CriSTaL
CRITICAL STUDIES IN TEACHING & LEARNING

2021, Volume 9, Issue 1
DOI: 10.14426/cristal.v9i1.316

Data-Driven Predictions of Academic Success among College Students in Saudi Arabia

Omar J. El-Moussa, Runna Alghazo[#], and Maura A. E. Pilotti
Prince Mohammad Bin Fahd University
[#]Corresponding Author: ralghazo@pmu.edu.sa

(Submitted: 12 September 2020; Accepted: 27 January 2021)

Abstract

Predictors of college performance are markers of learners' characteristics that can be used to optimize admission, advising, counseling, and instruction. The present study focused on female and male graduates of a Saudi Arabian university that follows a USA general education curriculum. Saudi Arabia exemplifies a society in transition from a rigid patriarchal system to one that is more gender equitable. The study investigated the extent to which gender, high-school Grade Point Average (*hs-GPA*), and the *GAT* (equivalent to the *SAT I*) can predict *GPA* at graduation, as well as verbal, analytical, and quantitative competencies of graduates in Business, Engineering, and Law. In our study, females outperformed males on most measures except on the *GAT*. Gender differences in the choice of major were also found. For all, *hs-GPA* and *GAT* were poor predictors of academic success. Alternative measures were proposed along with the use of a data-driven approach for predicting students' performance at a given institution.

Keywords: academic success; Middle East; gender differences; Saudi Arabia; standardized tests; University *GPA*

Introduction

In higher education, college admission officers usually rely on cognitive measures, such as High School Grade Point Average (*hs-GPA*) and scores on standardized tests, such as the American College Testing (*ACT*), General Achievement Test (*GAT*), or Scholastic Assessment Test (*SAT*), to assess college students' readiness, upon which decisions regarding admission, conditional admission, or rejection rest (Kobrin, et al., 2008; Westrick, et al., 2015; Zwick, 2006). The same measures can be used by administrators to predict students' need for services, such as advising, counseling, and remedial education, to allocate available funds equitably, and by educators, advisors, and counselors to optimize ordinary practices and develop targeted interventions (Khoshaim and Ali, 2015).

In the Western world, reliance on standardized tests for admission purposes has been controversial for a variety of reasons, including culturally biased content, tendency to favour

affluent test-takers and discriminate against ethnic minorities, and uneven predictive validity of different facets of academic performance (Lawlor, et al., 1997; Murray, 2012; Rattani, 2016; Rothstein, 2004; Soares, 2015). Large-scale studies or reviews illustrating the extent to which standardized tests, such as the *SAT*, can predict college performance have not hampered controversies (Burton and Ramist, 2001; Mattern, et al., 2011; Sackett, et al., 2012; Shaw, 2015). However, they have encouraged higher education institutions to reexamine their practices and institute changes endorsing fairness, such as optional test score submission, and holistic review of students' pre-college records. In the Middle East, notwithstanding controversies, the assumed informational value of cognitive measures has generally remained high, even in the face of scarce or incomplete evidence regarding their ability to predict college success (Alghamdi and Al-Hattami, 2014; Alnahdi, 2015; Sulphey, et al., 2018). Doubts have emerged though. First, most of the available findings pertain to students whose majors involve health sciences, medicine, or nursing (Dabaliz, et al., 2017; Urlings-Strop, et al., 2013). Second, the psychometric properties of the standardized tests used to estimate knowledge and skill foundations, such as the *Qudrat* (also known as General Aptitude Test, *GAT*), may not be equivalent to other more well-known tests they purportedly intend to simulate. For instance, the *GAT*, which is administered in Arabic by the National Center of Assessment in Higher Education (NCAHE) of the Kingdom of Saudi Arabia (KSA) and is required of all high-school students who apply to a higher education institution, is assumed to measure analytical and deductive skills in the verbal and quantitative domains. It is considered equivalent to the *SAT I* since it assesses students' general ability for information processing regardless of any specific skill. Although equivalence is assumed, the psychometric properties of the *GAT* are not as well established as those of its English counterpart and its predictive validity is an unclear matter (Alnahdi, 2015; Alshumrani, 2007; Althewini, 2019; Dabaliz, et al., 2017; Dimitrov, et al., 2015; Sideridis, et al., 2015), especially when students' performance pertains to universities which rely on an imported (e.g., USA) educational model taught in a foreign language (e.g., English). Third, universities in the Middle East often use a composite score with varying weights given to different factors, such as *hs-GPA* and *GAT* grades, to determine the admission eligibility of a student. Usually, standardized test scores are given greater weight than *hs-GPA* (Siddiek, 2011). Furthermore, individual differences, whose relevance may not be universal, but rather unique to the population that a specific institution or program of study serves, are often overlooked. Differences may pertain to demographic factors, such as college major (Alghamdi and Al-Hattami, 2014) and gender, or motivational factors, such as self-efficacy (Bussey and Bandura, 1999; Kitsantas, et al., 2008), self-determination, engagement, and self-regulation (Gregg, 2009; Kappe and Van Der Flier, 2012; Richman, et al., 2014; Zepke and Leach, 2010), which are related to exerted effort and persistence. For instance, consider that in the Western world, standardized tests and school marks exhibit non-overlapping gender differences. Namely, female students have usually greater school marks regardless of the subject matter (Voyer and Voyer, 2014), whereas standardized tests tend to favour males, especially when quantitative scientific competencies are measured (Hedges and Nowell, 1995; for evidence beyond the Western world, see also Akpotor and Egbule, 2020). Null gender differences or recent

declines in the magnitude of the differences in standardized tests have also been reported (Else-Quest, et al., 2010; Feingold, 1988). In Saudi Arabia, however, evidence of gender differences tends to favour females on both standardized tests and *hs-GPA* (Alghamdi and Al-Hattami, 2014). Yet, the rationale for the selection of predictors and the weight given to each predictor at particular universities are not entirely clear. Similarly, institutions, and colleges within each institution, may conceptualize the outcome variable 'academic success' differently, such as cumulative *GPA* at graduation or earlier, or grades in courses considered key to success in a given major or profession (Alghamdi and Al-Hattami, 2014; Alnahdi, 2015; Al-Alwan, 2009; Al-Tamimi and Al-Shayeb, 2002; Sulphey, et al., 2018).

The motivation for the present study rests on a concern informally expressed by a variety of college admission officers, administrators, and educators in the Middle East regarding the need to identify the particular predictors of academic success at their institution. The concern generally stems from a sense of dissatisfaction with the formulas used for admission decisions, which may be replicas of those used by other institutions or result from the extant literature, often of Western import, where the contribution of mostly cognitive factors to academic success is recorded. The statement 'one size does not fit all in higher education' is frequently the likely conclusion of expressions of dissatisfaction with current practices. It is easy to understand this feeling. When the factors included in the admission formula adopted by a particular institution have low predictive validity for that institution, opportunities for sensible decisions, adequate advising, and early interventions are undermined.

Through a retrospective case study, we intend to offer a simple demonstration of how key factors that are often used for admission decisions of young men and women in the Middle East may be largely ineffective for a particular institution, especially when the institution adheres to a USA instructional model and curriculum taught in the students' second language (i.e., English). We focus on college students from a university in the Kingdom of Saudi Arabia (KSA) because the social context in which this model is asked to exist is one of transition, from inequities of educational and economic opportunities for men and women to one that attempts to level the field (Abu-Ghaida and Klasen, 2004; Al Alhareth, et al., 2015). Thus, it offers a unique window for assessing the extent to which gender differences in the choice of major and academic performance reflect past inequities in opportunities or the whiff of change.

The social context upon which the imported educational model sits can be characterized by a past of sustained gender inequity in educational and economic opportunities available to its youth (Alhareth, et al., 2015; Soekarba, 2019). For instance, until 2002, all levels of women's education were overseen by the Department of Religious Guidance, whereas the education of males was overseen by the Ministry of Education. Until 2001, women were considered an extension of their male guardians, unable to exist as independent beings. As a result, women's education was focused on creating good wives and mothers, as well as preparing them for jobs that were believed to suit their 'unique nature'. In recent times, KSA has attempted to rectify systemic inequities through formal decrees and massive investments of resources to ensure that opportunities are evenly distributed between women and men. The implementation of the KSA

2030 Vision, a strategic framework of economic and systemic changes to move the country into the global economy, has increased women's participation in the workforce, intending to promote fair, equitable, and supportive working conditions, and financial independence (Topal, 2019). Yet, education, including higher education, is still largely segregated by gender. Although there are more female than male students in higher education (Alhareth, et al., 2015), women remain mostly underrepresented in several fields, including science, technology, engineering, and mathematics (STEM; Alwedinani, 2016). The latter pattern mirrors that of the Western world, where gender differences in the choice of major are characterized by STEM fields being dominated by males (Finger, et al., 2020; Porter and Serra, 2020; Witherspoon and Schunn, 2020).

In KSA, the institutional changes that open a particular field to women are usually in the form of a slow, trickle-down series of announcements. For instance, after reforms were introduced allowing higher education institutions to offer law degree programs to women (2004), but before the Ministry of Justice started to grant licenses to female lawyers (2012), women who graduated from law school were employed as legal consultants in government and court offices, could not practice law in the courtroom, and could not own and operate law firms. Now, female lawyers can appear in court on behalf of men or women in domestic cases, such as divorce and child custody disputes, go before male judges to argue their cases, and work with other male lawyers (Malaver, 2017). Similarly, studying for a business degree is now popular among females in the wake of the institutional recognition, through relaxation of rules governing commerce, that owning a business can be a way for women to contribute to the country's economy (Basaffar, et al., 2018). Yet, evidence of underrepresentation of women still emerges if the Gender Parity Index (GPI) formula, which UNESCO (n.d.) uses to measure the relative access to higher education of men and women, is applied to the numbers of male and female freshman students during the last decade (2013 - 2018). Although substantial institutional changes and resources towards gender parity defined the selected decade, the gender parity indices for freshmen (i.e., the number of males over the number of females; Ministry of Education, n.d.) for STEM fields (.92), and business and law combined (.82) are still below the value that UNESCO uses to mark equity (.97).

Gender ‘is a socio-cultural construction that assigns 'appropriate' demeanor, attributes, opportunities, activities to men and women’ (Wariboko, 2018; 18-19). As with other socio-cultural constructions, gender roles are transmitted to the members of a society through the process of socialization. They may not only differ from one society to another but also be changed by circumstances within a cultural setting. Thus, the snapshot that the present study offers is merely a window into a world that is steadily changing. Yet, change is built upon a past of rigid gender segregation encompassing the public and private spheres of everyday life (AlMunajjed, 1997; Mayer, 2000). Gender segregation has existed within a quotidian of inequities in women's rights, resources, and opportunities, which recent institutional actions (i.e., decrees and investments in education) have attempted to ameliorate. These institutional actions cannot easily delete the impact of decades of marginalization on women's self-concepts, and more broadly, on stereotypical perceptions of women as less competent independent agents. Yet, it is

important to note that the motives behind individual educational and occupational choices are substantially shaped by the social, cultural, and economic structure in which they exist, thereby defining perceptions of available opportunities and resources (Williams and Wolniak, 2021). In a context in which the ability to act independently of the constraints of such a structure (i.e., agentic power) is severely limited, educational and career choices are largely driven by perceived opportunities, thereby making institutional actions key to change. In KSA, institutional actions can be said to have fostered a healthy debate regarding rights, roles, and possibilities, permitted women to explore a host of previously forbidden opportunities (e.g., a career in engineering), and solidified their determination for gender parity. Evidence suggests that higher education, such as obtaining a bachelor's degree, has a positive impact on economic and career outcomes (Long, 2010; Perna, 2005; Roksa and Levery, 2010), and more broadly, creates an educated workforce that suits the needs of a knowledge-based economy (Haigh and Clifford, 2011). Over the recent past, in many parts of the world, including the USA and KSA, women have thrived in higher education, but these gains have yet to translate into more robust and sustainable equity in educational and career choices as well as in females' labor market participation (Naseem and Dhruva, 2017). For these women, 'the road has begun to rise to meet them', but their journey is merely at its starting point.

According to the *gender stratification hypothesis* (Baker and Jones, 1993), gender differences in academic success reflect inequities in educational and economic opportunities available to the youth of a given society. Whether KSA's recent investments in gender equity have not only more evenly distributed opportunities but also stimulated changes in behavior may be measured by the extent to which gender differences in higher education exist in the choice of major and academic performance of young women and men. Although top-down attempts at systemic changes to increase the intellectual, economic, and political participation of women are undeniably important, the way opportunities are received by the recipients determines critically the success or failure of top-down interventions. According to Princess Ameerah al-Taweel, women's past deprivation of opportunities may reinforce their determination to prove themselves, thereby making them capable of reversing gender inequity trends in their favour (cited in Soekarba, 2019). Accordingly, traditional gender stereotypes may be thought of powerful motivators not for preserving the status quo (Bussey and Bandura, 1999), but for speeding up systemic change coming from the top. Consistent with the *gender stratification hypothesis*, which predicts a decline in gender differences as a function of a country's investment in equity, and the hypothesized motivational power of past inequities, which leads to predict a reversal, middle school female students in KSA have been reported to outperform their male counterparts (Abdourahmane, 2019). Yet, the same pattern has been found in Iran, Jordan, Oman, and Syria (Al-Sindi, 2013) where, arguably, less investment has been made, thereby suggesting that the drive for change at the bottom is critical.

Our work starts with the consideration of two composite predictors usually selected for college admission decisions of young women and men: the *GAT*, which is a standardized test, and high-school Grade Point Average (*hs-GPA*), which is a measure of earlier scholastic

performance. Academic success is measured as *GPA* at graduation or as performance in particular general education courses that cover verbal, quantitative, and analytical competencies. Thus, we use measures of academic success whose variability is restricted as they relate to a group of females and males who have already achieved some success as demonstrated by their attainment of an undergraduate degree. Within this selected group of female and male students, we intend to examine the choice of major and the performance of students who have handled a curriculum of Western import (i.e., USA) to determine (a) whether gender differences exist that replicate, minimize or even reverse past inequities, and (b) the extent to which *GAT* and *hs-GPA* predict academic success with such a curriculum. If these composite measures are found to be weak predictors of academic success, the ancillary goal is to conceptualize a data-driven strategy for how to undertake the task of identifying more suitable factors. The latter may be a work plan more than a realized architecture, but one worthy of discussion. This study contributes to the extant literature by examining a neglected population (i.e., Middle Eastern students dealing with a foreign curriculum) in a society in transition, thereby highlighting the extent to which systemic change from the top may trickle down to its recipients.

Method

Sample

The Office of the Registrar of a university in KSA that follows a USA curriculum and pedagogy granted access to a list of graduates. The sample involved students who received provisional or regular admission at the university from three Colleges: Business, Engineering, and Law. It included graduates of a five-year period, starting five years after the university was established and became fully functional (2006 - 2008). Random numerical codes substituted any identifying information to ensure the anonymity of the data. Students of a nationality other than a country in the Middle East were discarded, thereby leading to a sample of 1594 graduates. Students (age range: 18 - 25) reported Arabic as their first language and English as their second language. For admission, English competency had to be demonstrated through standardized English proficiency tests (i.e., TOEFL, IELTS, or Aptis).

Instruments and Variables

Each student's code was accompanied by demographic indices, including gender and college major. To ensure that samples would be large enough to allow for meaningful statistical analysis, majors were organized by the college to which they belonged. Performance measures included *hs-GPA* (a composite measure of basic knowledge and abilities), *GAT* (a standardized composite measure of achievement), and the average scores in several general education courses whose primary learning objectives were verbal, analytical, and quantitative competencies (to be described in the next paragraph). The cumulative *GPA* obtained at graduation was also available.

The university was selected for our retrospective case study because of its curriculum, which conforms to the higher-education model of the USA. As such, it offered a unique opportunity to test the extent to which measures, such as *hs-GPA* and *GAT*, can accurately predict the success

of Middle Eastern students dealing with a foreign curriculum. At the selected university, the curriculum is organized into two components: courses in general education (i.e., Core) and courses in the students' selected major. English is the primary vehicle of instruction in Core courses with the exception of a set of 4 courses in Arabic and Islamic Culture which are taught mostly in Arabic. For all students, Core courses, which have been approved by the Texas International Educational Consortium (*TIEC*), pertain to three sets and are completed by all students: (a) verbal competency (including general written and oral communication, research writing, and professional communication), (b) analytical competency (including applications to technology, critical thinking, problem-solving, leadership, and teamwork), and (c) quantitative competency (including calculus, statistics, and algebra). Additional Core courses pertain to physical education, practice in metacognition, as well as Islamic and Arabic Culture, or are electives that focus on particular subjects (e.g., psychology, history, etc.). No information regarding these courses was included in the data set obtained from the Office of Registrar. It is important to note that the learning objectives, syllabus, and grading rubrics of each Core course are uniform across all sections, even though different faculty may impart instruction, thereby introducing uniformity in assessment and evaluation. Faculty are required to follow a student-centered model, according to which the educator serves as facilitator and guidance of students' largely independent pursuits. In this model, knowledge and skills are to be sought and evaluated by students. Important to note here is that in KSA, students at primary and secondary educational institutions are the recipients of a teacher-centered education, according to which the educator is the expert whose knowledge and skills are not to be questioned, and students are the repositories of such knowledge (Prensky, 2001; Rugh, 2002). Experience with this top-down pedagogical approach, also called 'the sage on the stage' model (King, 1993), creates an additional layer of challenges for first-year students who are expected to adjust to the more equalitarian learning approach promoted by the university.

Results

All results reported below are considered significant if $p < .05$. Results are organized according to topical issues.

Gender Differences in Selection of Major (as Defined by College)

We conducted a chi-squared (χ^2) test in each college to determine whether there were gender differences in the selection of a major (as defined by the college to which the major belongs). All colleges exhibited gender differences (see Table 1). Women were more numerous in Business, χ^2 (1, n = 837) = 81.39, $p < .001$, and Law, χ^2 (1, n = 157) = 4.64, $p = .001$. Men were more numerous in Engineering, χ^2 (1, n = 600) = 216.000, $p < .001$. Although the preponderance of males in Engineering conforms to evidence regarding females being underrepresented in STEM fields (Alwedinani, 2016; Finger et al., 2020; Porter & Serra, 2020; Witherspoon & Schunn, 2020), the preponderance of females in Business majors suggests that traditional stereotypes of gender roles are selectively crumbling under the weight of the 2030 Vision. Focus groups clarified these

patterns by indicating that women who may still be intimidated by STEM fields are eagerly venturing into entrepreneurial professions seeking sustainable financial independence. Women's reasons for selecting to major in law were mostly driven by the desire to take control over their affairs as well as assist relatives and others in need, thereby combining traditional nurturing roles with the need to be a competent and useful member of a collective still propelled by patriarchal forces.

Table 1: Descriptive Statistics

Colleges	*N*	*Female (%)*	*Male (%)*
Business $*^F$	837	65.59	34.41
Engineering $*^M$	600	20.00	80.00
Law $*^F$	157	58.60	41.40
	1594		

Note: $*^F$ Greater number of females. $*^M$ Greater number of males.

Gender Differences in Performance

Our analyses focused on measures that are often used to predict academic performance, such as *hs-GPA* (a composite measure of basic knowledge and abilities), and a standardized measure of achievement, *GAT*. Academic performance was conceptualized as a composite measure, such as *GPA* at the end of students' educational path, or as competency-specific measures, such as the average mark in courses devoted to verbal, analytical, and quantitative competencies.

If we were to consider gender differences without attention to the college major, we would be tempted to conclude that females had always higher scores than males. The exception would be the standardized test, *GAT*, whose scores did not appear to differ between male and female students. However, this conclusion would be misleading as college majors offered different patterns of data. Table 2 displays descriptive statistics and the results of one-way ANOVAs conducted within each college major with gender as the factor. In the College of Business, females had greater scores on a predictive measure (*hs-GPA*) and on all college-related measures, $Fs(1, 835) \geqslant 95.27$, $MSE \leqslant 118.67$, $p < .001$, $\eta p^2 \geqslant .102$. For these students, *GAT,* the other predictive measure, did not differ between males and females, $F = 1.48$, *ns.* In the College of Engineering, females had superior *hs-GPA* and scores on all other college-related measures, $Fs(1, 598) \geqslant 24.84$, $MSE \leqslant 111.61$, $p < .001$, $\eta p^2 \geqslant .040$, except for math and *GAT* scores, $Fs \leqslant 2.77$, *ns*. Students in the College of Law followed the pattern of those in Business. Namely, females had higher *hs-GPA* as well as scores on all college-related measures than males, $Fs(1, 155) \geqslant 15.70$, $MSE \leqslant 119.50$, $p < .001$, $\eta p^2 \geqslant .092$. For these students, there were no gender differences in the *GAT* scores, $F < 1$, *ns.* In sum, the *GAT* was the only measure for which male and female students had equivalent performance across all majors. Quantitative competencies among students in Engineering were also gender neutral.

Table 2: Descriptive Statistics: Mean and Standard Error of the Mean (in Parentheses) as a Function of Gender and College Major

College	*hs-GPA*	*Female %*	*Male %*	
Business		92.42 (0.29)	87.21 (0.47)	$*^F$
Engineering		94.61 (0.45)	90.96 (0.30)	$*^F$
Law		92.54 (0.71)	87.96 (0.81)	$*^F$
		93.19	88.71	
	GAT			
Business		73.04 (0.30)	73.72 (0.52)	*ns*
Engineering		75.89 (0.64)	75.97 (0.40)	*ns*
Law		73.92 (0.70)	74.06 (1.03)	*ns*
		74.28	74.58	
	Verbal Competency			
Business		85.45 (0.25)	79.65 (0.33)	$*^F$
Engineering		87.68 (0.53)	82.12 (0.23)	$*^F$
Law		84.85 (0.70)	79.11 (0.73)	$*^F$
		85.99	80.29	
	Analytical Competency			
Business		86.58 (0.26)	78.99 (0.34)	$*^F$
Engineering		88.33 (0.47)	81.86 (0.25)	$*^F$
Law		86.52 (0.69)	78.83 (0.73)	$*^F$
		87.14	79.89	
	Quantitative Competency			
Business		78.81 (0.35)	73.30 (0.40)	$*^F$
Engineering		79.55 (0.76)	78.21 (0.35)	*ns*
Law		79.03 (0.86)	73.59 (1.09)	$*^F$
		79.13	75.03	
	GPA			
Business		79.87 (0.48)	72.11 (0.60)	$*^F$
Engineering		77.19 (1.11)	71.81 (0.46)	$*^F$
Law		80.90 (1.22)	72.99 (1.21)	$*^F$
		79.32	72.30	

Note: $*^F$ Higher scores for female students.

Prediction of Academic Performance

We also asked the extent to which predictive measures, such as *hs-GPA* and *GAT,* and gender (coded as 0 = male and 1 = female) would make a unique contribution to holistic measures of academic performance (*GPA*), or competency-specific performance, such as grades in courses covering verbal, analytical, and quantitative competencies. Thus, we conducted regression analyses with *hs-GPA*, *GAT*, and gender as the predictors, and *GPA* or competency-specific performance as the outcome variable within each college major (see Table 3). The contribution of each predictor conveys valuable information to admission officers, university administrators,

faculty, and support service staff. To wit, it allows them to determine whether the behaviors and cognitions that define the predictor are to be thought of as an area of intervention for poor-performing students.

It is not surprising that cumulative *GPA* would be predicted by *hs-GPA* and *GAT* score across colleges as both are composite measures of performance similar to the cumulative *GPA*. The contribution of gender, though, suggests to admission officers, university administrators, faculty, and support service staff that the factors that might differentiate males and females in college (e.g., study habits, motivation, etc.) are to be thought of as areas of intervention to bridge the gap between males and females.

If quantitative performance in college courses is considered, all predictors seem to contribute with some glaring exceptions. Difficulties in math might not be predicted by the gender of students majoring in Engineering, and by the *hs-GPA* of students majoring in Law. Instead, *GAT* might not be used to forecast difficulties in verbal or analytical competencies of students majoring in either Engineering or Law, whereas it might be a useful predictor for students majoring in Business.

The last column of Table 3 provides more useful information if the goal is to determine the unique weight of each predictor for the purpose of determining the extent to which particular students' characteristics are relevant to academic success in a given area of study. It contains the percentage of unique variance that is accounted for by each predictor when the contribution of the other predictors is controlled. For instance, the percentage of 7.62 for *hs-GPA* is the coefficient of determination of the semi-partial correlation between *hs-GPA* and *GPA*. Namely, knowing a students' *hs-GPA* does not allow one to predict much of that student's cumulative *GPA* (less than 8%). Overall, the selected factors explain little of both the composite and competency-specific outcome variables. Yet, in most cases, gender, a demographic variable, remains a much more powerful predictor than cognitive measures, such as *hs-GPA* and *GAT*.

Table 3: Regression Analyses

College Major	*GPA*	*B*	*SE*	*Beta*	*t*	*Sig.*	*Part %*
Business	Constant	2.828	4.352				
	hs-GPA	.479	.050	.318	9.623	<.001	7.62
	GAT	.373	.047	.250	7.983	<.001	5.24
	Gender	5.517	.745	.228	7.409	<.001	4.49
Engineering	Constant	17.437	6.280				
	hs-GPA	.453	.069	.269	6.598	<.001	6.20
	GAT	.173	.051	.134	3.379	.001	1.64
	Gender	3.732	1.048	.139	3.560	<.001	1.82
Law	Constant	10.515	11.633				
	hs-GPA	.423	.129	.258	3.289	.001	5.24
	GAT	.341	.117	.217	2.922	.004	4.12
	Gender	6.031	1.736	.257	3.474	<.001	5.86
	Verbal Compt.						
Business	Constant	53.550	2.465				
	hs-GPA	.201	.028	.242	7.138	<.001	4.37

	GAT	.116	.026	.140	4.377	<.001	1.64
	Gender	4.832	.422	.361	11.459	<.001	11.29
Engineering	Constant	65.460	3.241				
	hs-GPA	.144	.035	.162	4.069	<.001	2.25
	GAT	.047	.026	.068	1.764	*ns*	
	Gender	5.035	.541	.354	9.307	<.001	11.77
Law	Constant	52.644	7.098				
	hs-GPA	.209	.079	.212	2.661	.009	3.53
	GAT	.109	.071	.116	1.535	*ns*	
	Gender	4.793	1.059	.340	4.525	<.001	10.24
	Analytical Compt.						
Business	Constant	53.229	2.596				
	hs-GPA	.175	.030	.190	5.877	<.001	2.72
	GAT	.143	.028	.157	5.129	<.001	2.07
	Gender	6.785	.444	.460	15.277	<.001	18.32
Engineering	Constant	71.006	3.433				
	hs-GPA	.077	.038	.081	2.038	.042	0.56
	GAT	.051	.028	.071	1.832	*ns*	
	Gender	6.193	.573	.409	10.808	<.001	15.76
Law	Constant	52.997	7.074				
	hs-GPA	.186	.078	.178	2.375	.019	2.50
	GAT	.128	.071	.128	1.805	*ns*	
	Gender	6.857	1.056	.459	6.496	<.001	18.58
	Quantitative Compt.						
Business	Constant	32.100	3.237				
	hs-GPA	.265	.037	.248	7.157	<.001	4.58
	GAT	.245	.035	.231	7.061	<.001	4.45
	Gender	4.295	.554	.249	7.754	<.001	5.38
Engineering	Constant	36.244	4.659				
	hs-GPA	.357	.051	.289	6.992	<.001	7.18
	GAT	.126	.038	.133	3.310	.001	1.61
	Gender	.046	.778	.002	.059	*ns*	
Law	Constant	41.628	9.398				
	hs-GPA	.115	.104	.091	1.104	*ns*	
	GAT	.295	.094	.246	3.135	.002	5.29
	Gender	4.959	1.402	.276	3.536	.001	6.76

Note: Business: $Rs^2 \geqslant .254$. Engineering: $Rs^2 \geqslant .125$. Law: $Rs^2 \geqslant .175$. The values in the column labeled "*Part %*" refer to the coefficients of determination of semi-partial correlations.

Discussion

The tale that the data of our retrospective case study appear to tell is that local measures of performance, such as *hs-GPA* and *GAT*, which are often expected to predict the academic success of students of Middle Eastern descent in college, might be of little use if the curriculum and the pedagogy are of foreign import. Our findings, however, are consistent with those of other studies whose data come from both local and foreign universities. For instance, Alshahrany (2017) reported that the *GAT* scores did not predict *GPA* at the time of graduation of Saudi students at USA universities. In a KSA university, Alghamdi and Al-Hattami (2014) found that *GAT* scores

were not a significant predictor of performance (as defined by third-year *GPA*) in the Colleges of Education and Applied Studies (defined as Humanities), and only modestly predicted performance in the College of Applied Medical Sciences. Instead, *hs-GPA* modestly predicted students' performance across all colleges. The small predictive validity of *GAT* and *hs-GPA* in our study is also consistent with that reported by Alkushi and Althewini (2020) for college grades of students majoring in Health Sciences (Dentistry, Medicine, and Pharmacy) at another KSA university, with the exception that they found *GAT* to play a much more relevant role than *hs-GPA*, whereas our data exhibited the opposite pattern. Thus, if current and past findings are taken together, there is sufficient evidence to indicate that *GAT* and *hs-GPA* predict little of the variance in composite or competency-specific performance (*GPA* or grades in specific courses) irrespective of the extent to which an institution relies on a foreign curriculum and pedagogy.

In our study, we also found that female students tend to outperform male students across the board. The exception was the *GAT* across all colleges and quantitative competency in the College of Engineering for which no gender differences were found. These findings not only are consistent with the *gender stratification hypothesis* but also support Princess Ameerah al-Taweel's proposal that women's past deprivation of opportunities may fuel their determination to prove themselves, thereby making them capable of reversing trends of gender inequity in their favour (Soekarba, 2019). Notwithstanding the variability in the contribution of gender across college majors, gender was found to make a difference in most subject matters, suggesting that targeted interventions need to be developed at the start of students' educational journey to promote performance in male students.

The extant literature suggests that motivational variables, such as self-efficacy (Bussey and Bandura, 1999; Kitsantas, et al., 2008), self-determination, engagement, and self-regulation (Gregg, 2009; Kappe and Van Der Flier, 2012; Richman, et al., 2014; Zepke and Leach, 2010), which are related to effort and persistence, might be much more relevant for admission decisions (especially for males) than cognitive measures that assess knowledge and skill preparedness. The challenge to overcome is that measuring knowledge and skill preparedness and developing interventions that compensate for weaknesses is much easier than creating opportunities for assessment of motivational factors as well as developing interventions that can improve students' motivation in academia. Yet, the imbalance between males and females in motivation to succeed academically might be the main source of males' lower performance (Cortright, et al., 2013; Kahn, et al., 2011).

Of similar importance are the gender differences in college major selection, with women underrepresented in Engineering (a STEM field) and men underrepresented in Business and Law, which suggest that targeted interventions may need to start earlier, at the time students are attending elementary, junior high, and high schools. Interestingly, gender differences in performance favouring females are accompanied by these students' reluctance to enter engineering, traditionally considered a male field of study and work. Thus, albeit KSA's massive investment in the equity of educational opportunities for men and women has produced some notable and worthwhile fruits, it has yet to deflate traditional stereotypes of professions as

suitable for one gender only as well as to overcome the scarcity of role models in such professions. Thus, in countries where resources are invested to ensure systemic changes in gender equity, the issue to be addressed is not access to opportunities, but the appeal that available opportunities have for young males and females. The evidence that emerges from KSA is one that can guide other Middle Eastern nations towards a more gender-equitable economy. It suggests that in addition to the need for systemic changes and related investment in resources, interventions must be developed to speed up the absorption of such changes by the recipients through, for instance, the promotion of attractive role models.

The lower performance of male students might be the accidental outcome of the gender equity policies introduced in academia. These policies have elevated women to suitable agents of economic progress, thereby requiring males to enhance their efforts in domains that they previously dominated undisturbed (Cortright, et al., 2013; Kahn, et al., 2011). According to Ladson-Billings (1994: 17-18), effective education is one that empowers 'students intellectually, socially, emotionally, and politically by using cultural referents to impart knowledge, skills, and attitudes'. Thus, the key to enhancing males' performance might rest in a restructuring of the curriculum and the pedagogy of higher education courses to promote academic success, cultural competence, and socio-political awareness across the board. Culturally relevant pedagogy (Ladson-Billings, 2006), which aggregates these factors into a unified didactic approach, might be particularly relevant to institutions using a foreign curriculum taught in a foreign language (Owens and Lane, 2014). An equitable and relevant pedagogy is based on the idea that women and men are exposed to curricula germane to their lives, thereby allowing for differences in some contents while preserving parity in learning outcomes.

Conclusion

Our study has limitations that will be addressed in future research. First, it is important to note that our selection of participants was purposely restricted to successful students since our goal was to identify the particular variables that contributed to or detracted from their success. The inclusion of students who failed to graduate might expand the range of variability of the outcome variables, thereby merely increasing the sensitivity of the selected predictive measures, or altering the identified patterns. Yet, the interpretation of either outcome would be questionable since the reasons for students who did not graduate were not available. To wit, students' absence from graduation rosters does not necessarily mean they were not able to graduate. They might have transferred to another institution or withdrawn for reasons other than any presumed inability to perform. Moreover, our operational definition of academic success, which entailed, besides college grades of students who graduated from college, their cumulative *GPA*, might further reduce the range of variability of the outcome variables, and thus their sensitivity to the predictive measures. The *GPA* of first-year students, including those being granted provisional and full admission, might offer evidence of a greater impact.

Second, measures based on the unique characteristics of the population and curricula that an institution serves might not serve well other institutions. The relevance of the results of the

present investigation, though, is not to be found in the extent to which they generalize to other institutions. On the contrary, their relevance is to the particular institution that supplies the data. Broadly speaking, their value resides in demonstrating the feasibility of a data-driven approach to issues that concern the effectiveness and sustainability of higher education practices. Consider, for instance, the aim of academic excellence, which is pursued by educational institutions around the world, and the need to ensure gender equity in the paths to excellence (i.e., equal opportunities and resources available to male and female students; Van den Branden, 2012). A data-driven approach, as exemplified here, might inform the decision-making actions of the administrators of a particular institution about the measures to take to realize both aims. At the very least, students grouped by key predictive measures allow administrators, educators, and counselors to identify which students might benefit from particular targeted educational interventions (Seifert, et al., 2017).

Third, if traditional cognitive measures do not predict much of students' academic success, what are the alternatives? A data-driven approach including motivational alongside cognitive measures is proposed whereby individual higher education institutions invest in the assessment of a variety of cognitive and motivational variables through testing and interviewing of applicants. The results of this broad-range assessment practice might offer valuable information to admission officers, counselors, faculty, and administrators for not only making sensible decisions regarding admission but also developing interventions targeting provisionally admitted students who display weaknesses in motivation as well as preparedness. The recommended approach requires that an institution determines the percentage of variance in academic success (as operationalized by the institution) accounted for by each of an array of empirically determined predictors so that the ones optimally defining the population it serves are identified. This valuable information is provided by the coefficient of determination of the semi-partial correlation of each predictor. Lastly, it might be helpful to assess general knowledge preparedness through standardized tests, such as the *SAT*, offered in the language upon which the university relies as the primary mode of communication (e.g., English), which might provide a better fit for the challenges that students will encounter in college. Yet, the demanding cognitive computations of a test taken in a second language, do not merely reflect the test taker's knowledge and skills. They are sensitive to processing load, past educational practices, as well as confidence in one's abilities (i.e., self-efficacy) which translates into effort and persistence (Pilotti, et al., 2019). A recognition of the impact of these factors on test performance might help understand the value of standardized tests as predictive measures.

The data-driven procedure advocated by the present research is a work in progress at the institution that produced the data for the current study as the socio-cultural environment, resources, opportunities, and students' responses to such resources and opportunities might change over time. It is a commitment though, as gender equity is a heartfelt goal for all its constituents.

Author Biographies

Omar J. El-Moussa is an educator whose research interests include the assessment of students' academic success in a multi-cultural context and the effectiveness of interventions for at-risk students. His education includes King Faisal University (KSA) and Nova Southeastern University (USA). He is currently the Vice-President of Student Affairs of Prince Mohammad Bin Fahd University.

Runna Al Ghazo is an educational researcher and rehabilitation counselor. She received her Ph.D. in rehabilitation from Southern Illinois University-Carbondale (USA). Her areas of interest include psychological and systemic variables that may contribute to students' academic success in Higher Education, educational technology, curriculum and instruction, leadership, and technical writing pedagogy.

Maura A. E. Pilotti is a cognitive psychologist whose research interests include learning and memory processes across the lifespan. Currently, her research focuses on the interrelations of memory, language, and emotion. She received her Ph.D. in cognitive psychology at the City University of New York (USA).

References

Abu-Ghaida, D. & Klasen, S. 2004. The costs of missing the millennium development goal on gender equity. *World Development,* 32(7): 1075-1107.

Akpotor, J. & Egbule, E. 2020. Gender difference in the Scholastic Achievement Test (SAT) among school adolescents. *World Journal of Education*, 10(1): 97-101.

Al Alhareth, Y., Al Dighrir, I. & Al Alhareth, Y. 2015. Review of women's higher education in Saudi Arabia. *American Journal of Educational Research*, 3(1): 10-15.

Al-Alwan, I.A. 2009. Association between scores in high school, aptitude and achievement exams and early performance in Health Science College. *Saudi Journal of Kidney Disease and Transplantation,* 20(3): 448-453.

Alghamdi, A.K.H. & Al-Hattami, A.A. 2014. The accuracy of predicting university students' academic success. *Journal of Saudi Educational and Psychological Association*, 1: 1-8.

Alkushi, A. & Althewini, A. 2020. The predictive validity of admission criteria for college assignment in Saudi universities: King Saud bin Abdulaziz University for health sciences experience. *International Education Studies*, 13(4): 141-149.

AlMunajjed, M. 1997. *Women in Saudi Arabia Today*. London, UK: Macmillan Press.

Alnahdi, G.H. 2015. Aptitude tests and successful college students: The predictive validity of the General Aptitude Test (GAT) in Saudi Arabia. *International Education Studies*, 8(4): 1-6.

Alshahrany, A.M. 2017. An Examination of the Relationship Among the General Aptitude Scores, English Proficiency Scores, and the Cumulative Grade Point Average (GPA) of Saudi Arabian

Undergraduate Students in American Universities during 2010-2015. Unpublished PhD diss., Howard University, United States of America.

Alshumrani, S. A. 2007. Predictive Validity of the General Aptitude Test and High School Percentage for Saudi Undergraduate Students. Unpublished PhD diss., University of Kansas, United States of America.

AlSindi, N. 2013. Single-sex Schooling and Mathematics Achievement in the Middle East: The Case of Iran, Syria, Jordan, and Oman. Unpublished PhD diss., Georgetown University, United States of America.

Al-Tamimi, H. & Al-Shayeb, A. 2002. Factors affecting student performance in the introductory finance course. *Journal of Economic & Administrative Sciences,* 18(2): 76-86.

Althewini, A. 2019. The predictive validity of standardized tests and English proficiency for Saudi Medical students' performance in Biology. *International Journal of Education and Literacy Studies*, 7(4): 158-162.

Alwedinani, J. 2016. *Gender and Subject Choice in Higher Education in Saudi Arabia*. Morrisville, NC: Lulu Press.

Basaffar, A.A., Niehm, L.S. & Bosselman, R. 2018. Saudi Arabian women in entrepreneurship: Challenges, opportunities and potential. *Journal of Developmental Entrepreneurship*, 23(02): 1850013.

Baker, D.P. & Jones, D.P. 1993. Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education*, 66(2): 91-103.

Barry, A. 2019. Gender differences in academic achievement in Saudi Arabia: A wake-up call to educational leaders. *International Journal of Education Policy and Leadership,* 15(15): 1-17.

Burton, N. W., & Ramist, L. (2001). *Predicting success in college: SAT® studies of classes graduating since 1980* (Report No. 2001-2). College Entrance Examination Board. Available at: https://files.eric.ed.gov/fulltext/ED562836.pdf (Accessed: 12 August 2020).

Bussey, K. & Bandura, A. 1999. Social cognitive theory of gender development and differentiation. *Psychological Review*, 106(4): 676-713.

Cortright, R.N., Lujan, H.L., Blumberg, A.J., Cox, J.H. & DiCarlo, S.E. 2013. Higher levels of intrinsic motivation are related to higher levels of class performance for male but not female students. *Advances in Physiology Education,* 37(3): 227-232.

Dabaliz, A.A., Kaadan, S., Dabbagh, M.M., Barakat, A., Shareef, M.A., Al-Tannir, M., Obeidat, A. & Mohamed, A. 2017. Predictive validity of pre-admission assessments on medical student performance. *International Journal of Medical Education*, 8: 408-413.

Dimitrov, D.M., Raykov, T. & Al-Qataee, A.A. 2015. Developing a measure of general academic ability: An application of maximal reliability and optimal linear combination to high school students' scores. *Educational and Psychological Measurement,* 75(3): 475-490.

Else-Quest, N.M., Hyde, J.S. & Linn, M.C. 2010. Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological Bulletin*, 136(1): 103-127.

Feingold, A. 1988. Cognitive gender differences are disappearing. *American Psychologist*, 43(2): 95-103.

Finger, C., Solga, H., Ehlert, M. & Rusconi, A. 2020. Gender differences in the choice of field of study and the relevance of income information. Insights from a field experiment. *Research in Social Stratification and Mobility*, 65: 100457.

Gregg, C.M. 2009. Self-Determination, Culture, and School Administration: A Phenomenological Study on Student Success. Unpublished PhD diss., University of Phoenix, US.

Haigh, M. & Clifford, V.A. 2011. Integral vision: a multi-perspective approach to the recognition of graduate attributes. *Higher Education Research & Development,* 30(5): 573-584.

Hedges, L.V. & Nowell, A. 1995. Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269: 41-45.

Kahn, J.S., Brett, B.L. & Holmes, J.R. 2011. Concerns with men's academic motivation in higher education: An exploratory investigation of the role of masculinity. *The Journal of Men's Studies,* 19(1): 65-82.

Kappe, R. & Van Der Flier, H. 2012. Predicting academic success in higher education: What's more important than being smart? *European Journal of Psychology of Education*, 27(4): 605-619.

Khoshaim, H.B. & Ali, T. 2015. Students' struggle with first-year university mathematics courses in Saudi Arabia. *College Student Journal*, 49(4): 588-598.

King, A. 1993. From sage on the stage to guide on the side. *College Teaching,* 41(1): 30-35.

Kitsantas, A., Winsler, A. & Huie, F. 2008. Self-regulation and ability predictors of academic success during college: A predictive validity study. *Journal of Advanced Academics*, 20(1): 42-68.

Kobrin, J.L., Patterson, B.F., Shaw, E.J., Mattern, K.D. & Barbuti, S.M. 2008. *The Validity of the SAT for Predicting First-Year College Grade Point Average* (College Board Research Report 2008–5). New York, NY: The College Board.

Ladson-Billings, G. 1994. *The Dreamkeepers: Successful Teachers of African American Children*. San Francisco,CA: Jossey-Bass.

Ladson-Billings, G. 2006. Yes, but how do we do it? Practicing culturally relevant pedagogy. In Landsman, J. & Lewis, C.W. (eds.) *White Teachers/Diverse Classrooms: A Guide to Building Inclusive Schools, Promoting High Expectations, and Eliminating Racism*. Sterling, VA: Stylus, 29-42.

Lawlor, S., Richman, S. & Richman, C.L. 1997. The validity of using the SAT as a criterion for black and white students' admission to college. *College Student Journal,* 31(4): 507-515.

Long, M. 2010. Changes in the returns to education and college quality. *Economics of Education Review,* 29(3): 338-347.

Malaver, M-J. 2017. Saudi Arabian women, the law, and moral time. *The Classic Journal*, *2.1*, 12 April. Available at: http://theclassicjournal.uga.edu/index.php/2017/04/12/saudi-arabian-women-the-law-and-moral-time/ (Accessed 12 August 2020).

Mattern, K. D., Shaw, E. J. & Kobrin, J. L. 2011. An alternative presentation of incremental validity: Discrepant SAT and HSGPA performance. *Educational and Psychological Measurement*, 71(4): 638-662.

Mayer, A. 2000. Benign apartheid: How gender apartheid has been rationalized. UCLA *Journal of International Law and Foreign Affairs*, 5: 237-338.

Ministry of Education. n.d. *Data and Statistics: Open Data.* Available at: https://data.gov.sa/Data/en/organization/ministry_of_education (Accessed 12 August 2020).

Murray, C. 2012. Abolish the SAT. In Soares, J. (ed.) *SAT Wars: The Case for Test-Optional College Admissions*. Place of Publication: Teachers College Press, 69-81.

Naseem, S. & Dhruva, K. 2017. Issues and challenges of Saudi female labor force and the role of Vision 2030. *International Journal of Economics and Financial Issues*, 7(4): 23-27.

Owens, T. L. & Lane, J. E. 2014. Cross-border higher education: Global and local tensions within competition and economic development. *New Directions for Higher Education*, 2014(168): 69-82.

Perna, L.W. 2005. The benefits of higher education: sex, racial/ethnic, and socioeconomic group differences. *Review of Higher Education,* 29(1): 23-52.

Pilotti, M.A., Aamir, S. Al Ghazo, R. & Al Kuhayli, H.A. 2019. Explicit conditional reasoning performance: An examination of educational past, processing load, and self-efficacy. *International Journal of Teaching and Learning in Higher Education*, 31(3): 402-412.

Porter, C. & Serra, D. 2020. Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3): 226-54.

Prensky, M. 2001. Why education and training have not changed? In *Digital Game-Based Learning*. New York, NY: McGraw-Hill, 1-21.

Rattani, S.A. 2016. SAT: Does racial bias exist? *Creative Education*, 7(15): 2151-2162.

Richman, E.L., Rademacher, K. N. & Maitland, T.L. 2014. Coaching and college success. *Journal of Postsecondary Education and Disability*, 27(1): 33-50.

Roksa, J. & Levey, T. 2010. What can you do with that degree? College major and occupational status of college graduates over time. *Social Forces,* 89(2): 389-415.

Rothstein, J.M. 2004. College performance predictions and the SAT. *Journal of Econometrics*, 121(1-2): 297-317.

Rugh, W.A. 2002. Arab education: Tradition, growth and reform. *The Middle East Journal*, 56(3): 396-414.

Sackett, P.R., Kuncel, N.R., Beatty, A.S., Rigdon, J.L., Shen, W. & Kiger, T.B. 2012. The role of socioeconomic status in SAT-grade relationships and in college admissions decisions. *Psychological Science*, 23(9): 1000-1007.

Seifert, T.A., Bowman, N.A., Wolniak, G.C., Rockenbach. A.N. & Mayhew, M.J. 2017. Ten challenges and recommendations for advancing research on the effects of college on students. *AERA Open,* 3(2): 1-12.

Shaw, E.J. 2015. *An SAT® Validity Primer*. College Board. Available at: https://files.eric.ed.gov/fulltext/ED558085.pdf (Accessed 12 August 2020).

Siddiek, A.G. 2011. Standardization of the Saudi secondary school certificate examinations and their anticipated impact on foreign language education. *International Journal of Humanities and Social Science,* 1(3): 57-64.

Sideridis, G.D., Tsaousis, I. & Al-harbi, K.A. 2015. Multi-population invariance with dichotomous measures: combining multi-group and mimic methodologies in evaluating the general aptitude test in the Arabic language. *Journal of Psychoeducational Assessment*, 33(6): 568-584.

Soares, J.A. 2015. *SAT Wars: The Case for Test-Optional College Admissions*. New York, NY: Teachers College Press.

Soekarba, S.R. 2019. Determinants of patriarchy in the Middle East: Hope for the 2030 Vision in a new Saudi Arabia. *Advances in Social Science, Education and Humanities Research,* 365: 193-200.

Sulphey, M.M., Al-Kahtani, N.S. & Syed, A.M. 2018. Relationship between admission grades and academic achievement. *Entrepreneurship and Sustainability Issues*, 5(3): 648-658.

Topal, A. 2019. Economic reforms and women's empowerment in Saudi Arabia. *Women's Studies International Forum*, 76: 102253.

UNESCO. n.d. *Gender Parity Index (GPI).* UNESCO's International Institute for Educational Planning. Available at: https://learningportal.iiep.unesco.org/en/glossary/gender-parity-index-gpi (Accessed 12 August 2020).

Urlings-Strop, L.C., Stegers-Jager, K.M., Stijnen, T. & Themmen A.P. 2013. Academic and non-academic selection criteria in predicting medical school performance. *Medical Teacher*, 35(6): 497-502.

Van den Branden, K. 2012. Sustainable education: basic principles and strategic recommendations. *School Effectiveness and School Improvement*, 23(3): 285-304.

Voyer, D., & Voyer, S.D. 2014. Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4): 1174.

Wariboko, O. 2018. Gender, terminologies and an appraisal of gender discourses in Africa in contemporary times. *Journal of African Studies and Sustainable Development,* 1(2): 17-31.

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M. & Schmidt, F. L. 2015. College performance and retention: A meta-analysis of the predictive validities of ACT® scores, high school grades, and SES. *Educational Assessment*, 20(1): 23-45.

Williams, T.M., & Wolniak, G.C. 2021. Unpacking the "female advantage" in the career and economic impacts of college. In Niemi, N.S. & Weaver-Hightower, M.B. (eds.) *The Wiley Handbook of Gender Equity in Higher Education*. Hoboken, NJ: John Wiley & Sons, Inc., 5-28.

Witherspoon, E. B. & Schunn, C. D. 2020. Locating and understanding the largest gender differences in pathways to science degrees. *Science Education*, 104(2): 144-163.

Zepke, N. & Leach, L. 2010. Beyond hard outcomes: 'Soft' outcomes and engagement as student success. *Teaching in Higher Education*, 15(6): 661-673.

Zwick, R. 2006. Higher education admission testing. In Brennan, R. (ed.) *Educational Measurement*. 4th edition. American Council on Education, Praeger, 647-679.