The impact of behavioural, cognitive and emotional dimensions of student engagement on student learning: the case of Azerbaijani higher education institutions

Zafar Hasanov*

International Training and Project Center,
Yusif Vazir Chamanzaminli Str 1996,
Baku, Azerbaijan
Email: zafar_hasanov@yahoo.com
*Corresponding author

Panayiotis Antoniou

Faculty of Education,
University of Cambridge,
184 Hills Road, Cambridge, CB2 8PQ, UK
Email: pa241@cam.ac.uk

Elchin Suleymanov

Baku Engineering University,
Hasan Aliyev Str 120, Khirdalan, Azerbaijan
Email: elsuleymanov@beu.edu.az

Vener Garayev

ADA University,
Ahmadbay Agha-Oglu Street 61, Baku, Azerbaijan
Email: vgarayev@ada.edu.az

Abstract: This study examines the relationship between student engagement and academic attainment. A random sample of 548 full time second-year and third-year undergraduate students from six higher education institutions in Azerbaijan was obtained. The study methodologically compares and merges two psychometric self-report measures, the National Survey of Student Engagement (NSSE) and the Burch Engagement Survey for Students (BESS), which capture three fundamentally different dimensions of engagement, namely behavioural, cognitive and emotional. The data were analysed using confirmatory factor analysis (CFA) and multiple linear regression. The findings of the study support the multidimensionality notion of engagement as suggested by the prominent student involvement theories and claim that all three dimensions are to a different extent statistically significant determinants of academic success. Regression results reveal significant relationships between students’ academic achievement and engagement both for one-dimensional and three-dimensional engagement models. Compared to...
The impact of behavioural, cognitive and emotional dimensions traditionally used measurements, the three-dimensional CFA model provides better explanation of academic outcomes. The results of this study could have implications for university academic staff and policymakers in utilising the benefits of each engagement factor to improve institutional quality.

**Keywords:** student engagement; behavioural engagement; learning outcomes; emotional engagement; cognitive engagement; higher education.


**Biographical notes:** Zafar Hasanov holds a Bachelor of Science in Economics with specialisation in Commercial Law from the Azerbaijan State University of Economy, Azerbaijan, Master’s in Public Administration from University of Central Florida, USA, Master of Philosophy in Education Leadership and School Improvement from Cambridge University, UK, and Master Certificate in Project Management from School of Business, George Washington University, USA. He has extensive experience in development and modernisation of tertiary and school education, fostering research and innovation, modernisation of education, and enhancement of entrepreneurship and business start-ups.

Panayiotis Antoniou is a Senior Lecturer in Educational Leadership and Evaluation, and member of Educational Leadership, Policy, Evaluation and Change (ELPEC) Academic Group at Faculty of Education, University of Cambridge. He is also a Coordinator of the MPhil/MEd in Educational Research (EdRes) and Official Fellow of the Darwin College. His research interests are in educational leadership, educational effectiveness and improvement at the teacher, school and system levels. He is also interested in teacher professional development and has been involved in research projects related with human resource management in education, and teacher and school evaluation.

Elchin Suleymanov is an author of one textbook, two lecture resources, four subject programs, 14 books and more than a 100 scholarly articles. He is an expert of the United Nations Food and Agriculture Organisation (FAO), The Korea Institute for International Economic Policy (KIEP), and World Economic Survey since 2015, as well as an author of several economic reports on Azerbaijan in several publications. His area of interest includes education management, education economy, higher education, quantitative modelling, public finance and energy economy.

Vener Garayev received his Bachelor of Arts in Political Science and International Relations from the Boğaziçi University in Turkey, and Master’s in Public Administration, and PhD in Public Affairs from University of Central Florida in the USA. He currently serves as an Assistant Professor of Public Administration at the ADA University in Azerbaijan. His main research interests are organisational behaviour, governance, partnerships, collaborations and networks.
1 Introduction

Following World War II, higher education (HE) shifted from an elite system to mass education with a focus on research and teaching (Anderson, 2016). In addition, enrolment of increasingly diverse student populations, and individual pursuits of international experience to fit the global workforce demand called for development of effective governance strategies that move beyond mere attraction of students (Quaye and Harper, 2014). The ensuing emphasis on placing students at the heart of education systems made the concept of engagement the ‘currency’ of institutional quality (Gibbs, 2014). As the main driver of learning, engagement does not only promise explanation of better scholastic achievement, self-development, interest, motivation or metacognition, but also serves as the principal instrument guiding leadership to reach desired educational outcomes (Lawson and Lawson, 2013).

Over the last decade, student engagement has become a significant area of academic research and university policy development. At the policy level, ‘student engagement’ is attracting increasing international attention as a core element of institutional learning and teaching strategies and through national student engagement surveys in most developed HE systems. The National Survey of Student Engagement (NSSE), firstly introduced in the USA in 2000, is now used in most other developed HE systems including Australia, Canada, South Korea, China, Japan, New Zealand, Mexico, Ireland, South Africa and the UK, and serves as an illustration of the trend (Coates and McCormick, 2014).

Nevertheless, according to Smyth and Fasoli (2007), a great number of young people are becoming disengaged from schools and universities, while educational policies fail to address this issue. Current policy preoccupations that emphasise accountability, greater student choice and a more prescriptive curriculum can present difficulties for young people, particularly those from underprivileged backgrounds. This is especially relevant nowadays, with increased use of ICT-based services, internet-based tools and mobile technological devices commonly used by a large number of HE students to perform their tasks and academic engagements (Elia and Poce, 2010). This, in turn, could indicate the need to further explore the concept of student engagement, especially in virtual environments where lack of motivation and self-efficacy is one of the main reasons for dropout.

Monitoring students’ satisfaction and engagement could be an important factor in guiding development of policies and strategies in HE, especially because today’s platforms allow students to perform many online activities, thus, generating huge amount of data that can be processed to get insights about the level of satisfaction on contents, services, community interactions and effort. This offers an opportunity to implement a non-intrusive and in-process evaluation of HE that complements traditional and time-consuming ways to collect feedback (Elia et al., 2019).

In this paper, we argue that universities and other higher education institutions (HEIs) need to evaluate and reconsider some of their practices and policies on maximising student engagement levels. Evaluation of HE students’ engagement rates and effectiveness of policies aiming to increase engagement rates has gradually gained increased significance in the international educational context. For example, until recently, Australian and New Zealand universities lacked data on students’ engagement in effective educational practices. In a study by Coates (2010), the foundations and development of the Australasian Survey of Student Engagement, which is the largest educationally-focused cross-institutional collection from students in Australasia, are
presented and the results from the 2008 survey are analysed to stimulate improvement of HE policies. This is important as, according to Coates (2010), a great percentage of HE students ‘seriously considered’ departing their institution before graduation. In this perspective, student engagement initiatives at the national, institutional and classroom level have emerged against a backdrop of rising participation rates and marketisation of HE (Macfarlane and Tomlinson, 2017).

Struggling from old-fashioned post-Soviet inheritance of socio-economic, political and educational structures, Azerbaijan has also taken significant steps towards modernisation of its HE system. The core policy reform of this transition comprises integration to the Bologna process by adopting international standards of European Space of Higher Education (ESHE) (Ministry of Education, 2014) as well as development of institutional evaluation mechanism that adheres to the principles of European Network for Quality Assurance (ENQA). Similar to other prominent and proven attempts such as NSSE, Faculty Survey of Student Engagement (FSSE), and Collegiate Learning Assessment (CLA) (Isakhanli, 2006), the principles of ENQA encompass student-centred learning and teaching, which integrates students’ and teachers’ voices into institutional quality evaluation (ESG, 2015).

Despite the fact that Azerbaijan has more than 167,000 students enrolled in 52 HEIs (The State Statistical Committee of the Republic of Azerbaijan, 2017), there is no nationally-accepted measure of student or teacher engagement. Although national quality assurance determinants echo various student self-report measures [quality of teaching staff (proportion of PhD degrees, teaching experience, etc.), published research, number of international partnerships, number of enrolled students, financial resources, availability of university electronic management system, university-industry relationship, student access to education resources, graduate employment, and student dropout rates (EHEA, 2015)], these evaluations do not take into account academic attainment. Thus, the quality assessments are limited in scope, as they are more focused on formal completion criteria (pass/fail) of the subjects or dropout rates (Coates, 2005). Similarly, even though student retention rates reflect the extent to which official course requirements have been fulfilled, they do not elaborate on specific details (e.g., the time required to complete the degree). Last, measurement of institutional quality based on employability outcomes can be subjective, as employment of new graduates is affected by a range of non-educational factors such as economic conditions, personal network and geographical location (Coates, 2005).

As students become more diverse, as HE becomes more competitive, and as the export of international education continues to grow, there is an intensified need for evaluation of the extent to which students are engaging effectively with university education (Coates, 2010). This article focuses on the importance of student engagement concept in developing a tool/scale of student success in particular, and HEI quality in general. Specifically, our aims is to develop and test an expanded instrument of student engagement construct on a sample of students from Azerbaijani universities with the purpose to develop a tool that better predicts student academic achievements. Section 2 focuses on definitions and dimensions of student engagement, followed by Section 3 with operationalisation of traditional vs. expanded student engagement models. Then, in Section 4, the results of confirmatory factor analysis (CFA) and regression tests are presented. Finally, the Section 5 links study results to the literature, providing suggestions for further research in the field (Section 6).
2 Student engagement: conceptual and methodological definitions

Student engagement has been defined in a number of ways (Trowler, 2010). It implies a series of conceptual commitments, teaching strategies and behavioural orientations expected of university students. In policy terms, it is driven by efforts to improve student completion and success rates at university, whilst pedagogically it is underpinned by a teaching philosophy that is linked with social constructivism (Macfarlane and Tomlinson, 2017). Many definitions of student engagement emphasise the importance of students being actively engaged in a participatory culture and experiencing an interactive approach to teaching (Newswander and Borrego, 2009). Although there is a lack of a settled or widely-agreed definition of student engagement, most well-cited definitions tend to draw on a framework composed of three dimensions: behavioural, psychological and socio-cultural (Fredricks et al., 2004).

In this perspective, the theoretical framework of this study is based on three theories, namely, student involvement theory, flow theory and self-determination theory. According to student involvement theory, “the greater the student’s involvement in college, the greater will be the amount of student learning and personal development” [Astin, (1984), p.528]. Astin’s (1984) conceptualisation of student involvement is answering the question of what and focuses more on students behaviours and actions rather than their thinking, feeling, and perceptions of learning experiences (Gunuc, 2014). Csikszentmihalyi’s (1990) flow theory, in turn, emphasises that people feel happy and genuinely satisfied if they are in a deep state of absorption in an activity involving their creative abilities and challenges their skills to a certain extent. The flow theory focuses on the question of how and explores students’ classroom learning experience and specific conditions under which they become engaged (Shernoff et al., 2003). Lastly, according to self-determination theory by Deci and Ryan (1985), when people are intrinsically motivated, they become goal-oriented and self-determined for accomplishment. As an aspect of motivation, intrinsically emanating psychological actions do not require any separate motivational consequences, as the central necessary reward is the interest and enjoyment gained from being engaged in activities (Brophy, 2013). Being a reflection of the emotional dimension of engagement, the self-determination theory answers the question of why. Thus, overall, the concept of ‘engagement’ in literature is understood and reflected through the lenses of individual behaviours, understanding and feelings.

The literature also highlights that the dimensions and factors of student engagement lack congruence. According to Kahu (2013), examination of all engagement features that constitute this complex construct in a single research is not possible. Student engagement is challenging to define, as it is a complex construct influenced by multiple factors. Accordingly, Fredricks et al. (2004) identify three dimensions of engagement, namely, behavioural engagement, i.e., students’ participation in education, including the academic, social and extracurricular activities of the school; emotional engagement, i.e., students’ emotional reactions in classroom and in school (a sense of belonging or connectedness to the school); and cognitive engagement, i.e., students’ investment in their learning (motivation and self-regulation). The authors also identify the factors that influence student engagement: teacher factors, school factors, student factors, family and community factors, and curriculum and resources factors, which include active participation and engagement in learning. Wang and Holcombe (2010) suggest that, whereas each type of engagement has different impact on students’ academic achievements, overall they are interlinked, thus, requiring a holistic approach to the
concept of engagement. Kahu (2013), however, asserts that, in striving to come up with a more holistic view of engagement, one of the commonly made mistakes is the lack of distinction between the antecedents and state of engagement, proposing that holistic engagement is a product of both a process and an outcome. According to Kahu (2013), a student with a holistic stage of engagement is someone who acknowledges herself/himself as part or autonomous member of the education institution (emotional), who is thoughtfully devoting in self-learning or demonstrating extra effort to go beyond common boundaries to gain deeper comprehension of complex learning tasks (cognitive) by proactively being involved in educational activities, both inside and outside the formal learning settings (behavioural).

Measuring student engagement has also been a difficult and complex task. Even educators who are skilled in engaging students in learning often lack understanding of the extent in which students fully and actively participate in the process of learning. In an effort to empower educators with tools and information to improve teaching and learning, a number of different measurement instruments have been designed to measure how students engage in educationally purposeful activities (Parsons and Taylor, 2011). Despite these attempts, it is often unclear which aspects of engagement are being measured, with some surveys focusing on single dimensions and others claiming to be a single general measure of engagement (Kahu, 2013).

In order to address this problem, a number of studies championed the development of alternative measurement tools of student engagement. The psychometric self-report instrument of NSSE (2009) is an example of a single-dimensional measurement of student engagement, increasingly used by the universities worldwide over the last two decades. In brief, the NSSE measures the extent to which students are engaged in educational practices, and how this contributes to learning and personal development (Kuh, 2002; Pike, 2006). The NSSE (2014) comprises four themes/benchmarks of engagement (formerly five) representing ten engagement indicators that emerged from “a blend of theory and empirical analysis” [Kuh, (2003), p.30]. These principal themes include: level of academic challenge, learning with peers, experience with faculty, and campus environment (NSSE, 2014). Due to increased use and nascent empirical evidence, the NSSE has been accepted as the primary assessment method of student engagement by a number of studies (Kahu, 2013). Despite being considered as the most prevalent attempt to measure engagement, the NSSE has been also criticised by some scholars. For example, Steele and Fullagar (2009) argue that the NSSE is too broad in scope, and is a survey of student educational experiences rather than a theoretical explanation of student engagement. Further, even though the NSSE contributes significantly to predicting students’ academic outcomes such as learning, critical thinking and practical skills in different institutional contexts (Carini et al., 2006; LaNasa et al., 2009), it primarily quantifies students’ involvement in specific actions and measures perceptions of their experience. This approach predominantly puts engagement responsibilities on leadership rather than instructional faculty (Burch et al., 2015).

Moreover, the NSSE benchmarks do not include engagement items of abstract/latent measures such as students’ individual thoughts, meta-cognitive understanding, perception of self-identity and autonomy, motivation or feelings towards education institution, peers, or learning environment. The lists of variables that constitute engagement concept theoretically expand beyond those that were used to create the NSSE (Pike, 2006). Burch et al. (2015) claim that the NSSE was developed to compare universities to one another, and therefore, aggregates student engagement to the college/university level, thereby
making it impossible to investigate course/class level engagement. Class or course-level engagement is of great significance for student learning, as it was found to explain the greatest percentage of variance in student achievement gains after student background characteristics (Antoniou and Kyriakides, 2011, 2013; Burch et al., 2015; Handelsman et al., 2005). Thus, there is a need to consider other engagement factors that contribute to learning.

Most of the theoretical and empirical attempts to describe student engagement agree on the convergence of three dimensions that comprise a broader understanding of this concept, namely, behavioural, emotional and cognitive dimensions (Burch et al., 2015; Conner, 2011; Fredricks et al., 2004; Kahu, 2013; Lawson and Lawson, 2013; Lester, 2013; Quaye and Harper, 2014; Trowler, 2010; Zepke and Leach, 2010). Focusing only on one or two of these dimensions would provide only partial explanation regarding possible antecedents and outcomes of engagement. Whilst the NSSE overemphasises engagement from primarily behavioural aspects (Kahu, 2013), it does not include indicators of other significant dimensions of engagement related with emotional and cognitive factors (LaNasa et al., 2009). Thus, a more in-depth understanding could be reached by conceptually and methodologically combining these perspectives to represent ‘one’ multidimensional construct. Findings of previous studies support that student engagement consists of more than just five dimensions, and further suggest that the concept should be understood in a more expanded, complex, and interdependent set of constructs (LaNasa et al., 2009).

At the same time, the emerging importance of productivity assessment of HEIs stresses the need for establishment of a better psychometrically-sound instrument that goes beyond traditional measures (Appleton et al., 2006; Gonyea, 2005; Gordon et al., 2008). In this context, this study employs a quantitative research methodology to compare and triangulate the NSSE and the Burch Engagement Survey for Students (BESS) instruments, to test their validity and reliability, and to empirically examine the impact of the behavioural, cognitive and emotional dimensions of student engagement on student learning. The rationale for this methodological triangulation rests upon recommendations by Gordon et al. (2008), LaNasa et al. (2007) and Pascarella (2006) to consolidate additional factors of engagement construct to the NSSE that would better predict students’ academic achievement, and improve understanding of this notion (LaNasa et al., 2009), as supported by the theories of student involvement. As Fredricks and McColskey (2012, p.779) state, “most current methods do not adequately capture the dynamic and interactive nature of engagement”, thus, “although the construct of student engagement has considerable promise, measurement issues should continue to be explored in order to fully realize this promise.” To sum up, “additional research is needed to fully understand the component structure of student engagement. Future research must seek to replicate these findings across and within institutional settings” [LaNasa et al., (2009), p.328] with different types, sizes and missions in order to understand how engagement varies across demographic student groups and changes over time (Krause and Coates, 2008).

3 Research methodology

In line with theoretical recommendations and study objective, this research triangulates the NSSE and BESS survey tools to acquire a more comprehensive understanding of
The impact of behavioural, cognitive and emotional dimensions

the multidimensional concept of student engagement and individual effects of complementary engagement types (behavioural, cognitive, emotional) on student success. In social research, triangulation refers to “combination of methodologies in the study of the same phenomena” [Denzin, (1973), p.297]. This process is followed by the in-depth analysis of survey questions that are expected to measure each separate engagement dimension.

Historically, student self-reported measures have been recognised as a credible and valid means for assessing academic development (Pohlmann and Beggs, 1974; Turner and Martin, 1985). Empirical evidence demonstrates that in surveys, college students are accurate and credible reporters of their actions, behaviours, college experience, and grades provided that questionnaires are well-structured and students have adequate information (Pace et al., 1985).

Measures that have been tested for their validity and reliability as well as proven for further use across diverse populations/contexts are crucial for social science, as development of new instruments is a time-consuming and expensive process (Harrington, 2008). Research further suggests that incorporation of existing measures can be cost-effective; nevertheless, to proceed further, a researcher is required to ensure the appropriateness of such technique of methodological triangulation and theoretical synthesis. This study adopted Harrington’s (2008) approach and focused on testing hypotheses via integration of existing theoretical models and operationalised instruments with CFA (Kline, 2015). One of the four primary functions of CFA – in addition to psychometric evaluation of measures, testing method effects, and testing measurement invariance – is to test construct validity of existing hypothetical construct applied to a new population (Brown, 2014; Harrington, 2008), namely to Azerbaijani HEIs context. According to Cronbach and Meehl (1955), construct validity refers to an assessment and operationalisation of theories (constructs) or a phenomenon that cannot be directly measured/observed (e.g., factors or latent variables) (Harrington, 2008).

3.1 Operationalisation of student engagement

3.1.1 Assumptions and operationalisation of NSSE

The NSSE psychometric assessment tool was designed to satisfy five fundamental conditions of self-reported measures (Kuh, 2002). Since 2000, more than 5.5 million students representing thousands of colleges across the world have extensively participated in the NSSE (2015). “The NSSE survey was designed by experts and extensively tested to ensure validity and reliability as well as to minimize non-response bias and mode effects” (Indiana University Center for Postsecondary Research, 2009). With the steadily growing number of HEIs participating in the NSSE, validity and reliability of this tool has been continuously field-tested in different contexts and environments (NSSE, 2015). Operationalisation stands for “developing operational definitions, or specifying the exact operations involved in measuring a variable” [Babbie, (2012), p.44]. Since 2014, the NSSE substituted formerly known five engagement benchmarks with four comprehensive themes of engagement. Previously used ‘enriching educational experiences’ benchmark is now reported separately as high-impact practices and not used to represent multi-dimensional nature of student engagement by the NSSE (2014). Accordingly, this study utilises four themes of engagement representing ten engagement indicators of the NSSE, namely:
1 level of academic challenge:
   IV1 higher-order learning
   IV2 reflective-integrative learning
   IV3 learning strategies
   IV4 quantitative reasoning

2 learning with peers:
   IV5 collaborative learning
   IV6 discussion with diverse others

3 experience with faculty:
   IV7 student-faculty interaction
   IV8 effective teaching practices

4 campus environment:
   IV9 quality of interactions

All ten NSSE engagement indicators were operationalised through ordinal scales. While 9 of these scales use four-option Likert scale (very often, often, sometimes, never), the ‘quality of interactions’ indicator uses a seven-option scale (ranging from 1 – poor to 7 – excellent). In order to obtain more detailed information regarding students’ perception, we have applied a five-option Likert scale (very often, often, sometimes, rarely, never) instead of four, for all engagement indicators.

3.1.2 Assumptions and operationalisation of BESS

Compared to the NSSE, the BESS is a relatively new model. Burch et al. (2015) report conducting BESS’s pilot testing followed by exploratory factor analysis with varimax rotation. Convergent validity issues were addressed, and reported factor items were adequate to truly represent the construct of engagement. Analysis results indicate that all BESS items (physical, emotional, cognitive in class and cognitive out of class) load on one construct of student engagement. The constructs of engagement were represented by four factors, each consisting of six indicators. Burch et al. (2015) report that internal consistency of four factors in the BESS model was tested through Cronbach alpha coefficients.

The BESS involves four components that measures emotional, physical, and cognitive (in-class and out-of-class) engagement. This study adopts only emotional and cognitive engagement components from the BESS, as most of the items on physical engagement dimension are less informative and overlap with the several items in the comprehensive ten indicators of the NSSE instrument. Cognitive engagement is grouped into two indices (in-class and out-of-class engagement), each measured by six questions:
IV11 cognitive in-class

IV12 cognitive out-of-class (Burch et al., 2015).

Emotional engagement scale was represented through six questions:

IV13 emotional feelings (Burch et al., 2015).

Similar to the NSSE, all three BESS indicators of engagement (cognitive in-class, cognitive out-of-class and emotional) were utilised with five-option Likert scale (very often, often, sometimes, rarely, never).

Overall, the questionnaire developed and used in this study comprises 13 independent variables (IVs) representing three dimensions, and one dependent variable (DV) of academic success. Figure 1 illustrates operationalisation of the student engagement concept.

Figure 1  Engagement dimensions and indicators (see online version for colours)

Notes: * Engagement indicators.
**Engagement dimensions.

3.2 Population and sampling

The total sample encompasses 43 different programs of study from six universities with diverse organisational profiles (see Table 1). The target population of this particular study comprised 2nd and 3rd-year undergraduate students, and the chosen sample represents 548 cases. Access to 4th-year students was not available, as students were mostly involved in their internship projects. Students of the 1st academic year were not included
in the sample, as their cumulative GPA was available only for one semester at the time of data collection (academic year consists of two semesters). The sample was composed of 43.4% females and 56.6% males. Of those, 56.2% were second-year and 43.8% third-year students. The total proportion of international students was 1.1% and nationals 98.9%. Taking into consideration that Azerbaijan is not a popular international destination for education, the low percentage of international students is a common character for most HEIs in the country.

Table 1  Profiles of universities included in the study

<table>
<thead>
<tr>
<th>Founded</th>
<th>Status</th>
<th>Budget ($)</th>
<th># of students</th>
<th># of faculty</th>
</tr>
</thead>
<tbody>
<tr>
<td>University A 1993</td>
<td>Private</td>
<td>6.2 million</td>
<td>4,300</td>
<td>422</td>
</tr>
<tr>
<td>University B 1990</td>
<td>Private</td>
<td>5.1 million</td>
<td>2,700</td>
<td>190</td>
</tr>
<tr>
<td>University C 2010</td>
<td>Public</td>
<td>2.1 million</td>
<td>1,760</td>
<td>183</td>
</tr>
<tr>
<td>University D 1921</td>
<td>Public</td>
<td>10.5 million</td>
<td>21,000</td>
<td>730</td>
</tr>
<tr>
<td>University E 1980</td>
<td>Public</td>
<td>2.5 million</td>
<td>2,500</td>
<td>350</td>
</tr>
<tr>
<td>University F 2006</td>
<td>Public</td>
<td>21 million</td>
<td>3,200</td>
<td>90</td>
</tr>
</tbody>
</table>

3.3 Data collection

3.3.1 Survey instrument

Data was collected by means of self-administered questionnaire containing 110 questions, which was distributed at six universities in Azerbaijan between April and June of 2017. The questionnaire was administered in two languages, English and Azerbaijani, based on the preference of national and international students. Translation of the questions into Azerbaijani was done with utmost possible accuracy respecting contextual meanings and relevance. More specifically, the questionnaire was translated into Azerbaijani following the approach proposed by Beaton et al. (2000). Accordingly, for the forward translation step, two independent translations of the original questionnaire were developed by two professional translators – native speakers of the Azerbaijani language fluent in English. A reconciled language version, on the basis of the two forward translations, and a report in English explaining translation decisions was developed. Further, for the backward translation step, the reconciled questionnaire in Azerbaijani language was translated into English by one professional translator – a native speaker of the Azerbaijani language fluent in English. The backward translation version and the original questionnaire were finally compared, and some minor discrepancies were resolved.

Translated questionnaire was validated and approved for correctness of translation or potential misinterpretation by an independent bilingual professor of social sciences (cultural validation) and an expert in English language translation (linguistic validation).

In order to secure a high response rate and to ease the data collection process, preference was given to hard-copy questionnaires rather than online administration (Cohen et al., 2007). The reason for this decision was based on various factors such as: likelihood for increased response rate, dearth of internet access at home and university required for participation in online questionnaire, logistical problems related to follow-up
The impact of behavioural, cognitive and emotional dimensions of the participants, cultural stereotypes of not taking online questionnaire as serious as hard copy, and inflexible bureaucratic procedures delaying access to students. The responses for most items were based on the five-point Likert scale ranging from 'very often' to 'never'. A number of other questions asked students for factual information such as GPA, desired highest education level, parental education level, university admission entrance scores, age and gender. Students were given a choice to opt-out and not to report their response, if they were not willing to do so.

Prior to the data collection phase, a pilot study was conducted among a small group of target students. Respondents were asked for specific issues such as the appropriateness of questions, clarity of instructions, and if they had any suggestions or comments. Based on the pilot study results, some of the items were modified to provide more clarity, and responses to the pilot survey were not included in the main sample data.

Overall, 900 questionnaires were distributed equally among six universities. Of those, 589 were completed (65% response rate), which could be considered as a satisfactory response rate (Babbie, 2012).

3.4 Data analysis

Following data collection process, the data were analysed through CFA to evaluate whether the measures of engagement construct are consistent with the theoretically-defined holistic portrayal of the construct. Using the SPSS AMOS (v. 23) (analysis of moment structures), we tried to exhibit whether the three dimensions model could be represented by one singular concept of student engagement. In addition, in line with the objective of this study, we also explored the relationship between a set of student engagement dimensions and students’ academic success.

4 Findings

4.1 Reliability and internal consistency

The next step in the analysis was to evaluate the reliability of each scale. Within the framework of this research, reliability of triangulated measurement of student engagement can be presented as the extent of internal consistency or homogeneity of the item scales within that survey instrument and the extent to which the results can be replicated across other people, times and settings. Internal consistency is used to judge reliability of an instrument by identifying how well items on a certain scale actually measure what they were supposed to measure (Trochim et al., 2015). This study employed Cronbach’s alpha reliability test. Nunnally and Bernstein (1994) argue that alpha reliability criterion of $0.9 \leq \alpha < 0.8$ is deemed as good and $0.8 \leq \alpha < 0.7$ as acceptable. Alpha values below 0.7 are considered poor and any values below 0.5 as unacceptable (Streiner, 2003). Initial internal consistency values measuring each engagement scale are illustrated in Table 2.

As can be seen from Table 2, the Cronbach’s alpha values for ETP, SE, CinCL, CoutCL and EE demonstrate values above 0.8, thus can be categorised as good. Coefficients for RIL, LS, QR, DWDO, SFI, and QOI also represent high values, $0.8 \leq \alpha < 0.7$, and can be regarded as acceptable. Despite being above minimum $\alpha$ threshold of 0.5, two of the engagement indicators, HOL and CL, initially reflected poor
coefficients. By removing items from HOL and CL which were found not to be appropriate and should be omitted to increase alpha coefficients, the internal consistency of the two indicators was improved. Particularly, item 3 in CL (Q.12 Come to class without completing readings or assignments) and item 1 in HOL (Q.30 Memorising course material) were removed from the respective scales. Table 3 presents alpha coefficients following the removal of the aforementioned items. As illustrated, all alpha coefficients exceed accepted ($\alpha > 0.6$) cut-off score. Items that were removed from respective scales (item 3 in CL and item 1 in HOL) were also not included during computation of IVs for HOL and CL engagement scales as well as in CFA or regression analysis.

**Table 2** Scale reliability coefficients before removing items

<table>
<thead>
<tr>
<th></th>
<th>HOL</th>
<th>RIL</th>
<th>LS</th>
<th>QR</th>
<th>CL</th>
<th>DWDO</th>
<th>SFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>N of items</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Cronbach’s alpha</td>
<td>.549</td>
<td>.757</td>
<td>.785</td>
<td>.773</td>
<td>.586</td>
<td>.784</td>
<td>.738</td>
</tr>
<tr>
<td>ETP</td>
<td>QOI</td>
<td>SE</td>
<td>CinCL</td>
<td>CoutCL</td>
<td>EE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N of items</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Cronbach’s alpha</td>
<td>.837</td>
<td>.786</td>
<td>.889</td>
<td>.899</td>
<td>.909</td>
<td>.892</td>
<td></td>
</tr>
</tbody>
</table>


**Table 3** Scale reliability coefficients after omitted items

<table>
<thead>
<tr>
<th></th>
<th>HOL</th>
<th>RIL</th>
<th>LS</th>
<th>QR</th>
<th>CL</th>
<th>DWDO</th>
<th>SFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>N of items</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Cronbach’s alpha</td>
<td>.786</td>
<td>.757</td>
<td>.785</td>
<td>.773</td>
<td>.608</td>
<td>.784</td>
<td>.738</td>
</tr>
<tr>
<td>ETP</td>
<td>QOI</td>
<td>SE</td>
<td>CinCL</td>
<td>CoutCL</td>
<td>EE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N of items</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Cronbach’s alpha</td>
<td>.837</td>
<td>.786</td>
<td>.889</td>
<td>.899</td>
<td>.909</td>
<td>.892</td>
<td></td>
</tr>
</tbody>
</table>


Given the small number of items representing each IV individually (N of items varying between 3 and 9), it was expected that gaining high internal consistency values would be difficult (Carmines and Zeller, 1979). Based on the alpha coefficients, it can be asserted that the scales of the NSSE and the BESS instruments represent significantly strong internal consistency.
### Table 4

Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>GPA</th>
<th>HOL</th>
<th>RIL</th>
<th>SFI</th>
<th>ETP</th>
<th>CL</th>
<th>LS</th>
<th>QR</th>
<th>DWDO</th>
<th>QOI</th>
<th>SE</th>
<th>CinCL</th>
<th>CoutCL</th>
<th>EE</th>
</tr>
</thead>
<tbody>
<tr>
<td>N valid</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mode</td>
<td>4</td>
<td>4.00</td>
<td>3.86</td>
<td>2.50</td>
<td>4.60</td>
<td>3.13</td>
<td>4.00</td>
<td>3.00</td>
<td>3.50</td>
<td>4.60</td>
<td>3.89</td>
<td>3.33</td>
<td>3.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>1.175</td>
<td>0.9466</td>
<td>0.7183</td>
<td>0.8278</td>
<td>0.5871</td>
<td>0.9316</td>
<td>1.0058</td>
<td>1.0055</td>
<td>1.3765</td>
<td>0.9168</td>
<td>0.8978</td>
<td>0.9244</td>
<td>0.9714</td>
<td></td>
</tr>
</tbody>
</table>

4.2 Descriptive statistics

After analysis of internal consistency reliability, descriptive statistics for the research variables are presented in Table 4.

The survey based on 548 responses also collected background information about students’ university entrance admission score (Q2), education year (Q3), gender (Q4), desired highest level of education (Q5), age (Q6), major (Q7), student status (Q8), and parents’ education (Q9). Accordingly, the average age of participant students was 19.78, while gender distribution was as 43.4% for males and 56.6% for females. Out of 548 respondents, only 13 students (2.4%) were international students, while in terms of class year, 308 (56.2%) were sophomore, while 240 (43.8%) were junior students. In terms of academic performance (out of 100), 59 students (10.7%) had a GPA below 61, only 42 students (7.7%) had a GPA above 90, and the remaining majority (447 students) had a GPA between 61 and 90. Collected background information was not used for multiple regression or CFA purposes.

4.3 Confirmatory factor analysis

4.3.1 Multivariate normality

Implementation of CFA test requires certain pre-conditions that need to be met in advance. One of the core assumptions is the multivariate normality of the involved variables in the data (Kline, 2015). Because the dependent and IVs are categorical using five-option Likert scale, and the sample size is big, normality test is not considered strictly compulsory. Nevertheless, each of the study variables was tested for their skewness and kurtosis values. Both skewness and kurtosis values for all variables were found to have acceptable cut-off range between –1 and 1 and –2 and 2, respectively (George and Mallery, 2010). As a next step, data imputation techniques were administered to substitute missing values in the sample data. Since IVs of the study are categorical, missing data was substituted with the most frequent value (mode) of each respective engagement indicator scale, as Williams (2011) suggests. From the total of 589 responses, 41 questionnaires had missing values varying between 5% and 19% (above acceptable threshold), which were not included in the sample data, as they could produce biased results. Of the remaining 548 responses, 17 cases (3% of the total sample) had 2%-5% missing values.

4.3.2 Sample size and adequacy

The second requirement for CFA is the size of the sample. Although there is no accepted ‘rule of thumb’ regarding adequate sample size for CFA (Kline, 2015), the generally recommended minimum number of sample is between 100 and 200 (Cohen et al., 2013; Wolf et al., 2013). In this research, we were able to secure data from 548 participants. Therefore, collected sample size meets CFA criteria. In order to determine how suitable the data is, a Kaiser-Meyer-Olkin (KMO) test of sampling adequacy was conducted. Accordingly, provided KMO value ranges between 0.8 and 1.0, which means that the sampling is adequate, according to Cerny and Kaiser (1977). Field (2013) considers KMO values between 0.7 and 0.8 as good. KMO values below 0.6 indicate that sampling
is not adequate for factor analysis; thus, correctional actions should be considered (Cohen et al., 2013). We also conducted Bartlett’s test of sphericity to determine if homogeneity of variances exists across the sample (Singh, 2007). Bartlett’s sphericity test compares correlation matrix and identity matrix for possible redundancies between variables. Sphericity test uses chi-square statistics and any $p$ value below 0.05 is considered statistically significant, suggesting suitability of the sample data for principal component analysis (Singh, 2007). Considering that current study utilises CFA, not exploratory factor analysis, conducting KMO Bartlett’s sphericity test is seen as an extra measure to assure adequacy of sample data for the analysis. The data met fundamental factorial analysis assumptions and satisfied recommended thresholds with a strong KMO coefficient of 0.924 and $\text{sig.} = 0.001$ value.

4.3.3 Bipartite correlation

As a preliminary step, in order to conduct CFA, SPSS bivariate correlation test was conducted both for the NSSE and the BESS engagement indicators separately, and between the indices of these two instruments. Because the sample data is mostly categorical, Spearman’s rho test was preferred to determine the strength of association among IVs. Table 5 illustrates inter-item correlation coefficients for 11 IVs (HOL, RIL, SFI, ETP, CL, LS, QR, DWDO, QOI and SE) operationalised through the NSSE. All of the $r$ coefficient values (except $\text{sig.} = 0.149$ for DWDO and ETP) are statistically significant and lower than 0.8.

The same test was performed for the BESS engagement indicators. Table 6 illustrates correlation coefficients for three IVs (CinCL, CoutCL and EE) of the BESS model. Results indicate that the $r$ coefficient values for cognitive and emotional engagement scales are also statistically significant and below the deemed threshold ($r \leq 0.8$).

4.3.4 CFA test

CFA test, using the AMOS (SPSS v. 23) software, was utilised to examine the construct validity of the newly-triangulated model composed of three second-order factors. Accordingly, while ‘behavioural’ factor combines ten computed IVs, ‘cognitive’ factor is represented by two computed IVs of in-class and out-of-class engagement. Since ‘emotional’ engagement dimension did not have any second-order factor, it was substituted with six items measuring students’ enthusiasm, energy, interest, pride, positive attitude and excitement. Figure 2 illustrates standardised estimates and correlation covariance among observed variables. In order to come up with a ‘good factor’ interpretation, the study followed Comrey and Lee’s (2013) and Hair et al.’s (1998) stringent cut-off categorisation. Comrey and Lee (2013) consider any factor loading ($\lambda$) value higher than 0.32 as poor, 0.45 as fair, 0.55 as good, 0.63 as very good, and 0.71 as excellent. Hair et al. (1998) emphasise that for any sample size over 350, a minimum factor loading value of 0.3 can be reported. The model shows that none of the standardised factor loadings was below 0.3. Overall, 8 out of 10 loadings in ‘behavioural’ dimension can be labelled between the range of good and very good. Compared to ‘behavioural’ dimension ($0.32 \leq \lambda \leq 0.67$), standardised estimate loadings for cognitive ($\text{CinCL} = 0.86$, $\text{CoutCL} = 0.80$) and emotional dimensions ($0.69 \leq \lambda \leq 0.82$) were much stronger, thus representing high convergent validity.
Table 5

<table>
<thead>
<tr>
<th>Spearman’s rho</th>
<th>HOL</th>
<th>RIL</th>
<th>SFI</th>
<th>ETP</th>
<th>CL</th>
<th>LS</th>
<th>QR</th>
<th>DWDO</th>
<th>QOI</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>RIL</td>
<td>Correlation coefficient</td>
<td>0.581**</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>SFI</td>
<td>Correlation coefficient</td>
<td>0.415**</td>
<td>0.440**</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>ETP</td>
<td>Correlation coefficient</td>
<td>0.378**</td>
<td>0.259**</td>
<td>0.190**</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>CL</td>
<td>Correlation coefficient</td>
<td>0.362**</td>
<td>0.406**</td>
<td>0.397**</td>
<td>0.184**</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>LS</td>
<td>Correlation coefficient</td>
<td>0.473**</td>
<td>0.416**</td>
<td>0.331**</td>
<td>0.398**</td>
<td>0.243**</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>QR</td>
<td>Correlation coefficient</td>
<td>0.390**</td>
<td>0.398**</td>
<td>0.446**</td>
<td>0.224**</td>
<td>0.257**</td>
<td>0.311**</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>DWDO</td>
<td>Correlation coefficient</td>
<td>0.271**</td>
<td>0.354**</td>
<td>0.241**</td>
<td>0.062</td>
<td>0.245**</td>
<td>0.140**</td>
<td>0.295**</td>
<td>1.000</td>
<td>-</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>QOI</td>
<td>Correlation coefficient</td>
<td>0.275**</td>
<td>0.239**</td>
<td>0.248**</td>
<td>0.360**</td>
<td>0.230**</td>
<td>0.232**</td>
<td>0.214**</td>
<td>0.117**</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
<tr>
<td>SE</td>
<td>Correlation coefficient</td>
<td>0.378**</td>
<td>0.273**</td>
<td>0.243**</td>
<td>0.528**</td>
<td>0.184**</td>
<td>0.362**</td>
<td>0.239**</td>
<td>0.204**</td>
<td>0.396**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
<td>548</td>
</tr>
</tbody>
</table>

Notes: **Correlation is significant at the 0.01 level (2-tailed).
The impact of behavioural, cognitive and emotional dimensions

Table 6  Bivariate correlation for BESS

<table>
<thead>
<tr>
<th></th>
<th>CinCL</th>
<th>CoutCL</th>
<th>EE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s rho</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CinCL</td>
<td>Correlation coefficient 1.000</td>
<td>N 548</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoutCL</td>
<td>Correlation coefficient .691** 1.000</td>
<td>N 548</td>
<td>N 548</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>EE</td>
<td>Correlation coefficient .714** .656** 1.000</td>
<td>N 548</td>
<td>N 548</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.</td>
</tr>
</tbody>
</table>

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

Figure 2  CFA model and standardised factor loadings (see online version for colours)
The model in Figure 2 shows that standardised factor loadings for DWDO and QI are poor and do not contribute significantly to the formation of behavioural engagement dimension compared to other variables in the same set. At first sight, it is intriguing to understand why this is the case. DWDO engagement indicator quantifies students’ interaction with people from diverse ethnic, economic, religious and political backgrounds and seeks to find the extent to which such interaction impacts their learning. It is a well-accepted fact that classroom and non-classroom experiences of diversity enrich and enhance the impact of universities on students’ productivity. Students that interact with a diverse spectrum of people, ideas, values and perspectives are cognitively challenging their personal and assumed views of truth and how their perceptions are different from others (Pascarella, 2006).

Considering the facts that more than 97% of the total population of Azerbaijan are identified as Muslim, less than 8% are ethnic minorities that are dispersed geographically (CIA Factbook, 2017), a large majority are in the same economic class and have same political views, it is not surprising to see how these elements do not have significant impact on students’ learning process. In other words, visibility or experience of these diversity elements in homogenous socio-economic environment of Azerbaijan is almost non-existent. For example, the proportion of international students in the whole sample size of 548 constituted only 1.1% (six students). Hence, during their education life, national students rarely get an opportunity to communicate with a diversity of international students. This is also true for other diversity elements (e.g., students with different sexual orientation). On the other hand, QI scale collects information about students’ perception of interaction with peers, academic advisors, faculty, service staff and others. The research by Mearns et al. (2007) highlights that, if teachers are perceived to be approachable and sensitive to needs, students tend to become more committed to their academic work and willing to express their thoughts. Constant academic support and peer interaction improves student-faculty and student-student collaboration, and strengthens the sense of institutional belongingness (Zepke and Leach, 2010). However, the process of learning in Azerbaijani context is strictly bounded to the classroom environment. Adhering to old-fashioned teaching heritage, academic staff at the universities often time put significant barriers to in-class and out-of-class interaction with students. The learning process takes place strictly in formal classroom environments and any other unplanned attempts for interaction outside the class are regarded as unwanted. Such student-teacher relationship is hierarchical. This is also applicable to other university staff and units. Nevertheless, the model shows that the cut-off score for DWDO and QI exceeded recommended minimum boundary (0.32); thus, we have decided to keep them in the model and regression analysis. In general, loadings of ‘cognitive’ and ‘emotional’ dimensions are above 0.63 cut-off value, thus, accepted as very good and excellent fits.

In terms of collinearity among engagement dimensions, the model shows acceptable correlation coefficients between behavioural-emotional (0.64) and behavioural-cognitive (0.69), but relatively high values for cognitive-emotional (0.91) latent variables. LaNasa et al. (2009) state that high correlation in CFA is a multicollinearity issue, and high collinearity leads to very large standard errors. Correlation between factors that exceeds 0.85 indicates poor discriminant validity. This result is somehow contradictory to multicollinearity test results, which describes low correlation coefficient values (0.656 ≤ r ≤ 0.714) among cognitive and emotional engagement indicators that explain less than 50% of variance (r²). In addition to bivariate correlation test, the results of VIF
The impact of behavioural, cognitive and emotional dimensions

Multicollinearity diagnostics in SPSS did not reveal any collinearity problem among IVs. Therefore, multicollinearity assumption of the CFA test is disregarded. The possible cause for CFA multicollinearity warning could be statistical distortion emerging from correlating computed engagement variables of ‘cognitive’ dimension with the items (not variables) of ‘emotional’ dimension, due to the fact that the latent variable of ‘emotional’ engagement did not have second-order factor.

4.3.5 CFA model fitness

In a CFA test, fit indices describe the extent to which proposed theory is acceptable and fits the data (Hooper et al., 2008). Fit indices explain the difference between observed covariance matrix and model/predicted covariance matrix (Moss, 2009). Hu and Bentler (1999) classify CFA fit indices in three categories: absolute fit indices (chi-squared test, RMSEA, GFI, AGFI, SRMR and RMR), relative fit indices (NFI and CFI) and parsimony fit indices (PGFI and PNFI).

In order to judge how good the hypothesised model is, we decided to include all of the mentioned indices. The model fit analysis was performed using SPSS AMOS software, and Table 7 shows fit indices result of the hypothesised model of this study.

Table 7  CFA model fit summary

<table>
<thead>
<tr>
<th>Chi-square</th>
<th>Absolute</th>
<th>Relative</th>
<th>Parsimony</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CMIN/DF</td>
<td>RMSEA</td>
<td>GFI</td>
</tr>
<tr>
<td>.000</td>
<td>4.471</td>
<td>.80</td>
<td>.880</td>
</tr>
</tbody>
</table>

Although the model’s chi-square value was bigger than acceptable threshold (sig. > 0.05), Hooper et al. (2008) assert that statistical significance test is sensitive to sample size, which means that the chi-square statistic nearly always rejects the model when large samples are used. Considering the large sample size composed of 548 cases, we disregarded chi-square value and accept model fitness. AMOS also reports statistical significance with CMIN and any value below 5 is deemed acceptable (Moss, 2009). All of the fit indices, with some on the border line, comfortably fit the data. While some of the indices show very modest difference from acceptable cut-off norms (e.g., GFI, CFI or NFI), we have considered them as ‘acceptable’, since there are also plethora of other empirical studies arguing for the feasibility of lower value standards for those indices.

Table 8  Bivariate correlation of NSSE and BESS

<table>
<thead>
<tr>
<th>NSSE</th>
<th>BESS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>NSSE Correlation coefficient</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
</tr>
<tr>
<td>BESS Correlation coefficient</td>
<td>.554**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>548</td>
</tr>
</tbody>
</table>

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

Lastly, Table 8 illustrates correlations between the composite NSSE and the BESS variables/dimensions. In accordance with the parsimony principle, one might question
why both of the measurement instruments are necessary. As Table 8 shows, alpha value of $r = .554$ expresses relatively low association. Provided that the NSSE and the BESS instruments were identical, the coefficient value would be expected to be strong. If this was the case, then it could be claimed that both of the instruments are not needed. Nevertheless, based on the given $r$ value, it can be assumed that these two instruments do not overlap completely and do not represent alternative measures of the same concept (Pedhazur, 1997). In other words, they measure different layers of the same construct.

4.4 Multiple linear regression

Two separate regression analyses were conducted in SPSS to test study Hypotheses H$_1$ and H$_2$. To explore the association between the NSSE engagement indicators and students’ academic outcome, linear regression analysis was chosen. The first analysis tested the explanatory power of NSSE indicators by regressing ten IVs of engagement (representing behavioural dimension) with one DV of GPA (Table 9). The second regression analysis tested explanatory power of hypothesised CFA model (incorporated engagement dimensions of the NSSE and the BESS) on students’ academic success (Table 10).

H$_1$ There is a positive relationship between student engagement and academic performance for both models (one-dimensional model comprising behavioural scale of engagement and three-dimensional model comprising behavioural-cognitive-emotional scales of engagement).

Table 9 Behavioural engagement regression model summary

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R$ square</th>
<th>Adjusted $R$ square</th>
<th>Std. error of the estimate</th>
<th>$B$</th>
<th>Std. error</th>
<th>$t$</th>
<th>$F$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.426$^a$</td>
<td>.181</td>
<td>.166</td>
<td>1.073</td>
<td>1.255</td>
<td>.331</td>
<td>3.796</td>
<td>11.873</td>
<td>.001$^b$</td>
</tr>
</tbody>
</table>

Notes: $^a$DV: grade point average.  
$^b$Predictors: (constant), SE, CL, DWDO, QR, QI, LS, HOL, SFI, ETP and RIL.

Table 10 CFA regression model summary

<table>
<thead>
<tr>
<th>Model</th>
<th>$R$</th>
<th>$R$ square</th>
<th>Adjusted $R$ square</th>
<th>Std. error of the estimate</th>
<th>$B$</th>
<th>Std. error</th>
<th>$t$</th>
<th>$F$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.453$^a$</td>
<td>.205</td>
<td>.185</td>
<td>1.060</td>
<td>1.162</td>
<td>0.334</td>
<td>3.478</td>
<td>10.582</td>
<td>.001$^b$</td>
</tr>
</tbody>
</table>

Notes: $^a$DV: grade point average.  
$^b$Predictors: (constant), EE, DWDO, QI, CL, QR, SE, HOL, SFI, LS, ETP, RIL, CoutCL and CinCL.

The regression results in Table 9 show that the model is statistically significant ($sig. = 0.001$). Overall, engagement indicators of behavioural dimension explain approximately 18% of variance in students’ academic attainment ($R$ square = 0.181). When evaluating standardised beta values ($sig. \leq 0.05$), the greatest influences on DV is caused by RIL (beta = 0.172), CL (beta = -0.130), DWDO (beta = 0.119), ETP (beta = 0.174) and QI (beta = 0.106).

H$_2$ The model comprising behavioural, cognitive and emotional scales of engagement provides more explanation of academic achievement than the model comprising solely behavioural scale of engagement.
The impact of behavioural, cognitive and emotional dimensions

Again, the $\text{sig.} = 0.001$ value indicates that the model is statistically significant. Compared to the first regression test, the $R^2 = 0.205$ demonstrates a modest, yet important improvement in the explanation of variance, increasing from 0.181 to 0.205. In other words, the three-dimensional model composed of behavioural, cognitive and emotional dimensions of engagement explain approximately 21% of students’ academic achievement. Based on this result, the researcher accepts $H_2$ and assumes that the three-dimensional CFA model better explains variance in student learning. Standardised beta values ($\text{sig.} \leq 0.05$) of the variables with highest influences on DV are RIL ($\beta = 0.156$), CL ($\beta = -0.116$), DWDO ($\beta = 0.128$), ETP ($\beta = 0.148$), CinCL ($\beta = -0.121$) and CoutCL ($\beta = 0.207$).

5 Discussion

It is important to acknowledge that many countries, as does Azerbaijan, include greater numbers of students in HE. As students become more diverse, as HE becomes more competitive, as the export of international education continues to grow, and as demand for greater numbers of capable graduates increases, there is an intensified need for sound insights on whether students are engaging effectively with university education (Coates, 2010). There is a growing need to understand how to engage students from enrolment through to graduation (Coates, 2005).

In light of the suggestions in the recent literature, this paper adds to the understanding of dimensionality of student engagement. Specifically, this study:

- articulated student involvement theories and their reflection on engagement dimensions
- developed a set of hypotheses
- empirically tested hypotheses through established methodological instrument in order to understand whether the engagement constructs are interdependent and load significantly on the construct they purport to measure.

Investigation results bound with the conceptual framework of this study demonstrated that the three theoretically and empirically distinct dimensions of student engagement (behavioural, cognitive and emotional) are inter-related, and constitute different aspects of the same concept (Fredricks et al., 2004; Kahu, 2013; Lawson and Lawson, 2013; Lester, 2013; Quaye and Harper, 2014; Trowler, 2010). The empirically derived CFA model indices represent a good fit, and compared to the traditionally-used NSSE measure, provides better explanation of students’ academic achievement, personal development, emotional and psychological states that affect their learning experiences. The present study not only points to the significance of the outcomes and benefits of possible engagement, but also emphasises the interplay of antecedents that lead to high engaging classrooms and campuses. These results were not meant to undermine the NSSE or promote the BEES instruments, but rather to convey additional and/or combined insight.

In addition to the above-mentioned finding, this study reinforces the claim regarding consideration of multidimensional engagement construct as an imperative factor leading to better authentic learning (Fredricks et al., 2004). Regression results adhere to the conceptual premise derived from the theoretical synthesis, emphasising that the very act
of being engaged is not only a process-related activity or task/behaviour, but also a product of students’ emotional and cognitive-affective conditions (Burch et al., 2015). Accordingly, the difference between the results of the regression tests for one-dimensional (behavioural) and three-dimensional (behavioural-cognitive-emotional) models reveals that compared to the NSSE measurement, multidimensional engagement model derived in this study provides a better explanation of academic success.

The findings of this study could have important implications for university leadership, faculty administrators, instructors, policymakers and students, overall contributing to social change. Our results could, therefore, be expanded to reconsider the importance of and the ways through which student engagement could be measured and taken into consideration in promoting development of a theory-driven and evidence-based model and quality management in HE (Antoniou and Mohan, 2017; Coates, 2010). Particularly, those responsible for the development of policy at the university level could raise important questions, based on the results of their student engagement level evaluations. Such questions could be related with the conditions that need to occur to improve student engagement and ways to overcome barriers to university engagement. What does the university do to improve student engagement in all three domains reported in this paper? How does the culture of a university help or hinder such attempts?

In terms of social change, the study results enlighten university leadership regarding the importance of students’ in-class and out-of-class psychological state that affect their performance, behaviour, achievement and well-being. Considering that HEIs are eager to demonstrate enhancement of institutional quality and added value (Krause, 2005), understanding the impact of individual engagement factors promises several improvements. Therefore, it is crucial to outline clear borders and distinguish key effects of each type of engagement on institutions, faculty and students. Such distinctions would empower institutional leaders and instructors with better practical understanding of what is lacking and how to improve it instead of simply answering where the lack is.

In addition, in response to continuous government pressure and push for accountability, universities need to demonstrate that they add value by improving (Krause, 2005) students’ life by contributing to their learning and self-development. Considering ever-growing diversity of students, institutional settings and learning environments, implementation of the universally-applicable quality assurance etalons or measurement standards eases general understanding of lacking areas. Empirical evidence from this study reinforces the significance of student engagement for efficacy of academic success that leads to satisfaction and lifelong learning. The pursuit of scholarship to methodologically advance current self-report assessment instruments is obvious. The NSSE has a crucial value for HEIs that aim to measure what students are doing and what they get in return from their universities (LaNasa et al., 2009). Nevertheless, based on the results of this study, it appears that, while promising overarching assessment of student engagement, the NSSE primarily captures the effects of educationally-beneficial activities and significantly isolates cognitive and emotional elements of learning. Although the founders of the NSSE approach student engagement as a concept of ‘deceptive simplicity’, the heightened inquiry of the scholarship concerning construct validity of this construct (LaNasa et al., 2009) is substantiated by this investigation.

In line with the objectives of this study and theoretical discussion, this investigation revealed that the notion of engagement is composed of an interplay of several critical dimensions. Limiting the focus to one or incorporation of two of these cornerstones by
The impact of behavioural, cognitive and emotional dimensions in institutions or policymakers provides only a partial picture of what happens inside classrooms, campus settings or students’ minds. The results of the regression analyses in this paper revealed that the psychometric self-report instrument consolidating fundamental in-class and out-of-class factors of cognitive and emotional engagement with students’ behavioural actions significantly improves explanation of learning outcomes.

The findings of this investigation support the claim that, at the same time with decomposing and adding finer sub-scales to the NSSE instrument (Gordon et al., 2008; LaNasa et al., 2007; Lerer and Talley, 2010; Pascarella, 2006), holistic consideration of contexts, dimensions, factors and domains of engagement adds significant value to the evaluation of institutional outcomes. Such in-depth information equips educational leaders with decision-making strategies for the early prevention of disengagement causes, rather than late intervention and mitigation of possible consequences.

Students’ engagement in HE cannot be taken for granted, and needs to be designed and led in effective ways. It is hoped that this study could stimulate discussions on how university students engage in high-quality learning. This may unfold in a basic fashion by adding the term ‘student engagement’ to the lexicon of a HE policy and practice or, more profoundly, by treating it as an “organizing construct for institutional assessment, accountability, and improvement efforts” [Kuh, (2009), p.5]. We therefore hope that this study could provide a solid framework to initiate changes in system and institutional thinking on how to improve students’ engagement in high-quality university education.

6 Conclusions

This study focused on the concept of student engagement with the purpose to advance methodological tools used to measure the concept. The study results, bound with the conceptual framework of this study, demonstrate that when the three theoretically and empirically distinct dimensions of student engagement, namely, behavioural, cognitive, and emotional, are added to the traditional NSSE instrument, the new expanded model bears a stronger model to measure the concept of student engagement. In particular, the cognitive and emotional elements of learning process were found to be important in measuring student engagement. The study, thus, reinforces the scholarly claims in favour of multi-dimensional models to measure engagement.

The study has a number of limitations. Covering only specific universities, programs, and year cohorts might represent only a partial picture of the student engagement concept. In addition, focusing only on Azerbaijan, which adds its own cultural aspects into the picture, is a second possible factor limiting generalisation of the findings to international contexts. Further, this study has been mainly of quantitative nature, lacking a qualitative insight through possible focus groups or one-to-one interviews. Last, being a cross-sectional study, there is no analysis of how student engagement concept evolves over time in response to institutional and environmental factors.

Further research could elaborate our understanding in terms of the ways through which student engagement impacts learning by utilising qualitative and longitudinal methodological designs. In addition, replication studies could be conducted in more countries to facilitate our attempts to provide a better understanding of the relationship between engagement dimensions. Such approaches could provide policymakers and educators with information and decision-making directions to proactively intervene and
prevent the causes of disengagement/alienation, rather than reactively mitigate or alleviate possible negative outcomes. In addition, it is significant to consider contextual differences when studying student engagement. The variation of learning cultures, teaching-student relationships, academic leadership, student agency and bonding values might reflect differently on each dimension. In this perspective, international longitudinal studies with repeated observations and measures could provide a more detailed understanding of the concept having stronger causal explanatory power.

References


Field, A. (2013) Discovering Statistics Using IBM SPSS Statistics, Sage, California [online] https://books.google.co.uk/books?id=en&hl=en&id=0Wk9MuBmAoC&oi=fnd&pg=PP2&dq=Discovering+Statistics+Using+SPSS&tots=LallHKWv0J&sig=D8v2XmsfOBgBtsp0r-xmonTk1ow.


The impact of behavioural, cognitive and emotional dimensions


