, serious multivariate outliers were screened in the dataset (Byrne, 2010; Kline, 2016; Pituch & Stevens 2016; Tabachnick & Fidell 2013). Accordingly, a total of 12 potential multivariate outliers was removed, resulting in the final CFA dataset of 235 cases. The multivariate kurtosis dropped from 144.009 to 98.299 and the critical ratio value dropped from 25.018 to 16.657, indicating a reasonable departure from multivariate normality given that maximum likelihood estimation is known to be quite robust against the violation of the normality assumption (Schermelleh-Engel et al., 2003) and that the

ML estimation method may still perform well when the data mildly depart from multivariate normality (Jackson et al., 2009). To address multivariate nonnormality, bootstrapping procedure was used to obtain bias-corrected confidence intervals around the estimates. As it is given as the most reasonable procedure to use in the face of multivariate nonnormal data while using the Amos program (Byrne, 2010, 2016), the Bootstrap ML and Bollen-Stine bootstrap procedures were requested, asking the Amos program to perform bootstrapping based on 2000 samples with bias-corrected 95 % confidence intervals. After performing bootstrapping, minimum was achieved with the chi-square value of 1236.138 and 431 degrees of freedom with the probability level of .000. When the Bollen-Stine bootstrap output was examined, the model fit better in 2000 bootstrap samples than it did in the original sample. Testing the null hypothesis that the model is correct, the bootstrap results indicated a poor fit to the data. In addition, the global model fit indices also indicated a poor fit of the initial hypothesized model to the data (e.g., GFI=.723, , CFI=.803, TLI=.788, RMSEA=.089). It was concluded that the initial model with 31 items was far from indicating at least an acceptable fitting model. All the factor loadings were statistically significant. However, in terms of the magnitude of the factor loadings, the standardized regression weights were examined, and accordingly five online behavioral engagement (OBE) items were removed due to their relatively weaker factor loadings on the OBE Factor (below .50) damaging the convergent validity of the OBE factor (Caskurlu, 2018). After the removal of the five OBE items (OBE8, OBE15, OBE5, OBE10, OBE20), all the items in the model had a standardized factor loading greater than .50, indicating strong factor loading and high convergent validity of each factor (Caskurlu, 2018). Table 4.1 presents the normal-theory ML unstandardized factor loadings, standard errors, critical ratios, statistical significance, and standardized factor loadings of the 26 indicator variables in the second model, Model 2 (see

Appendix M) . Table 4.2 presents the unstandardized factor loadings with bias-corrected 95% confidence intervals and Table 4.3 presents the standardized factor loadings with bias-corrected 95% confidence intervals. All the standardized factor loadings were strong (>.50) and statistically significant (p < .01), thereby indicating a reasonably good local model fit. However, the global model fit indices were still far from indicating at least an acceptable model fit (e.g., CMIN=918.990, DF=296, CMIN/DF=3.105, P=.000; CFI=.835, TLI=.818, IFI=.836; RMSEA=.095).

There were positive and moderate to strong correlations among the three factors of online student engagement (see Appendix M), as theoretically expected. The correlation between OBE Factor and OCE Factor was .406 (p<.001), the correlation between OBE Factor and OEE Factor was .659 (p<.001) and the correlation between OEE Factor and OCE Factor was .483 (p=.001). All these moderate to strong factor correlations between the three engagement dimensions were consistent with the student engagement theory (Fredricks et al., 2004, 2005) and prior research of online engagement (e.g., Kucuk & Richardson, 2019; Park & Yun, 2019; Sun & Rueda, 2012).

Post Hoc Model Modifications. Due to the still inadequate global fit of the hypothesized model to the sample data, it was reasonable to move into an exploratory mode and try to modify the hypothesized model in a logical and meaningful manner (Byrne, 2016). It should be noted that CFA as a factor analysis technique is "not strictly confirmatory" (Kline, 2016, p. 197) because it is not uncommon for many initially hypothesized CFA models to not fit the data, which leads the CFA researcher to specify a revised model and tests the revised model again with the same data (Kline, 2016).

Table 4.1

ML Unstandardized and Standardized Factor Loadings of the 26 Indicator Variables (N=235)

Item	Factor	B	S.E.	C.R.	*P*	*β*
OBE3	OBE	1.000				.619
OBE4	OBE	1.394	.221	6.299	<.001	.538
OBE6	OBE	1.685	.232	7.262	<.001	.668
OBE7	OBE	1.224	.175	6.995	<.001	.626
OEE1	OEE	1.000				.815
OEE2	OEE	1.005	.068	14.713	<.001	.816
OEE3	OEE	.874	.071	12.281	<.001	.716
OEE5	OEE	.994	.064	15.451	<.001	.843
OEE6	OEE	.915	.064	14.230	<.001	.797
OEE7	OEE	1.077	.073	14.789	<.001	.818
OEE8	OEE	.934	.061	15.236	<.001	.835
OEE9	OEE	.740	.063	11.790	<.001	.694
OEE14	OEE	.705	.068	10.379	<.001	.628
OEE15	OEE	.738	.070	10.545	<.001	.636
OEE18	OEE	.917	.078	11.704	<.001	.690
OEE19	OEE	.706	.066	10.736	<.001	.645
OCE6	OCE	1.000				.652
OCE7	OCE	.965	.096	10.054	<.001	.755
OCE8	OCE	.880	.096	9.150	<.001	.675
OCE10	OCE	1.294	.128	10.075	<.001	.757
OCE13	OCE	1.126	.110	10.192	<.001	.768
OCE14	OCE	.836	.086	9.757	<.001	.728
OCE15	OCE	.880	.101	8.667	<.001	.633
OCE17	OCE	1.023	.094	10.932	<.001	.840
OCE18	OCE	.956	.108	8.842	<.001	.648
OCE19	OCE	.941	.101	9.290	<.001	.687

Note: B=Unstandardized, S.E.=Standard Error, C.R.=Critical Ratio, β=Standardized

Table 4.2

Unstandardized Factor Loadings with Bias-Corrected 95 % Confidence Intervals (N=235)

Item	Factor	Estimate	Lower	Upper	*P*
OBE3	OBE	1.000	1.000	1.000	…
OBE4	OBE	1.394	.677	4.094	.001
OBE6	OBE	1.685	.931	4.807	.001
OBE7	OBE	1.224	.950	1.606	.003
OEE1	OEE	1.000	1.000	1.000	…
OEE2	OEE	1.005	.894	1.142	.001
OEE3	OEE	.874	.769	1.001	.001
OEE5	OEE	.994	.888	1.109	.001
OEE6	OEE	.915	.790	1.059	.001
OEE7	OEE	1.077	.969	1.226	.001
OEE8	OEE	.934	.836	1.029	.001
OEE9	OEE	.740	.622	.869	.001
OEE14	OEE	.705	.546	.874	.001
OEE15	OEE	.738	.581	.915	.001
OEE18	OEE	.917	.765	1.071	.001
OEE19	OEE	.706	.564	.868	.001
OCE6	OCE	1.000	1.000	1.000	…
OCE7	OCE	.965	.825	1.158	.001
OCE8	OCE	.880	.715	1.090	.001
OCE10	OCE	1.294	1.052	1.605	.001
OCE13	OCE	1.126	.906	1.412	.001
OCE14	OCE	.836	.676	1.044	.001
OCE15	OCE	.880	.705	1.099	.001
OCE17	OCE	1.023	.855	1.236	.001
OCE18	OCE	.956	.745	1.225	.001
OCE19	OCE	.941	.736	1.196	.001

Table 4.3

Standardized Factor Loadings with Bias-Corrected 95 % Confidence Intervals (N=235)

Item	Factor	Estimate	Lower	Upper	*P*
OBE3	OBE	.619	.289	.773	.001
OBE4	OBE	.538	.293	.765	.001
OBE6	OBE	.668	.440	.906	.002
OBE7	OBE	.626	.368	.792	.001
OEE1	OEE	.815	.750	.864	.002
OEE2	OEE	.816	.753	.867	.001
OEE3	OEE	.716	.630	.792	.001
OEE5	OEE	.843	.780	.886	.002
OEE6	OEE	.797	.731	.846	.001
OEE7	OEE	.818	.751	.870	.002
OEE8	OEE	.835	.768	.881	.002
OEE9	OEE	.694	.614	.763	.001
OEE14	OEE	.628	.513	.716	.001
OEE15	OEE	.636	.525	.723	.001
OEE18	OEE	.690	.596	.771	.001
OEE19	OEE	.645	.545	.725	.001
OCE6	OCE	.652	.551	.741	.001
OCE7	OCE	.755	.675	.817	.001
OCE8	OCE	.675	.584	.751	.002
OCE10	OCE	.757	.686	.817	.001
OCE13	OCE	.768	.684	.835	.001
OCE14	OCE	.728	.635	.806	.001
OCE15	OCE	.633	.535	.728	.001
OCE17	OCE	.840	.781	.888	.001
OCE18	OCE	.648	.538	.740	.001
OCE19	OCE	.687	.581	.777	.001

For post hoc modifications, standardized residual covariances and modification indices were utilized to improve the model fit. Standardized residual values greater than 2.58 are considered to be large residuals and would indicate any statistically significant discrepancy associated with the covariance between any item pairs (Byrne, 2016). Based on the standardized residuals greater than 2.58, six items (OEE15, OEE14, OEE19, OEE9, OEE18, OCE10) were removed from Model 2 and the estimates were re-calculated without these items. After the removal of these items, the standardized residuals were re-checked and no more residuals were found to be greater than 2.58, except for the items OBE4 and OBE 6, which were retained in the model due to the already limited number of OBE items. The chi-square value of the model dropped from 918.990 with 296 degrees of freedom (26 items) to 453.599 with 167 degrees of freedom and the probability level of .000 (20 items). The other model fit indices were also improved although still not indicating an acceptable model fit (e.g., GFI=.836, CFI=.895, TLI=.880, RMSEA=.086). These results indicated that the global fit of the three-factor model with the remaining 20 items improved although it was still not indicating at least an acceptable model fit.

Next, Modification Indices (MIs) were examined to identify any evidence of model misspecification (Byrne, 2016). To do so, 10 was specified as the threshold for modification indices in order to include only MI estimates equal to or greater than 10 in the Amos output (Byrne, 2016). As noted by Schumacker and Lomax (2016), researchers usually select the largest MI value. CFA researchers can specify error correlations to test their hypotheses regarding the variation commonly shared by the indicators in addition to the common variance accounted for by the factors (Kline, 2016). Following Byrne's (2010) recommendation, each error covariance was specified and incorporated into the model one at a time and then the parameter estimates were estimated after each error covariance was specified in the model. Based on the MIs at and above 10 (Byrne, 2016) and using the theoretical and logical sense, a total of four error covariances was specified in the model. The chi-square of the model decreased to 279.133 with 163 degrees of freedom (p=.000) and the CMIN/DF was 1.712, which was less than 2 and might indicate a reasonably good model fit (Tabachnick & Fidell, 2013). The model fit indices were also improved (e.g., GFI=.894, IFI=.958, TLI=.950, CFI=.957, RMSEA=.055 with p-close=.214). The standardized RMR, known as SRMR, of this model was also computed to be .0535, indicating an acceptable model fit as well.

Global Fit Summary. After performing the Bollen-Stine bootstrapping, minimum was achieved, indicating no estimation problems, with the chi-square value of 279.133 and 163 degrees of freedom with the probability level of .000, indicating lack of exact model fit. When the Bollen-Stine bootstrap output of the final best-fitting model was examined as another tool for evaluating model fit generally used in cases where the data have non-normality, it was found that the model fit better in 1998 bootstrap samples, and it fit worse or failed to fit in 2 bootstrap samples. Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap was statistically significant, $p = .001$. However, on the basis of the other fit indices (see Table 4.4), the final model with 20 indicator variables, three intercorrelated factors, and four error covariances was accepted as the final best-fitting model adequately representing the sample data (Byrne, 2016). Greater weight was given to the other fit indices than the Bollen-Stine when making a final judgment about the model fit because the sole purpose of the Bollen-Stine to begin with was to provide an alternative measure that is more likely to indicate better fit than the other normal theory methods when non-normality is present. In addition, it should be noted that not all model fit tests and indices necessarily have to agree and ultimately it is a judgment call on the researcher’s part (Kline, 2016). Figure 4.1 shows the final best-fitting model with the twenty OEHE items and three underlying online engagement factors, as was hypothesized and confirmed in this study.

Figure 4.1

Final Best-Fitting Model with 20 Indicator Variables (N=235)

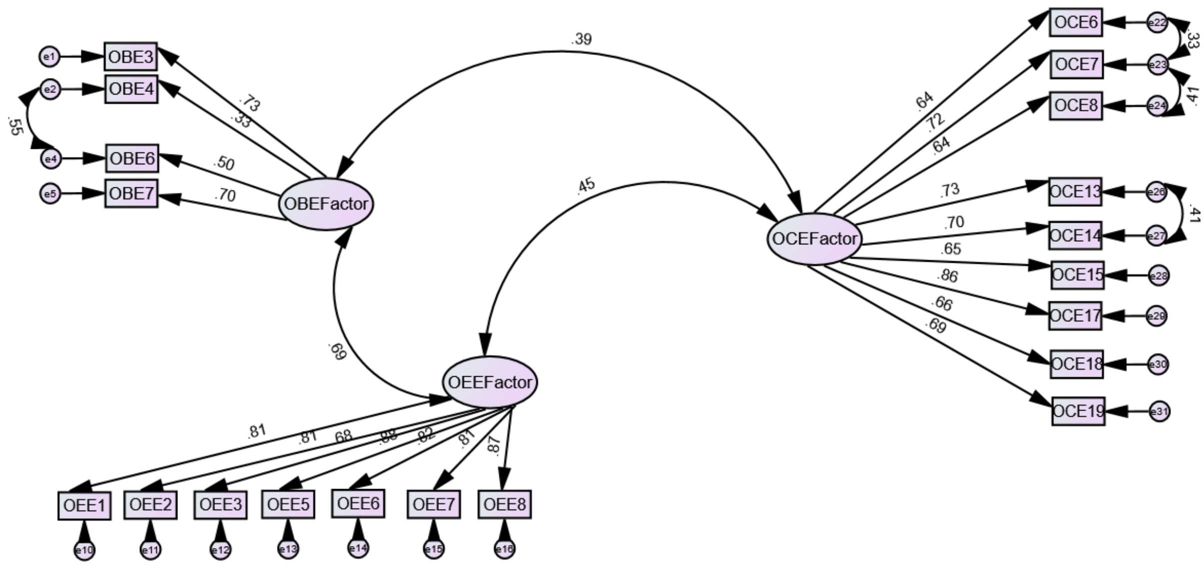


Table 4.4

Global Fit Indices of Final Best-Fitting Model (N=235)

Model-Fit Index	Value
CMIN	279.13
CMIN/DF	1.712
GFI	.894
AGFI	.864
PGFI	.694
NFI	.904
IFI	.958
TLI	.950
CFI	.957
SRMR	.053
RMSEA	.055

Note: CMIN=Chi-square ($\chi 2$); CMIN/DF=ratio of chi-square to degrees of freedom; GFI=Goodness-of-Fit Index; AGFI=Adjusted Goodness-of-Fit Index; PGFI= Parsimony Goodness-of-fit Index; NFI= Normed Fit Index; IFI=Incremental Fit Index; TLI= Tucker-Lewis Index (also called the Non-normed Fit Index); CFI= Comparative Fit Index; SRMR= Standardized Root Mean Square Residual; RMSEA= Root Mean Square Error of Approximation

The test of the null hypothesis that online student engagement is a 3-factor structure as depicted in Figure 4.1 produced a chi-square ($\chi 2$) value of 279.13, with 163 degrees of freedom and a probability of less than .0001, suggesting the lack of exact fit. However, given the sensitivity of the likelihood test to the sample size and its assumption of perfect fit in the population, this result was expected (Byrne, 2016). In addition, other goodness of fit indices may still indicate an acceptable or good model fit even when the chi-square test is statistically significant, which suggests that the hypothesized model can represent and explain the sample data very well, even if there is a statistically important discrepancy (Mulaik et al., 1989). The $\chi 2$/degrees of freedom ratio shown as CMIN/DF in Table 4.4 was found to be 1.712. According to Tabachnick and Fidell (2013), when the chi-square to degrees of freedom (CMIN/DF) ratio is less than 2, which was the case in this study, it may indicate a good-fitting model. However, due to the known limitations and problems associated with the chi-square statistic (including $\chi 2$/degrees of freedom ratio) resulting from its sensitivity to sample size and its assumption of perfect model fit, the overall model fit was assessed using several other alternative model fit indices. As absolute fit indices and incremental or comparative fit indices evaluate and represent model fit from different angles, researchers are highly recommended to report different fit indices representing several aspects of model fit (Whittaker, 2016). For example, Kline (2016) recommend reporting, as the minimum requirement, the chi-square statistics including degrees of freedom and the probability level, CFI, RMSEA including the 90 % confidence intervals and p value, and SRMR, all of which were reported in this study.

The GFI and the AGFI values, two commonly used absolute fit indices (Whittaker, 2016), were found to be .894 and .864 respectively, and values .90 or greater for these two indices indicate an acceptable model fit (Whittaker, 2016). Although both GFI and AGFI values found in this study indicated weaker fit, both values are known to be influenced by the sample size (Fan et al., 1999) and they assess the absolute or perfect fit of the model to the data. Therefore, additional model fit indices should be interpreted to evaluate several other aspects of model fit (Byrne, 2016). The comparative fit indices of NFI, IFI, TLI, and CFI values ranging from .904 to .958 all supported an acceptable model fit. Given that the values $\geq .90$ for these indices are conventionally treated as indicative of an acceptable fitting model (Whittaker, 2016), all these incremental fit indices suggested an acceptable model fit for the fit of the hypothesized model to the sample data.

The standardized RMR known as SRMR, another commonly used absolute fit index (Whittaker, 2016), was found to be .0535. According to Byrne (2016), a well-fitting model has a SRMR value of .05 or less. On the other hand, Hu and Bentler (1999) suggest a cutoff value close to .08 for SRMR and Kahn (2006) and Russell (2002) describe SRMR values of .08 or less as indicating good model fit to the data. In light of all these recommended cut-off criteria for the SRMR, the SRMR value of the final model was interpreted as indicating an acceptable model fit to the data.

The RMSEA value, another commonly used absolute fit index (Whittaker, 2016), was also considered with its confidence intervals and the test of closeness of fit (PCLOSE). A RMSEA value at or below .05 is indicative of close model fit and a RMSEA value up to .08 is indicative of reasonable errors of approximation (Browne & Cudeck, 1992; Byrne, 2016; Worthington & Whittaker, 2006). The RMSEA value for the final model was found to be .055. Researchers are strongly recommended to use and report the confidence intervals while reporting RMSEA values where "a very narrow confidence interval would argue for good precision of the RMSEA value in reflecting model fit in the population" (Byrne, 2016, p. 99). The 90 % confidence interval of the RMSEA value of the model ranged from .044 to .066, and the p-value for the test of close fit was .214. This confidence interval suggested that one could be 90% confident that the true RMSEA in the population would be between .044 (lower bound) and .066 (upper bound), which was a fairly narrow confidence interval supporting a good precision of the RMSEA value. The upper limit of the 90 % confidence interval was .066, which was still less than the .08 cutoff value suggested by Browne and Cudeck (1992) as indicating reasonable error of approximation, and the p-value for the test of close fit was .240, not indicating statistical significance. As a value of .05 was contained within the confidence interval of the RMSEA (.044 - .066) and the RMSEA value was accompanied by a p-value of .214 greater than .05, it also indicated at least an acceptable model fit (Whittaker, 2016).

Local Fit Summary. As shown by the unstandardized and standardized regression coefficients (factor loadings) of both ML estimates and the estimates with their bias-corrected 95 % confidence intervals, all the factor loadings were found to be positive as theoretically expected, and statistically significant at the .05 alpha level. Table 4.5 presents the normal theory ML unstandardized and standardized factor loadings of twenty variables in the final best-fitting model seen in Figure 4.1 (see Appendix N for the unstandardized and standardized factor loadings of twenty indicator variables with their bias-corrected 95 % confidence intervals). Since each OEHE item loaded onto a single factor, these standardized factor loadings are interpreted as correlation coefficients. All the standardized regression weights were strongly positive (>.50) except for the item OBE4 (.33), and all the standardized regression weights were also statistically significant (p < .01), thereby indicating a reasonably good local model fit. As for the .333 factor loading of the item OBE4 on the OBE factor, the magnitude of the loading, although not as strong as the others, was still considered acceptable because "a significant factor loading, as determined by a standardized coefficient of 0.30 or above, indicates that the item is a good measure of the underlying factor" (Wang et al., 2019, p. 116).

Factor Correlations. In the final best-fitting model seen in Figure 4.1, there were still positive and moderate to strong correlations among the three factors of online student engagement, as theoretically expected. The correlation between OBE Factor and OCE Factor was r=.392 (p=.001); the correlation between OBE Factor and OEE Factor was .689 (p=.001); and the correlation between OEE Factor and OCE Factor was .450 (p=.001). All these moderate to strong factor correlations between the three engagement dimensions specified in the final best-fitting model were consistent with the student engagement theory (Fredricks et al., 2004, 2005) and the prior research of engagement in both traditional and online learning environments (e.g., Kucuk & Richardson, 2019; Li & Lerner, 2013; Park & Yun, 2019; Rimm-Kaufman et al., 2015; Sun & Rueda, 2012).

Table 4.5

ML Unstandardized and Standardized Factor Loadings of the 20 Indicator Variables (N=235)

Item	Factor	B	S.E.	C.R.	*P*	*β*
OBE3	OBE	1.000				.731
OBE4	OBE	.731	.169	4.328	<.001	.333
OBE6	OBE	1.077	.168	6.411	<.001	.504
OBE7	OBE	1.159	.142	8.149	<.001	.700
OEE1	OEE	1.000				.814
OEE2	OEE	1.002	.069	14.530	<.001	.812
OEE3	OEE	.837	.073	11.519	<.001	.685
OEE5	OEE	1.043	.063	16.497	<.001	.884
OEE6	OEE	.936	.064	14.603	<.001	.815
OEE7	OEE	1.071	.074	14.547	<.001	.813
OEE8	OEE	.971	.061	16.030	<.001	.868
OCE6	OCE	1.000				.639
OCE7	OCE	.932	.083	11.271	<.001	.718
OCE8	OCE	.855	.101	8.445	<.001	.642
OCE13	OCE	1.088	.117	9.322	<.001	.727
OCE14	OCE	.820	.091	9.046	<.001	.701
OCE15	OCE	.916	.108	8.501	<.001	.647
OCE17	OCE	1.068	.102	10.517	<.001	.859
OCE18	OCE	.995	.115	8.660	<.001	.662
OCE19	OCE	.969	.108	8.990	<.001	.693

Note: B=Unstandardized, S.E=Standard Error, C.R.=Critical Ratio, β=Standardized

The final best-fitting model has four items for the underlying factor of online behavioral engagement, seven items for the underlying factor of online emotional engagement, and nine items for the underlying factor of online cognitive engagement (See Appendix O for the item content and descriptive statistics of each indicator variable in the final-best fitting model). Although the number of measures for the latent variable of online behavioral engagement was relatively lower than the other two factors, it still meets the criterion of having "at least three measures per factor or latent variable" (Russell, 2002, p. 1642). Kline (2016) similarly recommends a minimum of three to five indicators for each factor.

Tests for Convergent, Discriminant, and Criterion Validity

The second research question *RQ2: Assuming the hypothesized three-factor model of online student engagement is an adequately fitting model, does convergent, discriminant, and criterion validity evidence support the use of the scale factors as indicators of online student engagement in higher education contexts?* was answered satisfactorily.

Pearson Bivariate Correlations for Convergent Validity

Empirical evidence for convergent validity of the OEHE items and subscales was initially obtained from high factor loadings within each subscale in the context of the confirmed CFA model (see Figure 4.1). The standardized factor loadings of all the OEHE items within each subscale, except for one online behavioral engagement item, namely OBE4, were .50 and above, with majority of the item loadings in the .70s and .80s. Even the relatively less strong loading of the OBE4 item (.333) can be considered a reasonably good measure of the underlying factor (Wang et al., 2019).

Convergent validity can also be assessed by examining the relationships between the scores on a new measure of a particular construct or attribute and the scores on an existing and validated measure of the same construct, looking for moderate to strong and positive correlations between the two (Campbell & Fiske, 1959; Warner, 2008). For this purpose, the correlations between the mean scores of three subscales of the OEHE and the mean scores of the corresponding subscales of the online-modified version of Reeve's (2013) validated measure of student engagement (see Appendix I) were computed, looking for moderate to strong and positive correlations between the two sets of subscales. Table 4.6 presents the Pearson correlations between each OEHE subscale and each corresponding subscale of Reeve's (2013) classroom engagement measure. As shown in Table 4.6, the Pearson product-moment correlation coefficients between the mean scores of three subscales of the OEHE measure and the three corresponding subscales of Reeve's (2013) online-modified classroom engagement ranged from .17 to .32, indicating weak to moderate correlations, providing relatively weak albeit reasonable evidence for convergent validity of the OEHE subscales.

Table 4.6

Pearson Bivariate Correlations between OEHE Subscales and Reeve's (2013) Engagement

	ReeveBE	**ReeveEE**	**ReeveCE**
OBE	**.170****	.221**	.201**
OEE	.093	**.192****	.035
OCE	.173**	.240**	**.326****

Note: OBE=OEHE's online behavioral engagement, OEE=OEHE's online emotional engagement, OCE=OEHE's online cognitive engagement, ReeveBE=Reeve's behavioral engagement, ReeveEE=Reeve's emotional engagement, ReeveCE=Reeve's cognitive engagement
**Correlation is significant at the 0.01 level (2-tailed)
Listwise *N*=234

The overall mean score of the OEHE instrument and the overall mean score of Reeve's engagement measure were also computed and the Pearson correlation between the two was examined. There was a positive and significant correlation ($r=.242$, $p<.001$), although not strong, between the two scores, again providing relatively weak albeit reasonable evidence for convergent validity. One potential reason for these weak correlations is that although Reeve's (2013) engagement items were modified in wording to fit the online context, they were originally intended for traditional face-to-face classrooms, and they might not have worked as effectively even when modified to online learning environments. Another potential reason is that Reeve's (2013) student engagement items are quite general in wording and item content (see Appendix I), whereas the items of the OEHE subscales are much more specific in terms of behavior, emotion, and cognition regarding specific components of online courses such as discussion board activities and student-student and student-instructor interactions. Due to the highly specific wording and content of the OEHE items relative to Reeve's (2013) more general items, strong correlations might not have been obtained for convergent validity, although both measures theoretically converge on Fredricks et al.'s (2004) three core dimensions of engagement.

Pearson Bivariate Correlations for Discriminant Validity

Pearson bivariate correlations between the mean scores of the OEHE subscales and the mean score of a seven-item fixed ability subset of the 32-item version of the Epistemic Beliefs Inventory (EBI) (DeBacker et al., 2008; Schraw et al., 2002) (see Appendix J) were examined. For discriminant validity evidence, researchers look for very weak to non-existent correlations between the two measures intended to measure totally different constructs, namely engagement and fixed ability mindset in this study. Table 4.7 presents the Pearson bivariate correlations between each OEHE subscale and the EBI Fixed Ability (FA) measure.

Table 4.7

Pearson Bivariate Correlations between OEHE Subscales and EBI-FA Measure

OEHE Dimension	**EBI-FA**
OBE	**.051**
OEE	**-.093**
OCE	**-.119**

Note: OBE=OEHE's online behavioral engagement, OEE=OEHE's online emotional engagement, OCE=OEHE's online cognitive engagement, EBI-FA=Epistemic beliefs fixed ability mindset
**Correlation is significant at the 0.01 level (2-tailed). Listwise *N*=234

As shown in Table 4.7, the Pearson product-moment correlation coefficients between the three subscales of the OEHE measure and the EBI-FA were extremely weak and non-significant, providing strong evidence for discriminant validity of all three OEHE subscales. The Pearson correlation between the overall mean score of the OEHE instrument and the mean score of EBI was examined and an extremely weak and non-significant correlation was still found between the two (r= -.082, *p*=.211), similarly providing strong evidence for discriminant validity of the OEHE instrument.

Pearson Bivariate Correlations for Criterion-Related Validity

For this purpose, Pearson product-moment correlations between the mean score of each subscale of the OEHE and the mean score of an online-modified version of the six-item task value subscale of the Motivated Strategies for Learning Questionnaire Survey (MSLQ) (see Appendix K) were examined. Table 4.8 presents the Pearson product-moment bivariate correlations between each OEHE subscale and the task value measure.

Table 4.8

Pearson Bivariate Correlations between OEHE Subscales and Task Value Measure

	TV
OBE	**.179****
OEE	**.319****
OCE	**.382****

Note: OBE=OEHE's online behavioral engagement, OEE=OEHE's online emotional engagement, OCE=OEHE's online cognitive engagement, TV=Task value perception in an online course
**Correlation is significant at the 0.01 level (2-tailed)
Listwise *N*=234

As recommended by DeVellis (2017), some empirical relationship was sought between the OEHE subscales, and the task value measure used as a criterion measure. There was a positive and statistically significant relationship between each OEHE subscale and the task value. In addition, the Pearson correlation between the overall OEHE mean score and the task value was computed and similarly a positive and statistically significant relationship was found between the two ($r = .383$, $p < .001$) and it was a moderate relationship, providing reasonable evidence for criterion-related validity as well.

Reliability

The third and final research question *RQ3: Do the confirmed scale factors exhibit evidence of internal consistency reliability? If yes, to what extent?* was also answered satisfactorily. As evidence of reliability, internal consistency reliability, which indicates the extent to which a set of items within a scale are homogenous and the items are consistently measuring the same underlying construct, can be utilized (DeVellis, 2012; Pallant, 2020). The Cronbach's alpha values are the most common approach to estimating and reporting score reliability in social science research (DeVellis, 2012; Furr, 2011; Kline, 2016; Warner, 2013). Table 4.9 presents the descriptive statistics and the Cronbach's alpha (α) values of the twenty indicator variables in the final best-fitting three-factor model.

Table 4.9

Descriptive and Reliability Statistics of the 20 OEHE Items (Final) (N=235)

	Number of Items	Mean	SD	α
OBE	4	3.75	.871	**.698**
OEE	7	3.69	.841	**.931**
OCE	9	4.42	.497	**.898**
OEHE	20	4.03	.549	**.906**

OBE=OEHE's online behavioral engagement, OEE=OEHE's online emotional engagement, OCE=OEHE's online cognitive engagement, OEHE=OEHE instrument as a whole;
α= Cronbach's alpha reliability

Given that the Cronbach's alpha coefficients of .70 and above are generally considered to be acceptable or adequate, .80 and above very good, and .90 and above are considered excellent in social science research (Pallant, 2020), both the OEE subscale with seven items and the OCE subscale with nine items were found to have a very good to excellent internal consistency reliability, while the OBE subscale with only four items that passed the confirmatory factor tests had an almost adequate internal consistency (between .65 and .70 minimally acceptable, DeVellis, 2003), although still quite weaker compared to the other two OEHE subscales. The limited number of items in the OBE scale is one reason for the weaker internal consistency reliability since the number of items in a scale significantly influences its Cronbach's alpha reliability estimates (DeVellis, 2012). Another reason is associated with the less strong relationship between the OBE items as indicated by the CFA estimates previously. This less tight intercorrelation among the OBE items compared to the OEE and OCE items seems to have also influenced the internal consistency reliability of the OBE scale. However, the internal consistency reliability of the OEHE instrument as a whole with a total of 20 items (α = .906) was still very good, providing strong empirical evidence for the reliability of the OEHE as an instrument to measure online student engagement.

Chapter 5: Discussion and Conclusion

Overview

The purpose of this study was to develop and validate a new self-report survey of online student engagement in higher education, the Online Engagement in Higher Education (OEHE) instrument. Based on Fredricks et al.'s (2004) tripartite conception of engagement, the OEHE instrument was developed to test the three dimensions of student engagement (i.e., behavioral engagement, emotional engagement, and cognitive engagement) in online learning environments of higher education. The results of the study overall indicated that the OEHE is a reliable and valid instrument, albeit its limitations. By collecting and interpreting data about inclusiveness, bias, and accessibility perceptions of the students, this study also aimed to raise awareness about diversity, equity, and inclusion, DEI, issues that have been traditionally neglected in the instrument development studies (Cintron & Hagan, 2021).

In this chapter, a summary of the findings in relation to each of the research questions and a thorough discussion of the findings are presented. Then, study limitations and its implications for theory and practice are given, followed by the conclusions and recommendations for future research. As part of recommendations, DEI perspectives are also shared to raise awareness about the DEI issues that need to be considered in the instrument development process by future studies.

Summary and Discussion of the Findings

This study consisted of two major phases in developing and validating an assessment instrument to measure students' behavioral, emotional, and cognitive engagement in online courses of higher education by addressing the existing issues in the literature. The first phase was developing and validating the OEHE items through expert review and pilot cognitive testing. The second phase focused on investigating the three research questions. The findings to the three research questions are summarized below, followed by discussions.

Discussion of Results of Research Question One

Results regarding research question one *RQ1: Can a three-factor hypothesized model of online student engagement based on theory and prior research be confirmed in a validation sample of online students of higher education?* provided empirical evidence for the factorial validity of the OEHE instrument as one significant dimension of its construct validity. The a priori hypothesis for the factorial structure of the OEHE instrument was that online student engagement in higher education was comprised of three dimensions or factors and therefore three latent factors were used to assess online student engagement. To test this a priori hypothesis regarding the three-factor model of online student engagement, a series of CFAs were conducted by revising the model in order to be able to reach at least an acceptably fitting three-factor model of online student engagement. All revisions and modifications to the hypothesized model were made in light of both statistical evidence (e.g., global model fit indices, regression weights, modification indices) and theoretical sense (e.g., overlap in item content). After each refinement on the model (e.g., removal of items, specifying error covariances), both the model fit indices and the parameter estimates of the factor loadings within each hypothesized factor were carefully examined to judge the extent of the impact of the revisions and modifications on the model.

In the final best-fitting model with three factors as hypothesized and confirmed, there were a total of 20 indicator variables, with four items for the latent factor of online behavioral engagement, seven items for the latent factor of online emotional engagement, and nine items for the latent factor of online cognitive engagement. Although the number of measures for the latent variable of online behavioral engagement in the final best-fitting model was relatively lower than the other two factors, it still meets the criterion of having at least three items for each factor (Russell, 2002). In the final best-fitting model, there were positive, significant, and moderate to strong correlations among the three factors of online student engagement, as theoretically expected. All these factor correlations between behavioral, emotional, and cognitive engagement dimensions were consistent with the student engagement framework used in this study (Fredricks et al., 2004, 2005) and with prior research of student engagement in both traditional and online learning environments (e.g., Deng et al., 2020; Hoi & Le Hang, 2021; Kucuk & Richardson, 2019; Li & Lerner, 2013; Park & Yun, 2019; Rimm-Kaufman et al., 2015; Sun & Rueda, 2012). The relatively higher correlation between the online behavioral engagement and emotional engagement factors in this study can be explained by the considerable content overlap since the items in both subscales were asking the students about their online discussion interactions, whereas the online cognitive engagement items were also asking about overall online learning experiences besides online discussions.

Justification of Remaining Indicators of OBE in the Confirmed Model. The results of this dissertation research provide supporting evidence that behavioral engagement is one of the three hypothesized factors or dimensions of online student engagement, as theoretically hypothesized (Fredricks et al., 2004). The four indicator variables of online behavioral engagement in the final confirmed model are associated with student-student and student-instructor interaction behaviors happening during online discussions. This pattern of results is consistent with the relevant online learning literature stating that asynchronous online discussion boards or discussion forums are commonly used in online education and are the primary avenues in which students perform certain online learning behaviors through their interactions with their peers and their course instructor (Cheng et al., 2011; Ding et al., 2017; Goggins & Xing, 2016; Hew & Cheung, 2008; Kay, 2006; Lee & Recker, 2021; Salter & Conneely, 2015; Xie et al., 2006). Discussions taking place through discussion boards or forums have been regarded and shown to be an integral element of online learning environments by means of which students learn (Hew et al., 2010; Lee & Recker, 2021). Therefore, the discussion board participation behaviors measured by the OEHE instrument as indicators of online behavioral engagement in this study are noteworthy given that "students do not only learn from course materials and lectures, they also learn a great deal from interacting with one another" (Cheng et al., 2011, p. 254). Discussion board interactions provide this learning opportunity for online students. Through reading their peers' posts and participating in content-related dialogues or conversations with their peers as well as their course instructor during online discussions, online students can gain not only factual content knowledge but can also come up with new ideas and gain new insights into the course content (Cheng et al., 2011; Lee & Recker, 2021).

Measuring such discussion-based participation behaviors as indicators of behavioral engagement is also justified by the relevant prior research showing a positive relationship between participating in online discussions and academic achievement in the online course (Cheng et al., 2011; Lee & Recker, 2021). Measuring students' participation behaviors in online discussions as indicators of their engagement is also crucial and relevant because previous research suggests that "the presence of a forum transforms less engaged students into more engaged students" (Cheng et al., 2011, p. 261). Due to its engaging power, asynchronous online discussions are given in the relevant literature as an effective instructional tool or strategy to foster student engagement because online discussion board or forum conversations, interactions or dialogues are considered to be engaging learning activities for students in online learning environments (Salter & Conneely, 2015)

Measuring online behavioral engagement via such online discussion participation indicators is also consistent with the previous literature suggesting that student participation in asynchronous online discussions can foster and support higher-order thinking, problem solving, and active learning, and students need to participate in those online discussions so that the benefits of online discussions can be reaped by online students (Ding et al., 2017; Goggins & Xing, 2016; Salter & Conneely, 2015). Therefore, assessing the level of student engagement in online discussions is particularly important to identify whether online students are more likely or less likely to enjoy such positive outcomes of online discussions. Given that lack of or low participation in asynchronous online discussions is a widespread issue in online courses (Hew et al., 2010; Lee & Recker, 2021), measuring participation in online discussions as indicators of the level of behavioral engagement in online courses is also relevant and justified.

In addition, given that student engagement in online discussions is influenced by instructor participation as well as peer interactions (Ding et al., 2017), the OEHE items measuring instructor's involvement in online discussions as well as interactions with peers are also consistent with the relevant literature. Moreover, the current confirmed set of online behavioral engagement (OBE) items associated with students' participation behaviors in online discussions is also consistent with the previous research on student engagement, most of which associates behavioral engagement with students' participation in learning, which primarily takes place in asynchronous discussions in online learning environments (Cheng et al., 2011; Ding et al., 2017; Salter & Conneely, 2015). Further, the current OBE items measuring students' participation in content-related dialogues with the course instructor and with the peers, which encompass the behaviors of asking and answering questions, making comments, and offering suggestions during the natural course of these dialogues, rather than the behavior of posting only, and measuring participation in online discussions through the behavior of reading posts are also supported by the relevant literature. Dennen (2008) states:

> Discussion itself requires a pattern of call and response, with turn-taking and listening being as important as contributing thoughts to the dialogue. Is it possible that students might engage with the asynchronous discussion by reading, the online equivalent of listening? Of course. (p. 1625)

In support of the validity of the current OBE items measuring participation in content-related dialogues during online discussions, Goggings and Xing (2016) also point out that “true learning dialogue requires students to read and reflect in order to be part of a dialogue instead of just posting activities” (p. 242).

On the other hand, there were other behavioral indicators given in the traditional engagement literature, such as homework completion, studying, asking for help, or accessing course material (Bond et al., 2020) and those given in online learning literature, such as sharing additional knowledge resources with others (Park et al., 2016; Shackelford & Maxwell, 2012), taking notes (van de Sande et al., 2017; Watkins et al., 2015), completing online tasks and activities and submitting assignments online (Cerezo et al., 2016; Kokoc et al., 2021; Park et al., 2016) and getting feedback from the instructor (Watson et al., 2017). All of these indicators were theoretically hypothesized, based on the existing theoretical and empirical literature, and operationally used in the construction of the behavioral engagement items of the OEHE to measure other dimensions of online behavioral engagement in this study. However, these indicators could not pass the reliability and validity tests with the sample of this study, which needs further examination with follow-up and/or replication studies with a larger sample size.

This pattern of findings regarding the less well-defined status of the behavioral engagement dimension compared to the other two dimensions of online engagement in this study is consistent with the literature indicating that behavioral engagement as a construct is influenced by specific changes in specific contexts and it may greatly vary from context to context or even from moment to moment (Hoi & Le Hang, 2021; Nguyen et al., 2018). This is particularly relevant to the context of this study. Most of the online learning experiences were delivered to the participants in this study as fully asynchronous and mostly asynchronous online (70.2%) and for this reason their participation behaviors in online learning experiences could be "unstable over time and across specific activities" (Hoi & Le Hang, 2021, p. 1150). Considering that there were also important variations in the experiences of the participants, such as whether they had learning activities in their online course or whether they received feedback or not as evidenced by their responses to the course context questions, the participants might have used their general notions or interpretations of their behavioral engagement in online courses while judging and responding to this cross-sectional survey assessment of online behavioral engagement (Hoi & Le Hang, 2021). Findings from this study regarding the behavioral engagement dimension suggest that future researchers and scale developers should exercise caution by considering this context-specific and context-dependent nature of the behavioral engagement construct in online learning environments (Hoi & Le Hang, 2021).

Justification of Remaining Indicators of OEE in the Confirmed Model. The results of this dissertation research provide supporting evidence that emotional engagement is one of the three hypothesized factors or dimensions of online student engagement, as theoretically hypothesized (Fredricks et al., 2004). The seven indicator variables of online emotional engagement in the final confirmed model are also associated with student-student and student-instructor interaction behaviors happening during online discussions. This pattern of results is consistent with the relevant online learning literature stating that asynchronous online discussion boards or discussion forums are commonly used in online education and are the primary spaces in which online students are engaged in learning through their interactions with their peers and their course instructor (Cheng et al., 2011; Ding et al., 2017; Goggins & Xing, 2016; Hew & Cheung, 2008; Kay, 2006; Lee & Recker, 2021; Salter & Conneely, 2015). *Enjoyment* was confirmed in this study as the dominant academic positive emotion representing online students' emotional engagement associated with their interactions with their peers, their course instructors, and the course content during online discussions. *Curiosity* was also represented in the final best-fitting confirmed model by a single indicator variable assessing curiosity associated with peers' perspectives/experiences in each online discussion. This pattern of results is consistent with the previous literature indicating that "emotions are critically important for students' engagement with academic tasks" (Pekrun & Linnenbrink-Garcia, 2012, p. 278). These results are also consistent with the previous research suggesting that *enjoyment* is one of the positive-activating emotions most commonly reported and experienced by university students in academic settings and learning contexts (Pekrun et al., 2011; Pekrun & Linnenbrink-Garcia, 2012; Reindl et al. 2018). The use of the positive activating emotion of *enjoyment* to identify online students' emotional engagement is also relevant and important in the context of empirical research indicating the positive impact of *enjoyment* on academic performance and achievement (Reindl et al. 2018) and its positive relation with students' use of cognitive and metacognitive strategies in online courses for better learning (Artino & Jones, 2012). The use of *enjoyment* as the major emotion to measure students' emotional engagement in online learning environments is also supported by prior online learning research measuring emotional engagement in online discussions by using *enjoyment* associated with online discussions as an indicator of emotional engagement (e.g., Ding et al., 2017; Xu et al., 2020) as well as by similar measures of online emotional engagement using *enjoyment* as one of the positive emotions experienced by students in online learning environments (e.g., Hoi & Le Hang, 2021). Further, the use of enjoyment to measure emotional engagement in online learning environments is also important in the context of empirical research showing that *enjoyment* as a positive academic emotion contributes positively to students' online learning experiences (Tempelaar et al., 2012).

The use of *curiosity* as another indicator of online emotional engagement in the final OEHE instrument is also consistent with the previous literature that *curiosity* is one of the epistemic emotions triggered by cognitive incongruities and commonly experienced by students in academic settings and learning contexts (Litman & Spielberger, 2003; Pekrun et al., 2017; Pekrun & Linnenbrink-Garcia, 2012). The use of *curiosity* as an indicator of online emotional engagement in the final OEHE in this study is also justified by the fact that the measurement of the epistemic emotions in the relevant literature has primarily concentrated on measuring *curiosity* (Pekrun et al., 2017). The use of *curiosity* in the final OEHE instrument to measure emotional engagement is also relevant and important by the prior research indicating that academic learning and performance of students can be strongly influenced by epistemic emotions such as curiosity (Pekrun et al., 2017).

On the other hand, there were two other positive achievement emotions related to success and failure, namely *pride* and *hope* (Pekrun et al., 2017*)*, used in the conceptualization and operationalization of online emotional engagement in this study, but the indicator variables associated with these two emotions could not pass the reliability and validity tests with the sample used in this study. As indicated in the relevant research literature, one explanation is that although the achievement emotions may be similarly experienced in online learning environments like in traditional classrooms, certain distinct characteristics of online learning environments (e.g., lack of immediate contact, individual accountability, self-regulation, reliance on technology) may have differentially influenced the students' appraisals of control (i.e., their beliefs regarding how effectively they can control and determine the outcomes of their online learning experiences) and value (how important or valuable their online learning experiences and outcomes are to them) and therefore their appraisals of the specific achievement emotions of *hope* and *pride* in their specific online courses (Daniels & Stupnisky, 2012). In other words, the achievement emotions of *hope* and *pride* may have been influenced by those appraisals quite differently than the emotions of *enjoyment* and *curiosity* given that "the unique combination of students' control and value appraisals predicts their emotions" (Daniels & Stupnisky, 2012, p. 223). This area also needs further examination with follow-up and/or replication studies with a larger sample size.

More research is definitely needed to examine all these academic achievement emotions in online learning environments since most of the empirical studies examining these emotions have traditionally examined them in face-to-face learning environments (Artino, 2010; Artino & Jones, 2012; Daniels & Stupnisky, 2012; Tempelaar et al., 2012). Finally, the fact that the online emotional engagement items associated with studying, completing tasks and assignments, participating in learning activities, and getting feedback did not pass the reliability and validity tests was consistent with the failure of similar behavioral engagement items to pass the reliability and validity tests with the sample in this study, which similarly needs further examination with follow-up and/or replication studies with a larger sample.

Justification of Remaining Indicators of OCE in the Confirmed Model. The results of this dissertation research provide supporting evidence that cognitive engagement is one of the three hypothesized factors or dimensions of online student engagement, as theoretically hypothesized (Fredricks et al., 2004). The nine indicator variables of online cognitive engagement in the final confirmed model are associated with students' interaction behaviors happening during online discussions (e.g., explaining in depth during online discussions) and with their use of cognitive and metacognitive strategies to learn and understand the online course content (e.g., making connections between ideas, reflecting on one's own understanding of course concepts). This pattern of results is consistent with the relevant online learning literature stating that asynchronous online discussion boards or discussion forums are commonly used in online education and are the primary spaces in which online students are engaged in learning through their interactions with their peers and their course instructor (Cheng et al., 2011; Ding et al., 2017; Goggins & Xing, 2016; Hew & Cheung, 2008; Kay, 2006; Lee & Recker, 2021; Salter & Conneely, 2015).

In addition, the use of cognitive and metacognitive strategies to measure cognitive engagement in online learning environments is also consistent with the relevant literature indicating the importance and value of cognitive engagement and cognitive and metacognitive strategies in traditional and online learning environments (Dennen, 2008; Fredricks et al., 2005; Greene & Miller, 1996; Henrie et al., 2015; Miller et al., 1996, Reeve & Tseng, 2011; Reeve, 2013; Sun & Rueda, 2012; Wolters, 2004). The assessment of students' use of cognitive and metacognitive strategies to measure online cognitive engagement in the OEHE instrument is also justified and important in the context of the previous research indicating that students' use of such learning strategies influences their online learning success (Lin et al., 2017; Yeh et al., 2019). The assessment of cognitive engagement associated with online discussions is also important and relevant to online learning environments given that previous research indicates that students' cognitive engagement in online discussions can contribute to their learning and knowledge construction in online learning environments (Ding et al., 2017; Putman et al., 2012; Zhu, 2006).

When compared to the construct representation achieved by the OBE and OEE subscales, this subscale of online cognitive engagement (OCE) more adequately represents online students' engagement not only during online discussions but also during learning in the overall course. There were still some OCE items that could not pass the reliability and validity test with the sample in this study but almost all of the lost items are still expressed by the remaining OCE items due to their very similar cognitive and metacognitive content. However, there was one indicator variable intended to measure the metacognitive strategy of planning one's study actions and it did not pass the reliability and validity tests in this study and no remaining OCE items represent its content. Given that planning one's actions or setting goals is one of the important self-regulatory metacognitive strategies used by successful learners in both traditional and online learning environments (Lin et al., 2017; Wandler & Imbriale, 2017; Zimmerman, 1990), this issue also needs further examinations by future studies with larger samples.

In light of the three central issues that the current study was intended to address, which were (a) insufficient systematic and fully blown scale development research using Fredricks et al.'s (2004) tripartite conception to measure online student engagement in higher education contexts, (b) construct irrelevance (i.e., not measuring the intended construct of engagement) and

(c) construct underrepresentation (i.e., not measuring all fundamental aspects of engagement), the results of this dissertation study were able to fully address the first two issues. Through a systematic and fully blown scale development research, this dissertation study used and confirmed Fredricks et al.'s (2004) three-dimensional engagement conception in the context of online learning. This dissertation study also developed and validated an online engagement instrument that measures the intended construct of interest, engagement per se, rather than motivational elements such as self-efficacy, intrinsic motivation or any other irrelevant constructs, addressing the issue of construct irrelevance. The third issue, construct underrepresentation, was not fully resolved in this dissertation study due to the relatively limited nature of behavioral and emotional engagement subscales measuring behavioral and emotional engagement of students associated with their interactions in online discussions rather than overall course engagement including the online discussions. Given that "more items lead to better construct representation" (Eisinga et al., 2013, p. 641), both OBE and OEE subscales need to have more items to better represent the respective constructs of online behavioral engagement and online emotional engagement. This limitation regarding construct representation is discussed in detail in Limitations and Recommendations for Future Research sections.

Given (a) the meaningfulness, feasibility, and statistical significance of the factor loadings, (b) the adequate global fit of the three-factor hypothesized model with acceptable fitting model fit indices, and (c) no serious indication of model misspecification checked through modification indices and standardized residual covariances, it was concluded that the three-factor model adequately represents the structure of online student engagement for undergraduate and graduate students taking online courses in higher education settings, despite the fewer number and less variety of items of online behavioral and emotional engagement than theoretically hypothesized and operationally constructed at the onset of the development of the OEHE instrument. The limitations and recommendations regarding these issues are discussed in the Limitations and Recommendations for Future Research sections.

Discussion of Results of Research Question Two

Results regarding the second research question *RQ2: Assuming the hypothesized three-factor model of online student engagement is an adequately fitting model, does convergent, discriminant, and criterion validity evidence support the use of the scale factors as indicators of online student engagement in higher education contexts?* provided empirical evidence of convergent, discriminant, and criterion validity for the use of the OEHE factors and the OEHE instrument as a whole to assess online student engagement in higher education. The correlation results were used to test for convergent, discriminant, and criterion validity of each OEHE subscale and of the OEHE instrument as a whole. The magnitudes of the factor loadings within each subscale were also used to judge the extent of convergent validity. The standardized factor loadings of all the OEHE items within each subscale, except for one online behavioral engagement item, namely OBE4, in the final best-fitting confirmed model were .50 and above, with majority of the item loadings in the .70s and .80s. Even the relatively less strong loading of the OBE4 item (.333) can be considered a reasonably good measure of the underlying factor (Wang et al., 2019). In terms of the correlational results, the Pearson product-moment correlation coefficients between the mean scores of three subscales of the OEHE measure and the three corresponding subscales of Reeve's (2013) online-modified classroom engagement instrument indicated weak but positive, and significant correlations, providing relatively weak evidence for convergent validity of the OEHE subscales. There was also a positive and significant correlation, although still not strong, between the overall OEHE score and the overall score of Reeve's engagement measure, again providing relatively weak evidence for convergent validity.

One potential reason for the weak correlations is that although Reeve's (2013) engagement items were modified in wording to fit the online context, they were originally intended for traditional face-to-face classrooms, and they might not have worked as effectively even when modified to fit online learning environments. Another potential reason is that Reeve's (2013) student engagement items are quite general in wording and item content (see Appendix I), whereas the items of the OEHE subscales are much more specific in terms of behavior, emotion, and cognition regarding specific components of online courses such as discussion board activities and student-student and student-instructor interactions. Due to the highly specific wording and content of the OEHE items relative to Reeve's (2013) more general items, strong correlations might not have been obtained for convergent validity, although both measures theoretically converge on the same conception of student engagement (Fredricks et al., 2004).

The correlations between the three subscales of the OEHE measure and the EBI-FA were extremely weak and non-significant, providing strong evidence for discriminant validity of all three OEHE subscales. The correlation between the overall mean score of the OEHE instrument and the mean score of EBI-FA was also examined and an extremely weak and non-significant correlation was also found between the two, similarly providing strong evidence for discriminant validity of the OEHE instrument. Discriminant validity coefficients were interpreted in comparison to convergent validity coefficients, which is the recommended approach to the interpretation of convergent and discriminant validity evidence (Hubley, 2014). Translating this principle into the context of this study, there were relatively higher and statistically significant correlations between the subscales of the OEHE instrument and the corresponding subscales of

Reeve's (2013) engagement scale as two measures of the same construct of engagement, providing reasonable, although not strong, evidence for convergent validity. On the other hand, relatively lower or negligible and non-significant correlations were found between the subscales of the OEHE instrument and epistemic beliefs fixed ability mindset measure as the measures of two different constructs, providing strong evidence for discriminant validity. In addition, a relatively higher and statistically significant correlation was found between the overall OEHE mean score and the overall mean score of Reeve's (2013) measure of engagement, while a relatively much lower or negligible and non-significant correlation was found between the overall OEHE mean score and the EBI mean score.

Regarding criterion validity evidence, there was a positive and statistically significant relationship between each OEHE subscale and the task value measure. In addition, there was a positive and statistically significant relationship between the overall OEHE mean score and the task value measure, providing reasonable evidence for criterion-related validity as well. Overall, the results of research question 2 indicated that the three OEHE scale factors as indicators of online student engagement in higher education contexts as well as the OEHE instrument as a whole had reasonable evidence of convergent validity and criterion validity, and strong evidence of discriminant validity.

Discussion of Results of Research Question Three

Results regarding the third and final research question *RQ3 Do the confirmed scale factors exhibit evidence of internal consistency reliability? If yes, to what extent?* provided empirical information about the reliability of each construct-validated subscale of the OEHE instrument as well as the reliability of the OEHE instrument as a whole. Given that the Cronbach's alpha coefficients of .70 and above are generally considered to be acceptable or adequate, .80 and above very good, and .90 and above are considered excellent in social science research (Pallant, 2020), both the OEE subscale with seven items and the OCE subscale with nine items were found to have a very good to excellent internal consistency reliability, while the OBE subscale with only four items that passed the CFA tests had an almost adequate internal consistency (between .65 and .70 minimally acceptable, DeVellis, 2003), although still weaker compared to the other two OEHE subscales.

One possible reason for the weaker internal consistency reliability of the OBE subscale could be associated with the relatively limited number of items since the number of items significantly influences the Cronbach's alpha reliability estimates (DeVellis, 2012). It should be noted that more items are needed to better represent the subconstruct of online behavioral engagement and also that "the primary way to make measures more reliable is to increase the number of items" (Eisinga et al., 2013, p. 641). Another reason could be associated with the less strong relationship between the OBE items as indicated by the CFA estimates previously. Less tight intercorrelations among the OBE items compared to the OEE and OCE items seem to have also influenced the internal consistency reliability of the OBE scale. On the other hand, the internal consistency reliability of the OEHE instrument as a whole with a total of 20 items was very high (α= .906), providing strong empirical evidence for the reliability of the OEHE as an instrument to measure online student engagement.

Considering the results of a series of CFAs, validity coefficients, and Cronbach's alphas, it was concluded with reasonable confidence that the final 20-item three-factor OEHE instrument can indeed measure the three dimensions of student engagement (Fredricks et al., 2004) in online learning environments of higher education, although its resulting behavioral and emotional measures are relatively limited in content domain when compared to its cognitive engagement measures. Both the reliability and validity evidence obtained from this dissertation study support the use of the OEHE subscales as reliable and valid indicators of student engagement in online courses of higher education and the use of the OEHE instrument as a promising assessment tool to measure student engagement in online courses of higher education from three core dimensions of behavior, emotion, and cognition. However, the findings of this study regarding the validity and reliability of the OEHE instrument are still open to future studies and further improvements.

Implications of the Study

The findings of this dissertation study have implications for theory and research, and implications for practice.

Implications for Theory and Research. Although there is a consensus on the multidimensionality and significance of student engagement in learning environments (Hoi & Le Hang, 2021), there is still no scholarly consensus on the exact definition, operationalization, and assessment of student engagement (Fredricks, 2015, Reeve et al. 2020). Researchers investigating student engagement in online learning environments have conceptualized and examined this elusive construct in many ways, using different terms depending on their theoretical frameworks (Martin, Sun, & Westine et al., 2020). This study provides additional empirical evidence for the applicability of Fredricks et al.'s (2004) theory of student engagement to online learning environments via a multi-stage, systematic, and fully blown instrument development process. This dissertation study aimed to contribute to the consistency in defining and conceptualizing engagement constructs based on the three core dimensions of behavior, emotion, and cognition (Fredricks et al., 2004) by empirically testing them in online learning environments. The findings of this study indicating a three-factor student engagement model may be used to further explore Fredricks et al.'s (2004) meta construct theory of student engagement applied to online learning environments. In addition, the inconsistency in conceptualizations and assessments of engagement have also led to inconsistencies in research findings regarding what engagement is and how many dimensions it is comprised of (Fredricks, 2015). Assessment instruments that can reliably and validly measure engagement in learning environments, both traditional and online, are thus needed so that there is more consistency in findings in future research examining student engagement across learning contexts. This study may benefit the researchers in our field by offering them a promising measure of engagement that they can use to examine relationships between engagement and other constructs of their interest in online learning environments.

Implications for Practice. Online education continues to grow at a rapid rate at higher education institutions (Bolliger & Martin, 2021) and student engagement remains to be a significant parameter and indicator of successful online learning experiences (Bolliger & Martin, 2018; Kucuk & Richardson, 2019; Meyer, 2014; Robinson & Hullinger, 2008). Accurate assessments and evaluations of online student engagement are thus necessary for all stakeholders of online education to make appropriate and accurate evidence-based decisions about instructional design, instructional practices, and possible interventions in online courses to improve student engagement and so enhance and ensure the quality of students' online learning experiences (Czerkawski & Lyman, 2016; Dixson, 2015; Kozan & Richardson, 2014; Kucuk & Richardson, 2019). The OEHE was developed and validated as an assessment tool to measure student engagement in online courses of higher education. Practitioners ranging from course instructors to program coordinators to higher education institutions can use the OEHE instrument to obtain empirical information about the types and levels of engagement of the students enrolled in their online courses and programs. They can use the OEHE findings to make informed decisions about the performance of their courses and programs from the student engagement perspective and can take certain steps or design and implement interventions to improve student engagement accordingly (e.g., engaging course design elements). The OEHE is especially relevant to the assessment of student engagement in online learning through asynchronous online discussions where student-student, student-instructor, and student-content interactions mainly take place in online learning environments (Cheng et al., 2011; Ding et al., 2017; Goggins & Xing, 2016). Using the OEHE scores, online course designers and instructors can design, develop, and implement engaging tasks and activities to promote students' engagement behaviors, their emotional involvement with the learning process, and their cognitive engagement with the course content.

Implications for Inclusiveness, Diversity, and Accessibility in Online Classes. The demographic data collected in this study regarding students' perceptions and experiences of inclusiveness, diversity, and accessibility also have implications for scale development research and online education practice. This study aimed to raise awareness about the intentional incorporation of diversity, equity, and inclusion (DEI) issues in the entire instrument development process including item development, item validation, and sampling procedures (Cintron & Hagan, 2021). In addition, the fact that some students in this study reported that they sometimes, often, or always experienced different forms of bias against them in their online courses should draw attention to the issues of inclusive online course design in which such perceptions of bias can be minimized, if not eliminated, by course designers and course instructors through inclusive and engaging course components and expectations. Moreover, this study also provided empirical evidence that although the large majority of the students that participated in this study reported that they had not experienced any accessibility issues while taking online courses, some of them still reported that they had experienced such accessibility issues as lack of access to a computer, lack of access to hardware (e.g., headphones, webcam), lack of access to reliable Internet, lack of access to software (e.g., Microsoft Office), lack of accommodations for students with disabilities, and other similar accessibility issues. This finding also has implications for online educators, instructional designers, institutions of higher education offering online courses, and other policy makers in online education who should strive to ensure that those accessibility issues are optimally mitigated to ensure that online education is fully accessible to students of diverse backgrounds and life conditions (Linder et al., 2015).

Limitations of the Study

Despite best efforts to ensure rigor, this study still has its own limitations. Although the present results clearly support, as a priori hypothesized, a three-factor model of online student engagement in line with Fredricks et al.'s (2004) conception of engagement, it is appropriate to recognize several potential limitations.

Limitation Regarding Study Design and Sample Demographics. The generalizability of the findings regarding the reliability and validity of the OEHE instrument in this study might be limited by the cross-sectional nature of this study. One limitation of this study associated with the demographics of the participants in this study is that a large majority of the participants were female (80.8%, n=189) and only 16.7% (n=39) were male students. Another limitation associated with the demographics of the participants is that the proportion of graduate-level students (66.4%, n=156) was almost twice the number of undergraduate students (33.6%, n=79). In addition, most of the participants in this study were White (73.1%, n=171), while the other reported races/ethnicities were substantially lower. Therefore, future and/or follow-up studies should strive to recruit more equal numbers of participants of other genders, undergraduate and graduate-level students, and of different races/ethnicities to reduce gender, education-level, and racial/ethnicity bias in the sample and increase the generalizability of the findings. Future studies should also consider testing the reliability and validity of the OEHE instrument in longitudinal studies given that student engagement is very likely to change over time (Hoi & Le Hang, 2021).

Limitation Regarding Research Context and Sample Size. CFA as a member of the SEM family is a large sample statistical technique requiring large sample sizes to yield stable parameter estimates (Byrne, 2016; Worthington & Whittaker, 2006). After the removal of serious univariate and multivariate outliers and missing data cases on the OEHE variables, the sample size of this study used for CFA analyses was 235. There is no consensus on the ideal sample size for CFA and different authors recommend different sample sizes as adequate or minimal. For example, some researchers recommend that sample sizes smaller than 200 should not be used for SEM analyses while other researchers recommend using sample sizes of at least 100 and 200 cases (Worthington &Whittaker, 2006), which the study's sample size fits. However, future studies should strive to follow-up with a larger sample size because smaller sample sizes are known to decrease power and parameter stability.

Regarding the research context, the participants in this study came from various online courses that they took as a student enrolled in various majors or degrees at various institutions in the United States, which is desirable to enhance generalizability of the findings (Gall et al., 2007). However, the problem is that none of these different research contexts (i.e., course, program, university) was represented by a sufficient number of participants in this study, raising some validity issues (Gall et al., 2007). Therefore, future studies should test the reliability and validity of the OEHE instrument across such diverse contexts but with sufficient samples of each particular online learning context.

Limitation Regarding Item Correlations. Item correlations obtained in this study may emanate from sampling error or other reasons including similar wordings of items (Holmes, 2018; Warner, 2013). To reduce, if not eliminate, this potential limitation, a multi-stage fully blown scale development process was used where a thorough review of the relevant literature and the existing measures of student engagement was conducted as well as using a panel of experts to evaluate content and face validity of the items including their language and wording characteristics. As part of this systematic scale development process, pilot cognitive interviews were also conducted with potential respondents who read and responded to the OEHE items and evaluated their relevance, validity, and meaningfulness.

Limitations Regarding Likert Data, Data Analysis Method, and Software. A 5-point Likert-type scale was used in this study. Likert-type scales commonly used in social science research are widely regarded and treated as interval scales although there is still a long-standing debate in the relevant literature as to whether Likert-scales should be treated as interval or ordinal scales (Wu & Leung, 2017). A five-point Likert scale was used in this study and treated as an interval scale and the indicator variables as continuous because five levels often yield a response scale that behaves reasonably close to continuous although increasing the number of points to as high as eleven to make it a continuous measure is also recommended in the literature (Wu & Leung, 2017). The OEHE indicator variables in this study were also treated as continuous and the ML estimation was used with the main assumption that needs to be met being multivariate normality of the endogenous variables in the CFA model. Given that the ML estimation method has been shown to be robust to mild violations of multivariate normality (Jackson et al., 2009), the limitation associated with the CFA estimation method used in this study was also reduced, if not totally eliminated. Finally, IBM SPSS Amos Graphics 27 was used as the software to CFA-analyze the Likert-type data treated as continuous indicators. The software used in this study is known to have limitations regarding the correction procedures that can be used to address multivariate normality and Bollen-Stine bootstrap is the only option provided by Amos for evaluating the overall model fit and individual parameters providing the bias-corrected confidence intervals. This potential limitation of multivariate nonnormal data was therefore mitigated by using the bootstrap procedures provided by Amos as well. Also, an exploratory factor analysis (EFA) might need to be considered with a new sample to explore the OEHE items more thoroughly before conducting a confirmatory analysis to test and confirm the three-factor structure.

Limitation Regarding Online Engagement as Measured by OEHE Dimensions and Items. One major limitation of this study was that the results from the OEHE instrument regarding online student engagement were limited to the behavioral, emotional, and cognitive characteristics as measured by the remaining indicator variables in the final best-fitting 20-item OEHE model. During the item construction process, several indicators of behavioral, emotional, and cognitive engagement were constructed in line with Fredricks et al.'s (2004) three-dimensional conception of engagement and based on the relevant literature and existing measures of engagement. However, certain items from each hypothesized subscale were lost due to poor inter-item correlations, extremely nonnormal distributions, and the poor model fit (local and global) fit indices during the CFA, as thoroughly presented and explained in Chapter Four Findings. As a result of this loss of items, indicators of online behavioral engagement were limited to student-student and student-instructor interactions during online discussions only. Online emotional engagement indicators were also limited to the positive emotions of enjoyment and curiosity, with enjoyment being the most dominant emotion in this subscale. Also, emotional engagement was limited to students' emotional engagement associated with their interactions during online discussions only. Finally, online cognitive engagement indicators were more inclusive and representative of the originally intended content domain of online cognitive engagement with indicators of cognitive engagement associated with both online discussions and learning experience in the online course overall.

The most probable reason to account for these limitations in the resulting OEHE instrument is associated with the sampling strategy used in this study. Through snowball convenience sampling, the data were collected for the OEHE items in this study from multiple online courses, multiple programs, and multiple institutions with extremely small sample representation of each, and each case was embedded in a unique context with multiple unique features, expectations, structures, and other specific characteristics as evidenced by the responses to the OCC questions and the participant's qualitative comments about the OEHE items.

In this study, more than 100 online courses of different subject domains were reported by the participants enrolled in more than 50 different programs or academic disciplines at 51 higher education institutions. This sampling strategy was used to increase generalizability and external validity of the findings (Gall et al., 2007; Holmes, 2018). Using snowball sampling is also a good strategy to ensure high diversity in the study sample with better representativeness (Warner, 2013). However, considering that each online course or a program delivered at a specific higher education institution is one specific context with certain online learning components, requirements, and characteristics, not limiting the sample to a particular course or courses in a particular program or programs at a particular institution or institutions may have resulted in some validity issues because of too much variation and dissimilarity across various online courses, programs, and universities (Gall et al., 2007).

Given that engagement is a malleable state highly sensitive to specific contextual characteristics (Manwaring et al., 2017) and that it is a complex construct that "results from an interaction of the individual with the context and is responsive to variations in contextual characteristics" (Wang et al., 2016, p. 17), participants from such diverse online learning contexts and experiences are very likely to have interpreted and responded to the OEHE items in completely different ways, especially in terms of those online student behaviors such as studying the learning materials, completing tasks, participating in learning activities online or asking for help.

Considering also that there is some research evidence indicating the domain specificity of the construct of student engagement (Wang et al., 2016), the variations resulting from so many different subject domains in this study might have also influenced the participants' responses to the OEHE items. As can be understood from the participant's responses to the OCC questions and their open-ended comments, there was great diversity in the online learning experiences of the participants in this study, which is very likely to have yielded some validity issues regarding each dimension of the OEHE items. This inconsistency in contextual factors might also explain why *pride* and *hope* emotions in the subscale of online emotional engagement could not pass the reliability and validity tests in the sample of this study.

For all these reasons, a disclaimer should be noted that the current OEHE instrument developed and validated in this study is intended to be used to measure student engagement in online learning environments where asynchronous online discussions are the main component of online courses in which there are regular and structured student-student and student-instructor interactions during online discussions regarding the course content. Since the current OEHE instrument developed and validated in this dissertation study does not cover an exhaustive list of indicators of online student engagement, the results of the study should be interpreted with caution as the OEHE indicators may not be generalizable to all online learning contexts and settings. However, it should also be noted that such limitations resulting from measuring online engagement by a certain set of variables are similarly seen in similar instruments developed to measure online student engagement (e.g., Hoi & Le Hang, 2021). Finally, additional dimensions of student engagement such as social engagement shown empirically by other online engagement scale developers (e.g., Hoi & Le Hang, 2021) might need to be considered as a fourth dimension of engagement in online learning environments.

Limitation Regarding Data Collection during the Pandemic. The OEHE instrument and the other surveys and questions used in this study were all administered during a pandemic time when COVID-19 was still a major public health issue which profoundly influenced both traditional and online education at higher education institutions all over the world (Johnson et al., 2020). The pandemic-specific stress or other related challenges experienced by the participants may have influenced their responses to the OEHE items as well as the entire online survey.

Recommendations for Future Research

The development and validation of the OEHE in this study is an important step to systematically applying Fredrick et al.'s (2004) three-dimensional conception of engagement to online learning environments through a fully blown multi-stage scale development process. Therefore, the findings of this study contribute to our understanding of whether and to what extent Fredricks et al.'s (2004) behavioral, emotional, and cognitive engagement subdimensions would actually work in online learning environments. However, future validation studies of the OEHE instrument should continue to test the reliability and validity of the instrument. There are several suggestions that can be made to address the limitations of the current study since it is certainly necessary to continue research on the OEHE reliability and validity with follow-up studies. Below are suggestions for future research to make the OEHE instrument a more powerful and applicable instrument to measure online student engagement.

Diversity in Recruitment. Future validation studies of OEHE should recruit more male participants and participants of other gender identities, more undergraduates, and more participants of other races/ethnicities to reduce gender bias, education-level bias, and racial/ethnicity bias. In addition, the large proportion of White and female participants in this study likely influenced the proportion of gender and ethnicity across responses to inclusiveness, diversity, and accessibility questions. However, this does not change the fact that those White and female students that participated in this study still reported that they perceived bias against them in their online courses, which also calls for further scientific explorations in future studies.

Method, Data, and Analysis. The sample size of the current study was 235, which met several sample size guidelines given in the relevant CFA/SEM research literature. However, future and/or follow-up studies should strive to retain a much larger sample size and run a more complete set of data to obtain parameter estimates with more power and higher parameter stability. In this study, a five-point Likert scale was used, treating its data as interval and continuous. Future and/or follow-up studies can analyze the OEHE data as ordered categorical, using Diagonally Weighted Least Squares (DWLS), a CFA estimation method which is "specifically designed for ordinal data" (Li, 2016, p. 936) to see if different results would be obtained. In addition, the OEHE indicator variables were treated in this study as continuous, using the ML estimation with the main assumption that needs to be met being multivariate normality. Since Amos is known to be limited in the options available for dealing with the violation of multivariate normality, with the only option being Bollen-Stine bootstrap, future and/or follow-up studies are strongly recommended to use a more robust procedure such as Lavaan using R (Rosseel, 2012), where there is the option for Satorra-Bentler correction along with robust fit statistics, and parameter tests involving robust standard errors. Moreover, although the aim of this dissertation study was to test a theory and for that reason using CFA was justified, an exploratory factor analysis (EFA) might need to be considered with a new sample to explore the OEHE items more thoroughly before conducting a confirmatory analysis to test and confirm the three-factor structure. Furthermore, in this study, the sample was drawn from multiple institutions, multiple programs, and multiple courses, and so multiple contexts with each having an extremely small sample representation, yielding some validity issues regarding certain OEHE items not included in the final OEHE instrument. Therefore, future and/or follow-up studies should use a delimiter to limit the sample to a particular set of courses within a particular program at a particular higher education institution in order to ensure that all the participants in the sample have the same or very similar online learning experiences directly relevant to what the OEHE instrument was originally intended to measure as indicators of online student engagement. Limiting the sample to one such specific and directly OEHE-relevant context would help minimize confounding variability across courses and programs and decrease the validity issues experienced in the current sample of this study (Gall et al., 2007). Finally, to support the understanding and interpretation of the self-report data regarding students' behavioral engagement, learning analytics data (e.g., login frequency, number of posts) might also be considered for analysis and triangulation purposes.

Validity. In this study, content validity and face validity were established through expert reviews and pilot cognitive interviews, and construct validity was established through evidence of factorial/structural validity obtained in the CFAs and through reasonable evidence of convergent, discriminant, and criterion validity. Other dimensions of validity including consequential validity need to be established by future and/or follow-up studies. In addition, another measure of online engagement can be selected to use it for convergent validity purposes since Reeve's (2013) online-modified engagement scales used in this study for this purpose may not have adequately fit the online learning context. Regarding diversity, equity, and inclusion (DEI) validity issues that need to be considered in any instrument development process, serious attention should be given to the critical question of how instrument development researchers can ensure that their instruments are inclusive of all people and all social and cultural backgrounds and groups so that their instruments can help in providing inclusive, diverse, and representative assessments of individuals across diverse contexts (Cintron & Hagan, 2021). Through this dissertation study with empirical evidence about the participants' perceptions of inclusiveness, bias, and accessibility, the aim was to raise awareness about DEI validity issues that have been traditionally ignored in the instrument development process. Much work still remains to be done to ensure that DEI validity is fully considered during the instrument development process as well as other traditional forms of validity (Cintron & Hagan, 2021).

Timing of Data Collection. The OEHE data were collected from the students enrolled in a higher education institution in the United States during a pandemic time, namely the COVID-19 pandemic and this stressful and extraordinary period may have biased their responses to the OEHE items. Future and/or follow-up studies are strongly encouraged to test the reliability and validity of the OEHE instrument when the pandemic conditions are over to see whether different results would be obtained.

Although the present results support the tenability of a three-factor model of student engagement in online learning environments, the most important contribution of this dissertation study may be that it has raised a variety of intriguing questions for future studies. In terms of future research directions, it would be useful to extend the current findings by examining the issues indicated previously. Much work remains to be done before a full understanding of the extent to which Fredricks et al.'s (2004) conception of student engagement can be applied to online learning environments.

One noteworthy strength of this dissertation study was the solicitation and use of qualitative and open-ended data. The qualitative comments of the participants regarding their specific online learning contexts helped explain great contextual variations and inconsistencies across their responses, especially the OEHE items intended to assess behavioral engagement. If the qualitative and open-ended data had not been collected from the participants to learn more deeply about their unique online learning contexts, the quantitative data regarding the poor interitem correlations and the poor performances of the behavioral engagement items would not have been understood thoroughly. Therefore, future studies should seriously consider collecting such qualitative and open-ended data from the participants regarding their unique online learning contexts so that the quantitative data collected for reliability and validity can be interpreted accurately and meaningfully.

Conclusion

The goal of this dissertation study was to develop and validate an instrument that would assess student engagement in online learning environments of higher education, namely undergraduate and graduate-level online courses. This goal was accomplished through a systematic scale development and testing process with multiple stages and steps. The results obtained from this study indicate that the OEHE is a reliable and valid assessment tool that can be used to measure undergraduate and graduate students' behavioral, emotional, and cognitive engagement in online courses where asynchronous online discussions are a major component of student learning experiences. This study is significant and valuable in that it provided further empirical evidence for the applicability of Fredrick et al.'s (2004) three core dimensions of engagement within online learning environments by addressing the limitations of the existing measures of online student engagement in the extant literature. The limitations revealed in this study need to be addressed and fixed in subsequent follow-up studies with larger, more context relevant, and more criteria-fitting samples of online undergraduate and graduate students. Given that asynchronous discussions have an important place in online learning environments of higher education (Hew et al., 2010), this dissertation study may open the doors to further lines of inquiry regarding the applicability of Fredricks et al.'s (2004) three-dimensional student engagement within those online learning environments where online discussion boards and forums are used by instructors to foster and maintain student engagement and students are largely involved in interactions with their peers, their instructor, and the course content via online discussions (Salter & Conneely, 2015).

References

Ainley, M. (2012). Students' interest and engagement in classroom activities. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 283–302). Springer.

Allen, I. E., & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States*. Babson Survey Research Group and Quahog Research Group. https://files.eric.ed.gov/fulltext/ED541571.pdf

Allen, I. E., & Seaman, J. (2017). *Digital learning compass: Distance education enrollment report 2017*. Babson Survey Research Group, e-Literate, and WCET. https://files.eric.ed.gov/fulltext/ED580868.pdf

Allen, I. E., Seaman, J., Poulin, R., & Straut, T. T. (2016). *Online report card: Tracking online education in the United States*. Babson Survey Research Group and Quahog Research Group. https://files.eric.ed.gov/fulltext/ED572777.pdf

Almutairi, F., & White, S. (2018). How to measure student engagement in the context of blended-MOOC. *Interactive Technology and Smart Education*, *15*(3), 262–278. https://doi.org/10.1108/ITSE-07-2018-0046

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, *80*(3), 260–267. https://doi.org/10.1037/0022-0663.80.3.260

Anderson, E. (2017). *Measurement of online student engagement: Utilization of continuous online student behavior indicators as items in a partial credit Rasch model* (Publication No. 10259450) [Doctoral dissertation, University of Denver]. ProQuest Dissertations & Theses Global.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Motivation*, *44*(5), 427–445. https://doi.org/10.1016/j.jsp.2006.04.002

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, *45*(5), 369–386. https://doi.org/10.1002/pits.20303

Arbuckle, J. L. (2020). *IBM® SPSS® Amos™ 27 user's guide*. IBM.

Archambault, I., Janosz, M., Morizot, J., & Pagani, L. (2009). Adolescent behavioral, affective, and cognitive engagement in school: Relationship to dropout. *Journal of School Health*, *79*(9), 408–415. https://doi.org/10.1111/j.1746-1561.2009.00428.x

Artino, A. R. (2010). Online or face-to-face learning? Exploring the personal factors that predict students' choice of instructional format. *Internet and Higher Education*, *13*(4), 272–276. https://doi.org/10.1016/j.iheduc.2010.07.005

Artino, A. R., & Jones, K. D. (2012). Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *Internet and Higher Education*, *15*(3), 170–175. https://doi.org/10.1016/j.iheduc.2012.01.006

Asempapa, R. S. (2016). *Developing an instrument to assess teachers' knowledge of the nature of mathematical modeling and their attitude toward such modeling* (Publication No. 28004791) [Doctoral dissertation, Ohio University]. ProQuest Dissertations & Theses Global.

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel, 25*(4), 297–308.

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, *50*(1), 84–94.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497. https://doi.org/10.1037/0033-2909.117.3.497

Bentler, P. M. (2005). *EQS 6 Structural equations program manual*. Multivariate Software.

Bigatel, P. M., & Edel-Malizia, S. (2018). Using the "Indicators of Engaged Learning Online" framework to evaluate online course quality. *TechTrends*, *62*(1), 58–70.

Bigatel, P., & Williams, V. (2015). Measuring student engagement in an online program. *Online Journal of Distance Learning Administration*, *18*(2). https://www.westga.edu/~distance/ojdla/summer182/bigatel_williams182.html

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology*, *35*(1), 61–79. https://doi.org/10.1016/S0022-4405(96)00029-5

Bolliger, D. U., & Martin, F. (2018). Instructor and student perceptions of online student engagement strategies. *Distance Education*, *39*(4), 568–583. https://doi.org/10.1080/01587919.2018.1520041

Bolliger, D. U., & Martin, F. (2021). Factors underlying the perceived importance of online student engagement strategies. *Journal of Applied Research in Higher Education*, *13*(2), 404–419. https://doi.org/10.1108/JARHE-02-2020-0045

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education*, *17*(1), 2. https://doi.org/10.1186/s41239-019-0176-8

Bonny, A. E., Britto, M. T., Klostermann, B. K., Hornung, R. W., & Slap, G. B. (2000). School disconnectedness: Identifying adolescents at risk. *Pediatrics*, *106*(5), 1017–1021. https://doi.org/10.1542/peds.106.5.1017

Borup, J., Graham, C. R., & Velasquez, A. (2011). The use of asynchronous video communication to improve instructor immediacy and social presence in a blended learning environment. In A. Kitchenham (Ed.), *Blended learning across disciplines: Models for implementation* (pp. 38–57). IGI Global.

Borup, J., West, R. E., & Graham, C. R. (2012). Improving online social presence through asynchronous video. *Emotions in Online Learning Environments*, *15*(3), 195–203. https://doi.org/10.1016/j.iheduc.2011.11.001

Boston, W. E., & Ice, P. (2011). Assessing retention in online learning: An administrative perspective. *Online Journal of Distance Learning Administration*, *14*(2). https://www.westga.edu/~distance/ojdla/summer142/boston_ice142.html

Brault-Labbe, A., & Dube, L. (2010). Engagement scolaire, bien-être personnel et autodétermination chez des étudiants à l'université. *Canadian Journal of Behavioural Science*, *42*(2), 80–92.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Buelow, J. R., Barry, T., & Rich, L. E. (2018). Supporting learning engagement with online students. *Online Learning*, *22*(4), 313–340. https://doi.org/10.24059/olj.v22i4.1384

Byrne, B. M. (2005). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of Personality Assessment*, *85*(1), 17–32. https://doi.org/10.1207/s15327752jpa8501_02

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge.

Byrne, B. M. (2016). *Structural equation modeling with AMOS : Basic concepts, applications, and programming* (3rd ed.). Routledge.

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

Caskurlu, S. (2018). Confirming the subdimensions of teaching, social, and cognitive presences: A construct validity study. *The Internet and Higher Education*, *39*, 1–12. https://doi.org/10.1016/j.iheduc.2018.05.002

Caskurlu, S., Richardson, J. C., Maeda, Y., & Kozan, K. (2021). The qualitative evidence behind the factors impacting online learning experiences as informed by the community of inquiry framework: A thematic synthesis. *Computers & Education*, *165*, 104111. https://doi.org/10.1016/j.compedu.2020.104111

Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students' LMS interaction patterns and their relationship with achievement: A case study in higher education. *Computers & Education*, *96*, 42–54. https://doi.org/10.1016/j.compedu.2016.02.006

Chen, P.-S. D., Lambert, A. D., & Guidry, K. R. (2010). Engaging online learners: The impact of web-based learning technology on college student engagement. *Computers & Education*, *54*(4), 1222–1232. https://doi.org/10.1016/j.compedu.2009.11.008

Cheng, C. K., Paré, D. E., Collimore, L.-M., & Joordens, S. (2011). Assessing the effectiveness of a voluntary online discussion forum on improving students' course performance. *Computers & Education*, *56*(1), 253–261. https://doi.org/10.1016/j.compedu.2010.07.024

Chin, C.-L., & Yao, G. (2014). Convergent validity. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 1275–1276). Springer Netherlands. https://doi.org/10.1007/978-94-007-0753-5_573

Christenson, S. L., Reschly, A. L., & Wylie, C. (Eds.). (2012). *Handbook of research on student engagement*. Springer.

Cintron, D.W., & Hagan, E. (2021). *Diversity, equity, and inclusion validity arguments* [White paper]. Robert Wood Johnson Foundation, University of California, San Francisco. https://tinyurl.com/4ak7v5xb

Cleary, T. J., & Zimmerman, B. J. (2012). A cyclical self-regulatory account of student engagement: Theoretical foundations and applications. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 237–258). Springer.

Coates, H. (2005). The value of student engagement for higher education quality assurance. *Quality in Higher Education*, *11*(1), 25–36. https://doi.org/10.1080/13538320500074915

Coates, H. (2007). A model of online and general campus-based student engagement. *Assessment & Evaluation in Higher Education*, *32*(2), 121–141. https://doi.org/10.1080/02602930600801878

Connell, J. P., & Wellborn, J. G. (1991). *Competence, autonomy, and relatedness: A motivational analysis of self-system processes.* In M. R. Gunnar & L. A. Sroufe (Eds.), *The Minnesota symposia on child psychology, Vol. 23. Self processes and development* (p. 43–77). Lawrence Erlbaum Associates, Inc.

Corno, L., & Mandinach, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist*, *18*(2), 88–108. https://doi.org/10.1080/00461528309529266

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the motivated strategies for learning questionnaire. *Learning and Individual Differences*, *21*(4), 337–346. https://doi.org/10.1016/j.lindif.2011.03.002

Creswell, J. W. (2005). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (2nd ed.). Pearson.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Czerkawski, B., & Lyman, E. (2016). An instructional design framework for fostering student engagement in online learning environments. *TechTrends: Linking Research & Practice to Improve Learning*, *60*(6), 532–539. https://doi.org/10.1007/s11528-016-0110-z

Daniels, L. M., & Stupnisky, R. H. (2012). Not that different in theory: Discussing the control-value theory of emotions in online learning environments. *Internet and Higher Education*, *15*(3), 222–226. https://doi.org/10.1016/j.iheduc.2012.04.002

DeBacker, T. K., Crowson, H. M., Beesley, A. D., Thoma, S. J., & Hestevold, N. L. (2008). The challenge of measuring epistemic beliefs: An analysis of three self-report instruments. *Journal of Experimental Education*, *76*(3), 281–312. https://doi.org/10.3200/JEXE.76.3.281-314

Deng, R., Benckendorff, P., & Gannaway, D. (2020). Learner engagement in MOOCs: Scale development and validation. *British Journal of Educational Technology*, *51*(1), 245–262. https://doi.org/10.1111/bjet.12810

Dennen, V. P. (2008). Pedagogical lurking: Student engagement in non-posting discussion behavior. *Computers in Human Behavior*, *24*(4), 1624–1633. https://doi.org/10.1016/j.chb.2007.06.003

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Sage.

DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd ed.). Sage.

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.

Ding, L., Kim, C., & Orey, M. (2017). Studies of student engagement in gamified online discussions. *Computers & Education*, *115*, 126–142. https://doi.org/10.1016/j.compedu.2017.06.016

DiStefano, C., & Hess, B. (2005). Using confirmatory factor analysis for construct validation: An empirical review. *Journal of Psychoeducational Assessment*, *23*(3), 225–241. https://doi.org/10.1177/073428290502300303

Dixson, M. D. (2010). Creating effective student engagement in online courses: What do students find engaging? *Journal of the Scholarship of Teaching and Learning*, *10*(2), 1–13.

Dixson, M. D. (2015). Measuring student engagement in the online course: The Online Student Engagement scale (OSE). *Online Learning*, *19*(4), 1–15. https://doi.org/10.24059/olj.v19i4.561

Dixson, M. D., Greenwell, M. R., Rogers-Stacy, C., Weister, T., & Lauer, S. (2017). Nonverbal immediacy behaviors and online student engagement: Bringing past instructional research into the present virtual classroom. *Communication Education*, *66*(1), 37–53. https://doi.org/10.1080/03634523.2016.1209222

Duncan, T. G., & McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist*, *40*(2), 117–128.

Eisinga, R., Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, *58*(4), 637–642. https://doi.org/10.1007/s00038-012-0416-3

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, *54*(1), 5–12. https://doi.org/10.1037//0022-3514.54.1.5

Entwistle, N. J., & Entwistle, D. (1970). The relationships between personality, study methods and academic performance. *British Journal of Educational Psychology*, *40*(2), 132–143. https://doi.org/10.1111/j.2044-8279.1970.tb02113.x

Everett, D. R. (2015). Adding value: Online student engagement. *Information Systems Education Journal*, *13*(6), 68–76.

Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 56–83. https://doi.org/10.1080/10705519909540119

Ferrer, J., Ringer, A., Saville, K., A Parris, M., & Kashi, K. (2020). Students' motivation and engagement in higher education: The importance of attitude to online learning. *Higher Education (00181560)*, *83*(2), 317–338. https://doi.org/10.1007/s10734-020-00657-5

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, *59*(2), 117–142. https://doi.org/10.2307/1170412

Finn, J. D. (1993). *School engagement & Students at risk.* (ED 362 322; pp. 1–116). National Center for Education Statistics (ED). http://libraries.ou.edu/access.aspx?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED362322&site=ehost-live

Finn, J. D., Pannozzo, G. M., & Voelkl, K. E. (1995). Disruptive and inattentive-withdrawn behavior and achievement among fourth graders. *The Elementary School Journal*, *95*(5), 421–434. https://doi.org/10.1086/461853

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology*, *82*(2), 221–234. https://doi.org/10.1037/0021-9010.82.2.221

Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In S. L. Christenson & A. L. Reschly (Eds.), *Handbook of research on student engagement* (pp. 97–131). Springer.

Fletcher, E. C., D. Dumford, A., Hernandez-Gantes, V. M., & Minar, N. (2020). Examining the engagement of career academy and comprehensive high school students in the United States. *The Journal of Educational Research*, *113*(4), 247–261. https://doi.org/10.1080/00220671.2020.1787314

Fowler, F. J. (2014). *Survey research methods* (5th ed.). Sage.

Fredricks, J. A. (2015). Academic engagement. In J. D. Wright (Ed.), *The international encyclopedia of social and behavioral sciences* (2nd ed., Vol. 1, pp. 31–36). http://dx.doi.org/10.1016/B978-0-08-097086-8.26085-6

Fredricks, J. A., Blumenfeld, P. C., Friedel, J., & Paris, A. (2005). School engagement. In K. A. Moore & L. Lippman (Eds.), *What do children need to flourish?: Conceptualizing and measuring indicators of positive development* (pp. 305–321).

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. https://doi.org/10.3102/00346543074001059

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction*, *43*, 1–4. https://doi.org/10.1016/j.learninstruc.2016.02.002

Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report measures. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 763–782). Springer.

Fredricks, J. A., Wang, M.-T., Schall Linn, J., Hofkens, T. L., Sung, H., Parr, A., & Allerton, J. (2016). Using qualitative methods to develop a survey measure of math and science engagement. *Special Issue: Student Engagement and Learning: Theoretical and Methodological Advances*, *43*, 5–15. https://doi.org/10.1016/j.learninstruc.2016.01.009

Furlong, M. J., Whipple, A. D., St. Jean, G., Simental, J., Soliz, A., & Punthuna, S. (2003). Multiple contexts of school engagement: Moving toward a unifying framework for educational research and practice. *The California School Psychologist*, *8*(1), 99–113. https://doi.org/10.1007/BF03340899

Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. SAGE Publications.

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, *95*(1), 148-162.

Gall, J.P., Gall, M.D., & Borg, W.R. (2007). *Educational research: An introduction* (8th ed.). Pearson Education, Inc.

Gibbs, G. (2014). Student engagement, the latest buzzword. *Higher Education Times*. https://www.timeshighereducation.com/news/student-engagement-the-latestbuzzword/2012947.article

Gillen-O'Neel, C., & Fuligni, A. (2013). A longitudinal study of school belonging and academic motivation across high school. *Child Development*, *84*(2), 678–692. https://doi.org/10.1111/j.1467-8624.2012.01862.x

Glanville, J. L., & Wildhagen, T. (2007). The measurement of school engagement: Assessing dimensionality and measurement invariance across race and ethnicity. *Educational and Psychological Measurement*, *67*(6), 1019–1041. https://doi.org/10.1177/0013164406299126

Goggins, S., & Xing, W. (2016). Building models explaining student participation behavior in asynchronous online discussion. *Computers & Education*, *94*, 241–251. https://doi.org/10.1016/j.compedu.2015.11.002

Goodenow, C., & Grady, K. E. (1993). The relationship of school belonging and friends' values to academic motivation among urban adolescent students. *The Journal of Experimental Education*, *62*(1), 60–71.

Greene, B. A., & Miller, R. B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology*, *21*(2), 181–192. https://doi.org/10.1006/ceps.1996.0015

Groccia, J. E. (2018). What is student engagement? *New Directions for Teaching & Learning*, *2018*(154), 11–20. https://doi.org/10.1002/tl.20287

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Wiley.

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, *90*, 36–53. https://doi.org/10.1016/j.compedu.2015.09.005

Hew, K. F., & Cheung, W. S. (2008). Attracting student participation in asynchronous online discussions: A case study of peer facilitation. *Computers & Education*, *51*(3), 1111–1124.

Hew, K. F., Cheung, W. S., & Ng, C. S. L. (2010). Student contribution in asynchronous online discussion: A review of the research and empirical exploration. *Instructional Science*, *38*(6), 571–606. https://doi.org/10.1007/s11251-008-9087-0

Hilpert, J. C., Stempien, J., van der Hoeven Kraft, K. J., & Husman, J. (2013). Evidence for the latent factor structure of the MSLQ: A new conceptualization of an established questionnaire. *SAGE Open*, *3*(4), 1–10. https://doi.org/10.1177/2158244013510305

Hoi, V. N., & Le Hang, H. (2021). The structure of student engagement in online learning: A bi-factor exploratory structural equation modelling approach. *Journal of Computer Assisted Learning*, *37*(4), 1141–1153. https://doi.org/10.1111/jcal.12551

Holmes, J. L. (2018). *Integration of doctoral students in distance programs: An instrument validation study of educational doctorate students* (Publication No. 10784233) [Doctoral dissertation, Liberty University]. ProQuest Dissertations & Theses Global.

Hsu, H.-C. K., Wang, C. V., & Levesque-Bristol, C. (2019). Reexamining the impact of self-determination theory on learning outcomes in the online learning environment. *Education and Information Technologies*, *24*(3), 2159–2174. https://doi.org/10.1007/s10639-019-09863-w

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

Hubley, A. M. (2014). Discriminant Validity. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 1664–1667). Springer Netherlands. https://doi.org/10.1007/978-94-007-0753-5_751

Imlawi, J., Gregg, D., & Karimi, J. (2015). Student engagement in course-based social networks: The impact of instructor credibility and use of communication. *Computers & Education*, *88*, 84–96. https://doi.org/10.1016/j.compedu.2015.04.015

Jackson, C. R. (2018). Validating and adapting the Motivated Strategies for Learning Questionnaire (MSLQ) for STEM courses at an HBCU. *AERA Open*, *4*(4), 1–16. https://doi.org/10.1177/2332858418809346

Jackson, D. L., Gillaspy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, *14*(1), 6-23. https://doi.org/ 10.1037/a0014694

Jimerson, S. R., Campos, E., & Greif, J. L. (2003). Toward an understanding of definitions and measures of school engagement and related terms. *California School Psychologist*, *8*, 7–27.

Johnson, N., Veletsianos, G., & Seaman, J. (2020). U.S. faculty and administrators' experiences and approaches in the early weeks of the COVID-19 pandemic. *Online Learning*, *24*(2), 6–21. https://doi.org/10.24059/olj.v24i2.2285

Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *The Counseling Psychologist*, *34*(5), 684–718. https://doi.org/10.1177/0011000006286347

Kardash, C. M., & Amlund, J. T. (1991). Self-reported learning strategies and learning from expository text. *Contemporary Educational Psychology*, *16*(2), 117–138. https://doi.org/10.1016/0361-476X(91)90032-G

Kay, R. H. (2006). Developing a comprehensive metric for assessing discussion board effectiveness. *British Journal of Educational Technology*, *37*(5), 761–783. https://doi.org/10.1111/j.1467-8535.2006.00560.x

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(3), 333–351. https://doi.org/10.1207/S15328007SEM1003_1

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, *74*(7), 262–273. https://doi.org/10.1111/j.1746-1561.2004.tb08283.x

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed). The Guilford Press.

Kokoç, M., Akçapınar, G., & Hasnine, M. N. (2021). Unfolding students' online assignment submission behavioral patterns using temporal learning analytics. *Journal of Educational Technology & Society*, *24*(1), 223–235.

Kozan, K., & Richardson, J. C. (2014). Interrelationships between and among social, teaching, and cognitive presence. *The Internet and Higher Education*, *21*, 68–73. https://doi.org/10.1016/j.iheduc.2013.10.007

Kucuk, S., & Richardson, J. C. (2019). A structural equation model of predictors of online learners' engagement and satisfaction. *Online Learning*, *23*(2), 196–216. https://doi.org/10.24059/olj.v23i2.1455

Kuh, G. D. (2003). What we're learning about student engagement from NSSE. *Change*, *35*(2), 24.

Kuh, G. D. (2009). The National Survey of Student Engagement: Conceptual and empirical foundations. *New Directions for Institutional Research*, *2009*(141), 5–20. https://doi.org/10.1002/ir.283

Kuh, G. D. (2016). Making learning meaningful: Engaging students in ways that matter to them. *New Directions for Teaching & Learning*, *2016*(145), 49–56. https://doi.org/10.1002/tl.20174

Kumi, R., & Sabherwal, R. (2019). Knowledge sharing behavior in online discussion communities: Examining behavior motivation from social and individual perspectives. *Knowledge & Process Management*, *26*(2), 110–122. https://doi-org.ezproxy.lib.ou.edu/10.1002/kpm.1574

Lear, J. L., Ansorge, C., & Steckelberg, A. (2010). Interactivity/Community process model for the online education environment. *MERLOT Journal of Online Learning and Teaching*, *6*(1), 71–77.

Lee, J.-E., & Recker, M. (2021). The effects of instructors' use of online discussions strategies on student participation and performance in university online introductory mathematics courses. *Computers & Education*, *162*, 104084. https://doi.org/10.1016/j.compedu.2020.104084

Lee, J., Song, H.-D., & Hong, A. J. (2019). Exploring factors, and indicators for measuring students' sustainable engagement in e-learning. *Sustainability*, *11*(4). https://doi.org/10.3390/su11040985

Leung, S.-O. (2011). A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point Likert scales. *Journal of Social Service Research*, *37*(4), 412–421. https://doi.org/10.1080/01488376.2011.580697

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Li, Y., & Lerner, R. (2013). Interrelations of behavioral, emotional, and cognitive school engagement in high school students. *Journal of Youth & Adolescence*, *42*(1), 20–32. https://doi.org/10.1007/s10964-012-9857-5

Lin, C.-H., Zhang, Y., & Zheng, B. (2017). The roles of learning strategies and motivation in online language learning: A structural equation modeling analysis. *Computers & Education*, *113*, 75–85. https://doi.org/10.1016/j.compedu.2017.05.014

Linder, K. E., Fontaine-Rainen, D. L., & Behling, K. (2015). Whose job is it? Key challenges and future directions for online accessibility in US institutions of higher education. *Open Learning*, *30*(1), 21–34. https://doi.org/10.1080/02680513.2015.1007859

Litman, J. A., & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of Personality Assessment*, *80*(1), 75–86. https://doi.org/10.1207/S15327752JPA8001_16

Lomax, R. G., & Hahs-Vaughn, D. L. (2012). *An introduction to statistical concepts* (3rd ed.). Routledge/Taylor & Francis Group.

Maehr, M. L. (1984). Meaning and motivation: Toward a theory of personal investment. In R. Ames & C. Ames (Eds.), *Research on motivation in education: Student motivation* (Vol. 1, pp. 115–144). Academic Press.

Manwaring, K. C., Larsen, R., Graham, C. R., Henrie, C. R., & Halverson, L. R. (2017). Investigating student engagement in blended learning settings using experience sampling and structural equation modeling. *The Internet and Higher Education*, *35*, 21–33. https://doi.org/10.1016/j.iheduc.2017.06.002

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519–530.

Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya*, *B36*, 115–128.

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, *37*(1), 153–184. https://doi.org/10.2307/1163475

Martin, F., & Bolliger, D. U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning*, *22*(1), 205–222.

Martin, F., Stamper, B., & Flowers, C. (2020). Examining student perception of readiness for online learning: Importance and confidence. *Online Learning*, *24*(2), 38–58. ERIC. https://doi.org/10.24059/olj.v24i2.2053

Martin, F., Bolliger, D. U., & Flowers, C. (2021). Design matters: Development and validation of the Online Course Design Elements (OCDE) instrument. *The International Review of Research in Open and Distributed Learning*, *22*(2), 46–71. https://doi.org/10.19173/irrodl.v22i2.5187

Martin, F., Budhrani, K., Kumar, S., & Ritzhaupt, A. (2019). Award-winning faculty online teaching practices: Roles and competencies. *Online Learning*, *23*(1), 184–205. https://doi.org/10.24059/olj.v23i1.1329

Martin, F., Sun, T., & Westine, C. D. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & Education*, *159*, 104009. https://doi.org/10.1016/j.compedu.2020.104009

Messick, S. (1989). *Validity.* In R. L. Linn (Ed.), *The American Council on Education/Macmillan series on higher education. Educational measurement* (p. 13–103). Macmillan Publishing Co, Inc; American Council on Education.

Messick, S. (1995a). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Messick, S. (1995b). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, *14*(4), 5–8. https://doi.org/10.1111/j.1745-3992.1995.tb00881.x

Meyer, K. A. (2014). Student Engagement in Online Learning: What Works and Why. *ASHE Higher Education Report*, *40*(6), 1–114. https://doi.org/10.1002/aehe.20018

Miller, R. B., Greene, B. A., Montalvo, G. P., Ravindran, B., & Nichols, J. D. (1996). Engagement in academic work: The role of learning goals, future consequences, pleasing others, and perceived ability. *Contemporary Educational Psychology*, *21*(4), 388–422. https://doi.org/10.1006/ceps.1996.0028

Miserandino, M. (1996). Children who do well in school: Individual differences in perceived competence and autonomy in above-average children. *Journal of Educational Psychology*, *88*(2), 203–214.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105*(3), 430–445. https://doi.org/10.1037/0033-2909.105.3.430

Murillo-Zamorano, L. R., López Sánchez, J. Á., & Godoy-Caballero, A. L. (2019). How the flipped classroom affects knowledge, skills, and engagement in higher education: Effects on students' satisfaction. *Computers & Education*, *141*, 103608. https://doi.org/10.1016/j.compedu.2019.103608

Natriello, G. (1984). Problems in the evaluation of students and student disengagement from secondary schools. *Journal of Research and Development in Education*, *17*(4), 14–24.

Newmann, F., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. Newmann (Ed.), *Student engagement and achievement in American secondary schools* (pp. 11–39). Teachers College Press.

Nguyen, T. D., Cannata, M., & Miller, J. (2018). Understanding student behavioral engagement: Importance of student interaction with peers and teachers. *Journal of Educational Research*, *111*(2), 163–174. https://doi.org/10.1080/00220671.2016.1220359

Olivier, E., Galand, B., Hospel, V., & Dellisse, S. (2020). Understanding behavioral engagement and achievement: The roles of teaching practices and student sense of competence and task value. *British Journal of Educational Psychology*, *90*(4), 887–909. https://doi.org/10.1111/bjep.12342

Oncu, S., & Cakir, H. (2011). Research in online learning environments: Priorities and methodologies. *Computers & Education*, *57*(1), 1098–1108. https://doi.org/10.1016/j.compedu.2010.12.009

Ozan, O., & Ozarslan, Y. (2016). Video lecture watching behaviors of learners in online courses. *Educational Media International*, *53*(1), 27–41. https://doi.org/10.1080/09523987.2016.1189255

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (7th ed.). McGraw-Hill Education.

Park, Y., Yu, J. H., & Jo, I.-H. (2016). Clustering blended learning courses by online behavior data: A case study in a Korean higher education institute. *The Internet and Higher Education*, *29*, 1–11. https://doi.org/10.1016/j.iheduc.2015.11.001

Park, S., & Yun, H. (2018). The influence of motivational regulation strategies on online students' behavioral, emotional, and cognitive engagement. *American Journal of Distance Education*, *32*(1), 43–56. https://doi.org/10.1080/08923647.2018.1412738

Parnes, M. F., Suárez-Orozco, C., Osei-Twumasi, O., & Schwartz, S. E. O. (2020). Academic outcomes among diverse community college students: What is the role of instructor relationships? *Community College Review*, *48*(3), 277–302. https://doi.org/10.1177/0091552120909908

Paulsen, J., & McCormick, A. C. (2020). Reassessing disparities in online learner student engagement in higher education. *Educational Researcher*, *49*(1), 20–29. https://doi.org/10.3102/0013189X19898690

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology*, *101*(1), 115–135.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Students' Emotions and Academic Engagement*, *36*(1), 36–48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., & Perry, R. P. (2005). *Academic Emotions Questionnaire (AEQ): User's manual.* University of Munich, Department of Psychology. https://www.researchgate.net/profile/Javier-Sanchez-Rosas/post/How-can-i-apply-and-access-the-Academic-Emotions-Questionnaire-by-Pekrun-et-al-2005/attachment/59d6351c79197b8077992b77/AS%3A383031052390406%401468333127197/download/2005+AEQ+Manual.pdf

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*(2), 91–105. https://doi.org/10.1207/S15326985EP3702_4

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 259–282). Springer.

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The Epistemically-Related Emotion Scales. *Cognition and Emotion*, *31*(6), 1268–1276. https://doi.org/10.1080/02699931.2016.1204989

Picciano, A. G. (2019). *Online education: Foundations, planning, and pedagogy*. Routledge.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. K. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. University of Michigan.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, *53*(3), 801–813. https://doi.org/10.1177/0013164493053003024

Pituch, K. A., & Stevens, J. P. (2016). Structural equation modeling. In *Applied multivariate statistics for the social sciences* (6th ed., pp. 639–733). Routledge.

Putman, S. M., Ford, K., & Tancock, S. (2012). Redefining online discussions: Using participant stances to promote collaboration and cognitive engagement. *International Journal of Teaching and Learning in Higher Education*, *24*(2), 151–167.

Ramos, C., & Yudko, E. (2008). “Hits” (not “Discussion Posts”) predict student success in online courses: A double cross-validation study. *Computers & Education*, *50*(4), 1174–1182. https://doi.org/10.1016/j.compedu.2006.11.003

Redmond, P., Abawi, L.-A., Brown, A., Henderson, R., & Heffernan, A. (2018). An online engagement framework for higher education. *Online Learning*, *22*(1), 183–204. https://doi.org/10.24059/olj.v22i1.1175

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology*, *105*(3), 579–595. https://doi.org/10.1037/a0032690

Reeve, J., Cheon, S. H., & Jang, H. (2020). How and why students make academic progress: Reconceptualizing the student engagement construct to increase its explanatory power. *Contemporary Educational Psychology*, *62*, 101899. https://doi.org/10.1016/j.cedpsych.2020.101899

Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation & Emotion*, *28*(2), 147–169. https://doi.org/10.1023/B:MOEM.0000032312.95499.6f

Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, *36*(4), 257–267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Reindl, M., Tulis, M., & Dresel, M. (2018). Associations between friends, academic emotions and achievement: Individual differences in enjoyment and boredom. *Learning and Individual Differences*, *62*, 164–173. https://doi.org/10.1016/j.lindif.2018.01.017

Reschly, A. L., & Christenson, S. L. (2006). Prediction of dropout among students with mild disabilities: A case for the inclusion of student engagement variables. *Remedial and Special Education*, *27*(5), 276–292. https://doi.org/10.1177/07419325060270050301

Revere, L., & Kovach, J. V. (2011). Online technologies for engaged learning: A meaningful synthesis for educators. *Quarterly Review of Distance Education*, *12*(2), 113–124.

Richardson, J. C., & Newby, T. (2006). The role of students' cognitive engagement in online learning. *American Journal of Distance Education*, *20*(1), 23–37. https://doi.org/10.1207/s15389286ajde2001_3

Rimm-Kaufman, S. E., Baroody, A. E., Larsen, R. A. A., Curby, T. W., & Abry, T. (2015). To what extent do teacher–student interaction quality and student gender contribute to fifth graders' engagement in mathematics learning? *Journal of Educational Psychology*, *107*(1), 170–185. https://doi.org/10.1037/a0037252

Robinson, C. C., & Hullinger, H. (2008). New benchmarks in higher education: Student engagement in online learning. *Journal of Education for Business*, *84*(2), 101–109.

Rosseel, Y. (2012). lavaan: An R package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in personality and social psychology bulletin. *Personality and Social Psychology Bulletin*, *28*(12), 1629–1646. https://doi.org/10.1177/014616702237645

Ryan, R. M., & Deci, E. L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). University of Rochester Press.

Ryan, K., Gannon-Slater, N., & Culbertson, M. J. (2012). Improving survey methods with cognitive interviews in small- and medium-scale evaluations. *American Journal of Evaluation*, *33*(3), 414–430. https://doi.org/10.1177/1098214012441499

Salter, N. P., & Conneely, M. R. (2015). Structured and unstructured discussion forums as tools for student engagement. *Computers in Human Behavior*, *46*, 18–25. https://doi.org/10.1016/j.chb.2014.12.037

Salvo, S. G., Shelton, K., & Welch, B. (2019). African American males learning online: Promoting academic achievement in higher education. *Online Learning*, *23*(1), 22–36.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8*(2), 23–74.

Schmeck, R. R., Ribich, F., & Ramanaiah, N. (1977). Development of a self-report inventory for assessing individual differences in learning processes. *Applied Psychological Measurement*, *1*(3), 413–431. https://doi.org/10.1177/014662167700100310

Schraw, G., Bendixen, L. D., & Dunkle, M. E. (2002). Development and validation of the Epistemic Belief Inventory. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 103–118). Erlbaum.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Schumacher, S.C. (2018*). Measuring adult engagement in online education: Instrument development* (Publication No. 10813177) [Doctoral dissertation, Barry University]. ProQuest Dissertations & Theses Global.

Schumacker, R. E., & Lomax, R. G. (2016). *A beginner's guide to structural equation modeling* (4th ed.). New York: Routledge.

Shackelford, J. L., & Maxwell, M. (2012). Sense of community in graduate online education: Contribution of learner to learner interaction. *International Review of Research in Open and Distance Learning*, *13*(4), 228–249.

Sherblom, J. C. (2010). The computer-mediated communication (CMC) classroom: A challenge of medium, presence, interaction, identity, and relationship. *Communication Education*, *59*(4), 497–523. https://doi.org/10.1080/03634523.2010.486440

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, *50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924

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.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, *69*(3), 493–525. https://doi.org/10.1177/0013164408323233

Spurgeon, S. L. (2017). Evaluating the unintended consequences of assessment practices: Construct irrelevance and construct underrepresentation. *Measurement and Evaluation in Counseling and Development*, *50*(4), 275–281. https://doi.org/10.1080/07481756.2017.1339563

Sriram, R. (2014). *Five things not to do in developing surveys for assessment in students affairs*. https://www.naspa.org/images/uploads/main/5things-surveys_feb2014_web.pdf

Stone, C., Freeman, E., Dyment, J., Muir, T., & Milthorpe, N. (2019). Equal or equitable? The role of flexibility within online education. *Australian and International Journal of Rural Education*, *29*(2), 26–40.

Sun, J. C.-Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology*, *43*(2), 191–204. https://doi.org/10.1111/j.1467-8535.2010.01157.x

Tabachnick, B.G., & Fidell, L.S. (2013). *Using multivariate statistics* (6th ed). Pearson.

Tempelaar, D. T., Niculescu, A., Rienties, B., Gijselaers, W. H., & Giesbers, B. (2012). How achievement emotions impact students' decisions for online learning, and what precedes those emotions. *Internet and Higher Education*, *15*(3), 161–169. https://doi.org/10.1016/j.iheduc.2011.10.003

van de Sande, C., Abramson, J., & Judson-Garcia, J. (2017). An exploration of note-taking in an online calculus course. *Journal of Computers in Mathematics and Science Teaching*, *36*(1), 75–99.

Vayre, E., & Vonthron, A.-M. (2017). Psychological engagement of students in distance and online learning: Effects of self-efficacy and psychosocial processes. *Journal of Educational Computing Research*, *55*(2), 197–218. https://doi.org/10.1177/0735633116656849

Veiga, F. H., Reeve, J., Wenthzel, K., & Robu, V. (2014). *Assessing students' engagement: A review of instruments with psychometric qualities*. Repositório da Universidade de Lisboa. https://repositorio.ul.pt/handle/10451/18036

Voelkl, K. E. (1997). Identification with school. *American Journal of Education*, *105*(3), 294–318.

Wandler, J., & Imbriale, W. J. (2017). Promoting college student self-regulation in online learning environments. *Online Learning*, *21*(2), 1–16. https://doi.org/10.24059/olj.v21i2.881

Wang, M. (2010). Online collaboration and offline interaction between students using asynchronous tools in blended learning. *Australasian Journal of Educational Technology*, *26*(6). https://doi.org/10.14742/ajet.1045

Wang, M., & Fredricks, J. A. (2014). The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child Development*, *85*(2), 722–737. https://doi.org/10.1111/cdev.12138

Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Linn, J. S. (2016). The Math and Science Engagement Scales: Scale development, validation, and psychometric properties. *Learning and Instruction*, *43*, 16–26. https://doi.org/10.1016/j.learninstruc.2016.01.008

Wang, M.-T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, *47*(3), 633–662. https://doi.org/10.3102/0002831209361209

Wang, C., Hsu, H.-C. K., Bonem, E. M., Moss, J. D., Yu, S., Nelson, D. B., & Levesque-Bristol, C. (2019). Need satisfaction and need dissatisfaction: A comparative study of online and face-to-face learning contexts. *Computers in Human Behavior*, *95*, 114–125. https://doi.org/10.1016/j.chb.2019.01.034

Warner, R. M. (2008). *Applied statistics: From bivariate through multivariate techniques*. Sage Publications.

Warner, R. M. (2013). *Applied statistics: From bivariate through multivariate techniques (2nd ed.)*. Sage Publications.

Warren, A. N. (2018). Navigating assigned roles for asynchronous online discussions: Examining participants' orientation using conversation analysis. *Online Learning*, *22*(4), 27–45. https://doi.org/10.24059/olj.v22i4.1512

Watkins, R., Corry, M., Dardick, W., & Stella, J. (2015). Note-taking habits of online students: Value, quality, and support. *Quarterly Review of Distance Education*, *16*(3), 1–12.

Watson, F. F., Castano Bishop, M., & Ferdinand-James, D. (2017). Instructional strategies to help online students learn: Feedback from online students. *TechTrends*, *61*(5), 420–427. https://doi.org/10.1007/s11528-017-0216-y

Weinstein, C. E., Acee, T. W., & Jung, J. (2011). Self-regulation and learning strategies. *New Directions for Teaching and Learning*, *2011*(126), 45–53. https://doi.org/10.1002/tl.443

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In *Handbook of research on teaching* (3rd ed., pp. 315–327). Macmillan.

Wellborn, J. G., & Connell, J. P. (1987). *Manual for the Rochester Assessment Package for Schools*. University of Rochester.

Whittaker, T. A. (2016). Structural equation modeling. In K. A. Pituch & J. P. Stevens, *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS* (6th ed., pp. 639–733). Routledge.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–Value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology*, *96*(2), 236–250.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*(6), 806–838. https://doi.org/10.1177/0011000006288127

Wu, H., & Leung, S.-O. (2017). Can Likert scales be treated as interval scales?—A simulation study. *Journal of Social Service Research*, *43*(4), 527–532. https://doi.org/10.1080/01488376.2017.1329775

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. https://doi.org/10.2190/7BAK-EGAH-3MH1-K7C6

Xie, K., Heddy, B. C., & Greene, B. A. (2019). Affordances of using mobile technology to support experience-sampling method in examining college students' engagement. *Computers & Education*, *128*, 183–198. https://doi.org/10.1016/j.compedu.2018.09.020

Xu, B., Chen, N.-S., & Chen, G. (2020). Effects of teacher role on student engagement in WeChat-based online discussion learning. *Computers & Education*, *157*, 1–11. https://doi.org/10.1016/j.compedu.2020.103956

Yeh, Y.-C., Kwok, O.-M., Chien, H.-Y., Sweany, N. W., Baek, E., & McIntosh, W. (2019). How college students' achievement goal orientations predict their expected online learning outcome: The mediation roles of self-regulated learning strategies and supportive online learning behaviors. *Online Learning*, *23*(4), 23–41. https://doi.org/10.24059/olj.v23i4.2076.

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2019). Assigned roles as a structuring tool in online discussion groups: Comparison of transactional distance and knowledge sharing behaviors. *Journal of Educational Computing Research*, *57*(5), 1303–1325. https://doi.org/10.1177/0735633118786855

Young, S., & Bruce, M. A. (2011). Classroom community and student engagement in online courses. *MERLOT Journal of Online Learning and Teaching*, *7*(2), 219–230. https://jolt.merlot.org/vol7no2/young_0611.pdf

Zhoc, K. C. H., Webster, B. J., King, R. B., Li, J. C. H., & Chung, T. S. H. (2019). Higher Education Student Engagement Scale (HESES): Development and psychometric evidence. *Research in Higher Education*, *60*(2), 219–244.

Zhu, E. (2006). Interaction and cognitive engagement: An analysis of four asynchronous online discussions. *Instructional Science*, *34*(6), 451–480. https://doi.org/10.1007/s11251-006-0004-0

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, *82*(1), 51–59.

Appendices

Appendix A Candidate Items Selected from Expert Reviews for Pilot Cognitive Testing

5-point Likert-type scale: 1=Never, 2=Seldom, 3=Sometimes, 4=Often, and 5=Always

Candidate Items for Online Behavioral Engagement (BE)

1. I post to the discussion board for each online discussion.
2. I read the instructor’s posts on the discussion board.
3. I read my peers’ posts on the discussion board.
4. I respond to the instructor’s posts on the discussion board.
5. I respond to my peers’ posts on the discussion board.
6. I participate in content-related dialogues with the course instructor on the discussion board.
7. I participate in content-related dialogues with my peers on the discussion board.
8. I ask my peers and/or my instructor for help in this online course.
9. I study the course materials assigned to me by my online course instructor.
10. I complete the learning tasks assigned in this online course.
11. I complete the assignments in this online course.
12. I submit my assignments by due dates.
13. I incorporate the feedback I get into my course work.
14. I visit the online course site to access the course materials.
15. I participate in learning activities in this online course.
16. I miss assignment deadlines in this online course. (R)
17. I take notes while reviewing the course materials provided by my online course instructor.
18. I share resources with my peers in this online course.

Candidate Items for Online Emotional Engagement (EE)

1. I enjoy participating in conversations with others in online discussions.
2. I enjoy reading others’ posts on the discussion board.
3. I feel curious about the instructor's perspectives/opinions in each online discussion.
4. I enjoy contributing to the development of ideas in online discussions.
5. I feel curious about my peers' perspectives/opinions in each online discussion.
6. I enjoy sharing my responses to the prompts/questions on the discussion board.
7. I enjoy sharing my perspectives with others in online discussions.
8. I feel a sense of pride when I contribute to my peers’ learning in online discussions.
9. I feel a sense of pride when I contribute to the development of ideas in online discussions.
10. I enjoy studying the learning materials assigned to me by my online course instructor.
11. I enjoy completing the learning tasks assigned in this online course.
12. I enjoy getting feedback about my work in this online course.
13. I enjoy completing the assignments in this online course.
14. I enjoy participating in learning activities in this online course.

15. I feel curious about learning the content in this online course.
16. Overall, being an online learner is fun to me.
17. Studying the learning materials assigned in this online course is not enjoyable to me. (R)
18. I feel curious when learning new things in this online course.
19. I feel a sense of pride when I get positive feedback about my work in this online course.
20. Overall, I feel hopeful about my academic success in this online course.
21. Overall, I feel hopeful about improving my knowledge and/or skills in this online course.

Candidate Items for Online Cognitive Engagement (CE)

1. I create logical connections between my understanding and my peers' understanding of course concepts/ideas/issues while composing my discussion posts.
2. I create my own examples to express my understanding of the course concepts/ideas/issues during online discussions.
3. I reflect on others' posts on the discussion board.
4. Before attempting to explain my point of view, I try to understand what others already have written or said on the discussion board.
5. Before I post to the online discussion board, I try to understand the concepts/ideas/issues under study.
6. I put forth effort to explain the course concepts/ideas/issues in detail during online discussions.
7. I relate my prior knowledge to the content of this online course.
8. I relate the course content to my personal experiences.
9. I plan for my next work/task/assignment in this online course.
10. I draw conclusions from others' posts on the discussion board.
11. I question my understanding of the concepts/issues/ideas being taught in this online course.
12. Before I post to the online discussion board, I think about how I will express myself.
13. I check my own understanding of the course concepts/ideas/issues by thinking about others' posts on the discussion board.
14. Before I start studying for this online course, I plan what to do.
15. I reflect on my own understanding of the course concepts/ideas/issues during online discussions.
16. I put forth effort to comprehend the course concepts/ideas/issues during online discussions.
17. I put forth effort to understand different perspectives during online discussions.
18. I put forth no effort to make connections among the course concepts/ideas/issues during online discussions. (R)
19. I make connections among the course concepts/ideas/issues.
20. I create my own examples to better understand the course concepts/ideas/issues.
21. Before I start studying for this online course, I think about what I need to learn.
22. I reflect on the feedback I get in this online course.

Appendix B Initial IRB Approval Letter Before Cognitive Interviews

(July 27, 2021)

The UNIVERSITY of OKLAHOMA

Institutional Review Board for the Protection of Human Subjects

Approval of Initial Submission – Exempt from IRB Review – AP01

Date: July 27, 2021

IRB#: 13566

Principal Investigator: Murat Turk

Approval Date: 07/27/2021

Exempt Category: 2

Study Title: Developing and Validating a New Survey of Online Student Engagement

On behalf of the Institutional Review Board (IRB), I have reviewed the above-referenced research study and determined that it meets the criteria for exemption from IRB review. To view the documents approved for this submission, open this study from the *My Studies* option, go to *Submission History*, go to *Completed Submissions* tab and then click the *Details* icon.

As principal investigator of this research study, you are responsible to:

- Conduct the research study in a manner consistent with the requirements of the IRB and federal regulations 45 CFR 46.
- Request approval from the IRB prior to implementing any/all modifications as changes could affect the exempt status determination.
- Maintain accurate and complete study records for evaluation by the HRPP Quality Improvement Program and, if applicable, inspection by regulatory agencies and/or the study sponsor.
- Notify the IRB at the completion of the project.

If you have questions about this notification or using iRIS, contact the IRB @ 405-325-8110 or irb@ou.edu.

Cordially,

Lara Mayeux, Ph.D.
Chair, Institutional Review Board

Appendix C Qualtrics Online Survey Form for Cognitive Interviews

Introduction: Thank you again for agreeing to participate in this pilot cognitive interview. In this interview, after you respond to each of my candidate items, I will ask you a couple of questions about your understanding and interpretation of each item. As you have already provided your signed consent to participate in this interview, you can proceed to the next page to start the survey by clicking on the arrow below.

Directions: In this cognitive interview, you will see the survey directions and my candidate items that are designed to measure online behavioral engagement, online emotional engagement, and online cognitive engagement. You are asked to carefully read the survey directions for each set of items, read each item carefully, think about it, and then respond to it by choosing a response option provided. After you choose your response to an item, I will ask you a couple of questions about your understanding and interpretation of the item. You can proceed to the next page to see the survey directions and my candidate items by clicking on the arrow below.

Survey Directions for Behavioral Engagement Items and Behavioral Engagement Items

Survey Directions for Emotional Engagement Items and Emotional Engagement Items

Survey Directions for Cognitive Engagement Items and Cognitive Engagement Items

Appendix D A Summary of Major Item Issues and My Responses after Pilot Cognitive Testing

(adapted from Holmes, 2018)

Item#*	CIQ#	Item Issue	My Response
BE2	CIQ2,4	"instructor's posts" phrase was interpreted differently by the respondents	phrase clarified as "comment/ question posts"
BE3,5	CIQ2,4	The respondents asked about whether the item was about "all" of peers' posts, "most" or "some"	consulted the survey expert. The number/amount of posts was not the target of interest. No change was made.
BE4	CIQ2,4,5	One respondent commented "This one is more ambiguous…because the instructor may not always be speaking to me directly…I only respond if it is relevant to me…I wouldn't respond just for no reason"	item revised with "if he/she leaves a comment/question on my posts…"
BE8	CIQ2,4	The respondents made a distinction between "asking peers for help" and "asking instructor for help" and interpreted these as two separate behaviors	item split into two
		One respondent commented "I don't typically need assistance in my courses" and because of this selected "Seldom"	item revised with "If I need it" to clarify the target behavior to be rated
		One respondent asked "what kind of help?"	parenthetical example added
BE10,16	CIQ2,4	Two respondents interpreted "learning Tasks" as "assignments" or "homework"	parenthetical examples added and definitions provided in survey directions

BE19	CIQ2,4	One respondent commented "Resources could mean different things to different people". Another respondent commented "It is just additional".	item revised with "additional knowledge resources" parenthetical examples added.
EE8	CIQ2,4	One respondent asked "how am I helping them?" to contribute to peers' learning. The meaning of "contribute to peers' learning" was not clear to the respondent.	Item split into two as "answering peers' questions" and "helping peers' understanding with answers or explanations".
EE16	CIQ2,4	One respondent asked whether being an online learner was being compared against being an in-person learner	No change. No comparison was intended by the item.
		Two respondents interpreted the word "fun" differently as "having laughter" or "having fun while playing games etc."	"fun" was removed. Item revised with "enjoyable".
EE18	CIQ2,4	One respondent commented "When I read this, it almost makes me, like, compare it to an in-person course"	No change. No comparison was intended at all.
		The same respondent interpreted the item as curiosity about instructors or peers' human qualities	item revised to clarify curiosity as learning more about the content
EE20,21	CIQ2,4	Three respondents asked what was meant by the phrase "feel hopeful" and needed clarifications to understand the items	items revised with "have great hope that" Two new items written up with "am/feel quite optimistic about"
CE1	CIQ2,4	All four respondents asked if the item was about composing initial or reply posts, saying most students cannot see	item revised with "while replying to their posts"

		peers' posts before initial postings.	
CE3	CIQ2,4	One respondent asked "why would I reflect on others' posts?". The intended meaning was unclear to the respondent.	item revised with "reflect on my peers' responses to my posts"
CE4	CIQ2,4	The item's context of "attempting to explain my point of view" was unclear to the respondent.	item revised with "attempting to respond to a peer's post" to clarify the item context
CE8	CIQ2,4	One respondent misinterpreted "personal experiences" as "personal or private life"	"personal" was replaced with "own"
CE11	CIQ2,4	Two respondents interpreted "I question my understanding" as "doubting myself"	item revised with "ask myself questions about my own understanding"
CE12	CIQ2,4	One respondent misinterpreted "express Myself" as "expressing one's personality"	item revised with "express what I want to say"

CIQ=Cognitive Interview Question. CIQ-1=Do you think the survey instructions are clear and easy to understand? CIQ-2=What do you think this item is asking you about? Please explain. CIQ-3= Why did you choose the response you chose over the other response options? Please explain. CIQ4-= Do you think the item needs to be re-worded for clarity? If yes, how would you re-word it? CIQ-5 How relevant do you think this item is to you as an online student? Please explain. *See Appendix I for items submitted to pilot cognitive testing.

Appendix E Final List of 66 OEHE Items Used for Large-Group Data Collection

Online Behavioral Engagement Items

1. I post to the discussion board for each discussion in this online course.
2. I read my instructor's comment/question posts in online discussions on the discussion board.
3. I read my peers' posts during online discussions on the discussion board.
4. I respond to the instructor when he/she leaves a comment/question on my posts during online discussions on the discussion board.
5. I respond to my peers' posts during online discussions on the discussion board.
6. I participate in content-related dialogues with the course instructor when he/she leaves a comment/question on my posts during online discussions on the discussion board.
7. I participate in content-related dialogues with my peers during online discussions on the discussion board.
8. I ask my peers for help (e.g., clarification about an assignment) when I need it in this online course.
9. I ask my instructor for help (e.g., clarification about an assignment) when I need it in this online course.
10. I study the learning materials (e.g., readings, slides, lecture videos) of this online course.
11. I complete the learning tasks (e.g., doing the readings, watching the lecture videos, listening to the podcasts) of this online course.
12. I complete the assignments (e.g., tests, papers, projects) of this online course.
13. I submit the assignments by their due dates in this online course.
14. I read the feedback that I receive in this online course.
15. I incorporate the feedback that I get into my future work (e.g., next assignment) in this online course.
16. I visit the online course site (e.g., Canvas, Blackboard) to check for possible course updates (e.g., new file uploads, announcements).
17. I participate in the learning activities (e.g., discussion board activities, group presentations, case studies) of this online course.
18. I miss the assignment deadlines in this online course. (R)
19. I take notes while reviewing the learning materials (e.g., while reading a text, while watching a lecture video) of this online course.
20. I share additional knowledge resources (e.g., an interesting article, a useful YouTube video) with my peers in this online course.

Online Emotional Engagement Items

1. I enjoy discussing course concepts/issues/ideas with my peers during online discussions on the discussion board.
2. Reading my peers' posts during online discussions on the discussion board is enjoyable to me.
3. I enjoy reading the instructor's comment/question posts in online discussions on the discussion board.

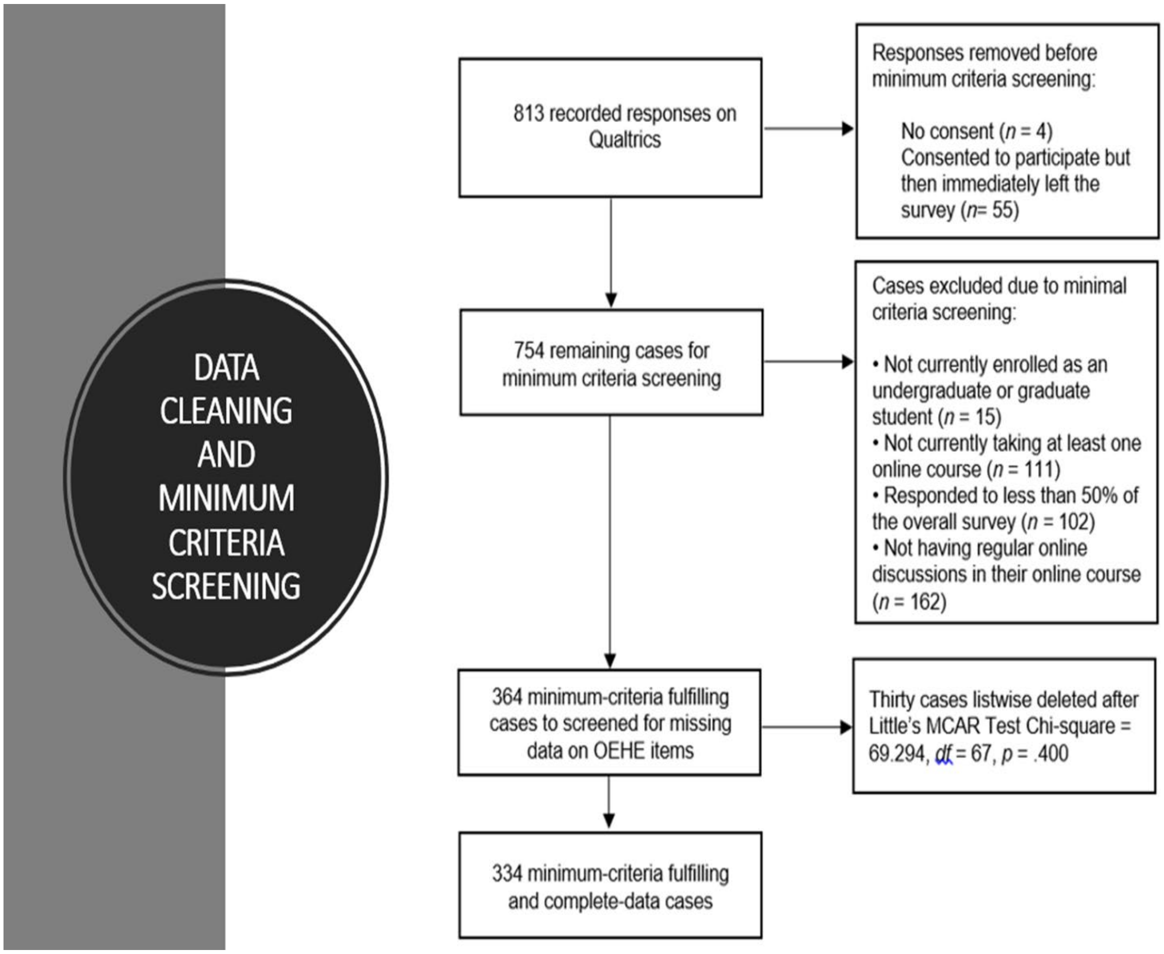
4. I am curious about the instructor's perspectives/opinions in online discussions on the discussion board.
5. I enjoy contributing to the development of peer dialogues in online discussions on the discussion board.
6. I feel curious about my peers' perspectives/experiences in each online discussion on the discussion board.
7. I enjoy sharing my responses to the instructor's discussion prompts/discussion questions during online discussions on the discussion board.
8. I enjoy sharing my perspectives/experiences with my peers during online discussions on the discussion board.
9. I am curious about further exploring the concepts/issues/ideas in this online course.
10. I would like to find out more about the concepts/issues/ideas being taught in this online course.
11. I feel proud of myself when I can answer my peers' questions during online discussions on the discussion board.
12. I am proud of myself when I can help my peers' understanding with my answers/explanations during online discussions on the discussion board.
13. I am proud of myself when I can contribute to the peer dialogues with my own perspectives/experiences/solutions during online discussions on the discussion board.
14. I enjoy studying the learning materials (e.g., readings, slides) of this online course.
15. I enjoy completing the learning tasks (e.g., reading a text, watching a lecture video) of this online course.
16. I like getting feedback about my work in this online course.
17. I like completing the assignments (e.g., doing a test, writing an essay) of this online course.
18. I like participating in the learning activities (e.g., pair work, small group discussions, online debates) of this online course.
19. I feel curious to learn more about the content in this online course.
20. Overall, taking this online course has been an enjoyable learning experience to me.
21. Overall, taking this online course has not been enjoyable to me. (R)
22. I feel curious about acquiring further information about the content in this online course.
23. I feel proud of myself when I get positive feedback about my work in this online course.
24. Overall, I have confidence that I will be academically successful in this online course.
25. Overall, I have confidence that I will be able to improve my knowledge and/or skills in this online course.
26. Overall, I am optimistic about my academic success in this online course.
27. Overall, I feel optimistic about improving my knowledge and/or skills in this online course.

Online Cognitive Engagement Items

1. I try to create connections between my understanding and my peers’ understanding of course concepts/ideas/issues while replying to their posts during online discussions.
2. I try to create my own examples to express my understanding of the course concepts/ideas/issues during online discussions on the discussion board.

3. I reflect on my understanding of peer responses to my posts during online discussions on the discussion board.
4. Before attempting to respond to a peer's post, I try to understand what my peer has actually said or written during online discussions on the discussion board.
5. Before I post to the discussion board in this online course, I try to understand the concepts/ideas/issues under study.
6. I put forth effort to explain the course concepts/ideas/issues in depth during online discussions on the discussion board.
7. I put forth effort to relate the content of this online course to my prior knowledge.
8. I put forth effort to relate the content of this online course to my own experiences.
9. I plan for my next task/next assignment/next submission (e.g., checking due dates, looking over directions) in this online course.
10. I ask myself questions about my own understanding of the concepts/issues/ideas in this online course.
11. Before I post to the online discussion board, I think about how I will express what I want to say.
12. I check my own understanding of the course concepts/ideas/issues by thinking about my peers’ posts on the discussion board.
13. I reflect on my own understanding of the concepts/ideas/issues in this online course.
14. I put forth effort to comprehend the concepts/ideas/issues in this online course.
15. I put forth effort to understand diverse perspectives during online discussions on the discussion board.
16. I put forth no effort to make connections between the course concepts/ideas/issues in this online course. (R)
17. I put forth effort to make connections between the concepts/ideas/issues in this online course.
18. I try to create my own examples to better understand the concepts/ideas/issues being taught in this online course.
19. I reflect on my understanding of the feedback that I get in this online course.

Appendix F Data Cleaning and Minimum-Criteria Screening



Appendix G Participants' Demographics (*N*=234)

Demographic Characteristics Variables	*Frequency (n)*	*Percent (%)*
Age Range (Years)		
16-25	87	37.1
26-35	58	24.8
36-45	48	20.5
46-70	41	17.5
Gender		
Female	189	80.8
Male	39	16.7
Non-binary/third gender	3	1.3
Prefer to self-describe	2	0.9
Prefer not to say	1	0.4
Race/Ethnicity		
American Indian/Alaskan Native	7	3.0
Asian	18	7.7
Black/African American	17	7.3
Native Hawaiian/Other Pacific Islander	1	0.4
White	171	73.1
Prefer to self-describe	17	7.3
Prefer not to say	3	1.3
Mexican, Mexican American, or Chicano/a/x	13	5.6
Puerto Rican	2	0.9
Another Hispanic, Latino/a/x or Spanish origin	8	3.4
Student Status		
Freshman	10	4.3
Sophomore	18	7.7
Junior	28	12.0
Senior	22	9.4
Post-baccalaureate	1	0.4
Master's student	108	46.2
Doctoral student	40	17.1
Post-doctorate	1	0.4
Other student status	6	2.6
Employment Status		
Full time	135	57.7
Part time	62	26.5
Not currently employed	34	14.5
Not applicable	3	1.3
Marital Status		
Single	115	49.1
Married	97	41.5
Divorced	11	4.7
Other	11	4.7

Appendix H Participants' Inclusiveness, Bias, and Accessibility Characteristics

Inclusiveness and Bias Characteristics (N=234)

Perceptions of Bias and Inclusiveness	Sometimes *n* (%)	Often *n* (%)	Always *n* (%)
Racial/Ethnic Bias in Online Classes	14 (6.0)	1 (0.4)	1 (0.4)
Gender Bias in Online Classes	17 (7.3)	2 (0.9)	0 (0)
Religion Bias in Online Classes	11 (4.7)	5 (2.1)	0 (0)
Socio-economic Bias in Online Classes	17 (7.3)	6 (2.6)	0 (0)
Underrepresentation in Online Classes	25 (10.7)	14 (6.0)	7 (3.0)
Not Belonging to Online Community	41 (17.5)	13 (5.6)	1 (0.4)
Opinions Not Respected or Valued by Online Course Instructors	17 (7.3)	4 (1.7)	1 (0.4)
Opinions Not Respected or Valued by Peers	25 (10.7)	4 (1.7)	0 (0)
Feeling Uncomfortable Sharing Perspectives in Online Classes	55 (23.5)	13 (5.6)	3 (1.3)
Feeling Dominated by the Mainstream Group in Online Classes	45 (19.2)	21 (9.0)	7 (3.0)
Feeling Uncomfortable with Peers' Language in Online Discussions/Interactions	20 (8.5)	0 (0)	1 (0.4)

Accessibility Issues Experienced by Participants (n=27)

Accessibility Issue	*n*
Lack of access to computer	3
Lack of access to hardware (e.g., mouse, headphones, webcam)	5
Lack of access to reliable internet	16
Lack of access to software (e.g., Microsoft Office)	4
Lack of accommodations for students with disabilities	3
Other	8

Note: Participants were able to indicate more than one accessibility issue; therefore, the total of responses may exceed 100%.

Appendix I Classroom Engagement Scale

(Reeve, 2013-Study 2)

Behavioral Engagement Items
When I'm in this class, I listen very carefully.
I pay attention in this class.
I try hard to do well in this class.
In this class, I work as hard as I can.
Emotional Engagement Items
When we work on something in this class, I feel interested.
This class is fun.
I enjoy learning new things in this class.
When I'm in this class, I feel good.
Cognitive Engagement Items
When I study for this class, I try to connect what I am learning with my own experiences.
I try to make all the different ideas fit together and make sense when I study for this class.
When doing work for this class, I try to relate what I'm learning to what I already know.
I make up my own examples to help me understand the important concept I study for this class.

Adapted Items for the Online Learning Context
<u>5-point Likert-type scale</u>
1=Strongly Disagree, 2=Disagree, 3=Neither Agree nor Disagree, 4=Agree, 5=Strongly Agree

Behavioral Engagement Items
When I'm in this online course, I can focus.
I pay attention in this online course.
I try hard to do well in this online course.
In this online course, I work as hard as I can.

Emotional Engagement Items
When we work on something in this online course, I feel interested.
This online course is fun.
I enjoy learning new things in this online course.
When I'm in this online course, I feel good.

Cognitive Engagement Items
When I study for this online course, I try to connect what I am learning with my own experiences.
I try to make all the different ideas fit together and make sense when I study for this online course.
When doing work for this online course, I try to relate what I'm learning to what I already know.
I make up my own examples to help me understand the important concept I study for this online course.

Appendix J Fixed Ability Subscale of the 32-item version of the Epistemic Beliefs Inventory (EBI)

(DeBacker et al., 2008; Schraw et al., 2002)

<u>5-point Likert-type scale</u>
1=Strongly Disagree, 2=Disagree, 3=Neither Agree nor Disagree, 4=Agree, 5=Strongly Agree

Original Item Numbers

5. Some people will never be smart no matter how hard they work.

8. Really smart students do not have to work as hard to do well in school.

12. People cannot do too much about how smart they are.

15. How well you do in school depends on how smart you are.

17. Some people just have a knack for learning, and others do not.

26. Smart people are born that way.

32. Some people are born with special gifts and talents.

Appendix K Task Value Subscale of the Motivated Strategies for Learning Questionnaire (MSLQ)

(Pintrich et al., 1991, 1993)

<u>7-point Likert-type scale</u>
1=Not at all true of me 2 3 4 5 6 7=Very true of me

1. I think I will be able to use what I learn in this course in other courses.

2. It is important for me to learn the course material in this class.

3. I am very interested in the content area of this course.

4. I think the course material in this class is useful for me to learn.

5. I like the subject matter of this course.

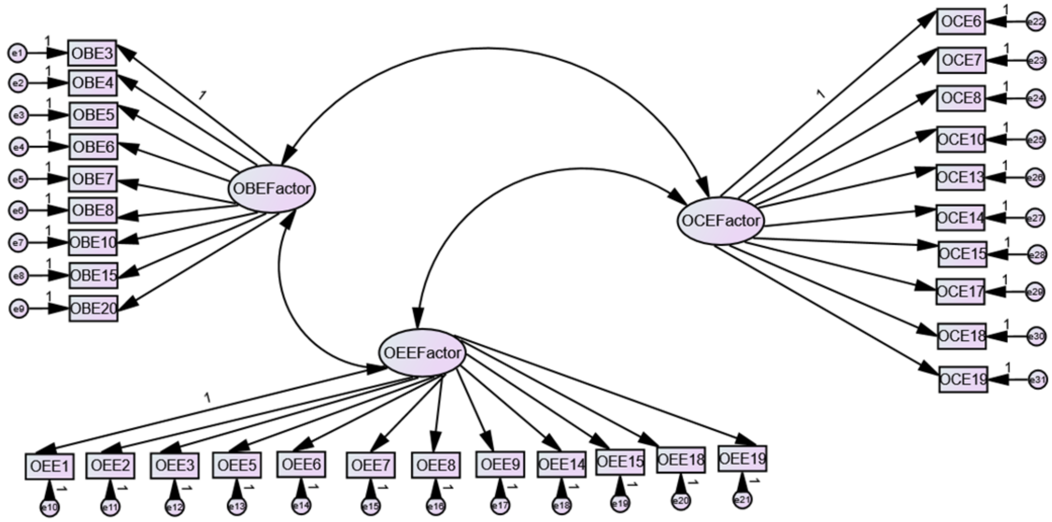
6. Understanding the subject matter of this course is very important to me.

Adapted Items for the Online Learning Context

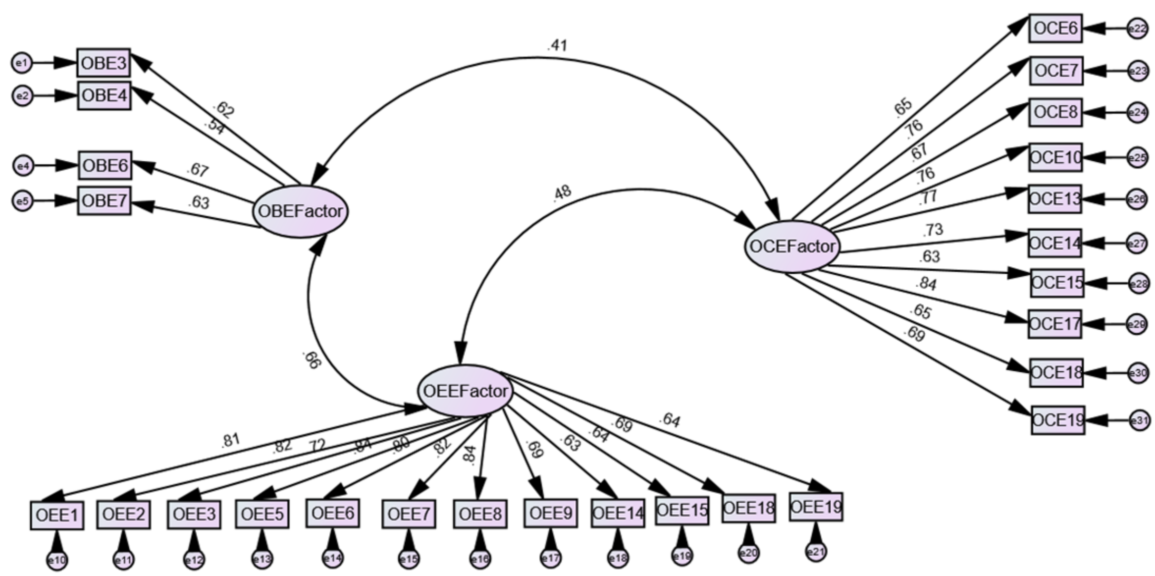
1=Not at all true of me 2 3 4 5 6 7=Very true of me

1. I think I will be able to use what I learn in this online course in other courses.
2. It is important for me to learn the course material in this online class.
3. I am very interested in the content area of this online course.
4. I think the course material in this online class is useful for me to learn.
5. I like the subject matter of this online course.
6. Understanding the subject matter of this online course is very important to me.

Appendix L Initial Three-Factor Model of Online Student Engagement with 31 Items



Appendix M Model 2: Three-Factor Model with 26 Indicator Observed Variables (N=235)



Appendix N Unstandardized Factor Loadings with Bias-Corrected 95 % Confidence Intervals (N=235)

Item	Factor	Estimate	Lower	Upper	*P*
OBE3	OBE	1.000	1.000	1.000	…
OBE4	OBE	.731	.382	1.119	.001
OBE6	OBE	1.077	.669	1.554	.001
OBE7	OBE	1.159	.920	1.447	.001
OEE1	OEE	1.000	1.000	1.000	…
OEE2	OEE	1.002	.887	1.134	.001
OEE3	OEE	.837	.721	.960	.001
OEE5	OEE	1.043	.939	1.170	.001
OEE6	OEE	.936	.817	1.078	.001
OEE7	OEE	1.071	.957	1.209	.001
OEE8	OEE	.971	.878	1.078	.001
OCE6	OCE	1.000	1.000	1.000	…
OCE7	OCE	.932	.784	1.150	.001
OCE8	OCE	.855	.692	1.062	.001
OCE13	OCE	1.088	.873	1.366	.001
OCE14	OCE	.820	.660	1.032	.001
OCE15	OCE	.916	.732	1.150	.001
OCE17	OCE	1.068	.890	1.283	.001
OCE18	OCE	.995	.776	1.280	.001
OCE19	OCE	.969	.751	1.235	.001

Standardized Factor Loadings with Bias-Corrected 95 % Confidence Intervals (N=235)

Item	Factor	Estimate	Lower	Upper	P
OBE3	OBE	.731	.612	.830	.001
OBE4	OBE	.333	.181	.475	.001
OBE6	OBE	.504	.346	.642	.001
OBE7	OBE	.700	.545	.812	.002
OEE1	OEE	.814	.755	.865	.001
OEE2	OEE	.812	.745	.865	.001
OEE3	OEE	.685	.589	.769	.001
OEE5	OEE	.884	.841	.917	.001
OEE6	OEE	.815	.757	.866	.001
OEE7	OEE	.813	.731	.870	.002
OEE8	OEE	.868	.816	.903	.002
OCE6	OCE	.639	.537	.728	.001
OCE7	OCE	.718	.632	.792	.001
OCE8	OCE	.642	.553	.727	.001
OCE13	OCE	.727	.636	.812	.001
OCE14	OCE	.701	.594	.788	.002
OCE15	OCE	.647	.539	.738	.001
OCE17	OCE	.859	.796	.907	.002
OCE18	OCE	.662	.549	.756	.001
OCE19	OCE	.693	.581	.787	.001

Appendix O Final Version of the OEHE (20 Items) (N=235*)*

Item Code	Item Content	Mean	SD
OBE3	I read my peers' posts during online discussions on the discussion board.	4.12	.881
OBE4	I respond to the instructor when he/she leaves a comment/question on my posts during online discussions on the discussion board.	3.47	1.412
OBE6	I participate in content-related dialogues with the course instructor when he/she leaves a comment/question on my posts during online discussions on the discussion board.	3.47	1.375
OBE7	I participate in content-related dialogues with my peers during online discussions on the discussion board.	3.97	1.066
OEE1	I enjoy discussing course concepts/issues/ideas with my peers during online discussions on the discussion board.	3.60	1.017
OEE2	Reading my peers' posts during online discussions on the discussion board is enjoyable to me.	3.40	1.022
OEE3	I enjoy reading the instructor's comment/question posts in online discussions on the discussion board.	3.83	1.012
OEE5	I enjoy contributing to the development of peer dialogues in online discussions on the discussion board.	3.77	.978
OEE6	I feel curious about my peers' perspectives/experiences in each online discussion on the discussion board.	3.89	.952
OEE7	I enjoy sharing my responses to the instructor's discussion prompts/discussion questions during online discussions on the discussion board.	3.60	1.091
OEE8	I enjoy sharing my perspectives/experiences with my peers during online discussions on the discussion board.	3.76	.927
OCE6	I put forth effort to explain the course concepts/ideas/issues in depth during online discussions on the discussion board.	4.30	.755
OCE7	I put forth effort to relate the content of this online course to my prior knowledge.	4.48	.629
OCE8	I put forth effort to relate the content of this online course to my own experiences.	4.50	.643
OCE13	I reflect on my own understanding of the concepts/ideas/issues in this online course.	4.33	.722
OCE14	I put forth effort to comprehend the concepts/ideas/issues in this online course.	4.59	.565
OCE15	I put forth effort to understand diverse perspectives during online discussions on the discussion board.	4.43	.684
OCE17	I put forth effort to make connections between the concepts/ideas/issues in this online course.	4.46	.600
OCE18	I try to create my own examples to better understand the concepts/ideas/issues being taught in this online course.	4.28	.726
OCE19	I reflect on my understanding of the feedback that I get in this online course.	4.49	.675

Note: OBE=Online Behavioral Engagement, OEE=Online Emotional Engagement, OCE=Online Cognitive Engagement