Syracuse University

SURFACE

Dissertations - ALL SURFACE

December 2020

Assessing Grit and Conscientiousness to Predict Medical Student Success on the National USMLE Step 2 Clinical Knowledge Exam

Jennifer C. Welch
Syracuse University

Follow this and additional works at: https://surface.syr.edu/etd

Part of the Education Commons

Recommended Citation

Welch, Jennifer C., "Assessing Grit and Conscientiousness to Predict Medical Student Success on the National USMLE Step 2 Clinical Knowledge Exam" (2020). *Dissertations - ALL*. 1286.
https://surface.syr.edu/etd/1286

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

Abstract

In an effort to consider additional or alternative variables that predict medical school success and to better identify qualified and more diverse medical school candidates to serve their communities, this study investigated the impact of two noncognitive traits, conscientiousness and grit, on the United States Medical Licensing Examination (USMLE) Step 2 Clinical Knowledge (CK) national examination. The assessment used to measure conscientiousness was Costa and McCrae's (1992) NEO-FFI-3 and the assessment used to measure grit was Duckworth and Quinn's (2009) Grit S scale. This study also examined the correlation of sociodemographic (gender, age, race/ethnicity, socioeconomic background, and first-generation college status) and academic variables (undergraduate major, undergraduate science GPA, MCAT scores, and master's degrees and post-baccalaureate coursework) on USMLE Step 2 CK performance. The conceptual framework that informed this study was Sternberg's (1999) theory of successful intelligence which considers life circumstances and individual strengths to achieve one's goals. Additionally, Yosso's (2005) cultural wealth theory emphasized the need to consider variables that recognized the talents, strengths, and experiences that underrepresented students bring with them to educational environments.

The sample eligible for this study was from two cohorts of medical students from Northeast Medical University, a public college of medicine, located in an urban city of approximately 150,000 people. The College of Medicine is the largest of four colleges that comprise a comprehensive academic medical center. At the time of this study there was approximately 666 students enrolled in Northeast Medical University's College of Medicine. One hundred and ninety-two students (N = 192) participated in this study. Students completed both instruments (the NEO-FFI-3 at the point of admission and the Grit-S during their clinical years of medical school). Sociodemographic and prior

academic information were retrieved from participant medical school application. The relationship among these variables on the dependent variable, USMLE Step 2 CK, were assessed using descriptive and multivariate statistical analysis. The grit and conscientiousness variables were then regressed on the USMLE Step 2 CK national examination scores.

The results of this study demonstrated that participants from higher socio-economic backgrounds (p = .023), those from educated families (p = .046), males (p = .040), White/Asian students (p < .001), and those under the age of 26 (p = .011) performed significantly higher than their peers on the USMLE Step 2 CK. Findings indicated that prior college undergraduate science GPAs (p < .001) and MCAT percentile scores (p < .001) were significantly predictive of Step 2 CK performance, with moderate effect sizes (r = .335, r = .378) respectively. The magnitude of the effect sizes within this study were likely impacted by the small sample size (Slavin & Smith, 2009; Turner, Paul, Miller, & Barbey, 2018) and also by the lack of persistence of those deemed ineligible for participation in this study. Conscientiousness was shown to be significantly predictive of USMLE Step 2 CK (p = 0.002) and also predictive of undergraduate science GPA. These results further revealed that grit had no significant impact on USMLE Step 2 CK performance. Underrepresented students (p < .001) and those 26 and older (p < .006) scored significantly higher on grit than their peers.

Chi square analyses were completed and showed that this sample was not representative of this medical school or the national profile of matriculated medical students. This was a significant limitation in this study and the results should not be generalized to this institution or other medical schools. Based on the need to offset the anticipated national physician shortage, meet the needs of a more ethnically and racially diverse patient population, lessen residency burnout, and retain more medical students during their long and arduous medical education, the

results of this study suggest the need for continued research on measures other than cognitive skills to predict success in medical school and beyond. Considering non-cognitive skills in the admissions process will promote a more diverse physician population and potentially allow for the recruitment of students able to persist through the narrow medical education pipeline, stay the course and eventually become desperately needed physicians within the United States.

Keywords: conscientiousness, grit, medical education, medical students, medical student success, non-cognitive variables, USMLE Step 2 CK, regression analyses

Assessing Grit and Conscientiousness to Predict Medical Student Success

on the National USMLE Step 2 Clinical Knowledge Exam

by

Jennifer C. Welch

B.A., State University of New York College at Potsdam, 1992
M.S., Syracuse University, 1998

Dissertation
Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Higher Postsecondary Education.

Syracuse University
December 2020

Copyright © Jennifer C. Welch, 2020
All Rights Reserved

Acknowledgements

I am so humbled by the support, guidance, and encouragement from so many amazing people throughout my scholarly journey. I am extremely grateful to the medical students that agreed to participate in this project. The remarkable medical students I have worked with over the years are the reason I do what I do – they inspire me each and every day and keep me motivated and excited about my profession. They have gone on to become compassionate, competent, and attentive physicians making a difference in this world. To Dr. Hassan Naqvi, Dr. Jasmin Sahbaz, Dr. Syed Hassan, Dr. Sylvia Cardounell, Mr. Kevin Damri, MSIV, and Ms. Marissa Rice, MSIII a very special and heartfelt thank you! To my colleagues, thank you for the encouragement and assistance, especially Ms. Isabelle Appler, Ms. Barbara-Ann Mitchell, Dr. Christopher Morley, Ms. Cristina Pope, and Ms. Tanya Milazzo. A special thank you to Dr. Julie Rawls White, thank you for always being my rock when I wanted to give up. Thank you for the encouragement, the help, and for allowing me time to focus on this project. Your guidance and "push" got me to the end and I am eternally grateful! I work at a pretty fantastic place and am so thankful for my colleagues and the friendships I have built.

To my committee members I am forever indebted to each of you. Your kindness, generosity, insight, and intelligence got me here – thank you! Dr. Qiu Wang, thank you for always being my cheerleader and helping to make statistics a bit more understanding for me! To Dr. Lauren Germain, an excellent teacher, advocate, and colleague. I can't thank you enough for always being a text or a walk away, your advice and support helped to keep me focused and persevere. To Dr. Cathy Engstrom, my chair, from day one you have been by my side, encouraging me to think and write in ways I never had before. You have been the consummate mentor to me. When life got in the way, you listened and encouraged me to keep going and I could not appreciate you more.

Finally, to my family, there are not enough words for me to adequately thank each and every one of you. To my sisters, Kim, Leslie, Jessica, Kelly, Shelley, and Toots, thank you for the support. To my parents, Mom, Dad, Gary, and Cindy, who would have thought? Thank you for the love, support, and babysitting over the years! I love you all dearly. To my children, Emily, Matthew, and Ryan, who were little when I started this, I couldn't be prouder that I get to be your mom. Thank you for supporting me and understanding when I was "dissertating". I love each of you to the moon and back. Ryan – let's go fishing! I could not have done any of this without my husband, Dennis. He was my rock, he supported me and always encouraged me to keep going. He cooked many dinners, watched movies alone, took the kids where they needed to go, and took them away for weekends at a time so that I could write! When I no longer believed in myself, he believed in me and for that I am deeply thankful. I love you more, D!

I would like to dedicate this dissertation to my sister, Kelly, who passed away in September 2019 and to my stepmother, Cindy, who passed away in June 2020. I wish that they could have been here to celebrate my accomplishment. I hope that they are watching over me and are proud of what I have done.

TABLE OF CONTENTS

ILLUSTRATIVE MATERIALS

CHAPTER ONE: INTRODUCTION

Although perseverance, grit, and communication skills have always been important in patient care, the current pandemic the world is facing, demonstrates the need for physicians to possess these qualities as they have been called upon to take on heroic roles. While their medical training is in fact essential, perhaps even more critical are their resiliency, grit, compassion, and level of conscientiousness to be there for their patients as they put their own lives on the line to care for others in a time of tremendous need and uncertainty. Given this landscape it is evident that physicians must possess medical knowledge, but now more than ever, non-cognitive traits such as conscientiousness and grit are vital. These traits allow them to put the interests of their patients first, yet these qualities are not captured by MCAT scores or undergraduate GPAs, the traditional gateways to medical education.

For more than 100 years, the cognitive factors of MCAT scores and GPA have been the primary consideration when selecting the next generation of physicians (Witzburg & Sondheimer, 2013). Although recently it has been the non-cognitive traits of professionalism, resilience, grit, perseverance, and conscientiousness that have been identified as being of critical importance (Blumenthal, 2020; Evans, 2020; Han & Salles, 2020; Zivot, 2020).

Recently, many retired physicians have volunteered to help in any way they could. As physicians, they have dedicated their lives to serving others and many felt compelled to help their colleagues. Similarly, many fourth-year medical students were eager to get started and were allowed to graduate early in order to help alleviate the shortage of health care workers across the country, especially in those areas most impacted by this health care crisis (Abrams & Ducharme, 2020; Redford, 2020).

The importance and relevance of this study could not be timelier. The purpose of this study was to explore how non-cognitive factors, specifically conscientiousness and grit, correlated with one key, essential component of medical school success, namely clinical knowledge as assessed by the United States Medical Licensing Examination (USMLE) Step 2 Clinical Knowledge exam at one public medical school in the Northeast (USA). These two qualities appear necessary to withstand the rigorous medical school curriculum, testing processes, and this current health care crisis; however, limited empirical evidence exists. Pursuing medical school is an often overwhelming and stressful period for students; this stress may be related to the workload and expectations related to their academics, time away from their friends and loved ones, or financial stressors related to the significant cost of a medical education (Lim et al., 2014; Saipanish, 2003; Shah et al., 2010; Sherina et al., 2004; Smith et al., 2007). These burdens may impact subsets of students in different ways depending upon their academic preparation, their familial support system, or their socioeconomic backgrounds (Hadinger, 2017; Henning et al., 1998). Traditional medical school admissions policies do not consider these non-cognitive factors in their admissions decisions, with the greatest weight placed on the traditional measures of undergraduate grades and MCAT scores.

This chapter outlines the current and anticipated state of our physician workforce, the traditional medical school curriculum, and the testing medical students are required to complete. A brief summary of traditional and other factors that are considered in medical school admission decisions and their correlation with medical school success are provided, particularly in light of the need for physicians to reflect the diversity of our communities in terms of race/ethnicity and socioeconomic background. Finally, the theoretical framework, problem statement, purpose and significance of the study, and key definitions are shared.

Current and Future Physician Workforce Needs

The journey to become a physician is long and arduous; it is not for the faint of heart. It takes considerable time, motivation, and a true dedication to the profession to make it through. Prospective students need to successfully complete the premedical requirements, prepare for and take the Medical College Admissions Test (MCAT), and apply to medical school. During the most recent application cycle, there were 53,371 applicants to medical school, with 21,869 matriculants, offering a 41% acceptance rate (AAMC, 2020a). Once accepted, students typically spend four years in medical school, take two national examinations, then apply to residency. Residency experiences range between three to seven years, depending on the field students chose to specialize in. At the end of their residency students take another standardized exam. Some students may secure a fellowship, and take more national board exams in their area of specialization. The minimum educational commitment to pursue a career as a physician is eleven years. Resiliency and grit are a must when it comes to pursuing this professional path (Bassett, Brosnan, Southgate, & Lempp, 2018; Wald, 2015; Wendland, 2010).

Currently, the US has too few physicians that are well trained, some burn out in residency, and others may not make it through their undergraduate medical education. Therefore, the earlier it can be determined what qualities are necessary to continue on this pathway and identify who needs support along the way, the easier it may be to address the real societal problem to graduate more physicians.

The changing demographics within the United States and the anticipated drop in the physician workforce (Dall, West, Chakrabarti, & Iacobucci, 2017; Kirch & Patelle, 2017), will significantly increase the demand for physicians, particularly for physicians of color (Association of American Medical Colleges, 2019). The United States Center for Labor Statistics (2019) anticipates a 7% growth in the demand for physicians by 2028. This rate is faster than the average growth rate for all other occupations. This need is mainly due to physician retirements and an increased demand in healthcare services because of the growing and aging population within the US. The fundamental problem which needs resolving is ultimately increasing the number of trained physicians (Kirch & Patelle, 2017; Lakhan & Laird, 2009), particularly those who will serve in under-served communities (Eden, Berwick, & Wilensky (Eds.), 2014; Petersen, Hutchings, Shrader, & Brake, 2011).

In addition to the demonstrated need for more physicians, the physician workforce requires more demographic diversity in order to improve patient outcomes in underserved communities (Norcini et al., 2008). As the demographics of the United States have changed over time and the medical field has evolved, allowing for a more holistic, individualized approach in the medical student selection process is needed (Conrad et al., 2016). Considering additional or alternative variables that are strong predictors of medical school success will allow admissions committees to better serve the diverse communities within the United States (Bore, Munro, & Powis, 2009; Cohen & Steinecke, 2006; Kirch & Petelle, 2017; Koenig et al., 2013; Reed et al, 2012; Witzburg & Sondheimer, 2013). Assessing supplemental variables at the point of admission presents the opportunity for a fairer assessment when considering underrepresented students, first-generation college students, and non-traditional students (Sedlacek, 2011). According to the 2010 US Census, the minority population in the US was 35.9% (United States Census Bureau, 2018), yet according to the American Medical Association (2019), fewer than 10% of the physicians in the United States were Black/African American, Hispanic/Latino, or Native American. Mirroring the demographic composition of our country would be ideal as patients want physicians who look like them and understand their specific health care needs (Marrast et al., 2014; Norcini et al., 2008; Reede, 2003;

Walker et al., 2012). As underrepresented students tend to perform lower on standardized tests than their peers (Orfield & Wald, 2000; Reiter et al., 2012; Steele, 1999) allowing for complementary predictors of academic success may greatly benefit the diversity of the workforce and the profession overall, while also addressing health care disparities within the United States (Marrast et al., 2014). Schmitt et al. (2009) advocated that considering noncognitive factors, specifically a noncognitive assessment in the selection criteria, would allow admissions committees to identify students who are able to manage the multiple demands that come with pursuing a medical degree, while also leading to a more diverse student population.

Similarly, medical educators must recognize that over time 45% of physicians exhibit signs of burnout and consider leaving the medical profession (Cortez et al., 2020; Leigh et al., 2002; Shanafelt et al., 2012; Zuger, 2004). Medical education and/or the curriculum itself may also be contributing to this outcome and medical student well-being, as students feel a lack of support, high stress, and a lack of control over their lives (IsHak, Nikravesh, Lederer, Perry, Ogunyemi, & Bernstein, 2013; Kushnir, Cohen, & Kitai, 2000; Santen, Holt, Kemp, & Hemphill, 2010; Williams, Tricomi, Gupta, & Janise, 2015). The high burnout rate for resident physicians is likely due to their overwhelming work responsibilities and the limited control they have over their graduate education (Thomas, 2004). Physician burnout could also be tied to area of specialty that students pursue (Martini, Arfken, Churchill, & Balon, 2004; Shanafelt et al., 2002). One study focused on the need for residents to further develop their social skills, indicating that these skills could be learned and could ultimately improve their patient care and combat their desire to leave the profession (Pereira-Lima, & Loureiro, 2015).

Although there is very little research available on medical student attrition, it is certainly a concern in medical education. Research studies at two foreign medical schools indicated a 5.7%

student attrition rate over ten-years and a 26% attrition rate over 30 years, respectively (Maher et al., 2013; Kruzicevic, Barisic, Banozic, Esteban, Sapunar, & Puljak, 2012). In both studies, prior academic qualifications and absenteeism predicted medical student continuation. Medical school attrition is a long-standing problem. Over 50 years ago, in 1966, Johnson and Hutchins identified medical student attrition in the US as one of the "most serious problems facing American medicine" (p. 1099). The authors further stated the need to "eradicate the causes of attrition by encouraging better professional preparation, more sophistication in the process of selection, and more effective teaching and counseling" (Johnson & Hutchins, 1966, p. 1099). Over the course of their ten-year study, 8.6% of matriculants did not graduate from medical school. According to the AAMC, the attrition rates for matriculated MD students from 2003-2013 was 3.1%, 1.8% for non-academic reasons and 1.3% for academic reasons (AAMC, 2018). Although, these numbers are better than they were half a century ago, given the need for physicians, they are concerning.

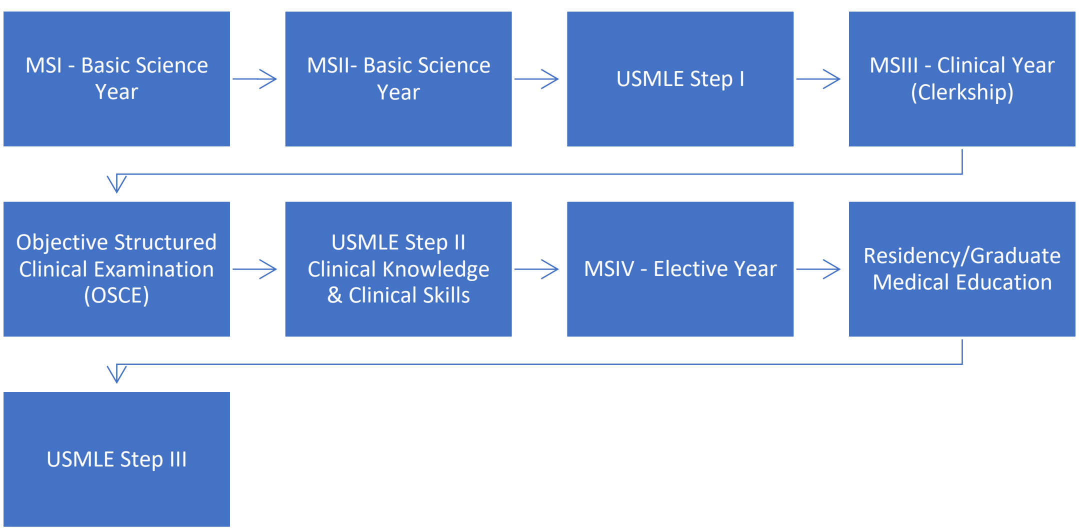
All of these factors together, the physician shortage, resident burnout, and medical student attrition could have devastating effects on the healthcare provided in the United States. Therefore, assessing particular personal qualities such as grit in the admissions process, may assist admissions committees in identifying candidates who will make it through their long and often-stressful undergraduate and graduate medical education, be successful on their standardized exams, and go on to thrive in their practice of medicine.

Curriculum and Testing in Medical School

The first two years of the traditional medical school curriculum, often referred to as the basic science years, are primarily pre-clinical with a transition into clinical training for the final two years. Students must be hardworking, focused, and dedicated to reaching the end goal of providing patient care. The intensity and volume of material that students are required to learn during their basic science coursework may leave some students struggling to remember why they entered medical school in the first place. Much of this time is spent in class, often large lecture-type classes, taking multiple choice exams and regurgitating memorized information (Dyrbye et al., 2006).

The competency exams administered by the USMLE mirror this curriculum transition, with the Step 1 exam typically following the pre-clinical educational phase and Step 2 occurring in the middle of the clinical years, typically after year three (often referred to as the clerkship year). During the clerkship year, students rotate through required clinical experiences in internal medicine, surgery, pediatrics, family medicine, psychiatry, radiology, neurology, and obstetrics/gynecology. Figure 1. provides a general overview of the traditional curriculum for undergraduate medical training. Once students enter their graduate medical training, most often referred to as residency, they will spend three to seven years completing further training in the particular specialty they have chosen. After the first year of residency, students complete their final USMLE assessment, Step 3.

Figure 1: Common Medical School Curriculum Flow.



Per the USMLE, the Step 1 exam focuses on the pre-clinical basic sciences and aims to test the student's knowledge of the mechanisms of health, disease, and treatment (USMLE, 2020a). After completing clerkships during the third year, students will typically take an in-house

Objective Structured Clinical Examination, referred to as an OSCE, which tests their knowledge and their clinical skills and is further described below. Following this, students take the Step 2 exam, which is composed of two parts: Clinical Knowledge (CK) and Clinical Skills (CS) components (USMLE, 2020b, 2020d). The CK component is a standardized multiple-choice exam that tests the student's ability to solve clinical cases by applying medical knowledge, assessing patient-centered skills, and their understanding of clinical medicine on patient care (Brandt et al., 2013). Step 2 CK focuses on health promotion and disease prevention and ensures that students are committed to clinical sciences and basic patient-centered skills, including competency in ethics and professionalism that is critical to patient care. The Step 2 CK exam is distinct from the Step 1 exam in that it requires a higher level of clinical problem-solving skills. It will be discussed in greater detail in Chapter Two. The CS component is graded on a pass/fail basis and uses standardized patients to ensure that students can interact with patients in an effective manner (Cuddy et al., 2007). The USMLE Step 3 exam is taken after the first year of residency and focuses on the diagnosis and treatment of patients and assesses whether or not the resident is able to practice medicine unsupervised (USMLE, 2020a).

The separate sections of the Step 2 exam were intended to assess student knowledge (CK) as well as the application of what they have learned (CS). This was done because Western students have not always felt that their training prepared them for the shift from knowledge to application. Eikeland et al. (2014) interviewed students in their third year of medical school and found that "throughout medical school, academic skills are prioritized over humanistic knowledge, and that this is an important part of their understanding of the physician's role" (p. 5). Similarly, students noted that time constraints, a strong emphasis on evidence-based medicine, and biomedical knowledge have led them to give "soft" skills a lower priority. Students in the above study pointed out that their clinical years led to increased cynicism due to patient overload and the bureaucracy of medicine, while their patients became just another number or interesting case. The culture of medical school is becoming hyper-focused on teaching medical content compared to teaching students the practice of medicine. In a recent article, *Trading Places: When Doctors Become Patients*, Dr. Rana Awdish stated her "education had taught her how to treat disease. But it didn't prepare her to treat the person" (Kalter, 2019, pg. 2). By contrast, cultures that emphasize the human side of medicine in their schooling do not see a drop in these soft skills when students enter their clinical years. Kataoka et al. (2009) showed on average, Japanese medical students improved these skills during their clinical years.

One strategy to address this concern of inadequate soft skills has been to assess student development prior to sitting for the USMLE Step 2 through school-specific Objective Structured Clinical Examinations (OSCE) utilizing standardized patients. The OSCE is intended to simulate the Step 2 examination environment and assess the student's clinical knowledge, problem-solving ability, and communication skills prior to taking the national examination (NBME, 2010). Yet because the OSCE is school-specific and graded by faculty and standardized patients employed at each particular medical school (Kaufman, Mann, Muijtjens, & van der Vleuten, 2000) questions arise regarding the reliability of this non-standard assessment as well as the effort of the students taking this assessment.

Further, as noted above, Step 3 is the final assessment in the USMLE testing sequence. It assesses physician knowledge and skills and leads to the license in which physicians are able to practice medicine independently, without supervision. This exam not only assesses medical knowledge but also the management of patient care, and communication and professionalism skills of the physician (USMLE, 2020a). These skills allow trust-based relationships between physicians and patients to develop which often leads to optimal clinical outcomes and greater compliance with treatment plans (Del Canale et al., 2012).

As the field of medicine has changed from being autonomous and individualistic to a patient-centered, team-based approach where physicians collaborate with one another and other health professionals to best meet the needs of their patients, the personal qualities that physicians possess become of paramount importance (Davis et al., 2005; Monroe et al., 2013). Allowing for complementary predictors of academic success may greatly benefit the physician workforce need as well as the profession overall. The need for valid measures at the point of admission to assess non-cognitive traits that can predict student success continues to complicate the medical school admission process.

Admission to Medical School

The process of medical school admission poses high stakes for both applicants and medical institutions. Applicants may have spent years, if not decades, and thousands of dollars in preparation for a chance to be considered by an admission committee. On the other hand, admission committees are the gatekeepers of medicine. Their purpose is to identify applicants from diverse backgrounds who will grow into mature, empathetic, and competent physicians worthy of representing the institution from which they graduate as well as the medical profession as a whole. In this endeavor, admission committees consider several factors. Grade point average (GPA) and Medical College Admission Test (MCAT) scores are often considered measurable and proven indicators of one's academic ability. Letters of recommendation, volunteer experience, healthcare exposure, and research projects are examined to ensure applicants have made an informed decision regarding their future career. Finally, the personal statement and interview are used to measure desirable personal characteristics, although measuring these characteristics can be skewed, subjective, biased, and time-consuming (Albanese et al., 2003).

Medical school admission and academic progression throughout medical school have long relied heavily on grades and standardized test scores to admit and make judgments about medical student's competence and success (Kulatunga et al., 2002; Salvatori, 2001). Traditional cognitive measures used to evaluate prospective medical students, while important, provide an incomplete assessment of a successful medical student and/or physician; assessing non-cognitive qualities allows for a more equitable evaluation of students from a wider range of backgrounds (Sedlacek, 2004). In a study conducted more than forty years ago, Roman Jr et al. (1979) suggested that utilizing objective cognitive criteria "is less predictive of minority student performance than of medical school applicants in general. Furthermore, the rigid application of cognitive performance criteria may be frankly exclusionary of qualified minority group applicants" (p. 664). Over many years, researchers have argued that underrepresented students are less likely to perform as well academically or on standardized examinations relative to their non-minority peers due to structural racism, inequality, stereotype threat, and discrimination (Johnson, Smith, Tarnoff, 1975; Sacks, 2007; Sedlacek, 1977; Sedlacek, 2004). Further, Roman Jr et al (1979) suggested that non-cognitive factors are just as important as cognitive factors, and perhaps even more so, for this group of students and if not considered in the admissions process will severely limit the racial and ethnic diversity of the physician workforce.

According to two studies, undergraduate grade point averages (UGPA) and Medical College Admission Test (MCAT) scores have been found to predict success in the first two years of medical school (Huff et al., 1999; Salvatori, 2001) and generally correlate with how students perform on the USMLE Step 1, the national exam students must pass following their second year of medical school.

However, some studies have found that these traditional measures do not correlate with the clinical grades students earn during their third and fourth years or on the USMLE Step 2 exam (Roth, Riley, Brandt, & Seibel, 1996; Salvatori, 2001; Witzburg & Sondheimer, 2013). On the contrary, two studies examined how MCAT scores correlated with medical school performance and the later professional performance of students at the Uniformed Services University's F. Edward Hebert School of Medicine and found that MCAT scores had a weak correlation with Step 1 and GPA performance, and the study agreed that no correlation was found between MCAT scores, USMLE Step 2 scores, and PGY-1 (first year of residency) performance (Donnon et al., 2007; Saguil et al., 2015). Additionally, by observing patient care encounters and reviewing residency performance evaluations, Saguil et al., (2015) and Sade et al. (1985) found that MCAT scores and undergraduate college GPAs did not gauge how good of a physician, based on patient perceptions, students would become later on.

In response to the concerns about the limited aspects of doctor effectiveness predicted solely through cognitive measures, in 1996 the Royal College of Physicians and Surgeons of Canada implemented their CanMEDS program, now used at all 17 Canadian medical schools. CanMEDS requires its students to demonstrate competency as the following: professional, leader, communicator, scholar, collaborator, health advocate, and medical expert (Fortin, Kealey, Slade, & Hanson, 2016). Since the implementation of CanMEDS, the Canadians have introduced two new medical admissions assessment measures—an interview process referred to as the Multiple Mini-Interview (MMI) in 2002 and the CASPer situational judgment test in 2014. These assessments are aligned to the CanMEDS framework (Altus Assessments, 2020). Both of these efforts demonstrate Canada's desire for the physicians they educate to be medically knowledgeable, but to also possess other personal skills and predispositions necessary to meet the needs of the patients they serve.

Concurrently in the US, the Association of American Medical Colleges (AAMC) suggested medical schools begin placing greater emphasis on assessing personal characteristics when selecting future medical students (Albanese et al., 2003). They argued that aside from intelligence, the following qualities were necessary to make a good physician: communication skills, integrity, altruism, self-management, multi-tasking, ability to think on one's feet, interpersonal skills, the ability to work as a team, and a strong ethical and moral compass (Kenny et al., 2003; Lumsden et al., 2005). In 2010, the AAMC formally directed US medical schools to use a holistic approach when assessing applicants. These efforts were meant to look beyond GPAs and MCAT scores and assess an applicant's life experiences and personal attributes.

In 2013, the AAMC created the Experiences, Attributes, and Metrics (E-A-M) Model (Appendix A) to assist admission committees with balancing their consideration of candidates' experiences, attributes, and metrics equally across the applicant pool by providing individualized consideration to each candidate (Arredondo, 2015; Harris et al., 2018; Lancaster et al., 2020; Wros & Noone, 2018). Using the EAM Model and considering each facet in combination with another allows admission committees to better assess the various dimensions of their applicants and evaluate them through various lenses. Assessing a student by only considering their academic metrics disregards much of their personal narrative. Being able to consider their life experiences, physical abilities, educational background or the languages spoken in their household paints a much more vivid picture of the applicant being considered. Using the EAM Model allows the admission committee to recognize the various abilities students bring to medical school when they have faced challenges and overcome adversity in their lives. These challenges often allow students to develop traits that are critical to success and therefore should be valued in the admission process (Kahlenberg, 2004; Razack et al., 2009; Witzburg & Sondheimer, 2013). Although the EAM Model does recommend various attributes to consider such as race, faith, and family status, others such as values and beliefs, perspectives, and maturity appear to be very subjective and therefore are inconsistent and perhaps only informally assessed by individual admission committees.

In 2002, due to concerns about the personal and biased nature of the traditional interview, McMaster University's Michael G. DeGroote School of Medicine created a highly structured interview process, known as the Multiple Mini-Interview (MMI). By utilizing multiple interviews, the design of the MMI allows for several perspectives regarding an applicant's candidacy. In an effort to consider non-cognitive factors in the admission process, approximately 30% of US allopathic medical schools have adopted Canada's multiple mini-interview (MMI) process (Axelson et al., 2010).

According to Eva et al. (2009), using a structured interview in the admissions process, where interviewers ask the same questions of all candidates, proves beneficial in assessing non-cognitive skills such as communication, empathy, integrity, and compassion, all of which are necessary to provide patient-centered health care (Monroe et al., 2013). These efforts have shown promise, and in recent years admission committees have begun to focus on some non-cognitive factors in selecting who to accept into medical school (Monroe et al., 2013). As admission committees attempt to assess these assets and gain further insight into an applicant's life experiences, struggles, and 'story,' these qualities can be difficult to quantify (Reiter et al., 2012). Often, admission committees rely not only on the interviews, but also personal statements and letters of recommendation in an attempt to assess an individual's attributes. These stories or experiences of "distance traveled" often reflect examples of grit, perseverance, and resilience. Medical school admission committees assess how these personal qualities have helped applicants reach their goals despite the challenging situations they have faced, including different levels of success, access to support, and social disparities (Ray & Brown, 2015; Stoffel & Cain, 2018).

There is some preliminary evidence of the importance of the qualities of conscientiousness and grit in performing clinical care and the pressures associated with it. Recent events related to our front-line physicians caring for COVID-19 patients while putting them and their family's health at risk validate the importance of these qualities, including the personality trait of conscientiousness defined by Costa and McCrae (1989) as being careful or diligent in one's work or life. The authors contend that conscientious people tend to be efficient and organized, want to perform their tasks well, and take their obligations to others seriously. Those who exhibit conscientious behaviors are those who strive for achievement and are competent, dutiful, and trustworthy, as well as assertive, altruistic, self-disciplined, and deliberate (Costa & McCrae, 1989). An individual's motivation may provide a plausible explanation as to why conscientiousness is the best predictor of academic success in medical school, undergraduate college, and graduate programs (Furnham et al., 2003; Lievens et al., 2009; Mann, 1999; Wiggins et al., 1969; Wolfe & Johnson, 1995). This trait has been found to be the strongest predictor of academic success in college students (Noftle & Robins, 2007; Poropat, 2009; Trapmann et al., 2007). Haight et al. (2012) reported that resilience and conscientiousness were more likely to produce physicians with good clinical and interpersonal skills who showed a greater level of humanism toward their patients (Hojat et al., 2013).

Conscientiousness has been found to consistently predict students' success in the basic science years of medical school (Ferguson et al., 2003; Lievens et al., 2002). Doherty and Nugent (2011) attributed conscientiousness to long-term success throughout a student's medical training and future job performance. Likewise, conscientiousness has been identified as a crucial predictor of job performance in medicine, as well as in other professions (Barrick et al., 2001; Behling, 1998; Hurtz

& Donovan, 2000; Tett et al., 1994). Magee and Hojat (1998) found that physicians who were nominated by their supervisors as positive role models had significantly higher levels of conscientiousness as measured by the Neuroticism, Extraversion, Openness (NEO) Personality Inventory.

Similarly, given the years of medical education and training required for acquiring a Medical Doctorate, it is of institutional interest to ensure matriculating students stay the course and acquire the skills necessary to provide skilled healthcare. Duckworth and Quinn's Short Grit Scale is intended to measure "the perseverance and passion for long-term goals" (Duckworth & Quinn, 2009; Robertson-Kraft & Duckworth, 2014; Underdahl, Jones-Meineke, & Duthely, 2018). By using tools such as the grit scale, those who are determined to be "gritty" have been shown to persevere and stay focused on long-term goals, which is necessary when dealing with the challenges of pursuing a medical education (Ray & Brown, 2015; Silvia et al., 2013). Grit is another trait that might help navigate the extraordinary and unique academic demands of medical school (Ray & Brown, 2015). Grit is the tendency to sustain interest in and effort toward one's long-term goals (Duckworth et al., 2007). Duckworth described grit as an individual having a goal that they care so deeply about that it gives meaning to almost everything they do, what some researchers have referred to as their "ultimate concern" (2016, p. 63). Duckworth (2016) stated:

> To be gritty is to keep putting one foot in front of the other. To be gritty is to hold fast to an interesting and purposeful goal. To be gritty is to invest, day after week after year, in challenging practice. To be gritty is to fall down seven times, and rise eight (p. 275).

Grit enables individuals to persevere and stay focused on long-term goals, critical attributes required to meet the rigors of medical school (Ray & Brown, 2015; Silvia et al., 2013). Grit has been shown to predict academic success for undergraduate students at an Ivy League institution, the University of Pennsylvania (Duckworth et al., 2007) but empirical evidence is warranted to examine the importance of this quality in predicting medical school success.

Recently, the grit construct has been challenged in education (Anderson, Turner, Heath, & Payne, 2016; Christopoulou, Lakioti, Pezirkianidis, Karakasidou, & Stalikas, 2018; Datu, Yuen, & Chen, 2017; Perry, 2016; Ris, 2016; Socol, 2014). Tewell (2020) suggested that the construct of grit furthers a deficit approach when considering under-represented students (Tewell, 2020). Mehta (2015) posited that focusing on grit, was another attempt to 'blame the victim' instead of focusing on larger systemic issues such as race and social and economic equality. Although the critiques surrounding grit have been primarily related to overall academic performance and/or the test performance of underrepresented and lower socioeconomic students (Love, 2019; Mehta, 2015; Ris, 2016; Tewell, 2020), Mehta's (2015) suggested that the focus on grit may be misleading as it insinuates that students just need to be 'grittier'. Mehta (2015) recommended that educators find ways to develop, support, and encourage individual students to find their own purpose and cultivate their own unique passions. Despite these recent challenges, grit has been shown to be linked to greater academic achievements, requiring determination and perseverance to achieve long-term goals (Duckworth et al., 2007; Strayhorn, 2013), which is particularly relevant to the pursuit of a Doctor of Medicine degree. Therefore, with these challenges in mind, assessing grit at the point of admission would provide an opportunity for admission committees to identify applicants with a high level of grit who may have previously demonstrated perseverance, commitment, and grit in their lives. The medical school application is not conducive to significant expansion of a student's narrative, details about metaphorical distance traveled, injustices and/or economic hardships faced, or cultural capital implications, all of which may demonstrate grit, are difficult to fully appreciate in a 5300-character count personal statement. Perhaps seeing a higher level of grit would trigger admission committees to dive deeper into the backgrounds of these applicants. Assessing these applicants based more holistically on their backgrounds and what they can ultimately bring to the medical school and the profession, will ultimately lead to a more diverse physician workforce.

Theoretical Framework

Collecting, analyzing, and presenting data within particular theoretical lenses have allowed researchers to further frame their research (Bogdan & Biklen, 2006; Towne & Shavelson, 2002) while also allowing readers to better understand the researchers' approach in developing their research questions, the methodology used, and the analysis conducted. The assumptions of two theories undergird this study.

First, Sternberg's Theory of Successful Intelligence (1999, 2005) was a useful framework for this research. This theory of holistic intelligence claims that a person's proficiency or success is linked to the circumstances within their lives where they focus on their strengths in order to ultimately reach their long-term goals (Sternberg, 1999, 2005). Sternberg advised that one's intelligence, and therefore success, should not be assessed by standardized examinations alone because they only describe the cognitive aspect of one's intelligence (Sternberg, 1999, 2005). Specifically, Sternberg explained the need to define and examine an individual's intelligence broadly, particularly for those students from underrepresented backgrounds. This theory of intelligence argues that cognitive assessments alone cannot sufficiently capture or measure ability and intelligence for students who have different strengths and life experiences (Sternberg, 1999, 2005). Examining a student's cognitive and noncognitive attributes more broadly may be more indicative of student strengths and weaknesses (Kamenetz, 2015). This broader assessment could impact admission decisions, allowing for a more diverse student body and therefore, physician workforce.

In addition, Yosso's Cultural Wealth Theory provided insight into the importance of recognizing the assets that Communities of Color possess and bring to educational environments (Yosso 2005). Yosso (2005) argued that the deficit view of students of color that projects that students often lack the skills, knowledge, and competencies valued in a dominant White society needs to be critiqued for how it reproduces privilege for the middle- and upper-class White elite. The traditional notions of cultural capital that are valued often are captured in college and medical school admissions criteria such as standardized test scores, undergraduate GPA, and undergraduate college selectivity. However, Yosso (2005) advocated that scholars instead "focus on and learn from the array of cultural knowledge, skills, abilities, and contacts possessed by socially marginalized groups that often go unrecognized and unacknowledged" (p. 69). Some of these assets or "cultural wealth capital" (p. 69) that she identified include aspirational, navigational, social, linguistic, familial, and resistant capital. These forms of capital draw on the knowledge students of color bring with them from their homes and communities into the classroom. Some of these forms of capital relate to the non-cognitive variables identified in this study. For example, aspects of community cultural capital that Yosso (2005) identified that directly related to perseverance or grit and/or conscientiousness, include navigational capital as critical assets that students of color bring with them to higher education to help them achieve high levels of academic success; they have learned knowledge and skills for negotiating educational systems that privileges Whites and must combat constant stress and racist conditions. She tied this asset to the construct of resilience which has been shown to correlate with grit, a non-cognitive variable used in this study (Hans & Salles, n.d.). Hard work and conscientiousness also might capture some of the critical skills and dispositions that students of color bring to help them maneuver educational environments that are unwelcoming and even hostile. Further, linguistic capital may be important in medical admissions, education, preparation, and success, as students figure out the jargon, the language and how they must represent themselves in the profession. Students of color might bring their linguistic capital to their clinical skills in serving patients and communities of color too. An additional asset may be the cultural significance of family, in terms of resources and strengths, particularly as mothers have played a critical role in supporting and validating their children, their determination to complete their life goals, not only for themselves, but also for their families. Grit and conscientiousness, reflect some of the capital that Yosso refers to that expands beyond persevering in the face of adversity. Yosso's work reinforced my desire to study a broader set of variables that might capture the rich and diverse assets that students of color bring to medical education, the physician workforce, and the communities they serve.

Garcia and Guerra (2004) argued that deficit thinking permeates our society, especially within our schools. The authors demonstrated that a change in educator views related to race, gender, and class was necessary, as was a "critical examination of systemic factors that perpetuate deficit thinking and reproduce educational inequities for students from nondominant sociocultural and linguistic backgrounds" (p. 155). Changing the focus to an anti-deficit model and considering what diversity could bring to the field of medicine would allow students to learn from one another and consider the practice of medicine from various perspectives and backgrounds. Further, strategically applying a holistic review framework in order to widen the lens through which admission committees view potential applicants will promote consistency, equity, and fairness for applicants. That said, admission committees must consider the downstream effect, as student experiences are defined by the path that applicants have taken to get where they are and the context in which their experiences have taken place. It is critical to understand that students from privileged backgrounds may have opportunities to create new experiences for themselves while other students must go above and beyond to achieve similar results. In order to further leverage the benefits of diversity and inclusion on the student's ability to meet the criteria for admission, these criteria need to be refocused and the lens through which these determinations are made widened.

Sternberg's Successful Intelligence theory and Yosso's Cultural Wealth Theory support a more holistic approach in the assessment of potential medical students. These theories provide a framework that helps this research to substantiate the need to identify and assess various factors in the medical school admission process beyond traditional cognitive measures. Doing so will hopefully increase the diversity of our medical school programs and physician workforce.

Problem Statement

There are significant challenges facing the medical profession including a shortage of physicians, a mostly homogenous physician workforce, and a more diverse population of patients to serve. Traditionally, medical education has prioritized scientific content knowledge in the student selection process throughout medical training (DeZee et al., 2012). However, the literature suggests that possession of hard scientific or medical knowledge is not enough to produce successful physicians. As the demographics of the United States have changed over time and medicine has evolved in terms of patients' desire to have physicians who demonstrate both intellectual and emotional intelligence has grown, a more holistic, individualized approach in the medical student selection process has been advocated.

Considering additional or alternative variables that are strong predictors of medical school success may equip admission committees to better identify qualified and more diverse medical school candidates who will then later serve their communities (Cohen & Steinecke, 2006; Kirch & Petelle, 2017; Witzburg & Sondheimer, 2013). The heavy weight placed by medical school admission committees on MCAT scores (Casey et al., 2016; Jerant et al., 2018) has been shown to disproportionately negatively impact under-represented candidates. Research has shown that standardized tests are biased and disadvantage particular groups of students (Casey et al., 2016; Hopkins, 2008; Kendi, 2018; Jerant et al., 2018; Sacks, 2007; Solórzano, 2008; Steele, 1999). For example, women have historically fared worse than men on standardized tests (Sacks, 2007; Sedlacek, 2004), and students of color, first-generation college students, and non-traditional students often do not test as well as their traditional-aged, white peers from educated, middle class families (Alameida et al., 2011; Bowen & Bok, 2016; Bowen & Rudenstine, 2003; Petterson et al., 2012). An equitable, valid predictor of academic performance, such as assessing non-cognitive traits, could lead to greater diversity among health care providers, which could also lead to a decrease in health disparities.

Identifying specific factors that can better represent the critical assets students need to bring to medical school but are not currently captured through standardized testing and science GPA is important. They should be cost and time effective, as well as reliable and valid measures that would be helpful in all phases of the medical school experience, including the clinical years and in practice. These measures would essentially augment qualitative data obtained during the medical school interview. The techniques physicians use to begin their practice of medicine are often developed and practiced during their third year of medical school, their clinical year. At this time students begin to interact directly with patients, their families, and the various professionals on their health care teams on a daily basis (Eva et al., 2009; Monroe et al., 2013). Students must learn to take a patient's history, develop a rapport with their patients, and work as a team while also remaining independent, thoughtful, and professional, and possessing tenacity and integrity. There are students who have the cognitive ability to make it through the first two years of their medical education, as the focus is on grades and test scores, yet when they get to their clinical year they lack the interpersonal skills necessary to fully engage their patients, perform poorly on their individual clinical assessments, struggle with the intense learning environment, long days and sometimes difficult encounters with their patients and/or preceptors, and may fail their USMLE Step 2 examination after three years of schooling (Chang et al., 2009; van Zanten et al., 2007). Therefore, the need is great for admission assessment measures to correlate with potential clinical success as well.

Purpose of this Study

The purpose of this study was to explore how specific non-cognitive factors, conscientiousness and grit, correlate with medical student success on the national USMLE Step 2 Clinical Knowledge (CK) examination at a public medical school in the Northeast. In this study, using a single institution dataset design, the relationship among the variables using descriptive and multivariate statistical analysis was assessed. Regression analyses were run in order to determine the relationship between the independent variables, and particularly the relationship of grit and conscientiousness on Step 2 CK performance.

The following four research questions were explored in this study:

Research Question 1: Do socio-demographic variables—specifically age, gender, ethnicity, socioeconomic status, or first-generation college status—correlate with success on the national USMLE Step 2 CK medical school examination at Northeast Medical University?

Research Question 2: Do academic variables—specifically MCAT scores, prior college science GPA, college major, post-baccalaureate course work, or graduate degree—correlate with success on the national USMLE Step 2 CK medical school examination at Northeast Medical University?

Research Question 3: Does conscientiousness, as measured by Costa and McCrae’s Neuroticism, Extraversion, Openness Five-Factor Personality Inventory (NEO-FFI-3) (1992), predict medical student academic success on the USMLE Step 2 CK medical school exam when controlling for demographics and prior academic history at Northeast Medical University?

Research Question 4: Does grit, as measured by Duckworth and Quinn's Short Grit Scale (2009), predict medical student academic success on the USMLE Step 2 CK medical school exam when controlling for demographics and prior academic history at Northeast Medical University?

Significance of this Study

Although science GPA and MCAT scores have correlated with success during the first two (basic science) years of medical school, as well as on the Step 1 exam (Donnon, et al., 2007; Muller & Kase, 2010), and the MCAT is considered a reliable measure of applicant academic competence (Albanese et al., 2003; McGaghie, 2002; Saguil et al., 2015), these assessments are unable to predict which applicants possess non-cognitive skills that may be correlated with success in the third and fourth (clinical) years of medical school or on the USMLE Step 2 CK examination. Previous research conducted regarding various psychosocial, non-academic predictors of academic success among graduate students, often with positive results (Kyllonen, 2012; Kyllonen et al., 2011; Lee et al., 2007). Yet little research exists in the literature about assessing non-cognitive factors in medical school admission and the correlation at distinct assessment points during the medical school experience. Researchers have suggested that including non-cognitive factors with cognitive factors when evaluating applicants might better predict how students perform in their clinical years, when they are interacting with patients and other health professionals (Ferguson et al., 2002; Sobowale et al., 2018). Similarly, Haight et al. (2012) found that conscientiousness was strongly related to various facets of clinical success including history taking, physical exam skills, patient rapport, and health care team rapport.

Predicting success in the clinical years and on Step 2 is imperative. Currently, cognitive assessments such as MCAT and undergraduate GPAs are not sufficient to predict medical students' clinical knowledge and skills (Kyllonen, 2012; Lemann, 2000; Sedlacek, 2001). This research contributes to this current body of literature by assessing whether non-cognitive measures, specifically grit and conscientiousness, serve to correlate with student performance on the USMLE Step 2 CK. Additionally, other student background measures, specifically age, socioeconomic status, first-generation college status, and advanced coursework, are investigated for their influence on student performance on this examination. There have been cases where students have made it through the basic science years of their medical education and been successful on Step 1 but were unable to successfully progress through their clinical years (Case et al., 1996; Fields et al., 2000). This disheartening situation is a physically, emotionally, and financially taxing process for students. Identifying these students earlier by using predictive assessment tools would allow faculty to work with individual students over a longer period of time to help them develop the qualities necessary to be successful on the USMLE Step 2 as well as clinically.

In the fall of 2019, the National Board of Medical Examiners (NBME) and the Federation of State Medical Boards (FSMB) announced that the scoring of the USMLE Step 1 would change from a three-digit numeric score to a pass/fail result (USMLE, 2019a). This change will take effect in January of 2022 and is the result of several years of discussion. The intention of the three USMLE examinations was to ensure that students met the standards for their medical license. Over the years, an overemphasis on the USMLE Step 1 by residency programs as a primary screening tool for candidate selection has placed students under significant pressure to meet particular thresholds based on specialty area, and has potentially limited diversity within specific specialty areas (Gardner et al., 2019; Prober et al., 2016). With this move to a Pass/Fail grading system for the USMLE Step 1 score, some have postulated this change will only force residency directors to further consider the results of the Step 2 CK examination in their screening process, continuing to put significant pressure on students and their test-taking abilities, influencing the specialties in which they go into, and continuing to negatively impact the diversity of students entering particular specialty programs (Chaudhry et al., 2020; Gardner et al., 2019). This change, and the implication that there may be increased reliance on the Step 2 CK score for residency screening, further suggests that assessing additional qualities students possess could be imperative for success on the USMLE Step 2 CK examination.

Finally, this study may contribute to improving the current medical school landscape characterized by a serious underrepresentation of students from minoritized backgrounds. As the profession seeks to increase and diversify the physician workforce, complementary predictors of academic success may greatly benefit medical programs. This research contributes to the understanding of medical students and practice implications that impact decision-making processes concerning admission to medical school. As future physicians need to better reflect the increasing diversity within our communities, the specific qualities these students bring to the medical profession must be part of the admissions process.

Definition of Relevant Terms

The following definitions seek to provide clarification to some of the medical terms, exams, and associations referenced frequently in the medical community and this research study.

AAMC - Association of American Medical Colleges (AAMC) leads the academic medical community. AAMC members include all 152 US accredited medical schools, 17 Canadian medical schools, 400 teaching hospitals and health systems, 51 Department of Veteran Affairs medical centers, and more than 80 academic societies (AAMC, 2020e).

AMCAS - American Medical College Application Service (AMCAS) is the centralized application processing service used by the AAMC. Most US medical schools use this service as the primary application service for their applicants.

***Medical Students* -** Students matriculated at a LCME (Liaison Committee on Medical Education) accredited medical school.

***USMLE Step 1* –** The first part of the national licensing examination for medical students. This exam is taken at the end of the basic science years, which normally follows the second year of medical school. It assesses foundational science concepts and the medical knowledge students have gained that are essential to the practice of medicine.

USMLE Step 2 - A licensing exam consisting of two sections, Clinical Skills (CS) and Clinical Knowledge (CK). Traditionally this exam is taken at the end of the third year of medical school, and unlike the USMLE Step 1, is more application-based than knowledge-based.

***Clinical Success*-** Student receives passing scores on USMLE Step 2 CK.

Summary

This chapter introduced the cognitive factors considered in the medical school admissions process and the recommendation from the AAMC to evaluate applicants from a holistic perspective, considering the "road that students have traveled" as well as their communication skills and personal characteristics. There are current challenges facing the medical profession including a shortage of physicians, changes in medical school admission processes, and a more diverse population of patients to serve. Based on these challenges, two measures were proposed to assess the non-cognitive qualities of prospective medical students during the admissions process, the conscientiousness trait from the NEO-FFI-3 (Costa & McCrae, 1992b) and grit from the Grit-S scale (Duckworth & Quinn, 2009) in predicting success on the USMLE Step 2 CK national examination.

Organization of the Chapters

This chapter reviewed information on current and future physician workforce needs, curriculum and testing in medical school, and admission to medical school. It identified the research to be conducted, the theoretical frameworks considered, and the purpose of this study. Previous research is considered in Chapter Two, specifically related to the predicted physician shortage, the landscape of knowledge and skills in medical education, and producing a diverse physician workforce with a particular focus on race/ethnicity, socioeconomic status, first-generation college status, gender, and the position of non-traditional students in medicine. Further, Chapter Two reviews medical school admission variables, the cognitive and the non-cognitive measures, as well as additional post-baccalaureate or master’s degree educational preparation, followed by a look into conscientiousness and grit viewed through a Cultural Wealth lens. Chapter Three provides the study rationale and research questions as well as an in-depth assessment of the constructs used in this study, followed by the research design, participants, data collection and analyses. In Chapter Four, the findings of this research are presented, and descriptive statistics of the variables used in this study are provided. A multivariate analysis explains the findings related to grit and conscientiousness as predictors of medical student success on the USMLE Step 2 CK. Chapter Five interprets the results of this data within the framework of relevant literature accompanied by potential implications for medical school admissions. Lastly, strengths and limitations of this study and recommendations for future research are reviewed.

CHAPTER TWO: REVIEW OF THE LITERATURE

This literature review provides an overview of predicted dramatic shortage of physicians to meet the medical needs of the U.S. communities, and huge inadequate representation of physicians from under-represented groups to meet the needs of communities of color. As background and context, an overview of the evolution and current assessments used to measure medical school success is shared. A description of who has access to medical school and how medical school experiences and success vary based on different under-represented groups is provided. The central variables (cognitive, non-cognitive, and post-college educational background) that are considered in current admissions processes and evidence of their predicative power in determining success are reviewed. Finally, two non-cognitive variables—grit and conscientiousness—that reflect an asset's perspective as argued by Yosso's Community Cultural Wealth Model (Yosso, 2005) is reviewed.

Predicted Physician Shortage

Research provides consistent evidence of a major physician shortage and in particular, significant inadequate representation of physicians from under-represented groups to meet the needs of communities of color. In a study by Kirch and Petelle (2017), projections indicate that by 2030 the United States will face a physician shortage between 40,800 and 104,900 physicians, with a possible shortage of 29,000 surgical specialists. Additionally, the population of the United States will increase in size and age, with the overall population increasing by 15% and the proportion of those over 65 increasing by 60%. Based on data from 2008, researchers found that adults over 65 are more likely to go to primary care appointments than people in younger age groups (Petterson et al., 2012). Because of the increase in the overall population, and the increase in people over the age of 65, researchers estimate that the number of primary care office visits will increase from 462 million in 2008 to 565 million in 2025 (Petterson et al., 2012). This would mean that the number of primary care physicians would need to increase by 51,880 (Petterson et al., 2012). Further, based on this anticipated physician shortage, providing care to marginalized groups making up an increasing amount of the US population is of the highest priority (Bradby et al., 2020; Petterson et al., 2012).

The COVID-19 pandemic has dramatically highlighted the inequities in our medical system as COVID-19 disproportionately affected communities of color. The need is great for physicians to serve in communities that historically have been made up of the most vulnerable and marginalized populations (Bradby et al., 2020; Filut et al., 2020; Kelly-Blake et al., 2018). Unfortunately, the current demographics of the physician workforce do not meet the needs of our heterogeneous population. Having culturally diverse providers may help reduce the racial disparities observed in the delivery of health care (Cohen et al., 2002). It is imperative to the health of our society that the diversity of our physician workforce grows exponentially to serve our increasingly diverse community (Institute of Medicine, 2004; Salsberg & Forte, 2002; Witzburg & Sondheimer, 2013). Sadly, this issue appears to be systemic, as the population growth in the US will occur within those populations that remain the least educated. Black, Indigenous, and People of Color students are disproportionally underrepresented at every stage of their education (Kelly, 2005). The Health Resources and Services Administration (HRSA) reviewed more than 35 studies and found when patients had racial or ethnic backgrounds and language in common with their health care providers, they utilized health care services more regularly and the quality of their health increased (HRSA, 2006). Creating more diversity among today's physicians will allow patients to be cared for by clinicians who look like them and share the same cultural and ethnic backgrounds. This will likely increase "communication, comfort level, or trust in patient-practitioner relationships and thereby improve partnership and decision making" (HRSA, 2006, p. 7). In addition, HRSA reviewed 17 distinct studies and found that physicians from underrepresented backgrounds were more likely to work in medically underserved and poverty-stricken areas, allowing greater access to health care for more vulnerable populations (HRSA, 2006). In a 2004 survey of Black Medicare recipients, Bach et al. found that 22% of those surveyed preferred to receive their care from African-American clinicians. Furthermore, in 2004, Sullivan predicted that “increasing diversity in the health care professions will improve health care access and quality for minority patients” (p. 13). Since that time, several published articles have reiterated the need to diversify the healthcare workforce to improve both the quality and access to healthcare (Aiken, 2011; Beacham et al., 2009; Buerhaus et al., 2009; Dapremont, 2011; Sutherland et al., 2012).

Assessment of the Knowledge and Skills in Medical Education: History and Current Landscape

In the United States, the examinations for medical licensure are governed by the National Board of Medical Examiners (NBME) and the Federation of State Medical Boards (FSMB). Three steps of the United States Medical Licensing Exam (USMLE) must be passed by medical students before they are eligible to practice in the United States with an unrestricted license. Step 1, which is typically taken after the second year of medical school assesses a student’s understanding of scientific concepts and the principles and mechanisms of medicine. Step 2 is broken into two parts: The Clinical Knowledge (CK) section completed after the third year of medical school assesses a student’ ability to apply medical knowledge and skills related to clinical science, and the Clinical Skills (CS) section, often completed during the fourth year, tests a student’s ability to interact with patients, perform physical examinations, and present their patient findings. Of these two sections, there is significant controversy surrounding the implementation, rationale, and validity of the Step 2 CS (Cuddy et al., 2016; Flier et al., 2016; Jayakumar, 2018; Lehman & Guercio, 2013; Mehta & Kramer, 2005). Step 3 is the final examination required for full medical licensure and is typically taken after a student’s first year of residency. This examination focuses on the student’s ability to apply their knowledge of both biomedical and clinical science and if passed, allows them to practice medicine unsupervised (USMLE, 2020a).

These examinations have undergone significant changes over the past several decades, as testing ideology and technology has changed. The United States Medical Licensing Examination Step 1 and Step 2 Clinical Knowledge, offered by the National Board of Medical Examiners, are often regarded as two of the most challenging milestones in medical school (Thundiyil et al., 2010). Many residency programs use the scores from the USMLE Step 1 and sometimes Step 2 CK, if the Step 1 score is not strong enough, as criteria for screening candidates (Chaudhry et al., 2020; Garder et al., 2019; Thundiyil et al., 2010). Finding early specific indicators of potential future failure on the USMLE Step 1 and the Step 2 CK would be of benefit for medical schools and for medical students, because they could be used to develop early interventions to solve deficiencies and help at-risk students to ultimately attain a passing score (Tamblyn, 1998; Veloski et al., 2000). Therefore, there is a growing interest in finding adequate predictors of success and failure around the USMLE, especially in light of the change to Step 1 scoring described in Chapter One.

Through a partnership between the NBME and FSMB, the USMLE was introduced in 1992. Step 1 was implemented first, followed by a single Step 2 exam and then the Step 3 exam in 1994 (National Board of Medical Examiners, 2015). In 2004, the NBME implemented a second part to the Step 2 exam, which became known as Step 2 Clinical Skills (CS). The Step 2 CS exam was modeled after an exam that had been created in 1998 for international graduates (Levine et al., 2017; USMLE, 2020d). The implementation of Step 2 CS necessitated a name change for the traditional multiple-choice exam designed to test the principles of clinical science essential to medical training from Step 2 to Step 2 Clinical Knowledge (CK) (USMLE, 2020b).

The USMLE Step 2 CK tests medical student's clinical knowledge, basic patient-centered skills, and their comprehension of clinical medicine that is necessary for patient care (United States Medical Licensing Examination, 2020c). There is a focus on health promotion and disease prevention. Step 2 CK ensures that students are committed to clinical sciences principles and basic patient-centered skills. The USMLE Step 2 CK is a one-day, 9-hour, examination, divided into eight 60-minute blocks (USMLE, 2020c). Students complete the CK examination at Prometric testing centers across the nation. According to the USMLE, beginning November 2020, the CK examination will include more questions that will specifically assess the following competencies: systems-based practice, patient safety, legal/ethical issues, and professionalism. Passing this exam allows students to practice medicine in a supervised environment. Below is an example of a competency-based question asked on the USMLE Step 2 CK examination (USMLE, 2020c):

> Competency: Professionalism Content Area: Social Sciences:
>
> Three days after hospitalization for diabetic ketoacidosis, an 87-year-old woman refuses insulin injections. She says that her medical condition has declined so much that she no longer wishes to go on living; she is nearly blind and will likely require bilateral leg amputations. She reports that she has always been an active person and does not see how her life will be of value anymore. She has no family and most of her friends are sick or deceased. On mental status examination, she is alert and cooperative. She accurately describes her medical history and understands the consequences of refusing insulin. There is no evidence of depression. She dismisses any attempts by the physician to change her mind, saying that the physician is too young to understand her situation. She says, "I know I will die, and this is what I want." Which of the following is the most appropriate next step in management?
>
> (A) Discharge the patient after she has signed an "against medical advice" form
> (B) Seek a court order to appoint a legal guardian
> (C) Offer insulin but allow the patient to refuse it
> (D) Admit to the psychiatric unit
> (E) Administer insulin against the patient's wishes

For further information, a content description and general information guide for Step 2 Clinical Knowledge (CK) is included as Appendix B. According to the USMLE, all exam content is organized by general principles and individual organ systems, with test questions being a part of 18 major categories (USMLE, 2020c). Nationally, the passing rate on the USMLE Step 2 CK in 2018 was 96%, and in 2019, it was 97% (USMLE, 2020e).

Since its implementation in 2004, the Step 2 Clinical Skills (CS) exam has been controversial. Even before its official initiation, its efficacy and value were called into question (Alvin, 2016; Jayakumar, 2018; Johnson, 2003). The Step 2 CS exam was introduced in order to identify medical students with poor communication skills or other interpersonal shortcomings, particularly as there has been a perceived decline in the clinical skills abilities of medical school graduates (Faustinella & Jacobs, 2018). Since the advent of Step 2 CS, students have been awarded pass/fail grades based on their interactions with standardized patients in this section (USMLE, 2020b). The NBME implemented the Step 2 CS exam with hopes that it would create a safeguard for the public by ensuring that graduating medical students possessed the clinical skills necessary to safely practice medicine (Gilliland et al., 2008; Hawkins et al., 2005; National Board of Medical Examiners, 2015). Competencies on Step 2 CS include information gathering, physical examination skills, and communication skills (FSMB & NBME, 2020). While raw scores of Step 2 CS are used to determine pass/fail status, these data are not released to the students (USMLE, 2020b). Residency programs currently receive the numeric Step 1 score which is used in many cases to screen applicants for residency positions, and the numeric Step 2 CK score which residency program directors often use for ranking their candidates in the residency match process (Gauer & Jackson, 2017; Swails, Aibana, & Stoll, 2019). Programs will then receive the Step 2 CS Pass/Fail result for students applying to their programs, yet sometimes this result arrives after students have been offered a residency position.

Many proponents point to the current body of literature which suggests that poorer clinical/communication skills are associated with a higher incidence of malpractice suits (Johnson, 2003; Waldman & Spector, 2003). If malpractice is used as a proxy for medical error, proponents for Step 2 CS may believe that evaluating physician communication will help prevent future medical errors. However, Johnson (2003) pointed out this rationale may not be entirely accurate. Other researchers suggest that malpractice suits are more strongly related to poor communication skills or perceived physician attitude than actual medical negligence (Ambady et al., 2002; Donaldson et al., 2000; Moore et al., 2000). Jayakumar (2018) suggested that the Step 2 CS pass/fail scoring system provides vague information that is not useful to students or residency directors. However, First et al. (2013) argued that implementation of Step 2 CS was important for ensuring patient safety; while the pass rate of the exam is high, the Step 2 CS exam is important for identifying those who are deficient. Ecker et al. (2018) added that ending Step 2 CS would devalue the inclusion of clinical skills in medical education. While the pass rate of the Step 2 CS is high, the authors suggested that the exam must be kept in place to prevent the small percentage who fail from potentially endangering patients.

One of the most important objectives of medical school is to prepare students for residency (Green et al., 2009; McGaghie et al., 2011; USMLE, 2020b). Although programs vary in the qualities they value, USMLE Step 1 and Step 2 CK exam scores are typically weighted heavily in residency placement as the goal of applicant selection is to choose individuals who will meet or exceed the expectations placed on them during residency and who will have a successful career after residency training (Taylor et al., 2005). Residency programs view USMLE scores as objective and consistent measures of medical student cognitive performance that provide an easy method of comparison of applicants from different medical schools (Harfmann & Zirwas, 2011; McGaghie et al., 2011; USMLE, 2020b).

A study conducted by Green et al. (2009) attempted to organize applicant variables based on importance for residency applications by surveying program directors across the United States. The authors concluded that the five most frequently used criteria in residency selection are clerkship grades, USMLE Step 1 scores, grades in senior electives pertaining to the specific specialty, number of honors grades, and Step 2 CK scores (Green et al., 2009). In the specialties considered less competitive, the USMLE Step 2 CK is ranked higher, while in the most competitive specialties, research experience is more highly valued (Green et al., 2009). Due to the controversy surrounding Step 2 CS and the very small number of students who fail this exam, this research will only include the assessment of how students perform on Step 2 CK. Step 2 CK will serve as a key proxy for medical student success.

Inadequate Medical School Pipeline to Produce a Diverse Physician Workforce: Access and Success in Medical School

Medical school matriculation has historically been skewed towards the privileged, with most matriculating students originating from well-educated, financially supported backgrounds (Nakae, 2014). Over the past several decades, medical school enrollment has shifted to include students from non-traditional backgrounds, such as those of low socioeconomic status (SES), students who are first in their family to attend higher education (FIF), and underrepresented minorities (URM) (Association of American Medical Colleges, 2020e). URM are defined by the Association of American Medical Colleges as students of Black, Mexican-American, Native American, or Puerto Rican origin (Association of American Medical Colleges, 2020e). While SES, FIF, and URM status are technically distinct demographics, people of low SES are more likely to be the first in their family to pursue higher education and are also more likely to be students of color. According to the Association of American Medical Colleges (AAMC), between 1978 and 2008, 75% of all medical school graduates were White. The most recent data available reports that of the 17,341 medical students who graduated from medical school in 2018, 12% of the graduates came from groups traditionally underrepresented in medicine (Association of American Medical Colleges, 2020c). As the field of medicine has historically been dominated by White men, dramatic changes in access to medical school and the medical school environment must be considered to transform the demographics within the medical profession to meet the needs of our changing society (Eden et al., 2014; Relman, 1989).

The broadly described "achievement gap" has been linked to racial and ethnic disparities in standardized test scores, specifically in considering medical students. The literature on the racial/ethnic achievement gap in standardized tests is extensive and important and some researchers have, in fact, asserted the need to move to conceptualize the "achievement gap" as an "opportunity gap," or an unfair advantage for more affluent, white student populations (Carter & Welner, 2013; Flores, 2007; Verstegen, 2015). This opportunity gap has been studied at various academic levels, including research on stereotype threat (Steele, 1998), including K-12 schools, college, STEM education, and graduate education (Carter & Welner, 2013; Flores, 2007; Verstegen, 2015). The challenge for medical education and admissions to recruit a diverse student population is one faced across all levels of education. In fact, diverse representation is needed across all health professions and STEM related fields given the demonstrated disproportionate impact on and return back to serving these communities of color (Grumbach, & Mendoza (2008); Smedley, Stith, Colburn, & Evans (2001).

An analysis of the literature investigating the impact of race/ethnicity, socioeconomic status, first-generation college student status, gender, and non-traditional/adult student status on both access and success in medical school follows.

Race/Ethnicity. Despite the United States becoming more diverse, the numbers of under-represented students graduating from medical school continues to be far less than what is needed to match the percentages within the population (Brummer et al., 2016). According to Cammarata (2010), US medical schools matriculate only a fraction of the number of minorities needed. Further, in 2008, the Kaiser Family Foundation communicated that there were 16,167 medical school graduates in the United States, with only 6.9% identifying as Black/African American and 7.3% identifying as Hispanic, while these groups represent more than 12% and 15% of the US population, respectively. A full decade later the numbers remain similar (AAMC, 2020a). Medical schools are still failing to reflect the communities they serve.

Not only are there not enough students of color applying to medical school (Cammarata, 2010), the Council on Graduate Medical Education (COGME) (as cited by McDade, 2019), a federal body commissioned by Congress, suggested that the problem starts much earlier than medical school:

> Increasing the number of under-represented minority students who successfully advance through the elementary, secondary, and post-secondary academic pipeline is the first step to enlarge the potential number of these students eligible to enter medical school...," COGME wrote in its 2005 report to Congress. (p. 13)

Cammarata (2010) stated that we are missing the mark on recruiting under-represented students into the field of medicine, suggesting that the real problem lies within their educational background long before they even consider a college education and argued that students are steered away from considering a medical education around the fourth grade.

The need is great to continue to educate under-represented students on the path to medical school, they must be shown that it is attainable. Pipelines must start earlier and provide significant advising and academic support (Cestone et al., 2018), fundamental reforms must start before middle school (Cohen et al., 2002), and students need to see people like them in these positions (Hill et al., 2018; Poole, 2020).

Yet once students get to medical school, these supports must continue and a different kind of support must begin. Once in medical school, students may face microaggressions as well as outright racism. Damon Tweedy, MD, author of the New York Times bestselling book, *Black Man in a White Coat, A Doctor's Reflections on Race and Medicine*, said in an interview:

> As a black man, a very tall black man entering the medical world, people see me and the last thing they're thinking is me as a doctor. As an early medical student, one of my first experiences was a professor asked me to fix the lights in his classroom, mistaking me in that capacity. Another incident as a young doctor where a patient said he didn't want an N-word doctor taking care of him, so there's all these difficult experiences you grapple with. Medical training is hard no matter who you are, and then there's this added element by virtue of being black (MedPage Today, para. 2).

Minority medical students have been shown to underperform academically relative to their peers in medical school (Woolf et al., 2013). This is often the result of educational background (Cohen et al., 2002), familial responsibilities (Rainey, 2001), imposter syndrome (Cohen et al., 2009; Mullangi & Jagsi, 2019; Villwock et al., 2016), and concerns about finances (Figueroa, 2014).

A study by Isik et al. (2017) acknowledged that minority students do not perform as well as their peers during medical school, and may have difficulty securing a residency in the field of their choice (Leyerzapf et al., 2015), yet Isik (2017) argues student motivation is a significant factor influencing their learning and ultimately their success in medical school; therefore, resiliency and perseverance is critical to their success at all levels of their medical training. This need for resiliency and perseverance is pervasive among students of color, particularly in today's climate. In a recent article by LaShyra "Lash" Nolen, an African American female medical student at Harvard Medical College, Ms. Nolan talks about how she has had to leave her "blackness at the door" throughout her entire educational experience and focus on the white comfort of her professors and peers, often at the expense of her own mental health, in order to maintain her professionalism. Students like Ms. Nolan want their white faculty members to "see my color. Because along with my melanated skin comes the reality of my unique struggles, the resilient stories of my ancestors and the strength of my community." (Nolen, 2020, para. 36).

Having a diverse student body not only benefits those from underrepresented groups, but race-sensitive, holistic admissions programs "enrich learning environments by giving all students the opportunity to share perspectives and exchange viewpoints with classmates from varied backgrounds" (Bowen & Rudenstine, 2003, p. 1). Despite the need for more diversity in medical providers and the current make-up of our society, the high stakes and highly traditional and narrow medical school admissions process continues to be hurdles for students from underrepresented racial groups.

White, Dey, and Fantone (2009) investigated the predictive validity of the MCAT with regard to subsequent clinical performance. They collected data from 8 classes of medical students at the University of Michigan between 1995 and 2004 with a sample size of 1,441. Variables of interest included UGPA, MCAT score, medical school exam grades, and Step 1 scores. Initially, it appeared as though there were a significant correlation between MCAT score and clinical clerkship performance among under-represented-students, but this effect disappeared when controlling for other variables such as pre-clinical medical school performance (White et al., 2009). Undergraduate GPA was weakly correlated with clinical performance for majority students (r=0.12), though there was no significant relationship between UGPA and clinical performance among the minority sample (r=0.09). Saguil et al., (2015) collected data from 340 students at the Uniformed Services University's F. Hébert School of Medicine. They found the MCAT was moderately associated with Step 1 score. It was also weakly associated with second year GPA, fourth year GPA, third year clerkship grades, and Step 2 CK score. There were no associations with Step 2 CS scores, and there was no significant association between the MCAT and Postgraduate Year One (PGY-1) of medical residency training evaluations. The authors hypothesized that the MCAT is most likely to be associated with Step 1 scores given the format of the exams. Both exams are multiple choice and emphasize the basic sciences. The lack of association between MCAT scores and clinical examinations, such as Step 2 CS, was suggested to be due to the fact that these clinical encounters are measuring more than knowledge (Saguil et al., 2015).

While not a perfect proxy, URM status is often correlated with SES status, so some of these studies have been included with hopes of extrapolating information regarding SES and medical student success. Jerant et al., (2018) investigated the medical school performance of students considered low SES or URM. The authors collected data from four consecutive class years (2011-2014) at UC Davis School of Medicine with a total of 402 participants. Measurements included undergraduate grade point average (GPA), Medical College Admission Test (MCAT) score, Multiple Mini-Interview (MMI) score, SES, URM status, Step 1 scores, Step 2 CK scores, and number of clerkship honors. Statistical analysis showed that neither URM status nor low SES was associated with Step 1 or Step 2 failure. URM status was not associated with Step 1 or Step 2 CK scores at all, though low SES was negatively associated with Step 1 score. Based on prior research in standardized test scoring, the authors hypothesized that the early negative association between SES and Step 1 scores may have been due to a limited opportunity for test preparation (Jerant et al., 2018). Step 1 scores have been shown to be correlated with MCAT scores, which are also correlated with SES and strongly influenced by test preparation opportunities (Gauer et al., 2016; Grbic et al., 2013). Interestingly, this negative relationship was not found with regard to SES and Step 2 CK scores. The authors were encouraged by this finding, suggesting that students of low SES were able to catch up with their peers and begin performing similarly to their higher SES counterparts on subsequent standardized tests (Jerant et al., 2018). The authors suggested that increased attention toward the racial biases faced by URM students has resulted in a conscious effort to reduce such discrepancies in the clinical setting (Jerant et al., 2018). While this finding may suggest that society is moving in a positive direction, it is impossible to say that implicit racial biases no longer affect URM students during clerkships (Jerant et al., 2018).

Socioeconomic class status. There is little existing research regarding the relationship between socioeconomic status and subsequent medical school performance. More literature exists regarding the relationship between URM status and academic success, although individuals from socioeconomically disadvantaged backgrounds are similarly underrepresented in the physician workforce (Grbic et al., 2013). While not a perfect proxy, URM status is often correlated with SES status. However, two studies that demonstrated the relationship between SES and medical school success are highlighted. Jerant et al., (2018) investigated the medical school performances of 402 UC Davis School of Medicine students considered low SES or URM from four consecutive class years (2011-2014). Measurements included undergraduate grade point average (GPA), Medical College Admission Test (MCAT) score, Multiple Mini-Interview (MMI) score, SES, URM status, Step 1 scores, Step 2 CK scores, and number of clerkship honors. Statistical analysis showed that neither URM status nor low SES were associated with Step 1 or Step 2 failure. URM status was not associated with Step 1 or Step 2 CK scores at all, though low SES was negatively associated with Step 1 score. Based on prior research in standardized test scoring, the authors hypothesized that the early negative association between SES and Step 1 scores may have been due to a limited opportunity for test preparation (Jerant et al., 2018). Step 1 scores have been shown to be correlated with MCAT scores, which are also correlated with SES and strongly influenced by test preparation opportunities (Gauer et al., 2016; Grbic et al., 2013). Interestingly, this negative relationship was not found with regard to SES and Step 2 CK scores. The authors were encouraged by this finding, suggesting that students of low SES were able to catch up with their peers and begin performing similarly to their higher SES counterparts on subsequent standardized tests (Jerant et al., 2018).

Cooter et al. (2004) Step 1 scores have been shown to be correlated with MCAT scores, which are also correlated with SES and strongly influenced by test preparation opportunities. The authors discovered that low SES was negatively associated with Step 1 and Step 2 scores, though the effect size was small (d = 0.30). There was no relationship seen between Step 3 scores, clerkship scores, or postgraduate clinical competence and SES. The authors concluded that SES affects performance in the preclinical years, though low SES students do not exhibit lower levels of clinical skill in the later years.

As the socioeconomic status (SES) disadvantaged indicator was implemented by the AAMC fairly recently in 2014, there appears to be very limited research on its predictive power. In 2015 Grbic et al. validated that the two-factor EO indicator utilized by AAMC was applicable to the national pool of medical students as a valid measure of identifying applicants from socioeconomically disadvantaged backgrounds. There was no available research on the transition to medical school, the medical school experience, challenges or pitfalls along the way, nor was there anything related to residency placement or patterns for those students who were coded as EO1 or EO2 (low socioeconomic status) on their AMCAS application. Further information describing EO1 and EO2 will be provided in Chapter Three.

Prior to the implementation of the SES disadvantaged indicator, there was and still is a Disadvantaged Status Indicator, which many medical schools have used to assess financial and educational limitations. Yet this indicator was and remains a confusing question for students, which therefore has led to inconsistent responses. In a qualitative study by Lowrance (2017), several concerns were noted regarding this indicator. Students needed to determine the meaning of the question without an official definition, as well as understand the perceived risks or benefits by answering this binary question on the application. According to Lowrance (2017):

> An applicant's resources and social capital are, in part, products of family background. The socioeconomic conditions an individual is born into determine, to a great extent, educational and social opportunities, the very factors that affect preparation for medical school and competitiveness in the admissions process. This means that many of the individuals for whom the Disadvantaged Status could benefit the most, are among the least likely [to] feel empowered to make use of the option (p. 277).

Lowrance (2017) goes on to add that almost all of the participants in their study considered the Disadvantaged Status Indicator as 'deficit language' and suggested that the language used could further marginalize vulnerable populations and therefore recommended changing the question to focus more on hardships and/or challenges faced. Based on the above, research utilizing the Disadvantaged Status Indicator would be difficult to quantify.

First-generation college. According to Sacks (2007), research has linked socioeconomic status and parental educational level to college success. This is particularly true of first-generation students, whose parents did not attend college; the evidence suggests that they are at a distinct disadvantage both in accessing a college education and in completing their degree (Lohfink & Paulsen, 2005; Pascarella et al., 2004; Pascarella et al., 2003; Pike & Kuh, 2005; Warburton et al., 2001). According to the American Medical College Application Service (AMCAS) (AAMC, 2020b) first generation is defined as anyone whose parents have not earned an associate degree or higher. According to Elam et al. (2007), first-generation students may not be as familiar with available resources, including technology and/or academic support services such as tutoring and test prep programs which are also often cost prohibitive to these households.

Although there is minimal research on the success of first-generation medical students, the stakes appear to be higher. Parents who have not pursued an undergraduate degree may be truly unaware of the competitiveness, the intense time commitment, and the cost of a medical education. First-generation students may also attend undergraduate programs that are not viewed as competitive enough to prepare them for the rigors of medical school, including community colleges (Elam et al., 2007). Not only may first-generation students interested in pursuing a medical school degree be unaware of these realities, they also may not know how to even go about applying to medical school.

A recent study published by Malau-Aduli et al. (2017) investigated the risk factors associated with academic difficulty in an Australian medical school. They surveyed 1,097 students, collecting demographic and academic information. Of the students surveyed, 26% were the first in their family to attend higher education. First in family (FIF) students were at a much higher risk for attrition.

Seventy-five percent of FIF students completed their education through to graduation, as compared to 83% of non-FIF students. The authors explained these effects by hypothesizing that a social sense of direction within the educational culture is inherited from parents. Children of university-educated parents have learned to develop strategic knowledge skills that benefit them in an academic setting, which their FIF peers may not possess. Without the opportunity to inherit and learn these skills from their parents, it is more difficult for FIF students to find success in academia (Malau-Aduli et al., 2017). This perspective reflects Bourdieu’s (1986) notion of various types of capital—cultural, social, and economic—that are acquired, accumulated, and passed down by privileged groups in society.

Brosnan et al. (2016) collected data from 22 FIF students at an Australian medical school via semi-structured, face-to-face interviews and concluded that FIF students had less access to these forms of capital. In terms of social capital, as they are less likely to be members of networks containing prominent individuals within the health care industry (Brosnan et al., 2016). Additionally, FIF students often feel out of their element after starting medical school, as they have not had as much social experience within the field as some of their more affluent peers (Brosnan et al., 2016). FIF and low SES students also have less access to economic capital (Brosnan et al., 2016; Stegers-Jager et al., 2015). These students frequently commented on the challenges of buying school supplies, eating out at restaurants with classmates, and affording their medical education. These additional challenges are likely an additional source of stress (Brosnan et al., 2016; Stegers-Jager et al., 2015). Many students also worked part-time jobs in addition to studying medicine, which reduced the amount of time they had available for studying. Cultural capital pertains to a sense of belonging within certain groups. FIF and low SES students reported feelings of dissonance between their “old self” and their “new self” (Brosnan et al., 2016). They reported stigmatization of higher education among their home communities, but they did not always feel they could relate with their medical school peers. Students reported feeling as though they did not belong in either world, at odds with both their current classmates and peers back home (Brosnan et al., 2016).

A year later, Southgate et al. (2017), published a qualitative analysis of this same data. The authors made a point to note that FIF and low SES often go hand in hand, and one is often an acceptable proxy for the other. In order to learn more about the experiences of these students within the medical education system, Southgate et al. (2017) further analyzed information from 22 medical students at an Australian medical school via interview. The authors discovered several common themes among FIF students. Many FIF students took years off prior to starting medical school, which they typically used to pursue additional coursework or join the workforce (Southgate et al., 2017). Students agreed that their journey to medical school was often more complex than that of their non-FIF peers. FIF students cited unhelpful academic guidance counselors, parents without experience, and dysfunctional family lives as poignant barriers to matriculation (Biggs et al., 1991; Southgate et al., 2017). Some students reported helpful encouragement from friends and family, while others faced indifference or doubt from their loved ones. Many students reported personal feelings of inadequacy, with frequent self-doubt plaguing their academic endeavors (Southgate et al., 2017). Many wondered if they were smart enough to pursue medicine, and others commented that it was a “big, sanctimonious thing” to be pursuing (p. 250). Many reported that feelings of self-doubt stemmed from a perception that medicine was not something they were entitled to pursue, regardless of prior academic successes. Once they had matriculated to medical school, many reported increasing levels of self-confidence as their education progressed, though imposter syndrome was common. While FIF students did report various unique difficulties, many felt that their backgrounds helped them interact with low SES patient populations. FIF students expressed a desire to remain humble, even after receiving their medical credentials. Several stated that they wanted to hold onto working-class culture, blending their sociocultural identity with their new professional identity (Southgate et al., 2017).

Although there is fairly limited research examining the effects of SES on medical school performance, the findings from the above recent publications have suggested that the three demographic variables of SES, FIF, and URM status may have independent effects on academic success in medical school, but additional research is needed to more completely understand these phenomena.

Gender. Over the past several decades, gender differences in medical school enrollment have begun to close, with women matriculating at rates closer to those of their male peers (Association of American Medical Colleges, 2020c). However, the existing body of literature shows that gender differences exist when it comes to medical student academic performance. Males have been shown to consistently outperform females on preclinical exams, including Step 1 and the MCAT (Cuddy, Swanson, & Clauser, 2008). However, this gender difference is thought to disappear, or even reverse, during the clinical years. Several studies suggest that women equal or outperform men in various clinical measurements, including Step 2 CK, Step 2 CS, and clerkship evaluations (Cuddy et al., 2007; Cuddy et al., 2011; Gauer & Jackson, 2018a). Several authors have offered explanations regarding gender performance in recent years, such as gender-based differences in basic science backgrounds, empathy levels, communication styles, or test anxiety levels, though no clear-cut mechanism has been revealed (Casas et al., 2017; Chapell et al., 2005; Colbert-Getz et al., 2013; Farooqi et al., 2012; Haist et al., 2003; Hall et al., 2009; Hojat et al, 2002a; Hojat et al, 2002; Roter et al., 2002; Yedidia, et al., 2003). Alternatively, some authors address the systemic issues related to females in medicine, the microaggressions that they face, and the privilege that male students have during their educational experience, particularly during their clinical years, specialty consideration, and residency placement (Babaria et al., 2009; Bickel, 2001; Bickel et al., 2002; Nora et al., 2002; Richman et al., 1992; Stratton et al., 2005; Wear et al., 2007) This body of work is particularly relevant when it comes to preparing individual students for the rigors of medical school and helping them develop the skill sets that will set them up for success. Additionally, if early exams are found to consistently under-predict the later achievement of women, admissions officers may need to take this into consideration during the selection process.

Gender differences on Step 1 and 2. Several consistent gender-based trends in Step 1 and Step 2 scores have emerged. A study by Cuddy et al. (2007) with a focus on the Step 2 CK exam collected data, including both academic and demographic variables, from 23,538 students across 136 medical schools. Their analysis showed that women scored higher than men in nearly all content areas of the Step 2 CK exam, with the exception of cardiovascular disorders. The differences increased after controlling for Step 1 scores, with women outscoring men in all disciplines. Step 1 scores were seen to be more strongly associated with subsequent Step 2 CK scores in men, though less so in women. The authors hypothesized that the association between Step 1 and Step 2 CK scores is stronger in men because of their stronger science backgrounds. While women may have a weaker basic science background, and thus lower Step 1 scores on average, they appear to catch up to, and even surpass, their male counterparts through the clinical years of medical school. In their discussion, the authors suggested that gender differences on Step 2 CK performance may be related to areas of interest, arguing that women are more likely practice obstetrics or gynecology than men and thus perform at a higher level in OB/GYN related areas on the Step 2 CK exam (Cuddy et al., 2007).

This pattern of exam performance was also reflected in a study conducted by Gauer and Jackson (2018a). Data was collected from 1,067 medical students at the University of Minnesota Medical School. The authors took note of undergraduate major, age, gender, MCAT score, Step 1, and Step 2 CK scores. As expected, male students scored significantly higher than females on both the Step 1 exam and the MCAT. After controlling for MCAT score, male students still outscored their female colleagues on the Step 1 exam, again suggesting that factors other than different basic science backgrounds may be influencing the gender performance disparity, though the authors did not speculate further. Gender was not associated with Step 2 CK performance, and the authors believed this signified that female students had effectively caught up to their male peers (Gauer & Jackson, 2018a).

In a recent study, Rubright et al. (2019) collected a large sample of data from 45,154 medical students enrolled across 172 medical schools in the United States and Canada. Each student had completed all three Step exams by the time data was collected, and a retrospective review was performed. The authors examined MCAT scores, undergraduate GPAs, Step 1, Step 2 CK, and Step 3 scores, as well as gender, race, English as a second language (ESL), and age. In line with the current literature, the authors found that females scored significantly lower on the Step 1 exam as compared to men. This trend was reversed for Step 2 CK, with women significantly outscoring men (Rubright et al., 2019). When adding MCAT and undergraduate GPA covariates to the model, the authors found that gender differences persisted, though other demographic effects such as race, age, and English as a second language were reduced or erased, suggesting that differences in gender performance could not be completely explained by differences in GPA or MCAT score (Rubright et al., 2019). Finally, in several studies, women have significantly outperformed male counterparts in clinical assessments (Cuddy et al. 2011; Swygert et al., (2012); van Zanten et al., 2007).

In a qualitative study by Babaria et al. (2009), the authors investigated how female medical students perceived their gender role during their clinical years of their medical education. Their findings included: female students struggling to define their roles and finding themselves participating in stereotypical gender roles; female students' workplace relationships tending to be more with nurses while male students developed relationships with attending physicians; female students being less likely to successfully navigate uncomfortable interactions with their supervisors; and female students encountering a gender learning curve while on the floors that shaped their view of themselves as future female physicians (Babaria et al., 2009).

Much research surrounding female medical students has been focused on the topics of mistreatment and sexual harassment (Nora et al., 2002; Richman et al., 1992; Stratton et al., 2005; Wear et al., 2007), particularly during the clinical years of their education. In a qualitative study conducted by Stratton et al. (2005), the authors demonstrated a sizeable number of female students reported that their specialty choice and ranking of residency programs was influenced by gender discrimination and sexual harassment. Further, Bickel et al. (2002), contended that there is a significant deficit of females in positions of leadership within academic medicine, which in turn has implications and limits the potential of female role models for female medical students, particularly in some specialties such as surgery (Bickel, 2001). Despite initiatives by groups such as Women in Medicine, these issues continue to impact the culture, educational environment, and the experiences of female medical students (Bickel et al., 2002).

Non-traditional, adult students. The average age at the time of matriculation for medical students in the United States is 24 years old, suggesting that many students take time off to work or pursue other endeavors before starting medical school. Some students take a significant amount of time to pursue a different career, a different degree, or some other option altogether before starting their medical education. Others apply through several cycles before receiving an acceptance. This study assessed age as a predictor of academic success in medical school. On one hand, these students often start their studies with more life experience and a more mature set of skills. However, many older students report a high level of unique non-academic obstacles, including commitments to family, financial distress, or a sense of isolation from their younger peers (Imlach et al., 2017). Several studies have attempted to shed some light on this complex relationship.

Ogunyemi and Taylor-Harris (2005) examined the interaction between age and Step 2 scores specifically. The author collected data from 171 students at Charles Drew University in an attempt to determine the factors that affect Step 2 CK exam scores. Ogunyemi and Taylor-Harris explored age, undergraduate GPA, MCAT scores, OB/GYN NBME exam scores, Step 1 scores, and faculty evaluations as possible predictors of Step 2 success. The author found that older students negatively correlated with Step 2 exam scores, but did not find a significant relationship between age and Step 2 exam failure (Ogunyemi & Taylor-Harris, 2005). Older students tended to score lower on the Step 2 exam than their younger peers, though they did not fail Step 2 at a higher rate (Ogunyemi & Taylor-Harris, 2005).

Additionally, Gauer and Jackson (2018a) examined the relationship between examinee age and performance on various standardized tests, including the MCAT, Step 1, and Step 2 CK exams. The authors collected data from 1,062 medical students at the University of Minnesota between 2007 and 2011. Significant differences were found between the Step 1 and Step 2 CK exam scores of traditional and non-traditional aged students, which were designated in this study as age 25 and older, with traditional-aged students outscoring non-traditional students (Gauer & Jackson, 2018a). The authors hypothesized that the emergence of an age-related performance disparity during these later exams may be due to external life events. They argued that older students are more likely to have commitments outside of medical school, including dependents or ill parents, and these commitments may reduce the amount of time they are able to spend on exam preparation (Gauer & Jackson, 2018a). Evaluating the impact of specific external life events on exam scores may shed more light on the observed interaction between age and academic success, though no studies have been identified that further clarify this relationship.

Kick et al. (2000) conducted a qualitative analysis of focus group interviews in order to gain a better understanding of the challenges faced by older learners, defined as those who would be 30 or older at the time of graduation. The authors found that older students had different learning strategies, had concerns about managing their home responsibilities, and felt that there was a lack of respect, based on their age, by the residents and attending physicians. Further, Kick et al. (2000) learned that older medical students reported higher levels of initiative and autonomy. They stated they felt more confident in their clinical setting than their younger peers, which they attributed to their prior work experience. Students with prior work experience tended to describe patient encounters as a learning experience, with less concern for perfect interactions. Older students reported better relations with professionals of their own age, even with traditional hierarchal boundaries in place (Chur-Hansen, 2003; Jurjus et al., 2017; Kick et al., 2000). Of these studies, none have really teased out the potential confounds that may truly exist for older students, such as additional time commitments outside of school, unique stressors, or specific challenges.

The differential access and success patterns of traditional White male students from those of non-traditional groups (URM, low SES, first-generation, women, and adult students) highlight the need to investigate the variables that have been used in admissions and their predictive power in medical school success.

Pre-Medical School Admission Variables as Predictors of Medical School Success.

This section will review studies that have examined cognitive and non-cognitive predicators of medical school success on the formal assessments such as Step 1 and Step 2 Clinical Knowledge. These predictors include the traditional cognitive measures such as undergraduate GPA (UGPA) and the Medical School Admissions Test (MCAT) and non-cognitive measures that capture qualities reflected in the AAMC's Experiences, Metrics, and Attributes (EAM) Model (Bills et al., 2019; Conrad et al., 2016; Milem et al., 2012).

Cognitive Assessment Measures: Undergraduate GPA (UGPA) and Medical College Admissions Test (MCAT). Traditionally, the medical school admissions process has focused heavily on pre-medical academic factors, such as students' undergraduate GPA (UGPA) and Medical College Admissions Test (MCAT) score. These cognitive factors are often combined with an assessment of the applicants' non-cognitive traits, which are evaluated through essays, letters of recommendation, and the interview process (Monroe, et al., 2013). Theoretically, this process selects students who possess the cognitive capacity to withstand the rigors of medical school, as well as the personality traits desired in a good physician. While several studies exist that support the use of UGPA and MCAT as primary predictors of medical school preparedness, there is large variation in the predictive validity of these factors. Typically, most authors have utilized pre-clinical variables, such as Step 1 exam scores or early medical school coursework, as indicators of medical school success (Basco et al., 2002; Donnon, et al., 2007; Gilbert et al., 2002; Julian, 2005). Relatively little work has gone into investigating the relationship between pre-admissions variables and subsequent clinical skills and/or Step 2 scores (Casey, et al., 2016; Ogunyemi & Taylor-Harris, 2005; Saguil, et al., 2015). There is also debate surrounding the validity of these pre-admission variables with regard to certain subgroups of students. Several studies have shown no association between MCAT score or

UGPA and medical school success among students who have completed a post-baccalaureate degree or other post-graduate work before starting medical school (Association of American Medical Colleges, 2020f; Giordani et al., 2001; Johnson et al., 2017; Kulesza et al., 2015; Sadik et al., 2017).

A study by Basco et al. (2002) investigated the relationship between institutional MCAT score and Step 1 performance. Institutional MCAT scores were determined by averaging the MCAT performance for all students enrolled at the institution over one year. The authors were interested in whether or not taking this additional variable into account would increase the predictive value of pre-admissions variables (Basco et al., 2002). They collected data from medical students matriculating to two different medical schools between 1996 and 1998, with a total sample of 16,954 applicants and 933 matriculants. Variables of interest included student demographics, institutional MCAT score, individual MCAT score, UGPA, and Step 1 score. Data analysis revealed a moderate correlation between individual science UGPA and Step 1 scores. There were also significant correlations between individual MCAT scores and Step 1 scores, including the verbal reasoning sub-score (Basco et al., 2002). The authors were surprised that the verbal reasoning score was a significant predictor of Step 1 score, as prior studies had found it to be unrelated (Basco et al., 2002). Using these data, the verbal reasoning score was related to the Step 1 score with $r=0.397$. The MCAT biological sciences score was most strongly correlated with Step 1 performance ($r=0.553$). Using institutional scores for analysis yielded similar results. In all, the individual MCAT scores were slightly more predictive of Step 1 scores than the institutional MCAT scores, though adding institutional MCAT scores to a model containing individual MCAT scores improved predictive power to a slight degree (Basco et al., 2002).

A study published by Gilbert et al. (2002) explored the relationship between MCAT score and USMLE performance. They also determined that UGPA, MCAT physical sciences and MCAT biological sciences scores were significantly predictive of Step 1 exam scores. MCAT biological sciences ($\beta=0.28$) and UGPA ($\beta=0.28$) were the strongest predictors of Step 1 performance (Gilbert et al., 2002). There was no correlation between Step 1 score and MCAT verbal reasoning or writing scores. With regard to Step 2, there was a significant association between exam performance and UGPA ($\beta=0.31$). There was no association between Step 2 performance and MCAT scores for any subcategories (Gilbert et al., 2002).

Later, in a study published by Ogunyemi and Taylor-Harris (2005), the relationship between preadmissions variables and Step 2 CK exam scores was investigated. The author was particularly interested in whether or not demographic factors played a role in Step 2 performance. Ogunyemi and Taylor-Harris collected data from students enrolled at Charles Drew University College of Medicine, which has historically matriculated a diverse student population. The author collected USMLE scores, MCAT scores, UGPA, and demographic data from 171 medical students rotating through the OB/GYN clerkship between 1992-2001. Among other findings, the author discovered a significant association between USMLE Step 2 exam score and UGPA ($r=0.287$), as well as between Step 2 score and MCAT score ($r=0.524$). Multiple MCAT attempts were associated with a failing Step 2 CK score, though UGPA did not independently predict failing Step 2 CK scores (Ogunyemi & Taylor-Harris, 2005). The author did not find any statistical association between race or gender and Step 2 CK score.

A study published by Julian (2005) reported a wide range of predictive power of UGPA and MCAT scores and medical school performance. Julian (2005) collected USMLE scores from more than 31,000 students across 125 medical schools and additional data from students entering 14 different medical schools between 1992 and 1993. The data collected from this cohort included preadmission MCAT and UGPA, medical school GPA, medical school academic difficulty, academic distinction, and Step exam scores. The highest corrected validity coefficient using MCAT/UGPA as predictors of medical school GPA was 0.81, which described 66% of the variance in the medical school GPA between students. The lowest corrected validity coefficient was found to be 0.53, which explained only 28% of the variance in the medical school GPA. The median corrected validity coefficient was 0.71 and accounted for 50% of the variance in the medical school GPA. When using MCAT in addition to UGPA, there was a 21% increase in variance, which suggested that models using both of the metrics are more powerful than a model using only UGPA or MCAT scores (Julian, 2005). In contrast, the author discovered that the MCAT alone may be sufficient in predicting USMLE performance (Julian, 2005). For Step 1, there was a validity coefficient of 0.49 when using UGPA alone, a coefficient of 0.70 when using MCAT alone, and a coefficient of 0.72 when using both. For Step 2, there was a validity coefficient of 0.44 for UGPA alone, a coefficient of 0.60 when using MCAT alone, and a coefficient of 0.63 when using both. These findings suggest that the MCAT alone is a fairly good predictor of USMLE Step 1 and Step 2 CK scores, and not much power is added by taking UGPA into consideration. The author also noticed general trends between academic difficulty, academic distinction, and MCAT score (Julian, 2005). As MCAT scores increased, demonstrated academic difficulty decreased. Similarly, as MCAT scores increased, a greater number of students received academic distinction. However, the author made a point to note that these trends are not absolute. Some students with high MCAT scores did go on to experience academic difficulty, just as some students with low MCAT scores received academic distinction. The author hypothesized that other student characteristics not measured by the MCAT may influence these outcomes (Julian, 2005). Given these findings, the author concluded that the predictive power of the UGPA with regards to medical school performance is more limited than that of the MCAT. They suggested that institutional differences in undergraduate grading may affect the usefulness of the UGPA, in contrast to the MCAT which is graded based on a national standard (Julian, 2005). Like many other authors, Julian also noted the similarity between the MCAT exam format and the USMLE exam format, which may explain some of the correlation between the two scores. The author further asserted that success on the USMLE is an important milestone in medical education, and thus the MCAT is a useful metric that should be preserved when evaluating incoming students (Julian, 2005).

In 2007, Donnon et al. published a meta-analysis of 23 studies which investigated the predictive validity of the MCAT with regard to medical licensing examinations. They selected studies published after 1991 that investigated the relationship between MCAT score and academic success at medical schools across the United States. The majority of studies reported on pre-clinical measures alone, including medical school GPA and Step exam scores. Similar to Julian's (2005) study, the authors found the biological sciences sub-score to be the most strongly associated with preclinical success, followed by the physical sciences sub-score and the verbal reasoning sub-score, in that order (Donnon et al., 2007). The pre-2015 MCAT was broken down into 3 sections, allowing for three sub-scores in Biological Sciences, Physical Sciences, and Verbal Reasoning. The totality of these numeric sub-scores provided the overall MCAT score for students. Prior to 2014, the MCAT also required a writing sample, which was scored on an alphabetic scale from "J" to "T". All studies used in the meta-analysis found no relationship between writing sample grade and USMLE Step scores. When investigating the predictive validity for each individual Step exam score, Donnon et al. (2007) found that the total MCAT had the largest predictive validity for Step 1 (r=0.66, explained 43.6% of variance). It had a medium validity for both Step 2 (r=0.43, explained 18.5% of variance) and Step 3 (r=0.48, explained 23% of variance).

Dunleavy et al. (2013) found similar results as the Donnon et al. (2007) meta-analysis about the impact of various cognitive preadmissions variables and medical students' unimpeded progress towards graduation. Unimpeded progress was defined as graduation within five years of matriculation with passing grades on first attempts at Step 1, Step 2 CK, and Step 2 CS. The authors collected data from students at 128 United States medical schools between 2001 and 2004. Preadmission variables of interest included UGPA and MCAT score. The authors created three predictive models: model 1 only considered UGPA, model 2 only considered MCAT score, and model 3 considered both. They then compared the predictive validity of each model. When considering UGPA, the authors discovered there was a positive association with UGPA and unimpeded progress until a GPA threshold of 3.50. After this marker, the likelihood of a student experiencing unimpeded progress levels off. When UGPA was low, there was more variability in the likelihood of a student experiencing unimpeded progress as compared to when UGPA was high. When considering MCAT scores, a similar trend emerged. There was a strong association between the likelihood of a student experiencing unimpeded progress and MCAT score, which increased until an MCAT score of 30 (80th percentile). After this point, the relationship leveled off. Again, there was more variance in the likelihood of a student experiencing unimpeded progress when the MCAT scores were low as compared to when MCAT scores were high. The authors also noted that the interquartile ranges were much smaller for the MCAT than for UGPA, indicating that the relationship between MCAT scores and unimpeded progress is more similar across schools when compared to the relationship between UGPA and unimpeded progress (Dunleavy et al., 2013). When combining the models, the authors discovered that the relationship between MCAT scores and unimpeded progress depends on UGPA (Dunleavy et al., 2013). For example, among students with low MCAT scores, the likelihood that they will experience unimpeded progress increases with higher UGPAs. This effect was much stronger among students with low MCAT scores as compared to students with high MCAT scores. The authors concluded that a combination of MCAT and UGPA is the most reliable predictor of unimpeded progress among medical students (Dunleavy et al., 2013). The authors concluded that MCAT scores are more powerful stand-alone predictors of unimpeded progress than UGPA (Dunleavy et al., 2013). They suggested that Step exams are more similar to the MCAT than the metrics used to determine UGPA, which explains the stronger relationship between Step scores and MCAT scores. They also noted that UGPAs are influenced by other factors outside of academic knowledge and skill, such as institutional variability or individual study habits (Dunleavy et al., 2013).

A team at Brown University developed a model which incorporated preclinical exam performance, USMLE Step 1 scores, and NBME subject scores to predict Step 2 Clinical Knowledge (CK) performance; ranges predicted by this model were accurate for 68% of students. This model was used to identify the need for educational intervention and to provide students with support services. Given this impact of Step 1 on Step 2 CK and the correlation of MCAT scores with Step 1 performance, the MCAT may have predictive value on Step 2 CK performance (Monteiro et al., 2017). The use of the MCAT as a predictor of USMLE Step 1 and Step 2 CK performance is rooted in the classification of these examinations as standardized tests. In a study conducted by Ogunyemi and Taylor-Harris (2005), the researchers reported that students who experienced difficulty on the MCAT were more likely to experience difficulty on the USMLE examinations. Further investigation is warranted to assess if students' performances on these exams are impacted by students' ability to take standardized tests.

More recently, Gauer et al. (2016) conducted a 5-year retrospective study on the predictive power of the MCAT. They were interested in determining whether the MCAT was a reliable predictor of Step 1 and Step 2 CK exam scores. They collected data from 1,065 students at the University of Minnesota Medical School between 2011 and 2015. MCAT scores were broken down into verbal reasoning (VR), biological science (BS), and physical science (PS) components. Multiple linear regression analysis revealed significant associations between the MCAT and Step scores. The MCAT explained 17.7% of the variance in Step 1 scores. The authors discovered that the Step 1 score increased by 3.548 points for every point on MCAT BS, 2.215 points for every point on MCAT PS, and 0.748 points for every point on MCAT VR (Gauer et al., 2016). Similar to prior studies, the authors suggested that the MCAT BS is most similar to the Step 1 exam, which strengthens the predictive power of the MCAT BS for most students. Using multiple linear regression analysis, the authors also discovered that the MCAT explained 12.0% of the variance in Step 2 CK scores. For each point on the MCAT BS, Step 2 CK score increased by 2.819. For each point on the MCAT PS and VR components, Step 2 CK score increased by 0.822 and 1.237 points respectively. In their discussion, they concluded that the MCAT was shown to be weakly to moderately associated with USMLE performance (Gauer et al., 2016). However, the authors noted that the predictive power of this exam is not very strong, and thus the MCAT should not be the only factor taken into consideration when evaluating students for admission (Gauer et al., 2016). Further, they questioned if the MCAT and USMLE scores were perhaps more related to test-taking abilities (Gauer et al., 2016). They suggested that a low MCAT score should not keep a student from pursuing medicine but could be used by institutions to flag a student needing more work with test-taking or study strategies (Gauer et al., 2016).

In the most recent studies discussed above, Dunleavy et al. (2013) and Gauer et al. (2016) suggest that the science sections of the MCAT, particularly the biological sciences section, are the greatest predictors of success on Step 1 and Step 2 CK, while the UGPA is moderately predictive of success on Step 1 and 2 CK.

Medical schools have a vested interest in ensuring that the MCAT is a valid measure of an applicant's ability to succeed in medical school and as a physician. To assess possible bias against Latino and African-American students, Davis et al. (2013) accessed data from the AAMC on MCAT examinees, applicants to US medical schools, and matriculants to US medical schools. MCAT scores for Black and Latino students were lower than scores for White students, which is consistent with the results of other graduate school admissions tests, including the GRE, GMAT, and LSAT (Davis et al., 2013). It has also been shown that women have historically performed worse than men on standardized exams (Sacks, 2007; Sedlacek, 1977; Sedlacek, 2004).

White et al. (2009) investigated the predictive validity of the MCAT with regard to subsequent clinical performance and discovered that MCAT scores are strong predictors of Step 1 scores among both minority (Black, Hispanic, and Native American) and majority (White or Asian) students. However, MCAT scores only correlated with first year medical school performance among majority students. There was no relationship between MCAT score and first year medical school performance among minority students (White et al., 2009). Instead, undergraduate GPA was a significant predictor of first year medical school performance among minority students.

A study by Saguil et al., (2015) also attempted to discern the relationship between pre-medical variables and subsequent clinical skills and found that MCAT scores were moderately associated with Step 1 score. They were also weakly associated with second year GPA, fourth year GPA, third year clerkship grades, and Step 2 CK scores. There were no associations with Step 2 CS or CIS scores. The lack of association between MCAT scores and clinical examinations, such as Step 2 CS, was suggested to be due to the fact that these clinical encounters are measuring more than knowledge (Saguil et al., 2015). Saguil et al. (2015) concluded that the predictive power of the MCAT is good with regard to pre-clinical examinations, though it is not sufficient when predicting clinical performance later in the medical school career.

Given the current body of literature, the MCAT appears to be a fairly reliable predictor of pre-clinical performance among traditional medical students. While the UGPA does not offer as much productive validity, several authors have concluded that the most accurate predictive models take both MCAT and UGPA into account.

As reviewed above, medical school admissions committees have long relied on the Medical College Admission Test (MCAT) scores and the undergraduate grade point average (UGPA) to select students for admission (Dunleavy et al., 2011). Such heavy weight on the MCAT has long disadvantaged students of color (Witzburg & Sondheimer, 2013), thereby introducing systemic challenges to efforts to increase the pipeline of these under-represented groups. Previous studies have suggested that continued reliance on traditional admission measures will lead to ongoing stifling of diversity within the medical profession (Prideaux et al., 2011; Reiter et al., 2012). Research conducted by Cantwell et al. (2010) assessed the congruency of using holistic review to promote diversity at a US medical school. According to Cantwell et al. (2010), medical education remains highly stratified by race and ethnicity, and "Higher education admission policies are ripe for this type of analysis given the seeming disconnect between what colleges report to value and the types of students they admit" (p. 29).

For example, White et al. (2009) collected data from eight classes of medical students at the University of Michigan between 1995 and 2004 with a sample size of 1,441. The authors discovered there was no relationship between MCAT score and first year medical school performance among minority students (White et al., 2009). Instead, undergraduate GPA was a significant predictor of first year success.

Several researchers have demonstrated that cultural test bias, socioeconomic privilege that comes with test prep assistance, and stereotype threat are all attributed to lower standardized test scores which disadvantage particular groups of students (Freedle, 2003; Hopkins, 2008; Kellow & Jones, 2008; Marrah, 2012; Rosner, 2001; Sacks, 2007; Solórzano, 2008, Steele, 1997; 1999). Despite significant research surrounding this topic, standardized test scores, (i.e., SAT and GRE) of under-represented students continue to be weak (College Board, 2016; Educational Testing Services, 2008; National Center for Fair and Open Testing, 2012; Kuncel & Hezlett, 2007; Pieterse, 2007; Ramsey, 1997; Roser, 1998; Weiss, 2001). A study by Akos and Kretchmar (2017) demonstrated that high school grades and standardized tests explained only 25% of student success in undergrad, leaving most of the variance of their success unexplained. Additionally, the large differences in standardized test scores between under-represented students and their non-minority peers led the researchers to acknowledge that using only grades and standardized test scores alone in the admission process would have detrimental effects on diversity and would overlook many ‘capable’ students (Akos & Kretchmar, 2017). If medical schools, and colleges in general, continue to use standardized tests as a requirement for admission and do not consider other non-cognitive factors in their assessment of students, these groups of students will continue to be underrepresented in the field of medicine.

According to a study by Davis et al. (2013), test bias was defined as “when deficiencies in a test itself or the manner in which it is used result in different meanings for scores earned by members of different identifiable subgroups” (p. 595). The authors contend that the MCAT exam is extensively reviewed for bias and irrelevant information that may be biased is removed from the exam. The administration of the MCAT exam, including testing instructions and testing time, is standardized (with the exception of accommodating students with disabilities), which minimizes the possibility of bias in administration procedure. The researchers demonstrated that URM test takers had lower MCAT scores than their peers; these differences were very similar to those on other standardized entrance examinations, including the GRE, LSAT, and SAT (Davis et al., 2013). The authors found no evidence that the MCAT exam was biased against these groups, yet they did find that parental education, family income level, growing up in a rural/urban/suburban neighborhood, and access to resources and educational materials could contribute to the reported differences in scores among racial and ethnic groups (Davis et al., 2013). Additionally, the parents of white medical students are more likely to have higher education levels than parents of URM students. URM students were less likely than white students to have at least one parent with a college or graduate degree, and were more likely to have been raised in lower-income households or single parent households. All of these factors have been shown to influence access to educational and occupational opportunities, which in turn may influence how one performs on the MCAT.

Although the authors above (Davis et al., 2013) have suggested that the MCAT examination is not biased against under-represented test takers, they demonstrate an influence on test scores due to parental education, family income level, and access to educational resources and opportunities. Therefore, continuing to use the MCAT as a factor in the admissions process may not be the best predictor of their medical school success. This is particularly true for under-represented students, first-generation students, and students from lower socioeconomic backgrounds.

Non-Cognitive Assessment Measures: Traditional and MMI Interviews. Non-cognitive factors can be assessed during an interview, as the interview is often used to identify those applicants who will struggle with the patient contact or teamwork part of the profession (Lemay et al., 2007).

Traditional interviews have been highly criticized over the years (Morris, 1999) as being too subjective and laden with personal and societal bias on the part of interviewers (Salvatori, 2001). According to Razack et al. (2009), performance on unstructured interviews is largely determined by context or chance, based on the questions that are asked, the commonalities among the interviewers and interviewees, and who was randomly assigned the “easy” versus “hard” interviewers. Interviewers using an unstructured interview process, where interviewers are free to ask their own, potentially random, questions can be greatly influenced by their own bias and/or the perceived positive or negative characteristics of the applicant (Eva at al., 2009; Kreiter et al., 2004). Failing to control for biases, not using standard questions, and using a non-structured interview process can lead to unreliable interview outcomes, influencing the overall impression and/or rating of an applicant (Rees et al., 2016). Morris (1999) concurred that the fairness of the non-structured interview is questionable. He pushed for carefully designed, highly structured interviews to improve the reliability and validity of the interview process and to minimize the subjectivity and social bias that often occurs in non-structured interviews. He further stressed that non-cognitive variables, such as perseverance, empathy, goal setting, and integrity be assessed during the medical school interview as they are the strongest predictors of clinical performance (Morris, 1999).

In 2009, Eva et al. proposed that using a structured interview in the admissions process, where interviewers ask the same questions of all candidates, proves beneficial in assessing non-cognitive skills such as communication, empathy, integrity, and compassion, all of which are necessary to provide patient-centered health care (Monroe et al., 2013). Due to concerns about the personal and biased nature of the traditional interview, McMaster University’s Michael G. DeGroote School of Medicine created a highly structured interview process, known as the multiple mini-interview (MMI) in 2002. The MMI has been deemed more psychometrically sound than traditional interviews as the traditional interview is often rife with context specificity and chance (Eva et al., 2004). Increasingly, medical schools and other health professions schools are moving away from traditional interviews and are offering the MMI approach to select future health care providers (Axelson et al., 2010). The MMI design allows for multiple perspectives regarding an applicant's candidacy and also allows admissions committees to further evaluate an applicant's non-cognitive qualities that will help to make them compassionate, empathetic, and patient-centered physicians. According to Razack et al. (2009), in a survey comparing the traditional medical school interview to the MMI, the MMI rated higher on perceived fairness and effectiveness as an interview tool by both applicants and interviewers. Both concluded that applicants were better able to showcase their individual strengths and non-academic aptitude for pursuing a career in medicine. According to Lemay et al. (2007), the MMI is a valid and reliable tool used to assess various non-cognitive attributes. A review of the research has demonstrated that the Multiple Mini-Interview process was a predictor of success on clinical performance and evaluations (Lemay et al., 2007; Reiter, Eva, Rosenfeld, & Norman, 2007; Roberts et al., 2009). Reiter et al. (2007) further argued that the MMI was able to predict clerkship (clinical education) performance during the third year of medical school, performance on the medical school Objective Structured Clinical Examination (OSCE) (which mirrors Step 2 Clinical Skills but provides the student with feedback) that many students take following their clerkship year (often their third year), and Canadian licensing examination performance.

However, Morris (1999) cautioned that even when using a structured interview approach, "The issue of subjectivity is strongly associated with criticisms based on the argument that the personal bias of interviewers may be a stronger influence on interview scores and decisions than the actual characteristics which are meant to be evaluated" (p. 474). Therefore, measuring personal attributes by *only* assessing them through the medical school interview may continue to be a highly subjective and very time-consuming process (Ehrenfeld & Tabak, 2000; Kulasegaram, Reiter, Wiesner, Hackett, & Norman, 2010). Having the ability to assess particular qualities utilizing standardized assessments that aid in the prediction of success for medical students during their clinical years, on Step 2, and later as a practicing physician may prove less costly and more reliable and valid than the medical school interview.

Post-baccalaureate coursework and master's degrees. In addition to cognitive and non-cognitive assessments used to assess and predict medical school admission and success, research has also investigated the impact of advanced education in the sciences. Over the past several years, the popularity of pre-medical school coursework has increased, with students participating in post-baccalaureate (PB) programs or obtaining their master's degrees prior to matriculation. Traditionally, these programs were started to create medical school pipelines for minority students in order to diversify the physician workforce (Agrawal, Vlaicu, Carrasquillo, 2005; Carnevale & Strohl, 2013; Ntiri, 2001). Historically, these students have faced a greater number of barriers while pursuing higher education than their non-minority peers and tended to have lower standardized test scores, more numerous undergraduate hardships, and access to fewer extracurricular opportunities than their non-minority peers. All of these factors tend to make successfully applying to medical school more difficult. In recent years, "academic enhancer programs" and "career changer" post-baccalaureate programs have begun to spring up across the United States (Andriole & Jeffe, 2012) and have changed the demographic of those pursuing these academic programs to prepare for medical school (Baum & Steele, 2017; Mullen et al., 2003). Based on these changes, there has been an increasing interest in investigating the efficacy of post-baccalaureate (Manusov et al., 2011) and master's programs on medical student success.

Reeves et al. (2008) investigated the efficacy of a post-baccalaureate program at the University of North Texas Health Science Center (UNTHSC) which selects students who have been unsuccessful in matriculating to medical school at least once before. The authors collected data from students enrolled in the program between 2000 and 2006. When comparing the GPAs from the first year of medical school, PB students outperformed their traditional peers by a small degree (Reeves et al., 2008). The PB students also outperformed their traditional peers on the COMLEX I exam, which is the osteopathic equivalent of the allopathic USMLE Step 1 exam, given at the end of the second year (542.3 vs 529.3). The authors concluded that the post-baccalaureate program at UNTHSC successfully prepared students for the rigors of medical school, despite apparent deficits in their initial applications (Reeves et al., 2008).

A more recent study published by Kulesza et al. (2015) evaluated the effectiveness of a post-baccalaureate program offered at the Lake Erie College of Osteopathic Medicine (LECOM). The authors collected undergraduate GPAs, MCAT scores, first year medical school GPAs (MS1), second year medical school GPAs (MS2), and COMLEX I exam scores from participants. Quantitative and qualitative data were collected from students graduating between 2013 and 2015. The two groups of students performed similarly on the COMLEX Level I exam, with non-PB students averaging a score of 530. Post-baccalaureate students had an average score of 508.9, though this difference was not statistically significant (Kulesza et al., 2015). Post-baccalaureate students also had a 100% first-time pass rate on the COMLEX I and a zero percent withdrawal rate within the first two years. Meanwhile, non-PB students had a lower first-time pass rates at 96.2% and 4.5% of the group withdrew or were dismissed during this time (Kulesza et al., 2015). The authors noted that pre-medical indicators, such as undergraduate GPA or MCAT scores, had very little correlation with PB students' COMLEX I score. However, medical school GPA during the first year was much more predictive of subsequent COMLEX I score. One hundred percent of PB students felt as though they were prepared for the evaluation methods used during medical school, while only 37% of traditional students felt as though they were prepared for the board-style questions typically used (Kulesza et al., 2015). Finally, PB students believed their programs helped them develop the time management, stress-coping, and study skills required to be successful in medical school. In general, the authors concluded that the LECOM post-baccalaureate program is effective in preparing students for medical school (Kulesza et al., 2015).

In reviewing a similar post-baccalaureate program, Epps (2015) saw slightly different results. In order to assess program effectiveness, they conducted an 11-year study from 2001 to 2011 at the Meharry Medical College School of Medicine, which implemented a post-baccalaureate program for students from disadvantaged backgrounds. The program is slightly longer than a year and takes place with three distinct phases. The initial phase offers an opportunity for students to review for the MCAT in a structured way. The second phase involves activities throughout the academic year, with instruction in biological, chemical, and physical sciences. The third phase allows students to become familiar with the medical school's first year curriculum, with the intention of smoothing their transition into the medical school class. Truncated courses in anatomy, biochemistry, and other relevant subject areas were offered. The sample consisted of 764 control students and 200 PB students. There was a significant difference between the control and intervention groups when analyzing the pass rates on the first attempt at the USMLE Step 1 exam (82% vs 74%). According to Epps (2015), this difference held true when analyzing data regarding pass rate on the first attempt of the USMLE Step 2 exam, with the control group passing at a much higher rate (90% vs 76%). There was no significant difference in the number of years required to graduate medical school between the control and PB groups. The difference in the performance of the two groups demonstrated that despite early intervention and academic enrichment the post-baccalaureate students, the student enrolled in the program still performed lower on their USMLE examinations.

Some of these post baccalaureate programs are targeted to increase the pipeline of students from under-represented groups. For example, Goode and Talbot (2016) investigated the efficacy of the Master's in Science in Medical Sciences (MSMS) offered at Western University of Health Sciences. The program strives to aid students in gaining admission to a post-graduate school programs in the health professions, with the ultimate goal of increasing the number of students who work in underserved areas. Specifically, the program aims to increase the number of URM students enrolled at the College of Osteopathic Medicine of the Pacific (COMP), which is housed at Western University. Less emphasis is placed on academic scores, as these students tend to apply with lower MCAT and undergraduate GPAs. When evaluating data from 2011-2017, the completion rate was 96%, with only 7 out of 146 students not completing the master's program. Of these graduates, more than 90% went on to a program in the health professions. Of the students interested in medical school, more than 95% of them went on to matriculate at an accredited medical school. The authors examined the preparedness of MSMS alumni matriculating to medical school by collecting incoming academic metrics (MCAT, undergraduate GPA) and medical school class rank. Similar to prior studies, there was no significant correlation between MCAT score and medical class rank for MSMS students. There was a significant relationship between MSMS program GPA and medical school class rank (r=0.45). In addition to validating the efficacy of the MSMS program, the authors reiterated that pre-medical admissions data is not always predictive of medical school performance for disadvantaged students (Goode & Talbot, 2016). Given the data collected from this study, they suggested that post-graduate coursework is a much more powerful predictor of medical school performance for these students (Goode & Talbot, 2016).

Similar results were obtained in a study by Sadik et al. (2017) that compared students who completed the master's coursework (N-32) with those who did not (N-558) at the Touro University Nevada College of Osteopathic Medicine (TUN). The researchers determined there was a strong correlation between master's GPA and COMLEX I score ($r = 0.49$) as well as COMLEX II scores ($r=0.5$). There was no correlation between MCAT scores and COMLEX I or II performance of master's students. Among the control group, there were very weak positive correlations between MCAT score and COMLEX I and II scores. Among both groups, the authors determined that COMLEX I and II scores were very strongly correlated with medical school GPA during the first and second years. Given this data, the authors concluded that master's or post-baccalaureate performance is more predictive of subsequent academic performance than traditional pre-admissions variables among students who have completed these types of programs (Sadik et al., 2017). They suggested that students with lower pre-medical scores (MCAT and undergraduate GPA) are still very capable of succeeding in medical school, and this probability of success is closely linked to their performance in their pre-medical, post-graduate studies (Sadik et al., 2017).

Giordani et al. (2001) explored the effectiveness of formal post-baccalaureate programs for underrepresented minority students prior to medical school entry. The authors collected data from students enrolled at the University of Michigan Medical School between 1993 and 1996. The authors noted that students who completed PB work tended to be older than traditional students and were more likely to be URM. When analyzing pre-matriculation data, Giordani et al. (2001) found there were significant differences between traditional and PB students. Traditional students had significantly higher undergraduate GPAs, though there was no significant difference in MCAT scores. However, the authors noted that non-University of Michigan PB students tended to score higher on the MCAT than those who completed their post-baccalaureate work at the University of Michigan (Giordani et al., 2001). There were no significant differences found between groups when first year medical school grades were analyzed. When comparing pre-admission scores with first year scores, the authors discovered that almost all pre-admission variables were strongly correlated with first year medical school performance among traditional students (Giordani et al., 2001). The authors concluded that non-traditional students demonstrated competency through the first year of medical school, despite lower pre-matriculation scores (Giordani et al., 2001). They did not find any evidence of poor academic performance among these students. In their discussion, the authors highlighted the role that non-cognitive factors may play in relation to student success (Giordani et al., 2001). For example, the UM post-baccalaureate program emphasized personal growth, study skills, coping techniques, and reading comprehension skills in addition to focusing on basic science content. The authors suggested that these other skills may have played a large role in the positive outcomes observed, and this may have been the driving force behind their academic success during their medical school (Giordani et al., 2001).

In all, the current body of literature suggests that pre-medical enrichment programs, including post-baccalaureate and master's programs, are effective when it comes to increasing medical school matriculation of non-traditional (Imlach, et al., 2017) and under-represented students (Grumbach & Chen, 2006). Additionally, it appears as though these students are able to perform at a level that is similar to their traditional classmates after matriculating in medical school, despite poorer pre-admissions variables.

Consideration of Non-Cognitive Variables—Conscientiousness and Grit—in Medical School Admissions Reflective of an Assets' Approach

The literature has highlighted the limitations of current traditional measures used in admissions for predicting success in the clinical years of medical school and on the USMLE Step 2 (Price et al., 1971; Sade et al., 1985); continuing to focus on these traditional measures will continue to have a negative impact on the diversity of the physician workforce. This diversity is multi-faceted, and includes not only ethnic diversity, but also gender, first-generation college status, socioeconomic background, and educational background and preparation. Adjusting the lens to consider more diverse backgrounds, strengths, and distances traveled will allow for anti-deficit consideration of students with a focus on the qualities they bring to the field of medicine based on their cultural background and life experiences. Continuing to only use measures that underscore the perceived limitations of students indicates that there is something wrong, or deficient, and in need of fixing, when in fact they have so much to bring to the table and offer the field of medicine. This change in focus is recognized in Yosso's (2005) work arguing for recognition and inclusivity of the rich and critical assets that Communities of Color bring to educational environments and to challenge a deficit lens that only recognizes knowledge and skills reflective of dominant White society. This approach presents the need for a change in one's mental models in order to effect the changes necessary in the medical school admissions process (Magzan, 2012). This change requires valuing assets that students bring with them from diverse cultural backgrounds that are critical to their navigation of an inequitable educational system. Two such constructs that are often cited as qualities critical to student success of those experiencing an educational system not tailored for them is conscientiousness and grit, which are introduced along with valid and reliable assessments to measure them. As discussed in Chapter One, these two traits may be characteristic of three types of Community Cultural Wealth Capital: aspirational, navigational and resistant.

Conscientiousness construct and medical school success. According to Roberts et al. (2005), the construct of conscientiousness is comprised of six factors listed in order of importance: industriousness, orderly, sense of self-control, responsible, traditional, and virtuous, particularly when used to assess one's dedication to their career. Further, Roberts et al. (2014) describe the conscientiousness construct as human capital in which individuals who are conscientious are self-controlled, responsible to others, hardworking, orderly and follow the rules, qualities that patients may want in their physicians. Conscientiousness involves being mindful of others, being highly organized, having a clear sense of direction, and persisting in reaching their goals despite boredom or distractions (Costa & McCrae, 1992b). Conscientious people tend to be successful, disciplined, high-achieving, dependable, and able to stay in control of stressful situations (Roberts et al., 2009). Similarly, Fayard et al. (2012) found that conscientiousness reveals replicable components, including "industriousness, responsibility, impulse control, and conventionality" (p. 2). Two of the most common domains of conscientiousness are orderliness and industriousness, followed by self-control and responsibility. Conventionality, decisiveness, formalness, and punctuality were also found to be common domains of this trait. Delay of gratification, ego control, effortful control, self-control, self-regulation, impulsivity, constraint, and grit are also associated with conscientious individuals (Roberts et al., 2014). Predicting success in the clinical years and on Step 2 is imperative. There have been cases where students have made it through the basic science years of their medical education and been successful on Step 1 but are unable to successfully make it through their clinical years (Case et al., 1996; Fields et al., 2000). That is a very disheartening situation and an incredibly taxing process, both emotionally and financially, for students. Researchers suggest that including non-cognitive factors with cognitive factors when evaluating applicants will better predict how students will perform in their clinical years, when they are interacting with patients and other health professionals (Ferguson et al., 2002; Sobowale et al., 2018). Similarly, Haight et al. (2012) found that conscientiousness was strongly related to various facets of clinical success including history taking, physical exam skills, patient rapport, and health care team rapport.

The Conscientiousness construct of the NEO-FFI-3 personality inventory (Costa & McCrae, 2010) was used in this study to examine potential attributes medical students possess that may correlate with USMLE Step 2 Clinical Knowledge results beyond that explained by undergraduate grade point averages and MCAT scores. While the NEO-FFI-3 construct has not been used to assess medical students at the point of admission, a study by Lievens et al. (2009) found that over the course of their medical school career, medical students' conscientiousness appeared to grow. Other studies have used the Conscientiousness Index (CI) to measure professionalism during the clinical years of undergraduate medical education (Chaytor et al., 2012; Finn et al., 2009; McLachlan et al., 2009). The authors state that the more conscientiousness medical students exhibit, the more professional competencies they display. Additionally, Lievens et al., (2009) stated that Conscientiousness, perhaps more than any other personality trait, may be considered an increasing asset for medical students over the course of their education. The psychometric properties of this instrument are reviewed in Chapter Three.

Studies have indicated conflicting findings regarding the relationship between the NEO-FFI-E personality inventory and medical school success. Griffin and Wilson (2012) compared MMI scores and personality traits of 868 medical applicants in Australia over three years. They discovered a significant positive relationship between MMI scores and both the conscientiousness and extraversion traits over all three years (Griffin & Wilson, 2012). Lievens et al. (2009) explored the association between medical student personality and academic performance throughout different stages of medical education. This longitudinal study took place in Belgium and followed 785 students across six universities. The authors discovered that the operational validity of extraversion, openness, and conscientiousness as predictors of student GPA all increased by the seventh year. Of these three traits, conscientiousness showed the greatest increase in correlation, with an operational validity that increased from 0.18 to 0.45 by the late stages of medical education (Lievens et al., 2009). The authors hypothesized that the observed increase in the apparent relevance of the conscientiousness trait may be due to the changing nature of the coursework. As in the United States, the later stages of the Belgian medical education become increasingly clinical, meaning that non-cognitive factors may become more relevant to success. While early academic performance may rely heavily on cognitive skills, non-cognitive personality traits, such as conscientiousness, may be critical in the later years (Lievens et al., 2009). They hypothesized that the attention to detail, organization, and diligence associated with conscientiousness plays a role in a student's ability to solve tough clinical problems. In addition to a higher capacity for clinical problem solving, the authors also suggested that conscientiousness is an "increasing asset" over the course of students' medical education and may predict persistence over the various stages of their professional education, while also bringing more honesty and integrity to their professional relationships (Lievens et al., 2009). While this study may provide an imperfect comparison to American medical students, the data suggest that personality scale measurements, particularly conscientiousness, become more powerful predictors of medical student success in their clinical years.

Similar findings were published by Haight et al. (2012). The authors performed a retrospective study using data from 175 medical students enrolled at the St. Louis University School of Medicine. Students completed the NEO-FFI, a shorter version of the revised NEO Personality Inventory (NEO-PI-R), after the third year of medical school. The authors subdivided the clinical evaluation scores into diagnosis planning skills (DPS) and interpersonal skills (IPS) to further parse out any subtleties. DPS included skills related to patient diagnosis and treatment, such as history taking, physical examination skills, diagnostic planning, and organizing data. IPS included communication, patient relationships, and care team relationships. The data showed a positive correlation between conscientiousness and performance on both the DPS and IPS aspects of clinical evaluation (Haight et al., 2012). This correlation supports the proposed explanation described above that was presented by Lievens et al. (2009), in which they suggested that conscientiousness may bolster both clinical problem-solving skills and interpersonal relationships. Extraversion was positively correlated with the IPS aspect of clinical evaluations only, as well as the number of Humanism in Medicine Honor Society nominations. The authors were also granted access to the students' MCAT scores and compared the predictive value of the MCAT to their personality characteristics and concluded that this pair of factors offers two different types of insight. Higher MCAT scores were positively correlated with success on pre-clinical exams, Step 1, and NBME subject exams, but there was no correlation between MCAT score and clinical skills (Haight et al., 2012). As already mentioned, NEO-FFI results were correlated with Humanism in Medicine Honor Society nominations, though they were not predictive of pre-clinical exam scores, Step 1 scores, or NBME subject test scores (Haight et al., 2012).

The relationship between conscientiousness and clinical skills was again supported by a study published by Sobowale et al. (2018). This national study sent questionnaires to 960 third-year medical students enrolled in 24 medical schools in the United States. The authors found that conscientiousness was the best personality trait to predict academic success in the clinical realm of medical education. The data showed that highly conscientious students were more likely to receive the highest clerkship grades across all specialties and was an important factor in gaining membership in Alpha Omega Alpha (AOA) Honor Medical Society (Boatright, Ross, O'Connor, Moore, &

Nunez-Smith, 2017; Ferguson et al., 2002; Sobowale et al., 2018; Wijesekera, Kim, Moore, Sorenson, & Ross, 2019). Neuroticism was inversely related to conscientiousness across all clerkships, with high neuroticism scores associated with low clerkship grades across all specialties (Sobowale et al., 2018). Students who were graded well in the surgery clerkship exhibited high levels of extraversion, while students who scored highly in pediatrics exhibited high levels of openness. The authors suggested that conscientiousness becomes a particularly useful trait during the clinical years because the curriculum is less structured (Sobowale et al., 2018). Students must be persistent and engage in more self-directed learning to keep up with the rigors of the clinical years. Additionally, they also mention that the diligence and dependability associated with conscientiousness are essential for good patient care (Sobowale et al., 2018).

Grit construct and medical school success. According to Tough (2012), some students demonstrate more fortitude than their peers in completing college and succeeding in their jobs. Duckworth et al. (2007) stated what often sets these committed individuals apart from others is their attitudes and behaviors, otherwise referred to as grit and defined as “perseverance and passion for long-term goals” (p. 1087). In a 2018 speech, Duckworth argued that “…in order to build grit, you must develop your interests, practice, cultivate a purpose such as serving others….Gritty people do not let setbacks hold them back and always have something to prove.” Grit, as a nonacademic attribute, is captured by standardized cognitive tests (Duckworth et al., 2007; Tough, 2012). Duckworth et al. (2007) differentiated grit from conscientiousness, stating that although grit may intersect with conscientiousness, grit emphasizes a long-term commitment to pursuing one’s goal and does not include the self-control and dependable traits that are found in conscientious people.

The Grit-S scale is an eight-item scale designed by Duckworth and Quinn (2009) to measure grit. Its psychometric properties are described in Chapter 4. Grit has been shown to be correlated with Step 2 CK performance. Miller-Matero et al. (2018) examined the relationship between grit, as measured by Duckworth's Grit survey, and medical school performance, measured by number of years in medical school, class rank, and USMLE Step 1 and Step 2 CK exam, in a retrospective study of 309 graduating medical students. While all students had a higher than average grit rating as compared to the general population, the students who completed school in four years had a higher average grit score than those who took five years to complete the program (Miller-Matero et al., 2018). There was no statistically significant correlation between grit level and Step 1 score. However, there was a significant positive correlation between grit and Step 2 CK grade, as well as grit and overall class rank (Miller-Matero et al., 2018). The authors suggested that grit may have been associated with Step 2 CK scores, but not Step 1 scores, because of the nature of each test (Miller-Matero et al., 2018). Step 1 tests basic science knowledge and relies heavily on memorization. Step 2 CK requires active application of knowledge and problem solving, as opposed to recall alone. The authors hypothesized it takes more practice and perseverance to learn how to solve the types of problems presented on Step 2 CK, and the students with higher grit scores are more likely to put forth this effort (Isenberg, Brown, DeSantis, Veloski, & Hojat, 2020; Miller-Matero et al., 2018).

Summary

There are many changes facing the current medical school climate: a need for more physicians, physicians able to serve a diverse population, and physicians who possess personality traits that enhance the physician-patient relationship. Cognitive measures alone simply do not assess who will become good clinicians. Physicians must stay the course, be resilient and focused on their educational goals and the health of their patients. Perhaps by assessing some of these non-cognitive qualities during the admissions process, we will be able to better predict which students will be successful on Step 2 Clinical Knowledge.

This chapter reviewed several areas in an effort to better understand the need for and context of this study. Literature regarding academic factors that influence medical student success as well as several diversity areas in medicine including ethnicity, first-generation college student status, socioeconomic status, gender, and non-traditional student age were appraised to guide this study. By examining previous research on medical student success, it was found that the majority of the literature examines cognitive and academic variables of this student population. Appendix C provides a summary table of the findings and methods from studies that focused on the key variables in this study. Following this literature review, the constructs of conscientiousness and grit were selected because both have previously assessed college students but neither have been used to predict medical student success on the national medical school examination Step 2.

CHAPTER THREE: METHODS

Using a single institution dataset design, the purpose of this study was to explore how specific non-cognitive factors, conscientiousness and grit, correlate with medical student success as defined by scores on the clinical knowledge (CK) section of the USMLE Step 2 versus traditionally cognitive-based variables, such as GPA and MCAT for 192 students at one state-funded college of medicine located in an urban, mid-size city in the northeast. Specifically, the predictive significance of medical students' scores on the Conscientiousness section of the NEO-FFI-3 (Costa & McCrae, 1992) and the Short Grit Scale (Grit-S) (Duckworth & Quinn, 2009) on passing USMLE Step 2 CK was examined. In this study, the relationship among the variables was assessed using descriptive and multivariate statistical analysis. Specifically, regression analyses were run in order to determine the relationship between the independent variables, particularly grit and conscientiousness, on Step 2 CK performance.

Quantitative data analysis methods were used, including descriptive and multivariate analysis, to answer the four research questions explored in this study:

Research Question 1: Do socio-demographic variables—specifically age, gender, ethnicity, socioeconomic status, or first-generation college status—correlate with success on the national USMLE Step 2 CK medical school examination at Northeast Medical University? The null hypothesis was socioeconomic variables – age, gender, ethnicity, socioeconomic status, or first-generation college status – did not predict medical school success as measured by the dependent variable, USMLE Step 2 CK.

Research Question 2: Do academic variables—specifically MCAT scores, prior college science GPA, college major, post-baccalaureate course work, or graduate degree—correlate with success on the national USMLE Step 2 CK medical school examination at Northeast Medical University? The null hypothesis was prior academic variables – MCAT scores, science GPA, college science major, post-baccalaureate course work, or a graduate degree – do not predict medical school success as measured by the dependent variable, USMLE Step 2 CK.

Research Question 3: Does conscientiousness, as measured by Costa and McCrae's Neuroticism, Extraversion, Openness Five-Factor Personality Inventory (NEO-FFI-3) (1992), predict medical student academic success on the USMLE Step 2 CK medical school exam when controlling for demographics and prior academic history at Northeast Medical University? The null hypothesis was conscientiousness does not predict medical school success as measured by the dependent variable, USMLE Step 2 CK.

Research Question 4: Does grit, as measured by Duckworth and Quinn's Short Grit Scale (2009), predict medical student academic success on the USMLE Step 2 CK medical school exam when controlling for demographics and prior academic history at Northeast Medical University? The null hypothesis was grit does not predict medical school success as measured by the dependent variable, USMLE Step 2 CK.

Research Setting

As previously described in Chapter One, this research was conducted at Northeast Medical University (a pseudonym), a public, semi-selective, college of medicine, located in an urban city of approximately 150,000 people. This college is the largest of four colleges that comprise a comprehensive academic medical center. The student body across the four colleges is comprised of roughly 1,500 students with approximately 666 students enrolled in the College of Medicine. The setting was selected as a convenience sample, as the researcher also serves as an admissions dean at this medical school in the US Northeast.

Of the 666 students that made up the College of Medicine at Northeast Medical University, 46% of the matriculated students identified as female, and 11% identified as under-represented minority students. This data as compared to national data, students eligible to participate in this study, and those who participated in this study can be seen in Table 3.1.

Table 3.1

Table 3.1: Enrollment Comparison: National, Northeast Medical University, Eligible Participants & Participants

	Enrollment	Male	Female	White/Asian	URM
National	21025	50.20%	49.80%	87.80%	12.20%
Northeast Medical University	666	54.00%	46.00%	85.00%	11.00%
Eligible to Participate	305	59.10%	40.90%	85.60%	13.00%
Study Participants	192	58.00%	42.00%	87.00%	13.00%

A chi-square was run and discussed in Chapter Four to assess the representation of this sample to that of the entire medical school student body at Northeast Medical University and with national matriculated medical students. Further, 14% of the medical students were 26 or older at the point of matriculation, 13 % of the students were first generation college students, and 38% were from low-socioeconomic backgrounds (EO1 and EO2). This information was not institutionally tracked at the time of this study; therefore, no other information is available. Over the past few years, more and more students matriculating into this medical school were first in their family to attend college and/or came from families that were from lower socioeconomic backgrounds.

Description of Participants

The study identified 305 enrolled students as eligible to participate in this study because they interviewed during the two admissions cycles where the NEO-FFI-3 was utilized on applicants' interview day. Of those 305, 12 students (or 3%) had left the institution by means of dismissal, withdrawal, transfer, and one student had passed away. Therefore, the Short Grit Survey was emailed to 293 students in the early spring of their 3rd and 4th years and were invited to participate in this research and asked to additionally complete the Grit-S Scale; 192 students completed the survey, representing a 65.5% participation rate. At the time the students completed the Grit-S scale, the students also consented to allow this investigator access to their academic records in order to collect demographic information as well as prior GPAs, MCAT scores, and USMLE test scores.

Table 3.2 is representative of all of the students who agreed to participate in this research. One-hundred-twelve students identified as male, 80 students identified as female, and no students identified as other. The age range of the students who participated in this research was 21-45. Twenty-six respondents self-identified with identities that under-represented in medicine (13%): 17 identified as Black/African American, six identified as Hispanic, and three identified as Native American. The students identified as first-generation college were those students who were the first in their families to attend college. Those who were considered low-socioeconomic were identified as socioeconomic disadvantaged (labeled as EO1 or EO2) on their AMCAS application (AAMC, 2020b).

Table 3.2

Table 3.2: Breakdown of Participants

Participants (N = 192)	Total (%)
Gender	
Female	80 (42%)
Male	112 (58%)
Other	0
Age	
Traditional Age (<26)	161 (84%)
Non-traditional Age (≥26)	31 (16%)
Race/Ethnicity	
URM (Under-Represented Minority)	26 (13%)
Majority students (Asian/White)	166 (87%)
Undergraduate Major	
Science Major	158 (83%)
Non-Science Major	34 (17%)
First in Family to Attend College	
First-Generation College	20 (10%)
Not First-Generation College	172 (90%)
Economic Status	
Low Socioeconomic (EO1, EO2)	61 (32%)
Higher Socioeconomic	131 (68%)
Post-baccalaureate (PB) Coursework	
Completed 24+ hours of PB	19 (10%)
Did not complete a PB	173 (90%)
Graduate Program	
Completed a graduate degree	23 (12%)
Did not complete a graduate degree	169 (88%)

Based on the small number of under-represented students and following the national norm in medical school, Asians were not considered as underrepresented in medicine for the purposes of this study (Nivet, 2010). Participant ethnicities were collapsed into two categories: White and Asian were considered together and called “majority” and those under-represented in medicine, such as Black/African American, Native American, and Hispanic, were grouped together and called “URM” for "underrepresented in medicine." As mentioned above, of the 305 medical students eligible for this study, 13% identified as underrepresented in medicine. During this same time period, the national proportion of underrepresented matriculated students was 12.2% (AAMC, 2020d).

Instrumentation

The Conscientiousness construct of the NEO-FFI-3 personality inventory (Costa & McCrae, 2010) and the short Grit scale (Duckworth & Quinn, 2009) were used in this study to examine potential attributes medical students possess that may correlate with USMLE Step 2 Clinical Knowledge results beyond that explained by undergraduate grade point averages and MCAT scores.

NEO Five Factor Personality Inventory and Its Conscientiousness subscale. The Five-Factor Personality Inventory (FFI) is a widely-used and empirically sound personality inventory developed by Costa and McCrae (1989) which measures five distinct personality traits: conscientiousness, extraversion, openness to experience, agreeableness, and neuroticism. Over the years it has gone through various iterations, resulting in a shortened assessment referred to as the NEO-FFI-3 (Costa & McCrae, 1992). The NEO-FFI-3 is used for assessing the Big Five personality traits: neuroticism (N), extraversion (E), conscientiousness, agreeableness, and openness (O), as mentioned previously (John et al., 2008). This assessment provides strong internal consistency as well as test-retest reliability (Costa & McCrae, 2010).

The NEO-FFI-3 construct employs a 5-point scale with options ranging from "Strongly Agree" to "Strongly Disagree" and questions such as "I keep my belongings neat & clean." The psychometric properties of the revised NEO Personality Inventory, NEO-PI-R, had been found to be generalizable across cultures, age, and various forms of measurement (McCrae et al., 2011). The test-retest reliability of the NEO-PI-R, over a six-year period, was found to be satisfactory with conscientiousness at .79 (McCrae & Costa, 1983). The findings demonstrated that this construct has good reliability and stability over time (McCrae & Costa, 2007). In 2010(a), McCrae and Costa discerned that the newer version of the NEO-PI-R, the NEO-FFI-3, was more reliable with a stronger factor structure. The shorter NEO-FFI-3 scale had consistent results with a Cronbach's alpha ranging from .87 to .95 across the five scales, indicating good approximations of the full survey (McCrae & Costa, 2010a).

The conscientiousness portion of the personality test has a total of 12 questions with the scores summed to get a final score. The higher the score, the more conscientious the person is determined to be (McCrae & Costa, 2010b) (Appendix D). Due to there being fewer questions, the internal consistency within each scale was smaller, but the NEO-FFI-3 conscientiousness rating was .83 (Sherry et al., 2007). In terms of the validity of this construct, related to conscientiousness only, several studies examined the effect of conscientiousness on academic success and found that it was a better predictor than only assessing standardized test scores for undergraduate college students (Busato et al., 2000; Conard, 2006; Conrad & Patry, 2012; Costa & McCrae, 2008; Kappe & van der Flier, 2012; Körner et al., 2015; McCrae & Costa, 2007; McCrae et al., 2011).

Although it is recommended that the NEO-FFI be used in its entirety in assessing all five personality traits, factor analysis has been used to determine independent and internally consistent questions within each of the domains, allowing for the assessment of only one construct at a time (Boyle et al., 2008). Doherty and Nugent's review (2011) concluded that conscientiousness can predict long-term success in medical education, identified by year-end grades and lower levels of attrition. Northeast Medical University decided to only use the conscientiousness subscale of the NEO-FFI-3 and not collect data from items measuring openness and agreeableness. Although openness has been linked to academic ability and divergent thinking (Goff & Ackerman, 1992; McCrae, 1996), its ability to predict academic achievement was not significant (Barrick et al., 2001;

Hough, 1992) and therefore was not utilized by Northeast Medical University. Similarly, the facet of agreeableness, which represents trust, altruism, modesty and tender-mindedness (Costa & McCrae, 1992b) can facilitate physician-patient relationships and has been positively associated with clinical performance in medical students (Gough et al., 1991; Shen & Comrey, 1997), but appears to be the weakest of the domains in terms of reliability (Caruso, 2000; Hahn, Gottschling, & Spinath, 2012) and therefore agreeableness was also not utilized by Northeast Medical University.

Therefore, Northeast Medical University utilized the facets of extraversion, neuroticism, and conscientiousness in its admissions process. Due to sample size concerns and the small number of participants, this researcher decided against adding the neuroticism and extraversion variables to this analysis as there was not enough power to analyze all three variables. Extraversion was not considered in this study as this trait can be viewed as culturally biased, particularly concerning Asian students, who may be regarded as scholarly and/or reserved (Hartocollis, 2018). Neuroticism was not considered in this study as previous research suggested that those individuals with a higher level of neurotic traits may have already found that their neurotic tendencies had "reduced their likelihood of gaining entry into medical school" (Enns, Cox, Sareen, & Freeman, 2001, p. 1040) and upon a cursory review of this facet, there were no matriculating students that performed poorly on this section of the NEO-FFI-3.

Conscientiousness has been associated with academic satisfaction in medical school (Lieberman, Stroup-Benham, & Peel, 1998), it has predicted academic success and job performance in medicine (Barrick & Mount, 1991; Behling, 1998; Hurtz & Donovan, 2000; Tett et al., 1994), and preliminary evidence demonstrated that conscientiousness correlates with academic and professional performance in medical school (Hojat et al., 2013) and it may capture qualities exhibited from a wider diversity of individuals thereby, reflecting a cultural wealth perspective. Medical education research generally suggests that among the Big Five factors, conscientiousness is the more reliable predictor of success during the first year of medical school and long-term success as a practicing physician, while also predicting humanism nominations (Hojat et al., 2013). Further, Magee and Hojat (1998), using the NEO PI-R, found that male and female physicians who were nominated as positive role models in medicine scored significantly higher on the conscientiousness factor compared to the general population. Based on all of the factors, the conscientiousness subset of the NEO-FFI-3 will be the only subset considered in this study.

Haight et al. (2012) found that scores on the MCAT correlated with academic examinations, whereas scores on conscientiousness correlated with indicators of clinical performance and humanism nominations. Further, a review conducted by Doherty and Nugent (2011) concluded that conscientiousness can predict attrition and year end grades in medical education.

Grit Construct and Grit Scale (Grit-S). The original grit theory posited that grit is one’s predisposition to be passionate and persevere when working toward long-term goals. Grit, as a noncognitive attribute, has been described as a perseverance quality of successful individuals not captured by standardized cognitive tests (Duckworth et al., 2007; Tough, 2012). A gritty person is one who single-mindedly works toward long-term goals, despite barriers and setbacks along the way, as well as maintains consistent interests over time (Duckworth et al., 2007). Further, in recent years grit has been identified as a predictor of success beyond intelligence and conscientiousness (Suzuki et al., 2015). Duckworth et al. (2007) differentiated grit from conscientiousness, stating that although grit may intersect with conscientiousness, grit emphasizes a long-term commitment to pursuing one’s goal and does not include the self-control and dependable traits that are found in conscientious people.

The Grit-S scale is an eight-item scale designed by Duckworth and Quinn (2009) to measure grit. The Grit Scale has predicted success for West Point United States Military Academy cadets, Scripps National Spelling Bee contestants, and college undergraduate students (Duckworth et al., 2007; Duckworth et al., 2010). This tool employs a 5-point scale where Very much like me (1) and Not like me at all (5) applied to statements such as “New ideas and projects sometimes distract me from previous ones” (Duckworth and Quinn, 2009). The scores from each question are added together and divided by 8 to obtain the overall Grit score. The higher the overall score, the “grittier” the person is (Appendix E).

At the November 2018 AAMC Conference in Austin, Texas, Angela Duckworth introduced the concept of grit to medical educators across the country. During her keynote address, Duckworth (2018) discussed her research with West Point US Military Academy cadets and stated that, for these students, being at West Point is often the first time in their lives where they are not at the top of the class. Instead, some will be average, and others will fail or drop out. This is much the same with medical school students. Most have been at the top of their class in high school, through undergraduate school, and perhaps even through a master’s program, but when they enter medical school, things change. They have successfully completed a rigorous undergraduate degree, performed well on the medical school entrance exam, and have been admitted to medical school. Now, they must successfully complete two years of rigorous academic coursework, successfully pass a national exam which assesses their basic science knowledge, complete a year of demanding clinical rotations, successfully pass two additional national examinations that assess their clinical knowledge and clinical skills, decide which specialty they want to pursue for their entire career, interview for a residency position (the location of which is largely out of their control), complete their fourth year of elective course work, and then begin a residency program where they will spend three to seven years learning specifics about the chosen specialty of their profession and then sit for yet another national exam. During this time, medical students must alter their lifestyle, schedule, and place some life experiences on hold to accommodate their medical education (Benner et al., 2010). In her plenary session remarks, Duckworth stated that grit and talent are not the same thing: “in order to build grit, you must develop your interests, practice, cultivate a purpose such as serving others, and have a growth mindset as defined by Carol Dweck, PhD” (Duckworth, 2018). Further, she stated that “motivation is internal; gritty people do not let setbacks hold them back and always have something to prove.” Therefore, grit is essential to meet the unique academic demands of medical school and Duckworth and Quinn’s Grit Scale (2009) is necessary to measure this non-cognitive quality.

According to Tough (2012), some students demonstrate more fortitude than their peers in completing college and succeeding in their jobs. Duckworth et al. (2007) stated that what often sets these committed individuals apart from others is their attitudes and behaviors, otherwise referred to as grit and defined as “perseverance and passion for long-term goals” (p. 1087). Grit has been investigated by studying the characteristics of successful individuals, observing diverse observable and tangible accomplishments, both in academia and business (Duckworth et al., 2007). Researchers have found that apart from intelligence, success can be achieved through persistence, hard work, and sustained interest (Duckworth et al., 2007).

The original Grit scale (Grit-O) was revised to remove the least predictive items (Duckworth et al., 2007) and renamed the Short Grit Scale (Grit-S) (Duckworth & Quinn, 2009). This the shorter Grit-S scale was determined to be psychometrically stronger than the original Grit-O scale and confirmatory factor analyses demonstrated that the Grit-S scale seemed to fit the data better (Duckworth et al., 2007; Rojas et al., 2012). The confirmatory factor analysis found that the Grit-S Scale showed acceptable goodness of fit indices using four independent samples as well as internal consistency ratings of .73 to .83 by using Cronbach's alpha (Duckworth & Quinn, 2009). Additionally, these studies found grit to be positively correlated with educational achievement, as evidenced by winners in the National Spelling Bee competition, grade point averages found within a sample of students from an Ivy League institution, and the retention of military cadets at West Point Military Academy. Collectively these studies provided evidence of predictive validity and test-retest reliability when using the Grit-S scale (Duckworth et al., 2007). According to Duckworth et al. (2007), based on its "superior psychometric properties [and] comparable predictive validity," the Grit-S scale is a reliable and valid tool to measure one's grittiness (p. 174).

A very recent study by Miller-Matero et al. (2018), surveyed graduating medical students from one medical school using the Grit-S scale, yet no research utilizing the Grit-S Scale at the point of admission to predict medical student success was revealed in this review. As the medical profession seeks to increase and diversify the physician workforce, medical schools must find alternate factors that can be used to predict success in medical school.

Based on the above, both instruments appear appropriate to assess the non-cognitive qualities of grit and conscientiousness in this sample of students.

Data Collection Procedures

Active third- and fourth-year medical students enrolled at Northeast Medical University were identified as those who had completed the NEO-FFI-3 as part of their admission to the medical school and had taken the USMLE Step 2 CK national examination. The Short Grit Survey was emailed to all students identified above. Students were requested to complete this short, voluntary Grit Survey and provide their student identification number. The purpose and participant expectations were described, and students were offered the opportunity to ask questions if they had any concerns about the use of these data. They were also notified that they would remain anonymous in the reporting of this research, and would be reported in aggregate. A reminder email was sent approximately one week after the initial email encouraging students who had not already completed the survey to do so. An announcement was also made during a March into Residency program, reminding graduating students of the survey. Participants were not compensated for their participation, although they were all entered into a random drawing to receive a $50 Amazon gift card for their participation. To increase participation, four random drawings, each offering a $50 Amazon gift card, were conducted. This incentive was a minor token of appreciation for participants' time and effort. Although this researcher is an administrator at this institution, she does not have grade influence over these students, minimizing influence upon their participation in this study. All students within this study participated willingly and provided their University ID number with their survey results.

Demographic and academic data for each student who submitted the Grit Survey and provided their student identification number were retrieved from the student data system at the medical school. These data included background information retrieved from their AMCAS application, including standing as a first-generation college student, undergraduate GPA, MCAT scores, ethnicity, and gender. The results from their Step 2 CK examinations were retrieved from the student's academic record within the student data system. All participants had completed the NEO-FFI-3 as part of the admissions process, and these data were available in the student information system as well. For this analysis, the results from the 12 Conscientiousness questions were used to obtain the overall Conscientiousness score from the NEO-FFI-3.

Once the students submitted the Grit Survey, their score was determined and a database was created. The Conscientiousness score and academic and demographic information were added to this database. Once completed, the student identification number was removed, and unique numbers were assigned to all participants. The original spreadsheet was kept secure and a key of original names and ID numbers with the assigned unique identifiers was kept in a hard copy format in a secure location. For security purposes, this information was not housed in the Cloud or on the Internet. This researcher was the only investigator able to access the raw data from the respondents. These precautions allowed for the confidentiality of the participants' data throughout the collection, analyses, and reporting of this research.

Independent variables

The independent variables for this study were grouped into three categories: non-cognitive variables, academic variables, and sociodemographic (background) variables, similar to groupings that had been used in previous studies **(**Duckworth & Quinn, 2009; Schmitt et al, 2009; Sedlacek, 2004; Strayhorn, 2013; Ting, 2009; Tracey & Sedlacek, 1984).

Sociodemographic variables. This study utilized five variables which were reviewed in Chapter 2, indicating that they correlate with student success in medical school. The background variables that were assessed included: age, first-generation college condition, gender, race, and socioeconomic status. These variables were drawn from the students' AMCAS applications.

Age - The age of the participants was determined by birth information provided on the AMCAS application when the student applied to medical school. According to the AAMC, the average age of students entering medical school was twenty-four (AAMC, 2020c). For this study, twenty-six or older at the time of matriculation was the age used to indicate the student as non-traditional in medicine (Jauhar, 2008).

First Generation College - Students self-identified their parental education level on their AMCAS application. Students with a parent/guardian whose education level was identified as “Less Than High School,” “High School Graduate” (high school diploma or equivalent), or “Some College but No Degree” were considered first-generation for the purposes of this study (AAMC, 2020b). This definition is the standard for first-generation students applying to medical school (Grabowski, 2018; Grbic et al., 2019; Williams et al., 2019; Winseman et al., 2018).

Gender - Students self-identified on their AMCAS application as Female, Male, or Other.

Race/Ethnicity- Applicants self-identify their race and ethnicity on their AMCAS application. Race/ethnicity options are American Indian or Alaska Native; Asian; Black or African American; Hispanic, Latino, or of Spanish Origin; Native Hawaiian or Pacific Islander; Other; and White (AAMC, 2020e; Petersdorf et al., 1989). Based on 2018 data from the AAMC, White students comprised 50-52% of each matriculating class, while Asian students made up 20-23%. Meanwhile, Other, which included unknown ethnicities and non-US citizens, made up approximately 8% of each year's matriculating class (AAMC, 2020d). For this study, there were no students who identified their ethnicity as Other. Matriculating students considered under-represented in medicine (URM) are Black or African American (approximately 8%), Hispanic/Latinx/Spanish and Native Hawaiian or Other Pacific Islander (just under 10%), and American Indian or Alaskan Native (less than 1%).

Socioeconomic Status- Socioeconomic status (SES) was defined as a measure of a person’s social standing or class measured by income and education level of the student’s parents. For this study, the SES disadvantaged indicators identified by the AAMC as EO1 and EO2 were used. Students whose parents earned less than a bachelor’s degree and were employed in service, clerical, skilled, or unskilled labor professions were identified as EO1 and those whose parents had earned a bachelor's degree or higher and were employed in service, clerical, skilled, or unskilled labor professions were identified as EO2. Students represented by EO1 or EO2 may have received Fee Assistance (from AAMC, Federal, or State), are high need, identify as disadvantaged, or have a family income level below $40,000 (Grbic et al., 2013; Grbic et al., 2015; Jerant et al., 2019; Lowrance, 2017).

Table 3.3

Table 3.3: Summary of Sociodemographic Variables Considered

Age	Traditional age- Less than 26 at the time of matriculation; Non-traditional age- 26 or older at the time of matriculation.
First-Generation College Students	Students with a parent/guardian whose education level is Less Than High School, High School Graduate or Equivalent, or Some College but No Degree.
Gender	Female, Male, or Other
Race/Ethnicity	American Indian or Alaska Native; Black or African American; Hispanic, Latinx or of Spanish Origin; Native Hawaiian or Other Pacific Islander will be considered Underrepresented in Medicine (URM), White and Asia will be considered majority students.
Lower Socioeconomic Status	Students identified as EO1 and EO2 by the AAMC

Academic variables. There were five specific academic variables considered within this study. The first was an earned graduate degree, which is coursework completed at the graduate level following a student's post-secondary academic program. For the sake of this research, only those who earned a graduate degree were included. The next academic variable included was the student's scores on the Medical College Admission Test (MCAT), a national standardized test required by most medical schools during the application process. The third variable was post-baccalaureate science coursework which is undergraduate coursework students complete after they have finished an undergraduate degree. Students typically complete post-baccalaureate work if they did not major in the sciences and need to complete their medical school prerequisites before applying to medical school. For the purposes of this study, post-baccalaureate science coursework was included if a student had completed 24-semester hours or more, which is equivalent to two full-time standard semesters of course work (National Center for Education Statistics, 2019). The student's prior undergraduate college major was the fourth academic variable to be assessed and identified as "science" or "non-science." The science category included the hard or natural sciences, including biology, chemistry, physics, astronomy, earth science, and applied sciences. The non-science category included social or soft sciences such as psychology, sociology, philosophy, history, and any other majors not considered natural sciences. Finally, the students' prior college science GPA was considered. This grade point average was calculated and verified by the American Medical College Application Service (AMCAS). It included all undergraduate biology, chemistry, physics, and math courses from each undergraduate college where credit was given leading up to a bachelors' degree.

Table 3.4

Table 3.4: Summary of Academic Variables Considered

MCAT	Results of the most recent Medical College Admission Test as reported by the AAMC. Overall percentiles were used as data points.
Prior College Science GPA (Biology Chemistry, Physics, Math - BCPM)	Includes all undergraduate biology, chemistry, physics and math courses from each undergraduate college where credit was received leading up to a bachelors' degree.
Prior College Major	The specific area of study that students studied during their undergraduate academic career. This variable was categorized as 'Science' and 'Non-Science.'
Post-Baccalaureate Coursework	Coursework taken after a bachelor's degree has been completed. Coursework was considered when 24 or more semester hours were completed.
Earned Graduate Degree in the Sciences	An earned a master's degree or PhD in the Sciences.

Non-cognitive variables. Non-cognitive variables are often qualitative areas of development that include self-concept, realistic self-appraisal, and leadership (Sedlacek, 2004; Sommerfeld, 2011). These characteristics included students' motivation, level of curiosity, resilience, perseverance, attitude, work habits, and social skills. For the purposes of this study, the two non-cognitive variables considered were conscientiousness and grit as they relate to medical students and their success on USMLE Step 2 CK. These non-cognitive variables were gathered by the Grit-S survey (Appendix E) and the specific questions related to conscientiousness in the NEO-Five Factor Personality Test (NEO-FFI-3) (Appendix D). The Grit-S Scale measures a student's self-reported level of grit (Duckworth et al., 2007) and the NEO-FFI-3 measures a person's five personality traits which include openness to experience, conscientiousness, neuroticism, extraversion, and agreeableness (McCrae & Costa, 2010a). For the purposes of this study, only conscientiousness was assessed through the NEO-FFI-3.

Dependent variables

The outcome variable in this research included the participant's score for the Clinical Knowledge (CK) section of the national USMLE Step 2 exam. Step 2 CK assesses students' ability to apply their medical knowledge and skills, as well as their mastery of the clinical aspect of medicine as they progress through their medical education and begin to deliver patient care. Detailed background information about this exam was covered in Chapter 2. Scoring for the Clinical Knowledge section was represented on a 3-digit scale, ranging from 1-300 (Kim & George, 2018). A failing score on this exam for all participants in this study was 208 or lower (USMLE, 2019b).

Data Analyses

Two phases of statistical analyses were conducted using SPSS-26. In Phase I, descriptive statistics were conducted on the independent variables and the dependent variable relevant to this research. The relationship between the variables was assessed by conducting a bivariate analysis. T-tests were conducted to determine if there were significant differences between the means of the variables. In Phase II, regression analyses were used to determine the relationship between the independent variables, and particularly the relationship of grit and conscientiousness on Step 2 CK performance. The research questions were used to guide this statistical analysis.

Known Limitations of this Study

There are several limitations to this study. First, the scores on the Grit-S Scale and NEO-FFI-3 are reliant on self-reported data by participants. Surveys are often used in educational research; participants may be influenced by the knowledge that they are being evaluated and therefore may not accurately report their "true" self (Fowler, 2013; Shechtman et al., 2013), and self-reported measures may easily be "faked" (Kyllonen, 2005, p. 3), often referred to as social desirability bias (Grimm, 2010). Although using tested measurements minimizes these concerns (Fowler, 2013), little is known about the social, cultural, and contextual factors that may influence grit in various contexts. Datu and McInerney (2017) indicate that the generalizability of grit has been criticized in different cultures, as little research has been conducted on this subject. According to Datu and McInerney (2017), there has also been limited research on how and why culture may influence perseverance of effort and consistency of interests. Further, the authors think some cultural psychological theories may help explain grit including the cultural dimension model, self-construal theory, the socioecological model, and the social axioms theory (Datu & McInerney, 2017).

Second, the timing of the surveys may be a potential limitation of the study. The timeframe that students were asked to complete the Grit-S scale was during their third and fourth years of medical school. Research conducted by Roberts et al. (2014), demonstrated that most personality traits are shaped by one's environment; conscientiousness, and related constructs like impulse control, continue to develop into adulthood and may change over time, and therefore these traits could potentially change over time and with maturity. Similarly, the data obtained for the Grit-S survey were collected by participants who knew they were being evaluated. Although surveys are often used in social science research, participants' responses to survey questions may simply be influenced by the fact that they know they are being evaluated (Rumrill et al., 1999).

Third, there may be other challenging variables that impact and affect a student's success in medical school which are beyond the scope of this study. These variables include illness, mental health, familial support, family obligations, financial health, etc. Similarly, additional qualities associated with quality care such as extraversion, empathy, and resilience may be appropriate traits to examine in future research.

Finally, there are certainly limitations with placing the underrepresented in medicine (URM) students into one category. There could be differences between the various ethnic groups, yet because of the limited numbers of students participating, the statistical significance would have been questionable. Limitations regarding the gender binary also exist, as these data were constrained by the database during that time period, which has since been rectified in the AMCAS application. Also, the results of this study are limited to a particular group of medical students at an individual medical school, therefore limiting the generalizability to students at other medical schools. Additionally, although I do recognize the possibility that my status at the medical school may have been a potential factor in student's willingness to participate in this research, I was very clear that their participation was fully voluntary and that it was not tied to any course grade or evaluation. Further, students who were on a leave of absence, withdrew from medical school, or were dismissed from medical school were not considered eligible as they did not have Step 2 CK data available prior to their departure. This certainly may have limited the statistical significance of my findings. This data would have been important to assess as these students specifically may have had low grit scores and low conscientiousness scores which could have strengthened the power of this assessment. These measurements would have informed the results of this study and increased the understanding of gritty and/or conscientious students enrolled in medical school. This would certainly be a topic for further research.

Summary

This section outlined the purpose, rationale, and questions that guided this research study. It also provided the methods in which the research questions were answered. A description of the study design, research setting, participant characteristics, the data collection process, and the management of this data were presented. Further, the rationale behind the instruments selected for this research, detail about each variable used in this study, and previous research utilizing the instruments proposed—the NEO-FFI-3 and the Grit-S scale—were explored. The two phases of research were described along with the specific statistical methods that were used. Finally, limitations of the study known prior to the analyses was presented.

CHAPTER FOUR: FINDINGS

The purpose of this study was to explore how specific non-cognitive variables, conscientiousness and grit, correlated with medical education success (defined as USMLE Step 2 CK score). This study assessed medical students at one academic medical center using two survey instruments, Duckworth and Quinn's Grit-S scale (2009) and McCrae and Costa's NEO-FFI-3 (2010a) to answer the research questions already presented in order to better understand medical student success on the USMLE Step 2 CK.

In Chapter Three the student participants and the instruments selected were described in great detail, followed by descriptions of the independent variables to be included in the analysis; sociodemographic variables, non-cognitive variables (the instruments used), academic variables, and finally a description of the dependent variable. In this chapter, the research questions were answered through specific phases of statistical analyses. In the first phase of the data analysis, descriptive statistics for each variable were used to describe the sociodemographic and academic characteristics of the study participants followed by a bivariate analysis which measured the strength of the relationship between the variables. Independent samples t-tests were performed to identify any statistically significant differences of the mean scores between participants. Finally, the effect size, Eta^2, was identified for each of the variables in order to observe any significant differences between them. While the independent samples t-test was used to determine if there were significant differences between the independent variables, effect size was used to determine if any of the differences were truly statistically meaningful (Tabachnick et al., 2007). According to Rosenthal (1996) correlation r can be treated as effect size, the greater an effect size the stronger the relationship will be. According to Cohen (1992), effect size can be presented as small, medium and large effect sizes, with the Eta^2 respectively reported as .01, .06, and .14.

In the second phase of the data analysis, the research questions were answered through linear, multiple, and biserial regressions to discover the factors that predicted success on the dependent variable, USMLE Step 2 CK. Throughout this research, the type-1 error parameters were set at $p < .05$ in order to identify significant relationships among the variables (Sprinthall, 2012). Regression analyses was utilized to answer the research questions. Regression has been identified as a powerful statistical test used in previous research related to medical student performance on the USMLE Step 2 CK examination (Dong, et al., 2014; Kleshinski et al., 2009; Simon et al., 2007).

The assumptions of multiple regression were analyzed. The first assumption, sample size, was met. According to Tabachnick and Fidell (2007), a sample size of 192 was acceptable. There were five sociodemographic variables (age, gender, first-generation college, race/ethnicity, and low socioeconomic status), five academic variables (MCAT, prior college GPA, prior college major, post-baccalaureate coursework, and master's degree), and two non-cognitive variables (grit and conscientiousness). Using N>50+8m, a sample size of 192 is sufficient for multiple regression (Tabachnick & Fidell, 2007). Next the relationship between each of the independent variables and the dependent variable was checked. Looking at each of the scatterplots, the relationship between each of the independent variables and the dependent variable, Step 2 CK, could be modeled by a straight line, which suggested that the relationship between the variables was linear.

The next assumption was to check for multicollinearity in these data. The data was checked to ensure that the independent variables were not too highly correlated. There was no correlation higher than 0.8. (Tabachnick & Fidell, 2007). According to Tabachnick and Fidell (2007), any correlation above 0.8 may be problematic. Multicollinearity was not an issue in this sample, as the highest correlation was r =.395. Further, in checking the collinearity statistics, the VIF score, which represented how well the variable is explained by other independent variables, should be well below ten. Is this sample, the highest VIF score was 2.21, while the lowest Tolerance scores was .452.

The next assumption to be tested was to see if the values of the residuals were independent (or uncorrelated) by using the Durbin-Watson statistic. For this assumption to be met, this value must be close to two (Tabachnick & Fidell, 2007). In this study, the Durbin-Watson value was 2.011, therefore this assumption has been met.

The assumption of homoscedasticity was tested to check the variation in the residuals and see if they were similar at each point in the model (Tabachnick & Fidell, 2007). This assumption considered the amount of error in the model. The variation of the residuals looked similar and appeared random, which indicated that the homoscedasticity in the model was acceptable and met this assumption.

The final assumption was checked to see if the values of the residuals were normally distributed (Tabachnick & Fidell, 2007). This assumption was tested by looking at the P-P plot of the regression model. In this plot, the dots were close to the diagonal line, which indicated that the residuals were normally distributed. Based on the above, the assumptions of linearity and normal distributions were checked and met.

Data Findings

This sample of participants was comprised from a pool of students who interviewed over two admissions cycles at Northeast Medical University using the NEO-FFI-3. These 305 students were asked to participate in this research study. As noted previously, twelve students (or 3%) of this 305 had previously left the institution. Therefore, the Short Grit Survey was emailed to 293 students in the early spring of their 3rd and 4th years and were invited to participate in this research and asked to additionally complete the Grit-S Scale; 192 students completed the survey, representing a 65.5%

participation rate. One hundred ninety-two students responded positively to participating in this research and completed the Grit-S survey. This study did not examine those students who were dismissed, withdrawn or on a leave of absence from the medical program. Those students that did participate reported varying levels of conscientiousness and grit on their respective surveys.

For those students who chose to participate in this study, 58% were men, 84% were under the age of 26, 67% were White, 20% were Asian, 9% were Black, 3% were Hispanic, and 1% identified as Native American. Thirteen percent of the participants were considered under-represented in medicine. Ten percent of the participants were first-generation college students, 32% were from low-socioeconomic backgrounds, 83% majored in the sciences as undergraduates, and 12% completed a graduate degree. The majority of participants were White males, under the age of 26, who were neither low socioeconomic status nor first-generation. At the time of this study there were more men (52.2 % in 2015 and 50.2 % in 2016) (AAMC, 2020c) than women enrolling in medical school nationally. The majority of participants were White, followed by Asian, and accordingly to national data, the average age of matriculating medical students was 24 (AAMC, 2020c). Similarly, as of 2018, the majority of physicians in the United States were male (64.1%) and White (56.2%) (AAMC, 2020d).

To assess whether the gender sample was representative of national matriculants and those that participated in this study, a two-by-two chi-square was conducted on the dichotomous gender variables. The distribution of gender in this sample was not statistically similar (df= 4, N = 192, Chi^2 = 5.032, p = .024) to the national gender breakdown of matriculated medical students. Further, a two-by-two chi-square was run to see that the participant sample did represent the population at Northeast Medical University, (df= 4, N = 192, Chi^2 = 1.103, p = .293).

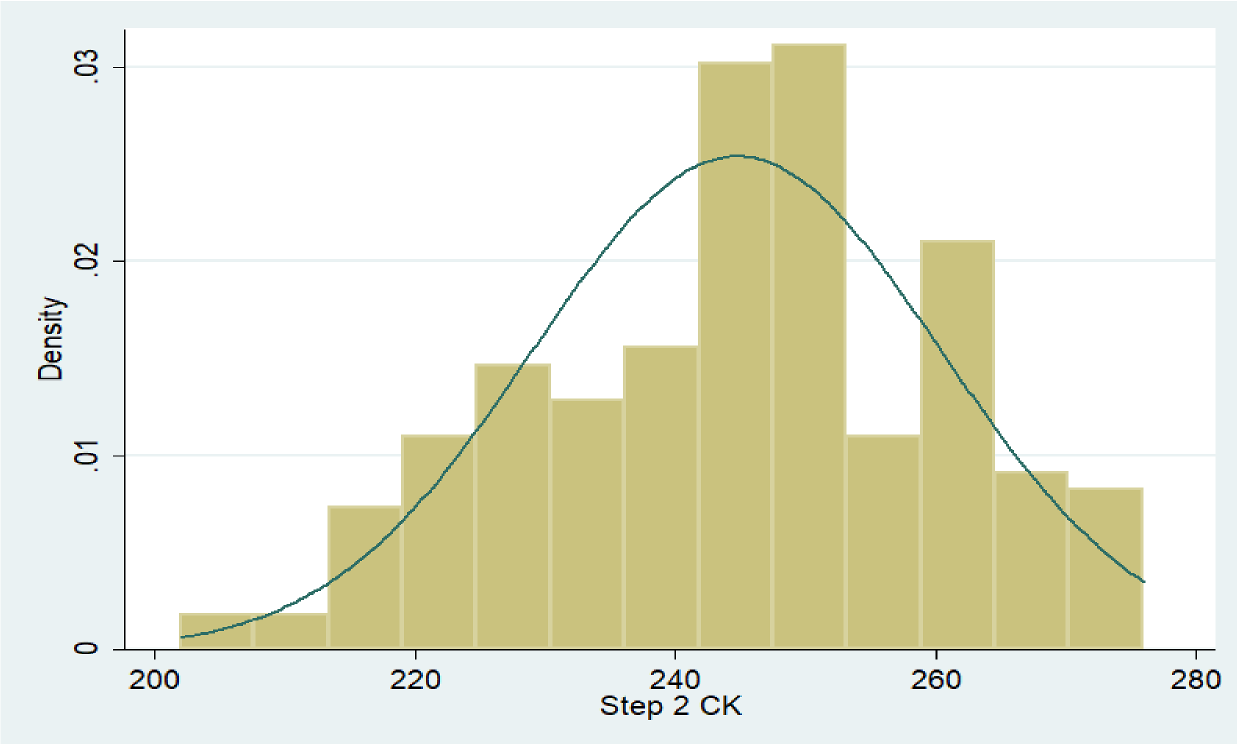
Further, to assess that the underrepresented student sample was representative of national matriculants, a two-by-two chi-square was conducted on the dichotomous variables URM and White/Asian. The chi-square test indicated that URM representation was not different when compared to national matriculants (*df*= 4, *N* = 192, Chi^2 = .032, *p* = .572). Likewise, when comparing URM study participants to the URM student body at Northeast Medical University, the sample is similar (*df*= 4, *N* = 192, Chi^2 = .258, *p* = .611). Except for gender, the sample of medical students who participated in this study did not reflect the national demographics of medical students across the US. The ethnicity/race demographic representation within the participant group is not representative of the national population nor the population at Northeast Medical University and therefore cannot be generalized across medical schools.

Descriptive Analyses of Sociodemographic Variables used in the Research Questions

Dependent Variable: USMLE Step 2 CK. The national medical school examination, USMLE Step 2 CK, was the dependent variable in this study. Students typically complete this examination between years three and four of their undergraduate medical education. This information was obtained from each consenting students' academic record. This examination had two components, the Clinical Knowledge (CK) section and the Clinical Skills (CS) section. Scoring for the Clinical Knowledge section was represented on a 3-digit scale, ranging from 1-300 (Kim & George, 2018). The failing score for the Clinical Knowledge section for all participants in this study was 208 or lower (USMLE, 2020b). The Clinical Skills (CS) section of this exam were scored as either Pass or Fail and was not included as a variable in this study.

To better understand the Step 2 CK examination dependent variable, the demographic variables of the participants were considered. Figure 2 shows the distribution of the USMLE Step 2 CK scores used in this research sample.

Figure 2. Histogram of Step 2 CK



N	Range of scores	Mean	Standard Deviation	Variance	Skewness	Kurtosis
192	204-276	244.729	15.708	246.743	-.284	2.53

The Step 2 CK scores for all participants (*N* = 192) in this study ranged from 204-276, with a mean score of 244.729 (*SD* = 15.708). Students who started with each of the cohorts and completed the NEO-FFI-3 as part of their admissions interview yet took a leave of absence, withdrew from the medical school, or were dismissed from medical prior to taking Step 2 CK (*N* = 12) were not included in this study as they had no Step 2 CK score to assess. These students left the medical program at various times and therefore these data were not considered in these analyses.

Research Question 1: Using sociodemographic variables, specifically age, gender, ethnicity, socioeconomic status, or first-generation college, to explore correlational relationships with success on the national USMLE Step 2 CK examination.

Descriptive/Correlational Statistics of Independent Variables. Two dichotomous age categories were identified: those students up to age 26, and those 26 and older. Similarly, ethnicity was broken down into two dichotomous categories: those considered underrepresented in medicine (Black/African American, Hispanic/Latinx, and Native American) and those considered to be the majority in medicine (Asian and White). Those who were considered underrepresented in medicine (URM) were grouped together for this research due to the small number of students ($N = 26$). Further, dichotomous categories for socioeconomic status and first generation were considered. No students identified their gender as other, so the dichotomous gender categories of male and female were used. Table 4.1 provides the descriptive statistics for the sociodemographic independent variables used in this study.

Table 4.1

Table 4.1: Descriptive Statistics of Step 2 CK Scores by Sociodemographic Variables (N = 192)

	N	Mean Step2CK	*Minimum*	*Maximum*	*SD*
Age					
26+	31	238.19	216	261	13.66
Below 26	161	245.99	202	276	13.80
Gender					
Male	112	246.70	204	276	15.76
Female	80	241.98	202	274	15.32
Race/Ethnicity					
URM	26	231.62	202	258	16.23
White/Asian	166	246.78	209	276	14.64
Economically Disadvantaged*					
Yes	61	240.97	202	267	14.72
No	131	246.48	212	276	15.90
First Generation College Student					
Yes	20	238.10	209	267	15.71
No	172	245.50	202	276	15.57

*Notes: Low SES is defined as EO1 or EO2 as defined by AAMC. EO1 identifies parental education levels as being less than a Bachelor's Degree and employed as an unskilled worker. EO2 identifies those with any level of education who are employed in unskilled jobs. These two classifications indicate socioeconomic disadvantage among applicants (AAMC, 2015).

As noted below in Table 4.2 below, students who were age 26 and lower earned significantly higher scores on their USMLE Step 2 CK exam (*M* = 245.99, *SD* = 13.80) than those who were 26 or older in this study (*M* = 238.19, *SD* = 13.66; *t* (190) = -2.57; *p* = .011, one-tailed). The effect size was calculated to determine if this was a meaningful difference between the test scores of those younger than 26 and those older than 26, and it was found to have a very small effect size (Eta^2 = .034) (Cohen, 1988; Leech, Barrett, & Morgan, 2011; Sprinthall, 2007), which demonstrated that the differences in Step 2 CK scores between these groups were unimportant. Students who identified as male earned statistically significantly higher Step 2 CK scores (M = 246.70, SD = 15.76) than their female counterparts (*M* = 241.98, *SD* = 15.32; *t* (190) = -2.07; *p* = .040, one-tailed), yet also had an extremely small effect size (*Eta*^{2} = .022), indicating that the true differences in their Step 2 CK scores were trivial. Those from higher socioeconomic backgrounds had a statistically significant Step 2 CK mean score (M = 246.48, SD = 15.90) compared to those from lower socioeconomic backgrounds (M = 240.97, SD = 14.72; t (190) = -2.29; p = .023, one-tailed) and again had an extremely small effect size (Eta^2 = .027), indicating that the difference in test scores was again trivial. Similarly, those students who were not considered first-generation college students had a statistically significant difference in Step 2 CK scores (M = 245.50, SD = 15.57), compared to first-generation students (M = 238.10, SD = 15.71; t (190) = -2.01; p = .046, one-tailed) and an extremely small effect size (*Eta*^{2} = .021), indicating inconsequential differences among mean test scores between these groups.

Table 4.2

Table 4.2: Differences in Means of Step 2 CK Scores by Participant Demographic Variables (N=192) via t-tests

Independent Variables		*Mean Step 2 (SD)*	*t-test*	*Df*	*Sig (p)*	*Effect Size (Eta^2)*
Total	192	244.7 (15.7)	…..	…..	…..	…..
Age						
26+	31	238.19 (13.66)	-2.57	190	.011*	.034
Below 26	161	245.99 (15.80)				
Gender						
Male						
Female	112	246.70 (15.76)	-2.07	190	.040*	.022
	80	241.98 (15.32)				
Race/Ethnicity						
URM	26	231.62 (16.23)	-4.84	190	<.001***	.011
White/Asian	166	246.78 (14.64)				
Low Socio-Economic Status						
EO1/EO2						
Not Low SES	61	240.97 (14.72)	-2.29	190	.023*	.027
	131	246.48 (15.90)				
First-Generation College Student						
	20	238.10 (15.71)				
Yes			-2.01	190	.046*	.021
	172	245.50 (15.57)				
No						

*Notes: $^+p<.10$, * $p<.05$, ** $p<.01$, *** $p<.001$*

To summarize, based on these analyses, females, under-represented students, low-socioeconomic students, and first-generation students were negatively correlated to Step 2 CK performance. More than just being negatively correlated to Step 2 CK performance, these differences were also statistically significant, especially with regard to those underrepresented in medicine, yet all variables had extremely small effect sizes indicating that the differences in means of the USMLE Step 2 CK scores were inconsequential (Slavin & Smith, 2009; Turner et al., 2018). Among medical students taking the USMLE Step 2 CK examination (N= 192), there was a statistically significant difference between underrepresented students (*M* = 231.62, *SD* = 16.23) and White/Asian students (*M* = 246.78, *SD* = 14.64), $t(190) = -4.84$, $p < .001$). Further, Cohen's effect size value ($Eta^2 = .011$) suggested an extremely small practical significance, therefore indicating that these differences were unimportant. Participants from higher socio-economic backgrounds, those who were not first in their families to attend college, and those whose ethnicities were either White or Asian performed higher on the USMLE Step 2 CK at Northeast Medical University. Those participants under the age of 26 and those that identified as male performed better on the USMLE Step 2 CK than non-traditional and female students. The difference in Step 2 CK scores between traditional (*M* = 245.99, *SD* = 15.80), $t(190) = -2.57$, $p < .011$) and non-traditional students (M = 238.19, SD = 13.66) was significant. Cohen's effect size value ($Eta^2 = .034$) suggested an extremely small practical significance in the mean differences demonstrating that this variable really had no impact on the Step 2 CK score.

Although there was a low effect size for underrepresented minority students ($Eta^2 = .110$), these data indicated that coming from a background underrepresented in medicine had a significant negative impact on how well a student performs on the National USMLE Step 2 CK exam. The normative scoring for the USMLE Step 2 CK was 195-300, with the national mean of 242 (SD = 17), (USMLE, 2019). Empirically, differences in USMLE Step 2 CK scores, even by a couple of points can make a huge impact on students and mean the difference between a very competitive residency such as orthopedic surgery or a residency in a less competitive specialty such as family medicine. The combination of demographic variables was significant (*M* = 250.586, $p < .001$). As seen in Table 4.3, when controlling for the independent variables, the largest, most significant adjusted difference in USMLE Step 2 CK score was reported in underrepresented students. URM students scored almost 15 points lower ($\beta = -14.86$, $p < .001$)

than the average Step 2 CK score. Following URM, older students scored almost 6 points lower (β = -5.85, $p < .045$), and female students scored almost 5 points lower on the USMLE Step 2 CK exam (β = -4.78, $p < .027$). While first-generation college students and students from lower socioeconomic backgrounds performed lower on Step 2 CK, (β = -4.19 and -2.11, respectively), these differences were found to be statistically non-significant. Considering these demographic variables together, explained 18% of the variance on USMLE Step 2 CK scores.

Table 4.3

Table 4.3: Predictors of USMLE Step 2 CK score, modeled via OLS Linear Regression with Sociodemographic variables

Independent Variable	*B*	*SE*	*t*	*P*	95 % *CI*
Constant	250.59	1.639	152.895	.000***	[247.353, 253.82]
Female	-4.78	2.146	-2.226	.027**	[-9.009, -.542]
URM	-14.86	3.314	-4.484	.000***	[-21.400, -8.323]
Low Socio-Economic	-2.11	2.621	-.806	.281	[-7.282, 3.060]
First-Generation	-4.19	3.876	-1.081	.045**	[-11.838, 3.458]
Non-traditional (26 or older)	-5.85	2.892	-2.023	.422	[-11.554, -.1430]

Model Summary: R^2 = .181; F=8.066, p<.001
*p < .10. * p < .05. ** p < .01. *** p < .001*

Research Question 2: Using prior academic variables, specifically MCAT scores, prior college science GPA, college major, post-baccalaureate course work, or graduate degree to explore correlational relationships with success on the national USMLE Step 2 CK examination.

Descriptive/Correlational Statistics of Independent Variables. Initially, post-baccalaureate coursework was considered in these analyses, but upon further investigation post-baccalaureate coursework was removed from this model as it was significantly highly correlated with those students 26 and older ($r=.422$, $p < 0.001$) and highly correlated with underrepresented students ($r = .173$, $p < 0.001$). Therefore, those students who pursued a post-baccalaureate program were likely the same students who were over the age of 26 and/or underrepresented in medicine in this study. Likewise, the literature suggested students who complete post-baccalaureate coursework tended to be older than traditional medical students and from underrepresented backgrounds Andre, 2020; Baill, Khallouq, Joledo, Jacobs, & Larkin, 2019; Dudkin, Bodek, Niiler, Geveke, Finneran, & Tierney, 2015; Wise, 2020).

To better understand the relationship between Step 2 CK and other independent academic variables, Table 4.4 shows that participant science GPA ranged from a minimum of 2.76 to a maximum of 4.0, (M = 3.63, SD = .255), while the overall MCAT percentiles[1] ranged from 27% to 100%, (M = 79.37, SD = 15.631). Three students who participated in this study did not have MCAT

[1] MCAT percentiles were used in this study as MCAT scoring changed in April, 2015. Prior to this date, the MCAT consisted of four sections, Verbal Reasoning, Biological Sciences, Physical Sciences, and a Writing Sample. The writing sample was graded using a letter scale which ranged from J to T, with T being the best. The other sections were scored numerically, offering individual section scores ranging from 1-15, and an overall composite score, ranging from 3-45. After April, 2015, the MCAT consisted of four sections; Chemical and Physical Foundations of Biological Systems, Critical Analysis and Reasoning Skills (CARS), Biological and Biochemical Foundations of Living Systems, and Psychological, Social and Biological Foundations of Behavior. Each of the four sections had individual scores ranging between 118-132 and a composite score that ranged from 472-528. Typically, MCAT scores are considered for up to three years Pre-April 2015 test scores were good for three years. Therefore, participants in this study had a mixture of pre-2015 scores and scores from the "new" MCAT and the only way to assess comparable was scores was to use the composite percentile scores.

scores as the MCAT was not required for the special pathway through which they entered medical school.

Table 4.4

Table 4.4: Descriptive Statistics of Science GPA & Overall MCAT Percentile

Independent Variables	N	Mean (*SD*)	Minimum	Maximum
Overall MCAT Percentile	189	79.37 (15.631)	27	100
Science GPA	192	3.632 (0.255)	2.76	4

Table 4.5 shows the *p* values for both dichotomous academic variables were very small ($p < .001$), therefore the chance of rejecting the null hypothesis was also particularly small.

Table 4.5

Table 4.5: Predictors of USMLE Step 2 CK score, modeled via OLS Linear Regression w/Academic Variables

Independent Variable	*B*	*SE*	*t*	*P*	95 % *CI*
Constant	168.35	14.552	11.569	.000***	[139.64, 197.056]
Overall MCAT Percentile	0.3	4.214	3.447	.000***	[6.211, 22.838]
Science GPA	14.53	0.069	4.348	.001***	[.164, .436]

Model Summary: $R^2 = .194; F=22.44, p=.000$

Based on the data above in Table 4.5, 14% ($R^2 = 0.143$) of the change in Step 2 CK scores was explained by MCAT percentiles; when adding science GPA ($R^2 = 0.194$), 19% of the change in Step 2 CK scores was explained. The strong positive relationships of the science GPA and MCAT percentiles on Step 2 CK can be seen in Appendix G and H, respectively. Appendix H represents collinearity between the science GPA and the MCAT percentile, therefore either variable would be sufficient in predicting success on the USMLE Step 2 CK.

Table 4.6 shows the mean Step 2 CK scores for those students who completed a master's program which were lower (*M* = 241.09, *SD* = 15.64; *t* (190) = -1.9; *p* = .237, two-tailed) than those who did not complete additional coursework (*M* = 245.22, *SD* = 15.70). An extremely small effect size (Eta^2 = .007) was found, indicating the impact on Step 2 CK for those who had completed additional coursework was irrelevant. Those participants whose undergraduate major was in the sciences (*M* = 245.53, *SD* = 16.01) performed slightly better, but did not reach statistical significance, on the USMLE Step 2 CK than those who were non-science majors (*M* = 241.03, *SD* = 13.83; *t* (190) = 1.52; *p* = 0.13, two-tailed) and again had a very small effect size (Eta^2 = .011), indicating that being a science major had little to do with USMLE Step 2 CK scores.

Table 4.6

Table 4.6: Descriptive Statistics of Step 2 CK Scores for Dichotomous Academic Variables

	N	Mean (*SD*)	*Minimum*	*Maximum*	*t-test*
Master's Degree	23	241.09 (15.64)	212	267	-1.9
No Master's Degree	169	245.22 (15.70)	202	276	.
Majored in the Sciences	158	245.53 (16.01)	202	276	1.52
Not a Science Major	34	241.03 (13.83)	217	267	.

Previous research demonstrated that students who pursued additional education prior to entering medical school tended to come from underrepresented and/or disadvantaged backgrounds (Epps, 2015; Giordani et al., 2001; Lipscomb et al., 2009; Reeves et al., 2008). As seen previously in Tables 4.2 and 4.3 these data demonstrated a strong correlation with lower Step 2 CK scores. Within this study, 42.3% of underrepresented participants either completed a post-baccalaureate or a master's program, while only 19.5% of their non-URM peers completed similar additional coursework. Students often pursue more education to combat a lower MCAT score or a lower undergraduate GPA which does not appear to correlate with success on the USMLE Step 2 CK in this study; however, low science GPAs and low MCAT scores do correlate with lower Step 2 CK scores.

As seen in Table 4.7, a point-biserial correlation showed that participants who did not complete a master's program and those participants that majored in the sciences performed better on the USMLE Step 2 CK exam, then their counterparts, yet these differences were found to be statistically non-significant.

Table 4.7

Table 4.7: Correlation between Step 2 CK Scores and Participant Dichotomous Academic Variables

	N	*df*	*Sig (p)*	Effect Size (*Eta^2*)
Master's Degree	192	190	.237	.007
Majored in Science	192	190	.13	.110

Notes. [a] point-biserial correlation for binary variables: Master's Degree and Majored in Science. It is statistically equivalent to the 2-group t-test results displayed in Table 4.6 (see Sheskin & Sheskin, 2011, pp. 1325-1328). $^{+}$ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$

In summary, research questions RQ1 and RQ2 have laid the foundation to run the regression analyses for RQ3 and RQ4 by providing an analysis of demographic and academic variables on USMLE Step 2 CK performance. In RQ3, variables included in RQ1 and RQ2 were added to the first and second blocks of the hierarchical regression to understand how *conscientiousness* contributed to Step 2 CK scores, given the demographics and academic background information. In RQ4, a hierarchical regression explored how *grit* contributed to our understanding of Step 2 CK performance given the demographics and academic background variables in the model.

Research Question 3: A three-block hierarchical regression used conscientiousness, as measured by Costa and McCrae NEO-FFI-3 (1992), to predict medical student academic success on the USMLE Step 2 CK medical school exam at Northeast Medical University when controlling for demographics and prior academics.

Descriptive/Correlational Statistics of Conscientiousness. As previously defined, conscientiousness is the personality trait of being diligent, deliberate, hard-working, persistent, and taking obligations to others seriously, qualities that patients may find important when looking for physicians to care for them and/or their families. The NEO-FFI-3 was utilized during the admissions process for Northeast Medical University. As noted in Chapter Three, only the conscientiousness facet was used in the research. Haight et al., (2012) found MCAT scores correlated with academic examinations, whereas scores on the conscientiousness facet correlated with indicators of clinical performance and humanism nominations; more specifically, conscientiousness predicted clinical skills. Not surprisingly, conscientiousness is also a strong predictor of academic performance at the graduate school level (Grehan et al., 2011). Therefore, it was used in this study to assess its impact on participants' USMLE Step 2 CK score.

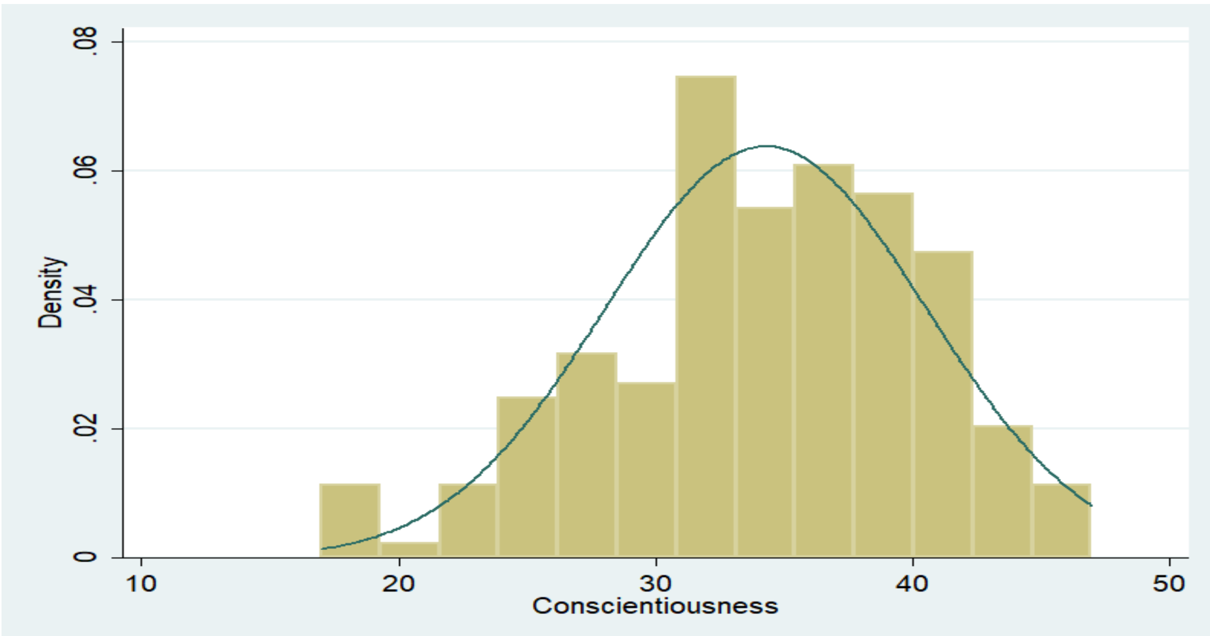
Table 4.8

Table 4.8: Descriptive Statistics of Noncognitive Variable Conscientiousness (N = 192)

	N	Minimum	Maximum	Mean	*SD*
Conscientiousness	192	17	47	34.29	6.25

Conscientiousness in medical students was assessed using the NEO-FFI-3 during the interview process at Northeast Medical University. The conscientiousness section of the NEO-FFI-3 is comprised of twelve questions related to order, dutifulness, achievement, and self-discipline (McCrae & Costa, 2010a). As demonstrated in Table 4.8, the conscientiousness scores of participants on the NEO-FFI-3 ranged from 17 to 47, with a mean score of 34 (SD = 6.25). The histogram distribution below in Figure 3 has a leftward skew (-.441) with most participants falling between 28 and 40.

Figure 3. Histogram of Conscientiousness



According to these data, women (*M* = 34.63, *SD* = 6.19) reported more conscientious behaviors than men (*M* = 34.05, *SD* = 6.31), but not at a statistically significant level (*p* = .534). While also not significant (*p* = .583), those from lower socioeconomic backgrounds (*M* = 34.66, *SD* = 5.28) reported more conscientious behaviors than those from higher socioeconomic backgrounds (*M* = 34.12, *SD* = 6.67). First generation participants (*M* = 34.80, *SD* = 4.86) were slightly more conscientious than their non-first-generation counterparts (*M* = 34.23, *SD* = 6.403) yet this difference was also not significant (*p* = .702). Conscientiousness shows no statistical significance (*p* = .804) for those 26 and older (*M* = 34.55, *SD* = 5.54) than their younger peers (*M* = 34.24, *SD* = 6.39). Prior to beginning this study, this researcher hypothesized that under-represented students would be more conscientious than their majority peers. These data did not support that hypothesis.

Although the numbers are small, Native American participants (*N* = 3) had a mean conscientious score of 34.67 (*SD* = 3.22), Black or African American participants (*N* = 17) followed with a mean of 33.88 (*SD* = 5.45), and Hispanic participants (*N* = 6) had a mean of 34.67 (*SD* = 4.37). Of those identified as majority students in this study, White students (*N* = 128) had mean conscientious scores of 35.00 (*SD* = 6.31), while Asian students had the lowest conscientious score with an average score of 32.00 (*SD* = 6.47). When comparing the two groups, underrepresented students (*N* = 26) had an average conscientious score of 34.09 (*SD* = 5.11) and majority students (*N* = 166) had an average conscientious score of 34.31 (*SD* = 6.45), with a $p = .872$, indicating there was a difference, but that difference did not reach statistical significance.

Further, the correlation of conscientiousness among academic variables demonstrated that neither science major ($p = .091$, $r = .098$) nor completion of a master's degree ($p = .105$, $r = -.091$) correlated with an individual's level of conscientiousness. While participants science GPA ($p < .001$, $r = .335$) and overall MCAT percentile score ($p < .001$, $r = .378$) indicating that an individual's level of conscientiousness had a moderate effect on USMLE Step 2 CK performance.

Hierarchical regression was run to investigate whether conscientiousness (added in the third block) further predicted USMLE Step 2 CK scores when controlling for sociodemographic independent variables and prior academic independent variables in the first and second blocks. The assumptions of linearity and normal distributions were checked and met.

Table 4.9

Table 4.9: Multivariate Correlation between Step 2 CK, Conscientiousness, and Individual Independent Variables

Independent Variable	Effect Size *(r)*	*Sig (p)*	*Df*
Conscientiousness	0.208	0.002***	189
Female	-0.160	0.014**	189
URM	-0.342	<0.001***	189
Low socioeconomic status	-0.175	0.008***	189
First generation college	-0.151	0.019**	189
Non-traditional (26 or older)	-0.192	0.004***	189
Science majors	0.098	0.091+	189
MCAT percentiles	0.378	<0.001***	189
Completed master's degree	-0.091	0.105	189
Science GPA	0.335	< 0.001***	189

Notes. $^{+}$ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$

These results in Table 4.9, revealed that individuals who reported more conscientious behaviors were statistically more likely to have had a higher science (BCPM) GPA from their undergraduate studies and were more likely to perform higher on the Step 2 CK. Prior data has shown that individuals with a higher level of conscientiousness tended to perform better in their undergraduate studies (Ivcevic & Brackett, 2014; Komarraju et al., 2009; Komarraju et al., 2011).

Regression Results with Conscientiousness in the Model

The Model Summary (Table 4.10) shows that the first block, demographic variables (age, gender, first-generation, and low socioeconomic), significantly predicted 18% (R^2 = .183) of the total variance in Step 2 CK scores, F (5, 183) = 8.204, $p < .001$. When the academic variables were added to the model (science major, science (BCPM) GPA, master's degree, MCAT percentile), it accounted for an added 25% of the variance (R^2 change = .069), F (4, 179) = 4.103, $p = .003$, which moderately predicted Step 2 CK scores. When conscientiousness was added in the third block of the equation, the variables slightly predicted Step 2 CK scores, F (1, 178) = 7.159, $p = .008$ and slightly improved the prediction by 3% over the demographic and academic variables, to 46% (R^2), and resulted in a small effect size (r = 0.208) (Cohen, 1988). This showed that conscientiousness had a statistically significant impact on Step 2 CK scores, indicating that the difference between the means was not due to chance. Yet because of the small effect size, an individual's score on the conscientiousness variable was not a meaningful variable to explain USMLE Step 2 CK scores when holding the other variables constant.

Table 4.10

Table 4.10: Model Summary of Hierarchical Regression with Conscientiousness on Step 2 CK

				Change Statistics				
Block	R	R^2	Std. Error of the Estimate	R^2 Change	F Change	df1	df2	*Sig.* (F Change)
1	.428	0.183	14.231	0.183	8.204	5	183	<0.001**
2	.502	0.252	13.772	0.069	4.103	4	179	0.003*
3	.530	0.281	13.541	0.029	7.159	1	178	0.008

Notes. ** Significant at the .01 level (p<.01). * Significant at the .05 level (p<.05). + Significant at the .10 level (p<.10)

1. Predictors: (Constant), 26andOver, FirstGeneration, Gender, Ethnic, LowerSES
2. Predictors: (Constant), 26andOver, FirstGeneration, Gender, Ethnic, LowerSES, ScienceMajor, BCPM GPA, CompleteMastersProg, Overall MCAT Percentile
3. Predictors: (Constant), 26andOver, FirstGeneration, Gender, Ethnic, LowerSES, ScienceMajor, BCPM GPA, CompleteMastersProg, Overall MCAT Percentile, Conscientiousness

When holding all independent variables constant, conscientiousness (M = 34.28, SD = 6.299) was positively correlated with Step 2 CK (M = 244.89, SD = 15.535), F(10,178) = 6.945, p < .002, adjusted R^2 = .24, representing a small effect size (Cohen, 1988) and a small impact on Step 2 CK performance. According to the change in R^2, adding conscientiousness to the model only represented 28% of the variance on Step 2 CK, leaving 72% of the variance unexplained by these analyses. This finding indicated that there is more to learn about what contributes to and/or predicts USMLE Step 2 CK.

Note, the science major was not significantly correlated with performance on the conscientiousness scale (p = .499), but science GPA was strongly significantly correlated with conscientiousness (p = .001).

Research Question 4: A three-block hierarchical regression used grit, as measured by Duckworth and Quinn's Short Grit Scale (2009), to predict medical student academic success on the USMLE Step 2 CK medical school exam when controlling for demographics and prior academics.

Descriptive/Correlational Statistics of Grit. Duckworth et al. (2007) postulated that grit is an individual's disposition to be passionate and persevere when working toward long term goals. Duckworth et al. (2007) posited that grit is different from other non-cognitive factors as it is a skill and is associated with an individual's educational achievement (Duckworth & Quinn, 2009). As medical school is a long and arduous endeavor, this hypothesis suggests that medical students are required to demonstrate grit. Those that have had a longer road or "traveled a greater distance" may have accrued grit to help navigate challenges that lay ahead including the preparation required to be successful in medical school and on the USMLE Step 2 CK.

The Grit-S survey was scaled according to Duckworth and Quinn (2009), and the scores were found to be within the ranges found in previous studies about Grit-S (Duckworth & Quinn, 2007; Strayhorn, 2013). Table 4.11 displays the basic descriptive statistics of the independent variable, grit.

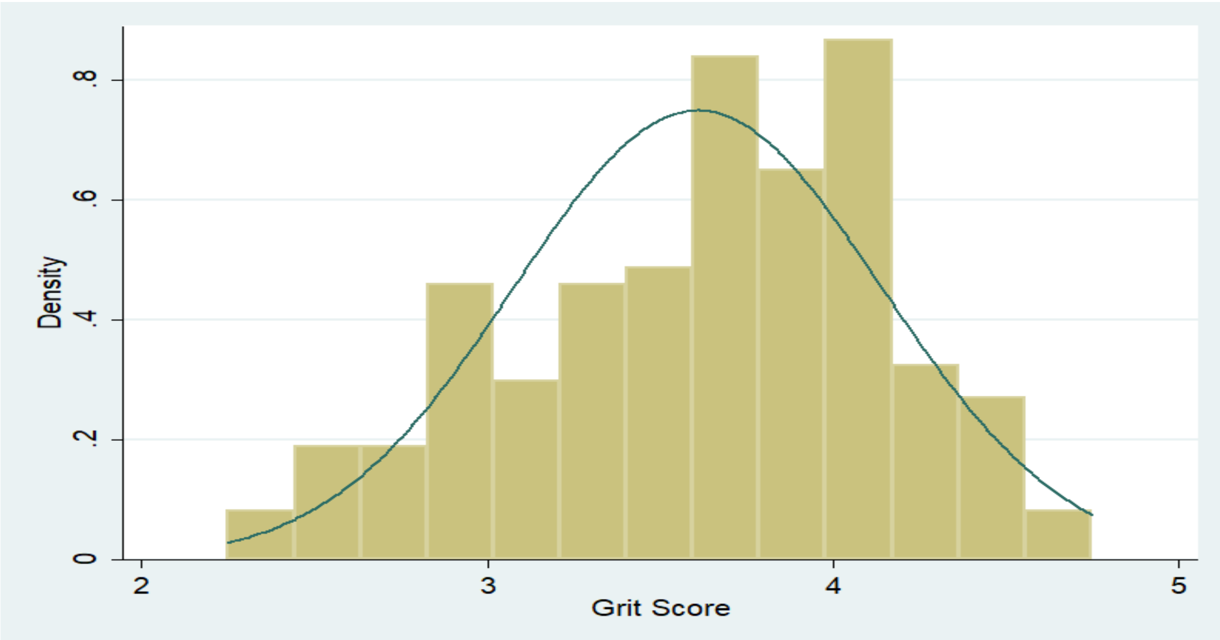
Table 4.11

Table 4.11: Descriptive Statistics of Noncognitive Variable Grit (N = 192)

	N	Minimum	Maximum	Mean	*SD*
Grit	192	2.25	4.75	3.61	.53

The histogram below skewed (-.441) to the right indicating that those who participated in this study had a higher than average level of grit. In this study (*N* = 192), the grit scores ranged from 2.250 to 4.750, with mean grit score of 3.61 (*SD* = .53; see Figure 4).

Figure 4. Histogram of Grit



According to these data, women (*M* = 3.675, *SD* = .530) were grittier than men (*M* = 3.562, *SD* = .531), but not to a significant degree (*p* = .148). While also not significant (*p* = .159), those from lower socioeconomic backgrounds (*M* = 3.68, *SD* = .524) were grittier than those from higher socioeconomic backgrounds (*M* = 3.57, *SD* = .533). First generation participants (*M* = 3.64, *SD* = .523) were slightly grittier than their non-first-generation counterparts (3.60, *SD* = .522) yet this difference was not significant (*p* = .759). Prior to beginning this study, this researcher anticipated that under-represented students would be grittier than their majority peers. The data verify the null hypothesis. Although the numbers are small, Native American participants (*N* = 3) had a mean grit score of 3.92 (*SD* = .260), Black or African American participants (*N* = 17) followed (*M* = 3.85, *SD* = .346), and Hispanic participants (*N* = 6) had the lowest mean (*M*= 3.67, *SD* .735). Those identified as majority students in this study, White students (N = 128) had mean grit scores of 3.61 (SD = .527), while Asian students had the lowest grit score with an average of 3.47 (SD = .571). When comparing the two groups, underrepresented students (N = 26) had an average grit score of 3.80 (SD = .465) and majority students (N = 166) had an average grit score of 3.58 (SD = .539), with a p = .062 indicating there was a difference, but did not reach statistical significance.

The correlation of grit among the sociodemographic independent variables showed that those 26 and older (p = 0.001) were statistically more likely to possess more grit than their peers, which makes sense as their greater life experiences and maturity would likely cause them to be grittier than their peers, yet this grittiness did not seem to affect their Step 2 CK score. There was a similar correlation with grit among those participants from low socioeconomic backgrounds (p = .079) and females (p = .074), while under-represented minority students appeared grittier (p = 0.031). First-generation student status (p = .380) was not at all correlated with grit. Further, the correlation of grit among academic variables demonstrated that science major (p = .496), masters-degree prepared participants (p = .340), nor science GPA (p = .435) correlated with an individual's grit. While a participant's overall MCAT percentile score was highly significantly correlated with grit (p = .005) there was a negative small effect size (r = -.205), indicating that grit is trivial to MCAT score performance.

Based on the information in Table 4.12, grit does not appear to correlate with Step 2 CK as it was not statistically significant (p = 0.21), and also had an extremely small effect size (r = .05), indicating that grit was unimportant on Step 2 CK performance. Grit did not correlate with success on this standardized examination, USMLE Step 2 CK.

Table 4.12

Table 4.12: Multivariate Correlation between Step 2 CK, Grit, and Individual Independent Variables

	Effect Size *(r)*	*Sig (p)*	*df*
Grit	0.059	0.207	189
Female	-0.149	0.020*	189
URM	-0.331	<0.001***	189
Low socioeconomic status	-0.164	0.012**	189
First generation college	-0.144	0.023**	189
Non-traditional (26 or older)	-0.183	0.006***	189
Science Majors	0.11	0.065+	189
MCAT percentiles	0.378	<0.001***	189
Completed Master's Degree	-0.086	0.118	189
Science GPA	0.331	<0.001***	189

Notes. $^{+}$ p < .10. * p < .05. ** p < .01. *** p < .001

There was a very significant (p<.001) negative effect size for URM students (r = -.331), with those 26 and older (p<.001) following closely behind (r = -183). Female students (p = 0.020), low SES students (p = 0.012), and first-generation college students (p = 0.023), were all significantly correlated with grit and had negative effect sizes, with URM students being the largest (r = -.331). The strongest and most statistically significant positive correlation with Step 2 CK scores continued to be the overall MCAT percentile (p < 0.001) and science GPA (p < 0.001).

Regression Results with Grit in the Model

The Model Summary below (Table 4.13) shows that the first block, demographic variables (age, gender, first-generation, ethnicity, and low socioeconomic), significantly predicted 16% (R^2 = .160) of the total variance in Step 2 CK scores, F (4, 184) = 8.754, p < .001. When the academic variables were added to the model (science major, science (BCPM) GPA, master's degree, MCAT percentile), it accounted for an added 9% of the variance (R^2 change = .085), F (5, 179) = 4.053, p = .002, which moderately predicted Step 2 CK scores. When grit was added in the third block of the equation, the variables slightly predicted Step 2 CK scores, F (1, 178) = 4.084, p = .045, and slightly improved the prediction by 2% over the demographic and academic variables, to 27% (R^2 = 0.262), and had a small to medium effect size, indicating that grit had a moderate impact on Step 2 CK performance (Cohen, 1988).

Table 4.13

Table 4.13: Model Summary of Hierarchical Regression with Grit on Step 2 CK

				Change Statistics				
Block	*R*	R^2	Std. Error of the Estimate	R^2 Change	*F* Change	*df*1	*df*2	*Sig.* (*F* Change)
1	0.4	0.16	14.93	0.16	8.754	4	184	<0.001**
2	0.495	0.245	13.831	0.085	4.053	5	179	0.002*
3	0.512	0.262	13.713	0.017	4.084	1	178	0.045*

Notes. ** Significant at the .01 level (*p*<.01). * Significant at the .05 level (*p*<.05). + Significant at the .10 level (*p*<.10)

1. Predictors: (Constant), 26andOver, FirstGeneration, Gender, Ethnic, LowerSES
2. Predictors: (Constant), 26andOver, FirstGeneration, Gender, Ethnic, LowerSES, ScienceMajor, BCPM GPA, CompleteMastersProg, Overall MCAT Percentile
3. Predictors: (Constant), 26andOver, FirstGeneration, Gender, Ethnic, LowerSES, ScienceMajor, BCPM GPA, CompleteMastersProg, Overall MCAT Percentile, Grit

Summary

This chapter reviewed the data analyses of Step 2 CK scores of 192 medical students from Northeast Medical University to answer the research questions. First the dependent and independent variables used in these analyses were described, including sociodemographic variables, academic variables, and the two non-cognitive tools utilized in this study: NEO-FFI-3 to measure conscientiousness, and Grit-S scale to measure an individual's grit. Then the relationships among the independent variables and the dependent variable were examined utilizing correlations. Finally, this researcher answered each of the research questions through multiple regressions in order to determine which factors were predictive of the dependent variable, scores on the USMLE Step 2 CK. Based on this research, those students who came from backgrounds considered underrepresented in medicine will not perform as well on the Step 2 CK as their non-minority peers. Not surprisingly, MCAT percentile and science GPA strongly predicted success on the USMLE Step

2 CK, those who majored in the sciences did as well. Likewise, students with a higher level of conscientiousness performed better on the Step 2 CK and grit moderately predicted Step 2 CK performance when controlling for sociodemographic and academic independent variables. Effect sizes for all independent variables were found to be small to extremely small in all analyses indicating that the independent variables in this study really do not matter when it comes to explaining how participants perform on the USMLE Step 2 CK examination. In the final chapter of this study, these results are discussed along with the strengths and limitations of this research, implications for medical school admissions processes, and suggestions for further research on this topic.

The Summary Table of Findings and Summary of Research Questions (Tables 4.14 and 4.15) present the key data reported in this chapter.

Table 4.14

Table 4.14: Summary Table of Findings

Analyses	**Findings (p < .05)**
Participant Demographics	· Sample is primarily White men under the age of 26 · Majority from educated families with higher income levels · Majority majored in the sciences and did not pursue additional education.
Dependent Variable	
USMLE Step 2 CK	· Range of scores were 204-276, with an average of 244.73 and an *SD* of 15.71 · Males (p = .040, Eta^2 = .022) scored statistically significantly higher than females, individuals under the age of 26 (p = .011, Eta^2 = .034) outperformed those aged 26 and higher, and White/Asian students (p < .001, Eta^2 = .011) outperformed their URM peers. Although statistically significant, the impact of these differences was trivial. · Students from higher socioeconomic backgrounds (p = .023, Eta^2 = .027) and those from educated families (p = .046, Eta^2 = .021) had statistically significantly higher Step 2 CK scores, yet the magnitude of Step 2 CK mean differences were negligible. · Participants with master's degrees (p < .237, Eta^2 = .007) scored lower on the USMLE Step 2 CK and was found to be statistically non-significant · Participants who majored in the sciences (p < .13, Eta^2 = .110) scored higher on Step 2 CK, yet did not reach statistical significance.
Independent Variables	

Conscientiousness	· Students scores ranged from 17-47, with an average score of 34.29 and a *SD* of 6.25
	- No statistically significant differences in conscientiousness when considering sociodemographics
	    · Females, lower SES, first-generation, and those over 26 reported more conscientious behaviors than their peers.
	    · URM students reported slightly lower conscientious levels than majority students
	- No statistically significant differences in conscientiousness were found when considering science majors and completed master's degrees

	- Students who reported more conscientiousness had statistically significantly higher science GPAs (p <.001, Eta^2 = .335) and higher MCAT percentiles (p < .001, Eta^2 = .378), with moderate effect sizes. · Conscientiousness levels were statistically significant to Step 2 CK performance (p < .002, Eta^2 = .208). Small effect size represented a trivial impact on mean Step 2 CK scores.
Grit	· Students scores ranged from 2.25-4.75, with an average score of 3.61 and a *SD* of .53 · Females (p < .020, Eta^2 = -.149), students from lower SES backgrounds (p < .012, Eta^2 = -.164), first-generation (p < .023, Eta^2 = -.144), URM students (p < .001, Eta^2 = -.331), and those over the age of 26 (p < 0.006, Eta^2 = -.183) were significantly statistically grittier than their peers. All had negative effect sizes. Grit was irrelevant for these variables. - No statistically significant differences in grit were found for those with master's degrees - Science major was found to be statistically non-significant - MCAT percentile and science GPA was highly correlated with Step 2 CK performance with a moderate effect size (p < .001, Eta^2 = .378) and (p < .001, Eta^2 = .331), respectively - Grit was not correlated with performance on USMLE Step 2 CK

Table 4.15

Research Question 1	
DV = Step 2 CK IV = Socio-demographics (Age, Gender, Ethnicity, SES, & First Gen)	· Students from higher income levels (p = .023), those from educated families (p = .046), males (p = .040), White/Asian students (p < .001), and those under the age of 26 (p = .011) performed significantly higher than their peers on the USMLE Step 2 CK.
Research Question 2	
DV = Step 2 CK	· Science GPA (p < .001, Eta^2 = .335) and MCAT percentiles (p < .001, Eta^2 = .378) were statistically significantly correlated to USMLE Step 2 CK performance with small to medium effect sizes
IV = Academics (Science GPA, MCAT percentiles, Master's degrees, post-bacc, science major)	· Science majors and master's degrees did not predict USMLE Step 2 CK
Research Question 3	
DV = Step 2 CK IV = Conscientiousness	· Conscientiousness is statistically significantly predictive of USMLE Step 2 CK performance (p = 0.002, Eta^2 = .208), with a small effect size, indicating that the magnitude of conscientiousness on Step 2 CK performance was trivial.
Research Question 4	
DV = Step 2 CK IV = Grit	· Grit is not predictive of USMLE Step 2 CK

CHAPTER FIVE: DISCUSSION

This research study explored whether the noncognitive variables of conscientiousness or grit could be used to supplement the academic variables that are considered for admission to medical school in order to predict student success on the national USMLE Step 2 CK examination. Over the last decade, medical schools have begun to consider a more holistic review of medical school applicants (Albanese et al., 2003; Monroe et al., 2013) in light of the predicted shortage of physicians (Lakhan & Laird, 2009; Salsberg & Grover, 2006) as well as the need to diversify the physician workforce to meet the needs of our increasingly diverse communities (Bore et al., 2009; Cohen & Steinecke, 2006; Conrad et al., 2016; Cooper, 1994; Gonzalez & Stoll, 2002; Kirch et al., 2012: Kirch & Petelle, 2017; Koenig et al., 2013; Reed et al., 2012; Witzburg & Sondheimer, 2013). Along the same lines, a holistic approach in medical admissions may lead to less resident burnout (Cortez et al., 2020; Leigh et al., 2002; Shanafelt et al., 2012; Zuger, 2004) and also retain more medical students (Maher et al., 2013; Kruzicevic, et al., 2012) over the course of their laborious undergraduate medical education journey. Several measures have been shown to predict a students' academic success, including their personality, motivation, and past experiences (Robbins et al., 2004). In order for medical schools to meet the workforce needs, a more flexible admissions process should be considered to allow for greater diversity in the workforce as well as to build a workforce who can serve their patients well both from a medical knowledge and a clinical skills perspective.

This study demonstrated that the traditional, cognitive factors of MCAT and science GPA showed the most significant results in terms of influence on USMLE Step 2 CK performance, yet the differences in the means were negligible. In terms of medical school admission and identifying students who will be successful academically as well as clinically, there is still much work to be done. This work is especially imperative in relation to the recruitment and retention of underrepresented students in order to meet the diverse physician workforce needs. Considering Sternberg's theory of successful intelligence and Yosso's asset approach to critical race theory in the admissions processes is crucial. In terms of admission, based on my practical experience, I believe that it is important that we push to identify variables that consider various ways in which to assess intelligence and the various assets that students, particularly those from diverse backgrounds, bring to medical school. This study attempted to consider additional ways to measure medical student success by utilizing the constructs of grit and conscientiousness, but were found to not be predictors of USMLE Step 2 CK success. I still believe that these non-cognitive variables, grit and conscientiousness, make a difference in the success of medical students. The results of this study may have been impacted by the specific instruments used, the small sample size, the high pass rate of Step 2 CK which ultimately produced limited variability, and perhaps by the individuals that were not included in this study. Therefore, identifying valid and reliable measures, better instruments, assessments, and sampling methods, must be found and utilized to move this important work forward and help to remedy the medical disparities across our country.

This correlational study was conducted at a College of Medicine at an Academic Medical Center in the Northeast. As part of the admissions process, the NEO-FFI-3 (Costa & McCrae, 1992) was utilized. For purposes of this study, as previously discussed, the conscientiousness section of the NEO-FFI-3 was the only section used in these analyses. Further, students completed the Grit-S survey (Duckworth & Quinn, 2009) during their third or fourth year of medical school. The primary purpose of this study was to ascertain how these non-cognitive factors influenced medical student success, as defined by passing scores on the USMLE Step 2 CK national examination.

This chapter analyzes the findings presented in Chapter Four. First, the participants in this study are discussed in relation to the national population of medical students. Then the discussion focuses on this study's descriptive statistics on sociodemographic and academic variables as well as the non-cognitive variables of conscientiousness and grit and the predictive success on the USMLE Step 2 CK. Next, the findings from the four research questions are analyzed to understand how they align with previous research studies. Following these analyses, implications of this study on medical student success are discussed. This study's strengths and limitations are presented and areas for future research are considered. Finally, closing thoughts about this work and medical student success are further discussed.

Variations Among Participants

Among the independent variables studied, there were some interesting differences noted in this research. When looking at the sociodemographic variables and the differences associated with conscientiousness, women in this study were significantly more conscientious then men ($p = .014$) (Keiser, Sackett, Kuncel, & Brothen, 2016); those from lower socio-economic backgrounds ($p = .008$) (Jackson et al., 2009) and those considered first-generation college students ($p = .019$) (Kessler, 2003) reported significantly more conscientious behaviors than their peers. For those participants 26 and older, it was found that they were slightly more conscientious than those under the age of 26 (Jackson et al., 2009), yet not significantly different. Similarly, students who identified as White were slightly more conscientiousness than their peers, yet not significantly so (Ivcevic, & Brackett, 2014). Students who identified as Asian were the least conscientious ethnic group in this study. When looking at academic variables and conscientiousness, neither those who completed a master's degree nor those who majored in the sciences correlated with a student's level of conscientiousness.

When considering sociodemographic variables and student's levels of grit, women demonstrated statistically significantly more grit than men ($p < .020$, $r = -.149$) (Miller-Matero et al., 2018), those from lower socio-economic backgrounds ($p < .012$, $r = -.164$) and first-generation college students ($p < .023$, $r = -.144$) were considered statistically significantly grittier than their peers. For those participants 26 and older, it was found that they scored statistically significantly higher on grittier behaviors ($p < 0.006$, $r = -.183$) than those under the age of 26 (Duckworth et al., 2007). Although not statistically significant, students who identified as Native American demonstrated slightly higher grit behaviors than their peers, and students who identified as Asian reported the fewest number of grit behaviors than the other ethnic/racial groups in this study. When all under-represented students were assessed together, they were significantly grittier ($p < .001$, $r = -.331$) than their White and Asian peers ($p = .000$) (Bowman, Hill, Denson, & Bronkema 2015; Hill, Burrow, & Bronk, 2016; Ivcevic, & Brackett, 2014; Miller-Matero et al., 2018). Although these sociodemographic variables demonstrated statistical significance, they each had negative small to moderate effect sizes indicating that the mean differences between groups were trivial when considering grit. When looking at academic variables and a student's grit, there was no statistically significant difference among those who earned a previous master's degree or majored in the sciences and their peers.

Direct comparisons of the various sociodemographic variables were made on the USMLE Step 2 CK exam. The analyses revealed that White/Asian men from wealthier, educated families out-performed their peers on their USMLE Step 2 CK exam (Alvarado, Capozza, Jackson, & Russell, 2008; AAMC, 2020h; Norcini, Boulet, Opalek, & Dauphinee, 2014; Richardson, 2017). The results of this study were similar to Ogunyemi and Taylor-Harris's findings (2005), reporting a negative correlation between age and Step 2 CK score. However, the finding that older students performed significantly lower than their younger peers on the USMLE Step 2 CK was contrary to a Kleshinski et al. study (2009), which found that age was a significant positive predictor of Step 2 CK scores.

Direct comparisons were also assessed among the independent academic variables on the dependent variable, USMLE Step 2 CK. MCAT scores and undergraduate science GPA were strong predictors of participants' USMLE Step 2 CK scores. These findings were consistent with previously reported studies (Donnon et al., 2007; Julian 2005; and Kleshinski et al., 2009).

Although participants who majored in the sciences performed better on Step 2 CK, this was not a significant predictor of success. This study revealed that having completed a master's degree had no impact on Step CK performance, which was similar to prior studies showing no association among students who completed post-baccalaureate coursework or other graduate work and medical student success (Giordani et al., 2001; Johnson et al., 2017; Kulesza et al., 2015; Sadik et al., 2017).

As grit has previously been defined by Duckworth and Quinn as "trait-level perseverance and passion for long-term goals" (2009, pp. 166), finding a higher level of grit among under-represented students was not surprising. These students tend to be very intrinsically-motivated and often work autonomously toward their long-term goals, particularly regarding their academics (Tough, 2012). This finding is consistent with similar studies assessing grit on the academic success of Black/African American students (Akos & Kretchmar, 2017; Powell, 2013; Strayhorn, 2013). Similarly, in a recent study, Reed and Jeremiah (2017) referenced the success of the Fisk-Vanderbilt Masters to PhD Bridge program and contributed its minority student success to mentorship and student grit, although grit was not empirically assessed.

The results of each research question are addressed in the following section.

Research Questions and Findings

The first two research questions were explored using correlational statistics of the independent variables on the dependent variable, USMLE Step 2 CK. Research questions Three and Four were investigated using hierarchical regression to discover variables that are predictive of medical student academic success on the USMLE Step 2 CK. The results of this study are organized by research question and discussed in the context of the relevant literature.

Research Question 1. Do socio-demographic variables—specifically age, gender, ethnicity, socioeconomic status, or first-generation college status—correlate with success on the national USMLE Step 2 CK medical school examination at Northeast Medical University?

The null hypothesis was that socioeconomic variables – females, those 26 and older, under-represented students, low socioeconomic status, or first-generation college status – did not predict success as measured by the dependent variable, USMLE Step 2 CK. Based on the correlational analyses, the null hypothesis was accepted for females, those 26 and older, under-represented students, low-socioeconomic students, and first-generation students were-negatively correlated and statistically significant to Step 2 CK It was rejected for the following variables: male students, non-minority students, students coming from households with educated parents, and higher socioeconomic students (those not identified as EO1 or EO2). Some of these results align with a recent study by Rubright et al. (2019) of 45,154 medical students. The authors found that participants from higher socio-economic backgrounds, non-first-generation college participants, those who identified as either White or Asian, and those 26 and above performed higher than their corresponding group of peers on the USMLE Step 2 CK. Inconsistent with this study, however, they reported that women were predicted to have higher Step 2 CK scores than men (Rubright et al., 2019).

Within this study, participants from higher socio-economic backgrounds (p = .023), those from educated families (p = .046), males (p = .040), White/Asian students (p < .001), and those under the age of 26 (p = .011) performed statistically significantly higher on the USMLE Step 2 CK national examination than their peers. The differences may be similar to the differences one might expect on undergraduate college entrance exams, such as the on the SAT and/or ACT standardized examinations, by these same groups of students (Deil-Amen & Tevis, 2010; Espenshade & Chung, 2010; Meredith, 2008). Although the sample size was small, holding all other variables constant, these data indicated that coming from a background under-represented in medicine will have the biggest impact on how well a student will perform on the National USMLE Step 2 CK exam. These data build upon recent studies by Jerant et al. (2019) and Harrison (2019) which suggested that looking at a variety of factors is imperative to achieving greater diversity within the physician workforce of the United States. Similarly, these findings build upon the previously mentioned study conducted by Rubright et al. (2019) who assessed demographic differences and prior academic achievement on USMLE Step scores for all three examinations. The authors suggested that additional research must include identifying factors that may contribute to demographic differences in test performance while identifying students who may need additional intervention prior to taking their USMLE exams, and the need to widen the lens when assessing applicants for residency placement in order to meet diversity goals (Rubright et al., 2019). The authors suggested that, in addition to their academic backgrounds, the personal attributes and experiences of applicants be considered during the admissions process, which speaks to the value of the EAM Model previously discussed in Chapter One.

Research Question 2. Do academic variables—specifically MCAT scores, prior college science GPA, college major, post-baccalaureate course work, or graduate degree—correlate with success on the national USMLE Step 2 CK medical school examination at Northeast Medical University?

The null hypothesis was prior academic variables – MCAT scores, science GPA, college science major, post-baccalaureate course work, or a graduate degree – do not predict medical school success as measured by the dependent variable, USMLE Step 2 CK. Based on the correlational analyses, the null hypothesis was rejected for those participants with higher science GPAs and those with higher MCAT scores (percentiles). These results are consistent with previous research finding that MCAT scores and prior college science (BCPM) GPA were significantly correlated with USMLE Step 2 CK scores (Burk-Rafel et al., 2019; Ghaffari-Rafi et al., 2019; Rubright et al., 2019). The null hypothesis was supported for those students that completed master's degrees, post-baccalaureate programs, and those who majored in the sciences. Further, those who majored in the sciences did not perform significantly different on Step 2 CK. This finding is consistent with the Hirshfield et al. (2019) study which did not find any association between undergraduate major and Step 2 CK scores.

Based on the results from this study, pursuing a graduate degree prior to entering medical school actually has a negative impact on USMLE Step 2 CK scores. Similar to this study, Epps (2015) also demonstrated that participants who pursued additional graduate level college coursework had a significant negative impact on Step 2 CK success unlike those students who proceeded to medical school directly from their undergraduate programs. These findings are also consistent with a more recent study by Orozco Cortes (2019), where a negative association between pursuing a graduate degree and USMLE Step 1 scores was found. Based on these studies, it is suggested that students who completed master's degrees will not perform as well as those without the advanced degree on both the USMLE Step 1 and the USMLE Step 2 CK national examinations.

Research Question 3. Does conscientiousness, as measured by Costa and McCrae's Five-Factor Personality Inventory (NEO-FFI-3) (1992), predict medical student academic success on the USMLE Step 2 CK medical school exam when controlling for demographics and prior academic history at Northeast Medical University?

The null hypothesis was conscientiousness does not predict medical school success as measured by the dependent variable, USMLE Step 2 CK. Based on the correlational analyses, the null hypothesis was rejected for conscientiousness predicting performance on the USMLE Step 2 CK. In this study, conscientiousness was a strong predictor of Step 2 CK performance. This finding was similar to studies regarding graduate school success both in the classroom, on standardized tests, and in their professions (Grehan et al., 2011; McCredie & Kurtz, 2020; Walsh, 2020). Similarly, prior data have shown that individuals with a higher level of conscientiousness tend to perform better in their undergraduate studies (Ivcevic & Brackett, 2014; Komarraju et al., 2009; Komarraju et al., 2011). Based on the above studies and the findings of this research, assessing conscientiousness at the point of admissions may predict success on the USMLE Step 2CK.

Research Question 4. Does grit, as measured by Duckworth and Quinn's Short Grit Scale (2009), predict medical student academic success on the USMLE Step 2 CK medical school exam when controlling for demographics and prior academic history at Northeast Medical University? The null hypothesis was grit does not predict medical school success as measured by the dependent variable, USMLE Step 2 CK. Based on the correlational analyses, the null hypothesis was supported for those students who had higher grit scores on the Grit-S scale; as grit does not predict success on the USMLE Step 2 CK examination. In this study, there was a difference between the grit scores of underrepresented students and their White/Asian peers, but the difference did not reach statistical significance. Similarly, the study by Akos and Kretchmar (2017) also reported that when incorporating Duckworth's Grit-S scale into the admission process, that Grit-S scores among underrepresented students were not significantly statistically different than that of their peers.

While there is considerable evidence to support using the Grit-S scale to predict academic success (Duckworth & Quinn, 2009; Eskreis-Winkler et al., 2014; Miller-Matero, et al., 2018; Robertson-Kraft & Duckworth, 2014; Strayhorn, 2013), according to the results of this study, grit does not appear to correlate with success on the USMLE Step 2 CK exam. Contrary to the findings, Miller-Matero et al. (2018) demonstrated that grit was related to medical school academic performance, with statistically significant differences between grit scores and scores on the USMLE Step 2 CK. This assessment was based upon one cohort of graduating medical students at one medical school which was representative of the graduating medical students there and across the United States (Miller-Matero et al., 2018). Other studies (Han, 2018; Majeed et al., 2019) showed similar average grit scores to be predictive of success defined in other ways (e.g., persistence in medical school and lower levels of depression). In a recent study by Han (2018), the author found that medical students who possessed greater grit levels as measured by Duckworth's Short Grit Scale, Grit-S, were more likely to continue with their medical education. The author suggested that opportunities to foster grit within the medical school curriculum should be considered and be intentional. According to Han (2018), students that have a higher level of grit cannot see themselves being anything other than a doctor and will work to make that happen.

Summary of Results

This research builds upon the current literature related to medical student success, as measured specifically by success on the USMLE Step 2 CK national examination. This study reinforced previous studies that predicted a correlation between success on the USMLE Step 2 CK examination with prior academic variables, namely MCAT scores and undergraduate science GPA. Grit was not found to be a reliable predictor of success on the USMLE Step 2 CK. However, the results did show that the non-cognitive trait of conscientiousness correlated with higher undergraduate science GPAs and better scores on the USMLE Step 2 CK exam.

These findings are important in the field of medical school admission as utilizing a more holistic approach in the admissions process has been strongly recommended by the Association of American Medical Colleges (Conrad et al., 2016; Witzburg & Sondheimer, 2013). Assessing these variables allows for a more inclusive asset approach as argued by Yosso (2005) and further considering intelligence through Sternberg's (1999) lens would assist admission committees in furthering the notion of holistic admission. Allowing for the consideration of non-cognitive, measurable factors in the admission process may allow for a much needed, more diverse physician workforce, yet the right variables must be found. More research on identifying and studying which non-cognitive variables incorporate culturally-based assets that students may bring with them to medical school (e.g., languages, spirituality) (Yosso, 2005) as well as diverse forms of intelligence which are often based on life circumstances (e.g., practical, creative, analytical) (Sternberg, 1999) are warranted in order to address the need for a more diversified physician workforce. Recognizing these factors that can identify students who will most likely be successful will not only save time, but money and effort as well.

Additionally, this study further revealed that URM students, based solely on their ethnic backgrounds, did not perform as well as their White and Asian peers on the USMLE Step 2 CK examination.

Implications and Recommendations for Medical Education

My interest in medical student success stems from my background in medical school admissions and a desire to enhance the admission process to include non-cognitive factors that can be assessed to further predict student success in medical school. In addition, I am interested in educating a more diverse medical student body in terms of race/ethnicity, first generation college student status, age, and socio-economic status to better serve our diverse communities. This study investigated two specific variables—conscientiousness and grit—to assess their relevance to predict medical school success in regards to clinical knowledge on the standardized national examination, the USMLE Step 2 CK. It is important to note that this data set was not representative of the medical university nor that of the national data and therefore the findings cannot be generalized to other medical schools. This is huge limitation that is discussed later in this chapter.

This study found that undergraduate science GPAs and MCAT scores predicted medical student success on Step 2 CK, regardless of sociodemographic and other academic variables. This study did not support the hypothesis that students who received a master's degree prior to enrolling in medical school outperformed their peers on Step 2 CK. This points to the importance of an individual's undergraduate performance and standardized test scores, the traditional measures. Similarly, while post-baccalaureate and/or master's degree programs have helped students to enter medical school, there does not appear to be a correlation with these programs and performance on the Step 2 CK (Epps, 2015; Grumbach & Chen, 2006; Johnson et al., 2017; Orozco Cortes, 2019). The access that these programs provide is one of great importance and should not be minimized. Further research should specifically investigate post-baccalaureate and master's programs whose curriculums allow for matriculants to take classes directly with medical students and be academically evaluated the same way as the matriculated medical students (Gohara et al., 2011).

Within the educational community, it is time to continue a larger discourse about the role of standardized tests, including in the medical schools. In light of COVID-19, undergraduate colleges and universities (Fairtest.org) have reconsidered standardized tests in their admissions processes and have gone test optional for this application cycle. To date, over 70% of schools/colleges are not requiring these tests for the 2020-21 admissions cycle. This step will give administrators an opportunity to assess what is lost in predictors of academic success with the absence of this variable and what is gained. If universities find they see an increase in applications of students of color, increase their yield of these students, they can track their academic success compared to past entering classes. What institutions learn should be considered in light of medical education. This research study was narrowly focused on medical education, namely the MCAT and only one standardized test taken during an individual's medical school journey.

The challenge to create predictors of what constitutes medical student success and develop assessment tools during the medical school curriculum that truly measure medical students' competency and knowledge required as physicians is the same challenge faced across all levels of education, particularly in STEM education. We must address issues of racial inequity, especially in testing, across disciplines for Black, Latinx, and Native American students more broadly.

As the Association of American Medical Colleges (2016) has encouraged medical schools to use an individual and holistic approach when reviewing applications, the timing of this study is important especially considering the encouragement by the AAMC for using the Experiences, Attributes, and Metrics (EAM) Model for admission. Not doing so will continue to provide long-standing barriers for access from candidates from minoritized backgrounds and a dramatic void in serving our diverse communities. Often, URM students tend to also be the students who have pursued academic enhancement programs to get to medical school (Carnevale & Strohl, 2013; Ntiri, 2001). Eliminating the use of standardized tests and focusing on developing assessment tools that measure criteria identified in the EAM Model would allow admissions committees to utilize an inclusive assets approach. Future research could focus on these variables and others, including Sedlacek's (2011) non-cognitive variables such as non-traditional knowledge and leadership skills. Quite simply, medical school educators need to develop reliable and valid instruments to capture inclusive, assets of students of color and other under-represented groups (Yosso 2005) that correlate with medical school success.

Additionally, recognizing the formal ties that the testing business, NBME, has with medical school accreditation, LCME, and the AAMC is important to consider. USMLE Step 1, Step 2 CK and Step 2 CS, and Step 3 testing must be examined and considered in relation to the inter-connectedness of these agencies (Eaglen, 2017). Recognizing these relationships may demonstrate that success is defined by a "system" that is problematic for an equitable representation (Kendi, 2016) and a way to perpetuate some testing industries.

As alluded to in Chapter One, while grit has been advocated for in research on college students, there have been some recent challenges regarding the assessment of grit, particularly suggesting that grit continues to further a deficit approach when considering under-represented students (Tewell, 2020). Some researchers have argued that this intense focus on grit may again shift the lens towards looking at a deficit discourse focused on what is missing in the disposition or motivation of students of color (Love, 2019; Mehta, 2015; Ris, 2016; Tewell, 2020). According to Mehta (2015), the most significant critique surrounding grit is that focusing on an individual's grit may ultimately be a way to blame the students themselves instead of centering on the larger systemic issues related to economic, social, and racial justice that require grit from students from minoritized backgrounds.

Tough (2012) argued that the low-income, under-represented students have already demonstrated grit in their everyday lives. In fact, Love (2019), suggested that questioning the grit of African American students is 'deeply hurtful' as they already demonstrate significant grit each and every day in their fight against racism.

In retrospect, when looking at the Grit-S Scale through the lens of these critiques, the items in this survey do not get at those items that may be more reflective of what Yosso (2015) suggested if framed from a cultural wealth perspective and help URM students navigate inequitable, unwelcoming educational systems. The Grit-S has been used in very few analyses considering different cultures, ethnicities, and socioeconomic backgrounds (Datu, Yuen, & Chen (2017). The majority of research using the Grit-S instrument has been linked to Ivy League colleges/universities with highly selective, and majority White participants (Datu et al., 2017). Duckworth's Grit-S scale does not get at the strengths of "others," specifically how students have learned the skills necessary to negotiate educational systems that privilege Whites. Yosso (2015) argued that in order to truly be successful, students must bring their linguistic, resistant, or navigational wealth to their educational environment in order to deal with these oppressive structures. These are the attributes educators should try to measure.

Limitations of this Study

The findings of this study need to be interpreted with the limitations of the data in mind. Some of these limitations were identified previously in Chapter Three. This study was limited by focusing on a single medical school within the United States. Additionally, two years of matriculated students were used in this study, yet the sample size, particularly related to underrepresented students, was small. The biggest limitation of this study was the sample not being representative of the medical school or national medical student population, therefore it cannot be generalized even to the population at Northeast Medical University. At the time of this study Northeast Medical University matriculated its smallest percentage of female students (39%) in more than two decades. The university made significant changes to the admissions processes, by giving preference to females in the screening process, and matriculated 45% female students in 2016. Students from both of these cohorts were participants in this study. The following year, 2017, Northeast Medical University matriculated 50% of students who identified as female and has remained at roughly half of the matriculating class each year since. The female variable was most similar to the national number of matriculated students (p = .024) but not to the female populations at Northeast Medical University (p = .293), indicating that a higher percentage of males at Northeast Medical University chose to participate in this study. For those students who were under-represented in medicine, participants were not at all comparable to the national (p = .572) number of URM matriculants, nor those enrolled at Northeast Medical University (p = .611). This difference may be attributed to the those that were deemed ineligible to participate in this study. Those that were deemed ineligible included students that had matriculated in either the 2015 or 2016 cohort and were enrolled in a decompressed curriculum, had taken a leave of absence, taken time off to prepare for USMLE Step 1 and were now on a different timeline, and did not have USMLE Step 2 CK scores, at the time of this study. Others may not have felt they had time to complete the grit scale due to their academic schedule, they just weren’t interested, or they had withdrawn or been dismissed by the University, thereby making an already small number of students even smaller.

Another limitation of this study is the possibility that more conscientious students chose to participate, thus skewing the results. Grit and conscientiousness may be distinctive qualities that medical students possess as part of their character. Perhaps those with higher levels of conscientiousness enter medicine in the first place, or perhaps female and/or underrepresented medical students have demonstrated higher levels of grit to persevere in the White, male-dominated profession. This might be worth further investigation.

Similarly, the accuracy of both constructs requires that people are genuine about their behaviors. Social desirability bias refers to the tendency of participants in research studies to provide socially desirable responses instead of selecting responses that are truly reflective of their behaviors and feelings (Grimm, 2010). There is certainly the potential for participants to try to "game" the Grit-S and NEO-FFI- 3 constructs in order to make themselves appear grittier or more conscientious than they actually are (Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007; Orzeck, & Lung, 2005). Medical students are smart, the potential to not know what the "right", or socially desirable, answer is for the "I work hard to accomplish my goals" question on the Grit-S survey or "I am a hard worker" on the Conscientiousness section of the NEO-FFI-3, especially at the point of admission to medical school is questionable (Datu, Yuen, & Chen, 2017). Social desirability bias could certainly explain the lack of magnitude in this study and therefore the differences in the means of these constructs. Social desirability bias could also reduce the variability of student responses as well. Another reason for these findings could, in fact, be the instruments themselves. This does not mean that conscientiousness and grit are not important, but these constructs may not be adequate to assess this population of students.

An additional limitation of this study is the limited definition of academic success as defined as performance on the USMLE Step 2 CK. For those that took the exam in 2018, the national passing rate was 96%, while those that took the exam in 2019, had a national passing rate of 97% (USMLE, 2020e). The high pass rate on the USMLE Step 2 CK may have also explain the extremely small effect sizes found in this study. Also, due to the high success rate on Step 2 CK it was difficult to find variability in the mean scores that are meaningful.

The importance of the Step 2 CK examination may become a more highly valued outcome measure, particularly for residency placement, as the scoring of the USMLE Step 1 moves to Pass/Fail scoring system. In the middle of February 2020, the Federation of State Medical Boards (FSMB) and the National Board of Medical Examiners (NBME) announced that the USMLE Step 1 examination will be moving to a Pass/Fail model effective January 2022. This change, along with many medical schools moving their basic sciences years, years one and two, to Pass/Fail grading systems may put more reliance on the results of the Step 2 CK examination scores in residency selection. This may be particularly true during the residency application cycle when residency programs are looking for quantifiable differences in their applicants as they look to select the “best” candidates (Chaudhry et al., 2020).

Additionally, another limitation which was alluded to in Chapter Three, may be the timing of the Grit-S survey. The underlying premise of this study was to assess the use of additional non-cognitive factors during the admissions process that may help predict success in medical school, yet this study measured grit during the third and fourth year of medical school. Several studies have stated that grit can change over time, and that it can be modified and further developed with practice (Brady et al., 2017; Damgaard, & Nielsen, 2018; Kalesnikava et al., 2019; Wilson & Buttrick, 2016). Therefore, the finding that grit was not correlated with Step 2 CK success may have been related to timing and perhaps may be better assessed at the point of application. Further in regard to grit, perhaps the participants of this study had already demonstrated a significant level of grit prior to entering medical school based on their path to matriculation (i.e. during their undergraduate studies, in preparation for the MCAT, throughout their application process, or during their clinical experience, prior research or their volunteer or extra-curricular involvement, or prior life experiences and circumstances that have shaped their lives). If grittier students overall pursue a medical education in the first place, these data could be critical. As these data do not include those students who withdrew or were dismissed from Northeast Medical University, the data might be skewed in terms of the impact grit had on their persistence. Not including these data may be because the main source of variability could have been reflected in the students who left the institution. Also, grit may not reflect a quality that is required to effectively pass the USMLE Step 2 CK exam but may be an important factor to explain other hurdles, including persistence, in medical school. Therefore, future research is warranted.

Another significant limitation, alluded to above, includes the fact that 3.9% of those students that matriculated in 2015 and 2016, the cohorts in this study, left the institution and therefore these data were not considered in these analyses. Out of the twelve students who left Northeast Medical University, five students were dismissed, one transferred to another medical school, five withdrew from medical school, and one student passed away. Although these students completed the conscientiousness facet of the NEO-FFI-3, they were not included in this study because they did not complete the Grit-S scale nor did they have USMLE Step 2 CK scores to consider. The variability, particularly regarding grit and conscientiousness, might have been captured by the students who were no longer enrolled. This could be another reason for the low effect sizes when assessing these constructs, the magnitude may have gotten distilled because this group that would have added some variability was not considered. If both grit and conscientiousness were assessed at the point of admission, one could assess the correlation with these traits to see if more or less gritty and/or conscientious students were dismissed or withdrew from medical school or persevered and stayed the course.

The small sample size, particularly among certain variables (e.g. race/ethnicity, low socioeconomic, first-generation college, those 26 and older) may have had a significant impact relative to the power of the analyses (Slavin & Smith, 2009; Turner, Paul, Miller, & Barbey, 2018). It is possible that statistical significance could have been achieved with a larger sample size, especially given that extremely small effect sizes were found among multiple independent variables. The small effect sizes that were found among variables in this study may also be explained by the sample of who was accepted into this particular medical school. As noted earlier, participants in this study were not representative of the local or national medical school population.

Future Directions for Research

This study reinforced previous research findings about student characteristics that correlated with performance on the national USMLE Step 2 CK examination, while also exploring additional factors for consideration. As demonstrated, there was significant correlation between conscientiousness and several independent variables, including female students, lower socioeconomic students, first-generation college, and under-represented student status. While there was no statistically significant correlation between conscientiousness and age, those who majored in the sciences, science GPA, or MCAT percentile score, there was also no statistical correlation between having completed a master's degree (p = .105) and a person's conscientiousness. Individuals who pursue master's degrees or record enhancing coursework tend to be older and may have had more life experiences. One would expect that an individual's conscientiousness would change over time, as they are perhaps juggling more life responsibilities, including a job or familial responsibilities, which may increase their level of conscientiousness. Perhaps it is worth further investigation to better understand this population's correlation with conscientiousness (Roberts et al., 2017). Similarly, it may be worth examining the use of the conscientiousness construct and associated instruments in an effort to make sure they capture the behaviors that are relevant for medical students, while also exploring the impact of social desirability bias in this group of students and on the use of these instruments in the admissions process. I also recommend the development of instruments measuring the constructs of conscientiousness, grit, and a related construct-resilience-that include less obvious desirable responses.

Although current psychology research suggests that grit is not any different than the Big-Five personality trait of conscientiousness, several studies are looking at the effects of both grit and conscientiousness on academic and career performance (Camp et al., 2019; Hong & Lee, 2019; Samuelsen et al., 2019; Schmidt et al., 2020; Walsh, 2020; Werner et al., 2019). While conscientiousness includes an individual's perseverance, passion related to an individual's grit has been less defined (Dahl, 2016). Dahl (2016) demonstrated that grit was more correlated with an individual's background, or sociodemographic variables, while conscientiousness was more correlated with an individual's academic background. Perhaps additional research could utilize path analysis in order to evaluate the direct relationships between the independent variables and their impact on Step 2 CK scores. It could assess the direct relationships (paths) between background (demographic) variables, academic variables, and non-cognitive variables (grit and conscientiousness) and Step 2 CK scores similar to that of studies by Austin et al. (2005) and Mavis (2001). Using path analysis would assess the identification status of the model by measuring if it is just-identified or over-identified so that the model can be estimated (Keith, 2006). The model would need to be estimated to determine the direct effects and disturbances and then be added into the model. The direct, indirect, and total effects could then be calculated and interpreted while the effect, or causality, of the different variables on Step 2 CK could then be assessed.

Additionally, despite the small numbers of participants, it is certainly worth noting that the Step 2 CK scores, the grit scores and the conscientiousness scores showed significant differences for the URM students, low SES students, and first-generation students in relation to their peers. The disparities between the USMLE Step 2 CK scores appear very similar to the disparities one might expect in undergraduate standardized testing performance, SAT or ACT scores, for these particular populations of students (Terenzini, Cabrera, & Bernal, 2001). Future research could investigate the association between SAT, MCAT, and USMLE performance in greater depth in order to see if any correlation exists.

Kendi (2016) argued that by continuing to use standardized tests to measure intelligence and predict success we continue to perpetuate postracial ideology preventing admissions committees from appreciating how these standardized tests exclude the admission of black students. Kendi (2016) went on to say:

These [standardized tests] have failed time and again to achieve their intended purposes: measuring intelligence and predicting future academic and professional success. The tests, not the black test-takers, have been underachieving (para. 17).

Under-represented student doctors will need additional support and preparation, while learning how to take the standardized examinations that they will take throughout their medical career. If the MCAT and USMLE standardized examinations, and national specialty examinations continue to be the measures used to quantify medical performance, inequities will remain. Therefore, in order to recruit, prepare, educate, and train more diverse future physicians, the system must change. Medical schools must dedicate research to identify valid and reliable instruments to capture non-traditional variables that be better indicators of medical school success for students from minoritized backgrounds.

It also may be valuable to explore the demographics of the students who complete academic enhancement programs; an assumption could be that they are disproportionately from underrepresented backgrounds. Should that assumption hold true, it may be an argument for considering alternative opportunities to medical school matriculation as the master's degree shows no statistically significant impact on the USMLE Step 2 CK score. Similarly, regarding academic enhancement or enrichment programs, further investigating the types of students, the various curricula, and the support services in place for those who completed post-baccalaureate and/or master's degrees prior to medical school could provide some insight into the benefits (or not) of these programs on Step 2 CK scores.

Further, based on the extremely small and negative effect sizes found in this study of these two non-cognitive variables, grit and conscientiousness, this study indicated that neither trait ultimately mattered when it came to explaining how students performed on the USMLE Step 2 CK. Therefore, future studies should examine ways in which the medical school curriculum and/or experiences in medical school impacted performance on the USMLE Step 2 CK.

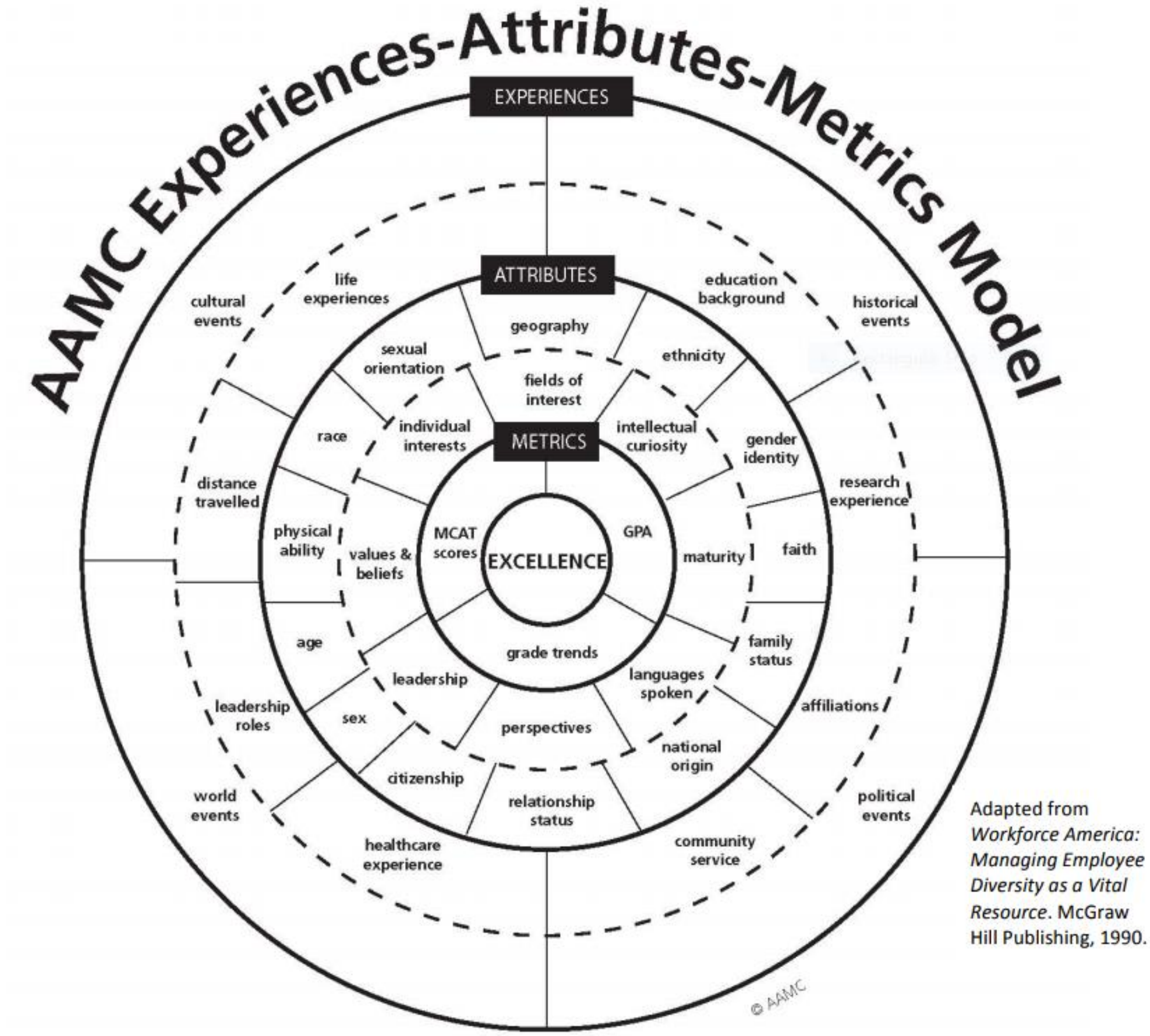
Additionally, future research could investigate the need for pre-medical students to major in the sciences in order to be successful in medical school. Some schools have started to consider backgrounds in the humanities (Muller, 2014; Muller & Kase, 2010) as well as competency-based admissions (Kerrigan et al., 2016; Kirch, 2012; Rolfe, Pearson, Powis, & Smith, 1995). While majoring in the sciences is helpful in terms of preparing for the MCAT and having a baseline knowledge of what you will be taught in medical school, it is difficult to know what a student's baseline science and social science knowledge was and whether it helped to prepare them to take the MCAT.

Chapter Summary and Conclusion

I have reviewed the current challenges facing medical education and suggest that medical schools further consider utilizing the Experiences, Attributes, and Metrics (EAM) Model that the AAMC created in 2014 (see Appendix A). This approach will provide context and suggestions when considering non-cognitive variables in the assessment of prospective medical school students. Considering non-cognitive variables has been identified in the literature as identifying clinically sound and knowledgeable physicians. Likewise, in this study, I specifically introduced two surveys to assess the non-cognitive factors of grit and conscientiousness in medical students. I used the Grit-S (Duckworth & Quinn, 2009) to assess students' grit and the Conscientiousness section of the NEO-FFI-3 Inventory to assess their level of conscientiousness (Costa & McCrae, 2008). I used hierarchical and logistic regression techniques and statistical methods to answer the research questions. This research contributes to an understanding of non-cognitive variables and success in medical school, providing potentially important implications for using different methods in the medical school admissions process to predict academic success. The findings of this study provide evidence for consideration of assessing conscientiousness in the medical school admissions process as it is significantly predictive of performance on the national USMLE Step 2 CK scores, yet has a small effect size, which may be impacted by the small sample size of participants. Assessing conscientiousness, in addition to the traditional cognitive measures of science GPA and MCAT scores, may provide a reliable and valid approach to evaluating future medical school student applications at least in regards to Step 2 CK performance.

Given the current pandemic facing our country and the strain it is putting on our healthcare providers, the importance of this study could not be timelier. This study assessed two qualities that this researcher considered necessary to be an effective physician within the world in which we currently live. Conscientious individuals are defined as careful, efficient, organized, competent, and deliberate (Costa & McCrae, 1989), although not considered statistically significant related to USMLE Step 2 CK performance, grit is defined as an individual who cares so much about something that it continues to provide meaning and drive them, day after day, week after week (Duckworth, 2016) and may play a role in who enters the profession in the first place. The COVID-19 pandemic has demonstrated the need for both of these characteristics in today's physician. Not only are these qualities necessary to get through the grueling medical school admissions process and the challenging medical school curriculum, but also now, as we have seen, the emotional trauma and endless days of being on the "front lines" of this pandemic. Never before have these qualities been more important in a physician than they are now. Based on this study, assessing the domain of conscientiousness at the point of admission should be considered due to its correlation with USMLE Step 2 Clinical Knowledge exam performance.

APPENDIX A



The AAMC Experiences-Attributes-Metrics Model is a central focus of the Holistic Review Project. The E-A-M model broadens the lens through which admission committee members can view applicants to recognize the varying dimensions and contexts that shape each candidate's identity. It is not meant to be a complete representation of all dimensions of an individual, nor is it intended to serve as a modified checklist. Schools have found it to be a helpful point of reference when assessing their admissions policies and practices.

APPENDIX B

Step 2 Clinical Knowledge (CK)
CONTENT DESCRIPTION & GENERAL INFORMATION
https://www.usmle.org/pdfs/step-2-ck/Step2CK_Content.pdf

A Joint Program of the Federation of State Medical Boards of the United States, Inc., and the National Board of Medical Examiners®

Copyright © 2020 by the Federation of State Medical Boards of the United States, Inc. (FSMB), and the National Board of Medical Examiners® (NBME®). All rights reserved. The United States Medical Licensing Examination®(USMLE®) is a joint program of the FSMB and the NBME.

This booklet will help you prepare for the Step 2 Clinical Knowledge (Step 2 CK) component of the United States Medical Licensing Examination® (USMLE®). Practice materials, which include Sample Test Items (PDF) and simulated web-based Tutorial and Practice Test Items, as well as other informational materials, are available at the USMLE website www.usmle.org. Examinees must also read the *USMLE Bulletin of Information* at www.usmle.org/bulletin.

IMPORTANT:
• The term *item* is used to describe a test question in any format.
• You **must** run the web-based Tutorial and Practice Test Items to become familiar with the test software **prior to your test date**.
• The tutorial provided at the beginning of the Step 2 CK Examination has fewer screens and less detailed information than the Step 2 CK web-based Tutorial and Practice Test Items on the USMLE website.
• The web-based Tutorial and Practice Test Items on the USMLE website include single multiple-choice questions, a sequential set of multiple-choice questions, a scientific abstract (a summary of an experiment or clinical investigation, accompanied by two or more questions), and items with audio findings.
Please visit the USMLE website www.usmle.org often to view announcements regarding changes in test delivery software, and to access updated practice materials. You must obtain the most recent information before taking any USMLE examination.
Step 2 CK consists of multiple-choice questions (MCQs) prepared by examination committees composed of prominent faculty members, teachers, investigators, and clinicians who make up the USMLE Test Material Development Committees. All committee members have recognized expertise in their respective fields. They are selected to provide broad representation from the academic, practice, and licensing communities across the United States and Canada.
• Step 2 CK is a one-day examination. It is divided into eight 60-minute blocks and administered in one 9-hour testing session. The number of questions per block on a given examination will vary but will not exceed 40. The total number of items on the overall examination will not exceed 318.

The examination also includes a minimum allotment of 45 minutes of break time and a 15-minute optional tutorial. The amount of time available for breaks may be increased by finishing a block of test items or the optional tutorial before the allotted time expires.
Step 2 CK assesses an examinee's ability to apply medical knowledge, skills, and understanding of clinical science essential for the provision of patient care under supervision and includes emphasis on health promotion and disease prevention. Step 2 CK ensures that due attention is devoted to principles of clinical sciences and basic patient-centered skills that provide the foundation for the safe and competent practice of medicine under supervision.
Test questions focus on the principles of clinical science that are deemed important for the practice of medicine under supervision in postgraduate training.

The content description that follows is not intended as a curriculum development or study guide but rather models the range of challenges that will be met in the actual practice of medicine. It provides a flexible structure for test construction that can readily accommodate new topics, emerging content domains, and shifts in emphasis. The categorizations and content coverage are subject to change. The best preparation for the examination is broad-based learning that establishes a strong general understanding of concepts and principles in the basic and clinical sciences.

Content Outline

All USMLE examinations are constructed from an integrated content outline. The outline is available on the USMLE website (www.usmle.org/pdfs/usmlecontentoutline.pdf).
Content is organized according to general principles and individual organ systems. Test questions are classified into one of 18 major categories, depending on whether they focus on concepts and principles that are applicable across organ systems or within individual organ systems. Sections focusing on individual organ systems are subdivided according to normal and abnormal processes, including principles of therapy. They include subcategories of specific disease processes. In most instances, knowledge of normal processes is evaluated in the context of a disease process or specific pathology.
Not all topics listed in the content outline are included in every USMLE examination. Overall content coverage is comparable in the various examinations that will be administered to different examinees for each Step. The Step 2 CK examination covers content of specific disease processes or pathology. The Step 2 CK system specifications are listed in Table 1.

Table 1: Step 2 CK System Specifications*

General Principles of Foundational Science	Immune System
Blood & Lymphoreticular System	Behavioral Health
Nervous System & Special Senses	Musculoskeletal System/Skin &
Cardiovascular System	Subcutaneous Tissue
Cardiovascular System	Respiratory System
Gastrointestinal System	Renal & Urinary System & Male
Pregnancy, Childbirth & the Puerperium	Reproductive

Pregnancy, Childbirth & the Puerperium	Female Reproductive System & Endocrine System
Breast	
Multisystem Processes & Disorders	

Biostatistics & Epidemiology/Population Health/Interpretation of Medical Literature
Social Sciences: Legal/Ethical Issues & Professionalism/Systems-based Practice & Patient Safety

Physician Tasks/Competencies: An additional organizing construct for Step 2 CK design is physician tasks and competencies. Each test item is constructed to focus on assessing one of the competencies listed in Table 2. Detailed information about the physician tasks and competencies outline is available at the USMLE website (www.usmle.org/pdfs/tcom.pdf).

Table 2: Step 2 CK Physician Tasks/Competencies Specifications*

Patient Care: Laboratory/Diagnostic Studies	Patient Care: Diagnosis
Patient Care: Prognosis/Outcome	Patient Care: Health Patient Care:
Pharmacotherapy	Maintenance/Disease Prevention
Patient Care: Pharmacotherapy	Patient Care: Clinical Interventions
Patient Care: Mixed Management	Practice-based Learning & Improvement
Professionalism	Systems-based Practice & Patient Safety

Each Step 2 CK examination covers content related to the traditionally defined disciplines listed in Table 3.

Table 3: Step 2 CK Discipline Specifications*

Medicine	Surgery
Pediatrics	Obstetrics & Gynecology
Psychiatry	

Examples of MCQs focused on each of the competencies and samples of topics from different areas of the content outline are shown below.

Competency: Patient Care: Diagnosis - Laboratory/Diagnostic Studies Content Area: Behavioral Health

A 17-year-old girl comes to the office for an examination prior to entering college. She reports that she feels well but is nervous about leaving home for the first time. She states that she has tried to diet to improve her appearance but that food restriction often "backfires" because she becomes hungry and then engages in episodes of binge eating. She reports a loss of control during these episodes, saying, "It's like I stop thinking at all and before I know it, I have eaten two pizzas." She induces vomiting several times during each binge and has developed a pattern of binging and purging every evening. She has no history of serious illness and takes no medications. She is 165 cm (5 ft 5 in) tall and weighs 57 kg (125 lb); BMI is 21 kg/m2. Vital signs are within normal limits. Physical examination shows dry mucous membranes, erosion of enamel on the lingual surface of the front teeth, and hypertrophy of the parotid gland.

Serum studies are most likely to show which of the following sets of findings in this patient?

Potassium Bicarbonate
(A) Decreased decreased
(B) Decreased increased
(C) Increased decreased
(D) Increased increased
(E) Normal decreased
(F) Normal increased

Answer: B

Competency: Patient Care: Diagnosis: Content Area: Musculoskeletal Sys/Skin & Subcutaneous Tissue
A hospitalized 57-year-old man has had severe progressive pain in his left knee since awakening 2 hours ago. He was admitted to the hospital 2 days ago for an acute myocardial infarction. Cardiac catheterization showed occlusion of the left anterior descending artery, and he underwent placement of a stent. Current medications include aspirin, metoprolol, lisinopril, simvastatin, clopidogrel, and heparin. Vital signs are within normal limits. Examination of the knee shows a large effusion. The knee is hot to touch and erythematous. He holds the knee in 30 degrees of flexion; the pain is exacerbated with further flexion or extension. Laboratory studies show:

Hematocrit 40%
Leukocyte count 13,000/mm3
Serum
Ca2+ 9.2 mg/dL
Urea nitrogen 15 mg/dL
Creatinine 1.0 mg/dL
Albumin 3.6 g/dL

An x-ray of the left knee shows calcification of the synovium. Which of the following is the most likely diagnosis?

(A) Deep venous thrombosis
(B) Gonorrhea
(C) Gout
(D) Hemarthrosis
(E) Pseudogout
(F) Septic arthritis

Answer: E

Competency: Patient Care: Health Maintenance/Disease Prevention & Surveillance Content Area: Pregnancy, Childbirth & the Puerperium

A 21-year-old woman comes to the office for counseling prior to conception. She is recently married and would like to conceive within the next year. She does not eat meat, fish, or dairy products and wishes to decrease the risks of her diet on her baby. Menses occur at regular 28-day intervals and last 5 days. She does not smoke or drink alcohol. She takes no medications. She is 157 cm (5 ft 2 in) tall and weighs 50 kg (110 lb); BMI is 20 kg/m2. Physical examination shows no abnormalities. Pelvic examination shows a normal appearing vagina, cervix, uterus, and adnexa.

Which of the following is most likely to decrease the risk of fetal anomalies in this patient?

(A) Adjusting diet to include more sources of protein during the first trimester
(B) Beginning folic acid supplementation prior to conception
(C) Calcium supplementation during the first trimester
(D) Iron supplementation during the first trimester
(E) Soy protein shakes throughout pregnancy and lactation

Answer: B

Competency: Patient Care: Management – Clinical Interventions Content Area: Immune System

A 10-year-old boy is brought for a follow-up examination 2 days after he was seen in the emergency department because of hives, hoarseness, and light-headedness. His symptoms began 15 minutes after he was stung by a bee and lasted approximately 60 minutes; they resolved before he was treated. He has been stung by bees three times over the past year, and each reaction has been more severe. Examination shows no abnormalities.

Which of the following is the most appropriate recommendation to prevent future morbidity and mortality from this condition?

(A) Avoid areas known to have bees
(B) Avoid wearing colorful clothing outside
(C) Carrying diphenhydramine tablets
(D) Carrying self-injectable epinephrine
(E) Seek immediate medical attention following any future sting

Answer: D

Competency: Practice-based Learning Content Area: Biostatistics

A cohort study is conducted to compare the incidence of adverse effects of a recently approved antihypertensive pharmacotherapy with that of conventional therapy. A total of 20,000 patients are enrolled. Twelve thousand are prescribed the recently approved therapy, and 8,000 are prescribed conventional therapy. Patients in the study and control groups are matched for baseline blood pressure, age, and gender. Data are collected from the records of the patients' ongoing clinical care. Results show that those receiving the newly approved treatment have twice the incidence of fatigue compared with those receiving the conventional treatment. The results are statistically significant (P=0.01).

Which of the following potential flaws is most likely to invalidate this study?

(A) Publication bias
(B) Selection bias
(C) Type I error
(D) Type II error

Answer: B

Competency: Professionalism Content Area: Social Sciences
Three days after hospitalization for diabetic ketoacidosis, an 87-year-old woman refuses insulin injections. She says that her medical condition has declined so much that she no longer wishes to go on living; she is nearly blind and will likely require bilateral leg amputations. She reports that she has always been an active person and does not see how her life will be of value anymore. She has no family and most of her friends are sick or deceased. On mental status examination, she is alert and cooperative. She accurately describes her medical history and understands the consequences of refusing insulin. There is no evidence of depression. She dismisses any attempts by the physician to change her mind, saying that the physician is too young to understand her situation. She says, "I know I will die, and this is what I want."

Which of the following is the most appropriate next step in management?

(A) Discharge the patient after she has signed an "against medical advice" form
(B) Seek a court order to appoint a legal guardian
(C) Offer insulin but allow the patient to refuse it
(D) Admit to the psychiatric unit
(E) Administer insulin against the patient's wishes

Answer: C

Competency: Patient Safety Content Area: Social Sciences
A 45-year-old woman is hospitalized for management of *Staphylococcus aureus* endocarditis with persistent bacteremia. The patient is discussed during interdisciplinary rounds, which includes physicians, nurses, pharmacists, and social workers. During rounds, a pharmacy student notices that the patient missed two doses of her scheduled antibiotic last week but is unsure why. The physician and nurse are unaware of these missed doses, and the student does not mention her observation.

Which of the following measures is most likely to improve communication within this interdisciplinary health care team?

(A) Conduct interdisciplinary rounds in a quieter location
(B) Encourage questions from all team members
(C) Implement a checklist for standardizing patient rounds
(D) Use computers during rounds to review medications
Answer: B

APPENDIX C

Summary Table of Key Variables

Key Variables	Authors	Summary of Findings	Methods
Race/Ethnicity	Isik, Wouters, Ter Wee, Croiset, & Kusurkar (2017)	Academic performance scores were not significantly different between the majority and ethnic minority groups when assessing motivation.	Cross-sectional study of all students in a Dutch medical school were invited to complete Academic Self-Regulation Questionnaire, measuring autonomous and controlled motivation. Linear regression analysis was used to determine the association between motivation and performance (GPA) to assess the mean differences of academic scores by Majority & URiM students.
Race/Ethnicity	White, Dey, & Fantone (2009)	MCAT scores are strong predictors of Step 1 score among minority and majority students. Undergraduate GPA was a significant predictor of first year medical school performance among minority students. They concluded that a more thorough investigation into the predictive validity of the MCAT and uGPA as admission criteria is needed.	Collected data from 8 classes of medical students with a sample size of 1,441. Variables included UGPA, MCAT score, medical school exam grades, and Step 1 scores. Bivariate Pearson's Correlation was used to assess the correlation of MCAT & uGPA scores on Step 1 scores.
Race/Ethnicity & Gender	Ogunyemi & Taylor-Harris (2005)	The author did not find any statistical association between race or gender and Step 2 CK score.	The author collected USMLE scores, MCAT scores, UGPA, and demographic data from 171 medical students and performed a retrospective record review. Pearson correlation and Spearman correlation coefficients were used. A logistic binary regression analysis was performed.
Race/Ethnicity & SES	Jerant, Henderson, Griffin, Talamantes, Fancher, Sousa, & Franks (2018)	URM status was not associated with Step 1 or Step 2 CK scores at all, though low SES was negatively associated with Step 1 score. Based on prior research in standardized test scoring, the authors hypothesized that the early negative association between SES and Step 1 scores may have been due to a limited opportunity for test preparation.	Collected data from four consecutive class years with a total of 402 participants. Measurements included uGPA, MCAT, MMI, SES, URM, Step 1, & Step 2 CK scores
MCAT/UGPA, Race/Ethnicity	Davis, Dorsey, Franks, Sackett, Searcy, & Zhao (2013)	Accessed data from the AAMC on MCAT examinees, MCAT scores for Black and Latino students were lower than scores for White students. Found that parental education, family income level, growing up in a rural/urban/suburban neighborhood, and access to resources and educational materials could contribute to the reported differences in scores among racial and ethnic groups.	Conducted a comprehensive review of the MCAT exam over four-years at one medical school. Examined four issues: (1) whether racial and ethnic groups differ in mean MCAT scores, (2) whether any score differences are due to test bias, (3) how group differences may be explained, and (4) whether the MCAT exam is a barrier to medical school admission for black or Latino applicants. Conducted logistic regression analyses to estimate the probability of success on the basis of students' MCAT total scores. We conducted the analyses separately by school and then pooled the results.

MCAT/UGPA	Saguil, Dong, Gingerich, Swygert, LaRochelle, Artino Jr, Coruess, & Durning (2015)	MCAT was moderately associated with Step 1 score and weakly associated with...Step 2 CK score, and Step 3 score. The predictive power of the MCAT decreases farther along on the medical education journey, as students grow as learners. Negative relationship was not found with regards to SES and Step 2 CK scores. Encouraged by this finding, suggesting that low SES students were able to catch up with their peers and perform similarly to their counterparts on subsequent standardized tests.	Collected data from 340 students at the Uniformed Services University's F. Hébert School of Medicine. Pearson correlation coefficients were used to explain the difference between the measures and the reported variance.
MCAT/UGPA	Gauer, Wolff, & Jackson (2016)	The authors discovered that the MCAT explained 12.0% of the variance in Step 2 CK scores. The MCAT was shown to be weakly to moderately associated with USMLE performance. The authors noted that the predictive power of the MCAT is not very strong, and thus the MCAT should not be the only factor taken into consideration when evaluating students for admission. They suggested that a low MCAT score should not keep a student from pursuing medicine but could be used by institutions to flag a student needing more work with test-taking or study strategies. Researchers suggest that the biological science sections of the MCAT, are the greatest predictors of success on Step 1 and Step 2 CK, while the UGPA is moderately predictive of success on Step 1 and 2 CK	Multiple linear regression, correlation, and chi-square analyses were performed to determine the relationship between MCAT component and composite scores and USMLE Step 1 and Step 2 CK scores from 1,065 students at one medical school over a 4-year period. Revealed significant associations between the MCAT and Step scores.
MCAT/UGPA	Casey, Palmer, Thompson, Laack, Thomas, Hartz, Jensen, Sandefur, Hammack, Swanson, Sheeler, & Grande (2016)	The authors discovered a strong relationship between MCAT score and Step 1 score. All sub test scores were significantly associated with Step 1 score. There was a moderate relationship between MCAT score and Step 2 score, with all sub scores except writing significantly correlated with Step 2 scores. uGPA also had a significant correlation with both Step 1 and Step 2 CK.	Correlations between measures were evaluated using the Pearson correlation coefficient. Multivariable linear regression models were fit using backward variable elimination to identify sets of independent predictors of USMLE 1 and USMLE 2 examinations.
MCAT/UGPA	Basco, Jr., Way, Gilbert, & Hudson (2002)	Investigated the relationship between institutional MCAT score and Step 1 performance. There were significant correlations between individual MCAT scores and Step 1 scores, including the verbal reasoning sub-score. Data analysis revealed a moderate correlation between individual science UGPA and Step 1 scores.	Subjects consisted data from medical students matriculating to two different medical schools, with a total sample of 16,954 applicants and 933 matriculants. Data were drawn from the longitudinal applicant and matriculant data sets maintained by both schools. Multiple regression with blockwise selection was completed. Multicollinearity was assessed and beta coefficients were adjusted using ridge regression. Cross-validation procedures were employed to estimate the shrinkage that results from using multiple regression for prediction and demonstrated that multiple regression results would be generalizable to a similar sample of students. Descriptive statistics were completed. Correlations with performance on the USMLE Step 1 were adjusted for restriction of range.

MCAT/UGPA	Gilbert, Basco, Blue, & O'Sullivan (2002)	UGPA, MCAT ps and MCAT bs scores were significantly predictive of Step 1 exam scores. MCAT bs and UGPA were the strongest predictors of Step 1 performance. There was a significant association between Step 2 CK performance and UGPA . There was no association between Step 2 performance and MCAT scores for any subcategories.	Multiple regression models calculated additional variance accounted for by the addition of the WS to a model containing grade point average and the other admissions test section scores. In multivariate analyses, when GPA and all MCAT scores were considered as predictors of licensing exam scores, the WS variable did not add to the ability to predicting the Step 1 or Step 2 scores.
MCAT/UGPA	Ogunyemi and Taylor-Harris (2005)	Discovered a significant association between USMLE Step 2 exam score and UGPA, as well as between Step 2 score and MCAT score. Multiple MCAT attempts were associated with a failing Step 2 CK score, though UGPA did not independently predict failing Step 2 CK scores. The researchers reported that students who experienced difficulty on the MCAT were more likely to experience difficulty on the USMLE examinations.	The author collected USMLE scores, MCAT scores, UGPA, and demographic data from 171 medical students and performed a retrospective record review. Pearson correlation and Spearman correlation coefficients were used. A logistic binary regression analysis was performed.
MCAT/UGPA	Julian (2005)	Reported a wide range of predictive power of UGPA and MCAT scores on medical school performance. Discovered that MCAT alone may be a fairly good predictor of USMLE Step 1 and Step 2 CK scores, and not much power is added by taking UGPA into consideration. Author suggested that institutional differences in undergraduate grading may affect the usefulness of the UGPA, in contrast to the MCAT which is graded based on a national standard.	Collected USMLE scores from more than 31,000 students over two cohorts of students from 14 different medical schools for predicting medical school grades and academic difficulty/distinction, while their peers from all of the U.S. medical schools were used to predict performance on USMLE Steps 1, 2, and 3. Regression analyses assessedthe predictive power of combinations of uGPAs, MCAT scores, and undergraduate-institution selectivity.
MCAT	Donnon, Paolucci, & Violato (2007)	Investigated the predictive validity for each individual Step exam score and found that the total MCAT had the largest predictive validity for Step 1 and medium validity for Step 2.	Published a meta-analysis of 23 studies which investigated the predictive validity of the MCAT with regard to medical licensing examinations.
MCAT/UGPA	Dunleavy, Kroopnick, Dowd, Searcy, & Zhao (2013)	The authors concluded that a combination of MCAT and UGPA is the most reliable predictor of unimpeded progress among medical students. The authors concluded that MCAT scores are more powerful stand-alone predictors of unimpeded progress than UGPA. They suggested that Step exams are more similar to the MCAT than the metrics used to determine UGPA, which explains the stronger relationship between Step scores and MCAT scores. Suggest that the science sections of the MCAT, particularly the biological sciences section, are the greatest predictors of success on Step 1 and Step 2 CK, while the UGPA is moderately predictive of success on Step 1 and 2 CK	Collected data from students at 128 United States medical schools over 3 years. Variables included UGPA and MCAT score. The authors created three predictive models: model 1 only considered UGPA, model 2 only considered MCAT score, and model 3 considered both. They then compared the predictive validity of each model. Logistic regression analyses were used to estimate the relationships between UGPA and MCAT total scores.

MCAT/UGPA	Monteiro, George, Dollase, & Dumenco (2017)	Researchers developed a model to predict Step 2 Clinical Knowledge (CK) performance. This model was used to identify the need for educational intervention and to provide students with support services. The MCAT may have predictive value on Step 2 CK performance. The MCAT as a predictor of USMLE Step 1 and Step 2 CK performance is rooted in the classification of these examinations as standardized tests.	Data was collected from one medical school over 3 years. Created a regression equation to predict a student's Step 2 CK score from previous academic indicators. Assessed correlations between the academic indicators and Step 2 CK performance. Ensured that only predictors with a statistically significant association with Step 2 CK performance would be included in the regression models.
Non-Traditional Students	Ogunyemi and Taylor-Harris (2005)	Examined the interaction between age and Step 2 scores specifically and found that older students negatively correlated with Step 2 exam scores, but did not find a significant relationship between age and Step 2 exam failure	The author collected USMLE scores, MCAT scores, UGPA, and demographic data from 171 medical students and performed a retrospective record review. Pearson correlation and Spearman correlation coefficients were used. A logistic binary regression analysis was performed.
Non-Traditional Students	Gauer and Jackson (2018a)	Examined the relationship between examinee age and performance on MCAT, Step 1, and Step 2 CK exams. Significant differences were found between the Step 1 and Step 2 CK exam scores of traditional and non-traditional aged students (age 25 and older), with traditional-aged students outscoring non-traditional students. The authors hypothesized that age-related performance disparity may be due to external life events (older students are more likely to have commitments outside of medical school)	The authors collected data from 1,062 medical students at one medical school over 4 years. Conducted independent-samples t-tests for undergraduate science vs nonscience majors, traditional-aged vs nontraditional-aged students, gender, and state of legal residency status to explore group differences in scores on the MCAT, USMLE Step 1, and USMLE Step 2 CK exams. Effect sizes (Cohen's d) were calculated for each comparison.
SES & FirstGeneration	Grbic, Jones, & Case (2013)	Step 1 scores have been shown to be correlated with MCAT scores, which are also correlated with SES and strongly influenced by test preparation opportunities	The primary data source for this research was the AAMC MSQ, an annual survey of entering medical students that includes self-reported data on combined parental income in the previous year. A total of 126,856 matriculants provided parental income data.
SES	Cooter, Erdmann, Gonnella, Callahan, Hojat, Xu (2004)	The authors discovered that low SES was negatively associated with Step 1 and Step 2 scores, though the effect size was small. Step 1 scores have been shown to be correlated with MCAT scores, which are also correlated with SES and strongly influenced by test preparation opportunities.	The sample used in the current study included 1,464 graduates from one medical school over 10 years. Income data were provided via the parents' tax returns that were submitted as part of the financial aid application process. The ANOVA approach for continuous dependent variables was used: chi-square tests and ztests for proportions for discrete variables were employed. Correlational methods were not used because of the skewed distribution of family income. Effect size estimates were calculated when statistically significant associations were observed.
SES	Jerant, Henderson, Griffin, Talamantes, Fancher, Sousa, & Franks (2018)	The early negative association between SES and Step 1 scores may have been due to a limited opportunity for test preparation. this negative relationship was not found with regard to SES and Step 2 CK scores. The authors were encouraged by this finding, suggesting that low SES students were able to catch up with their peers and perform similarly on subsequent standardized tests	Collected data from four consecutive class years with a total of 402 participants. Measurements included uGPA, MCAT, MMI, SES, URM, Step 1, & Step 2 CK scores

SES	Gauer, Wolff, & Jackson (2016)	Step 1 scores have been shown to be correlated with MCAT scores, which are also correlated with SES and strongly influenced by test preparation opportunities	Multiple linear regression, correlation, and chi-square analyses were performed to determine the relationship between MCAT component and composite scores and USMLE Step 1 and Step 2 CK scores from five graduating classes at one Medical School.
Gender	Cuddy, Swanson, & Clauser (2008)	Males have been shown to consistently outperform females on preclinical exams, including Step 1 and the MCAT	A series of hierarchical linear models was conducted predicting Step 1 scores. The sample included 66,412 examinees from 133 U.S. LCME-accredited medical schools/ campuses.
Gender	Cuddy, Swanson, & Clauser (2007)	Women scored higher than men in nearly all content areas of the Step 2 CK exam. Step 1 scores were seen to be more strongly associated with subsequent Step 2 CK scores in men, less so in women. The authors hypothesized that the association between Step 1 and Step 2 CK scores is stronger in men because of their stronger science backgrounds.	Collected data, including both academic and demographic variables, from 23,538 students across 136 LCME medical schools. Descriptive statistics were computed, and a series of examinees nested-in-schools hierarchical linear models was conducted.
Gender	Cuddy, Swygert, Swanson, & Jobe (2011)	Women equal or outperform men in various clinical measurements, including Step 2 CK, Step 2 CS, and clerkship evaluations	Data included demographic and performance information for examinees that took Step 2 CS. The sample contained 27,910 examinees, 625 standardized patient/ case combinations, and 278,776 scored patient encounters. Hierarchical linear modeling techniques were employed.
Gender	Rubright, Jodoin, & Barone (2019)	Females scored significantly lower on the Step 1 exam than men. This trend was reversed for Step 2 CK, with women significantly outscoring men. When adding MCAT and undergraduate GPA covariates to the model, the authors found that gender differences persisted, suggesting that differences in gender performance could not be completely explained by differences in GPA or MCAT score	Collected a large sample of data from 45,154 medical students enrolled across 172 medical schools in the United States and Canada. Students had completed all three Step exams and a retrospective review was performed. Completed hierarchical linear modeling of data on USMLE Step 1, and completing USMLE Step 3. Main outcome measures were computer-based USMLE examinations: Step 1, Step 2 Clinical Knowledge, and Step 3. Test-taker characteristics included sex, self-identified race, U.S. citizenship status, English as a second language, and age at first Step 1 attempt. Covariates included composite MCAT scores, uGPA, and previous USMLE scores.
Gender	Gauer & Jackson, 2018a	Male students scored significantly higher than females on Step 1 exam and MCAT. Basic science backgrounds may be influencing the gender performance disparity. Gender was not associated with Step 2 CK performance. The authors believed this signified that female students had effectively caught up to their male peers	The authors collected data from 1,062 medical students at one medical school over 4 years. Conducted independent-samples t-tests for undergraduate science vs nonscience majors, traditional-aged vs nontraditional-aged students, gender, and state of legal residency status to explore group differences in scores on the MCAT, USMLE Step 1, and USMLE Step 2 CK exams. Effect sizes (Cohen’s d) were calculated for each comparison.

Post-bacc/Graduate Degree	Reeves, Vishwanatha, Yorio, Budd, & Sheedlo (2008)	Comparing the GPAs from the first year of medical school, PB students outperformed their traditional peers by a small degree. They also outperformed their traditional peers on the COMLEX I exam, osteopathic equivalent of USMLE Step 1 exam.	Collected data from students enrolled in a post-baccalaureate program at a Health Science University over 6 years.
Post-bacc/Graduate Degree	Sadik, Woldemariam, & Wang (2017)	Researchers determined there was a strong correlation between master's GPA and COMLEX I score and COMLEX II scores. There was no correlation between MCAT scores and COMLEX I or II performance of master's students. the authors concluded that master's or post-baccalaureate performance is more predictive of subsequent academic performance than traditional pre-admissions variables among students who have completed these types of programs	Compared students who completed the master's coursework with those who did not at an Osteopathic Medical School.

APPENDIX D

NEO-FFI – Conscientiousness Only

Please read each statement carefully. For each statement select the answer that best represents your opinion.

I keep my belongings neat and clean.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I'm pretty good about pacing myself so as to get things done.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I often come into situations without being fully prepared.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I try to perform all the tasks assigned to me conscientiously.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I have a clear set of goals and work toward them in an orderly fashion.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

I waste a lot of time before settling down to work.

- ○ Strongly Disagree
- ○ Disagree
- ○ Neutral
- ○ Agree
- ○ Strongly Agree

I work hard to accomplish my goals.

- ○ Strongly Disagree
- ○ Disagree
- ○ Neutral
- ○ Agree
- ○ Strongly Agree

When I make a commitment, I can always be counted on the follow through.

- ○ Strongly Disagree
- ○ Disagree
- ○ Neutral
- ○ Agree
- ○ Strongly Agree

Sometimes I'm not as dependable or reliable as I should be.

- ○ Strongly Disagree
- ○ Disagree
- ○ Neutral
- ○ Agree
- ○ Strongly Agree

I am a productive person who always gets the job done.

- ○ Strongly Disagree
- ○ Disagree
- ○ Neutral
- ○ Agree
- ○ Strongly Agree

I never seem to be able to get organized.

- ○ Strongly Disagree
- ○ Disagree
- ○ Neutral
- ○ Agree
- ○ Strongly Agree

I strive for excellence in everything I do.

- ○ Strongly Disagree
- ○ Disagree
- ○ Neutral
- ○ Agree
- ○ Strongly Agree

Adapted and reproduced by special permission of the Publisher, Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Lutz, Florida 33549, from the NEO Five-Factor Inventory-3 by Paul Costa, Jr., PhD and Robert McCrae, PhD, copyright 1978, 1985, 1989, 1991, 2003, 2010 by PAR, Inc. Further reproduction is prohibited without permission of PAR, Inc.

APPENDIX E

Short Grit Scale

Directions for taking the Grit Scale: Here are a number of statements that may or may not apply to you. For the most accurate score, when responding, think of how you compare to most people – not just the people you know well, but most people in the world. There are no right or wrong answers, so just answer honestly. In the end, you'll get a score that reflects how passionate and persevering you see yourself to be.

1. New ideas and projects sometimes distract me from previous ones.

- Very much like me
- Mostly like me
- Somewhat like me
- Not much like me
- Not like me at all

2. Setbacks don’t discourage me. I don’t give up easily.

- Very much like me
- Mostly like me
- Somewhat like me
- Not much like me
- Not like me at all

3. I often set a goal but later choose to pursue a different one.

- Very much like me
- Mostly like me
- Somewhat like me
- Not much like me
- Not like me at all

4. I am a hard worker.

- Very much like me
- Mostly like me
- Somewhat like me
- Not much like me

- Not like me at all

5. I have difficulty maintaining my focus on projects that take more than a few months to complete.

- Very much like me
- Mostly like me
- Somewhat like me
- Not much like me
- Not like me at all

6. I finish whatever I begin.

- Very much like me
- Mostly like me
- Somewhat like me
- Not much like me
- Not like me at all

7. I am diligent. I never give up.

- Very much like me
- Mostly like me
- Somewhat like me
- Not much like me
- Not like me at all

8. I have been obsessed with a certain idea or project for a short time but later lost interest.

- Very much like me
- Mostly like me
- Somewhat like me
- Not much like me
- Not like me at all

Scoring:

For questions 2, 4, 7, and 8 assign the following points:

5 = Very Much Like Me
4 = Mostly Like Me
3 = Somewhat Like Me
2 = Not Much Like Me
1 = Not Like Me at All

For questions 1, 3, 5, and 6 assign the following points:

1 = Very Much Like Me
2 = Mostly Like Me
3 = Somewhat Like Me
4 = Not Much Like Me
5 = Not Like Me at All

Grit Scale Citation

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short grit scale (Grit-S). *Journal of Personality Assessment, 91*(2), 166-174. doi: 10.1080/00223890802634290

Duckworth, A.L., Peterson, C., Matthews, M.D., & Kelly, D.R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 9*, 1087-1101.

APPENDIX F

Subject line: Informed consent for Medical Student Survey

Good afternoon!

I am hoping you could take a few minutes to help me. As you may or may not know, I am currently working on my PhD and my dissertation research is on non-cognitive variables and success on Step 2 CK. As a medical student, your input is critical to my research.

If you would be willing to complete this short Grit-S survey (Duckworth & Quinn, 2009), it should take you no more than 5 minutes, and I would be very grateful.

My hope is to prove that there are non-cognitive variables that are just as important, if not more important, in some cases, than cognitive factors in predicting success in medical school. Therefore, if you choose to participate, you are giving me permission to access your AMCAS application and your student academic record through the student information system (Banner) for research purposes only. Therefore, I need your consent to report the data from this survey and review your academic record for MCAT and undergraduate GPA, and performance on USMLE Step 2 CK.

Please know, your confidentiality is of my utmost concern and you will not be named in my research. All information will be presented in aggregate. I will take steps to guard your privacy and be sure that your information is kept confidential. My records will not be accessible to anyone but me and I will keep them locked securely. I will assign you a unique identifier, and I will keep the coded log in a locked storage area. I will destroy all information, including consents and surveys, at the completion of this study. Your data, with your unique identifier will be entered into SPSS and will be kept on a password-protected laptop. I will be the only one with access to your survey and academic record. Whenever one works with email or the internet there is always the risk of compromising privacy, confidentiality and/or anonymity. Your confidentiality will be maintained to the degree permitted by the technology being used. It is important for you to understand that no guarantees can be made regarding the interception of data sent via the internet by third parties. As noted above, the information on this survey and from your academic records will be reported in aggregate form only, and your information will not be identifiable.

To thank you for your participation, I will enter you (based on your student ID so please make sure you enter it correctly) in a raffle. I will randomly select 4 students to each receive a $50 Amazon Gift Card.

The benefit of this research will help us to better understand other personal qualities to assess in order to help predict success on Step 2 CK. There are no direct benefits to you by participating in this study. The risks to you are minimal and there are no further surveys you will need to complete. Should you choose not to participate in this voluntary research that is fine too and you may also withdraw from this study at any time without penalty.

By choosing to participate in this voluntary study, you affirm that you are 18 years of age or older and give me your consent to access your AMCAS application and your academic record at SUNY Upstate Medical University.

Should you have any questions, please do not hesitate to contact me or stop in to see me.

The link to the survey is: https://www.surveymonkey.com/r/DQRTYZP

Thank you so much for your time!
My Best,

Jennifer
Jennifer Welch
Associate Dean, Admissions & Financial Aid

1212 Weiskotten Hall
766 Irving Avenue
Syracuse, NY 13210
welchj@upstate.edu

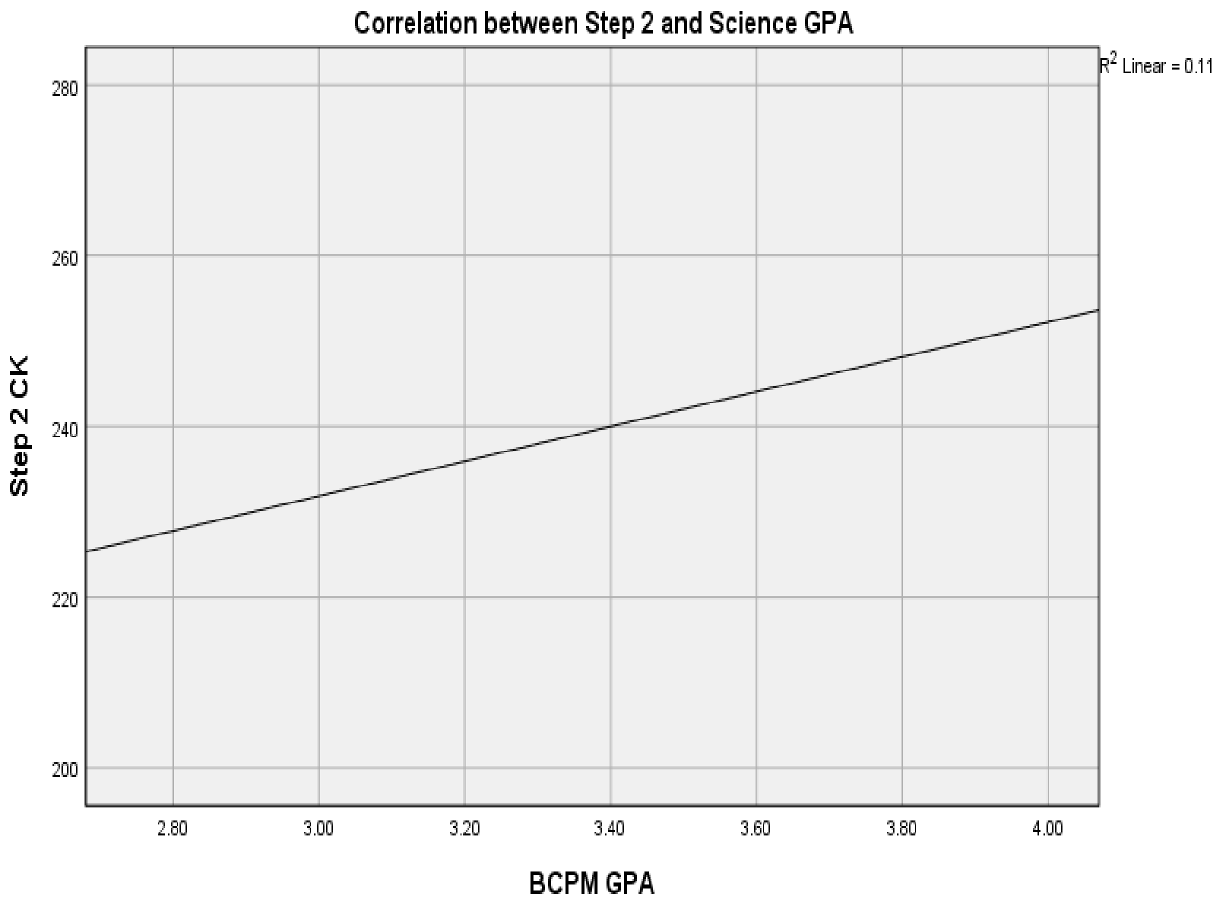
315-464-4816

Dissertation Chair:

Catherine Engstrom, PhD
Department of Higher Education, Syracuse University
cmengstr@syr.edu

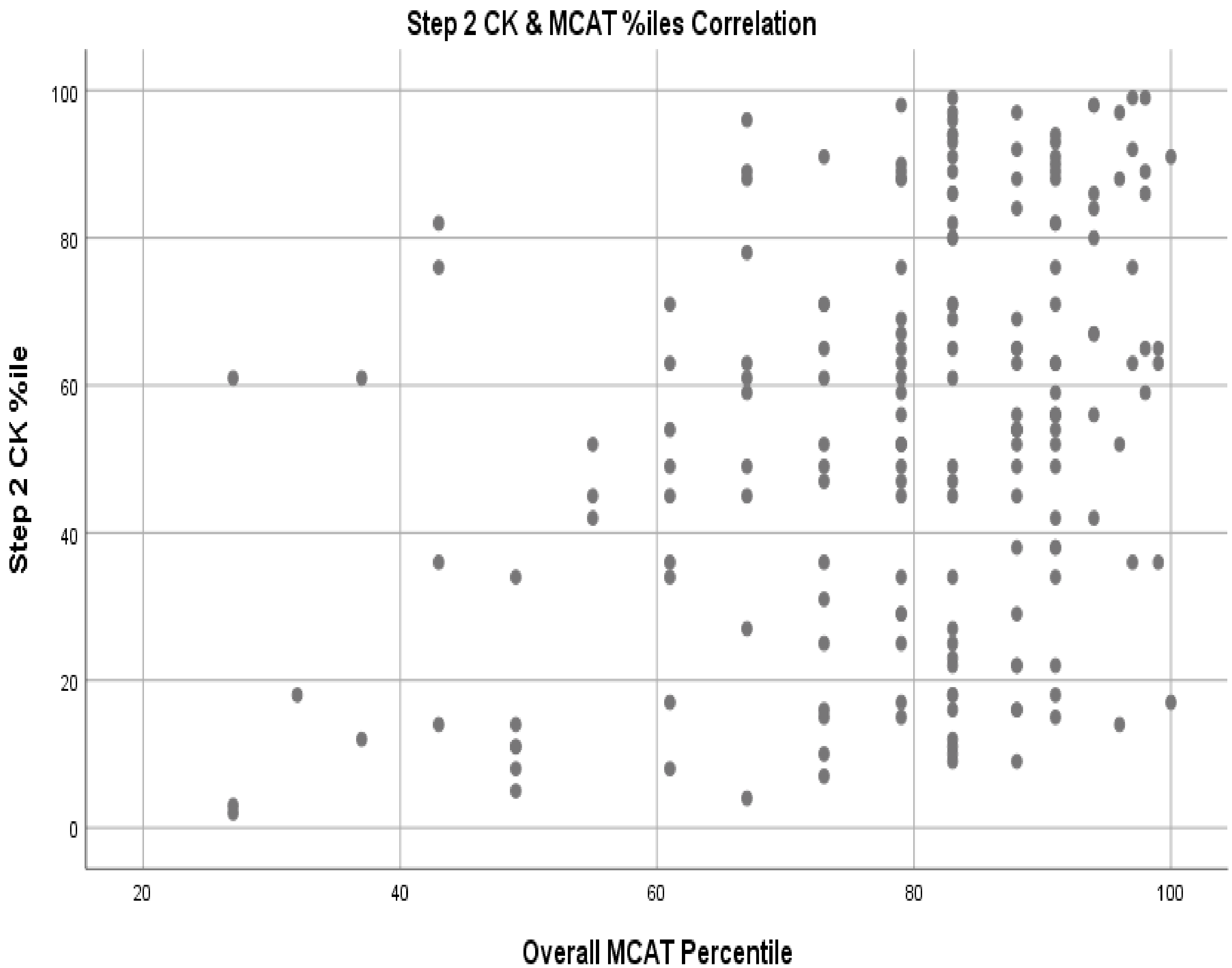
APPENDIX G

Correlation Between USMLE Step 2 CK and Science GPA



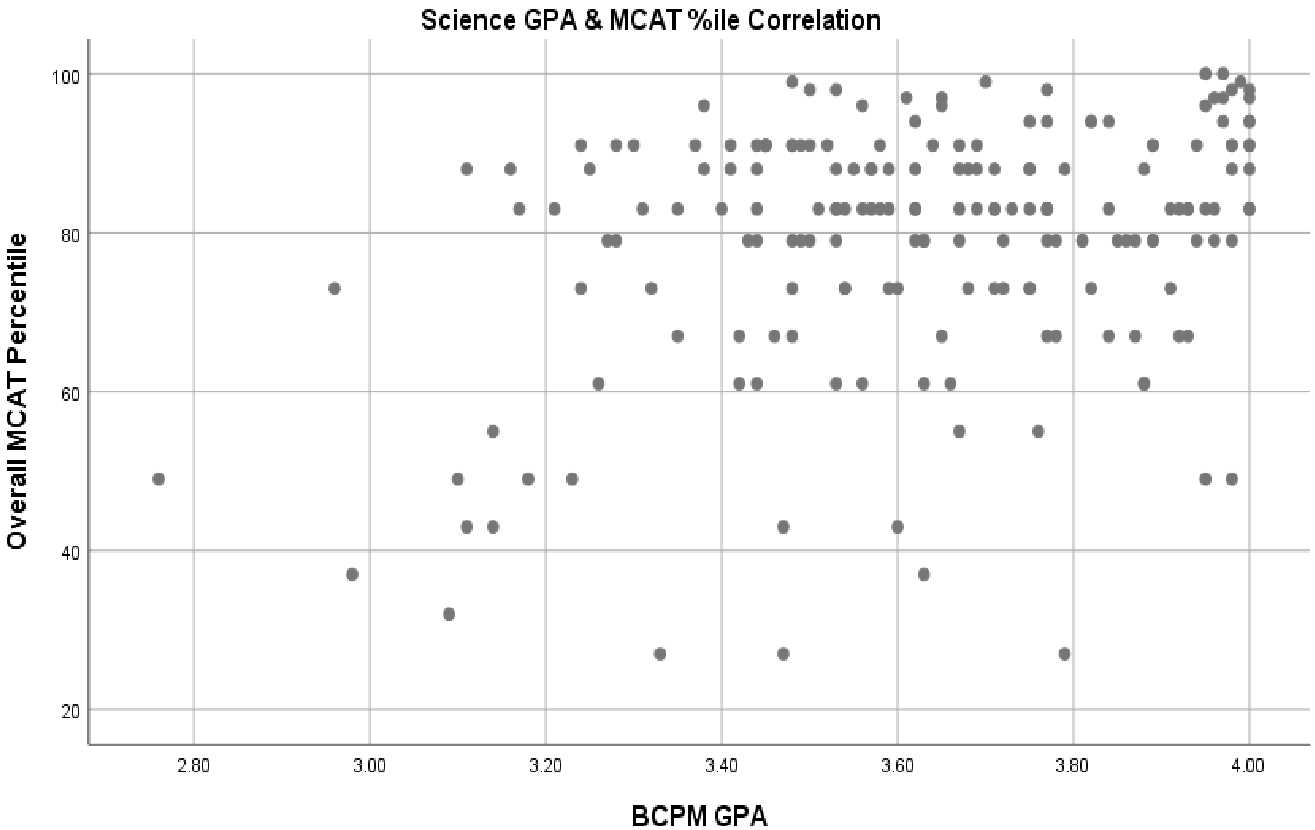
APPENDIX H

Correlation Between USMLE Step 2 CK Percentile and MCAT Percentile



APPENDIX I

Collinearity Between Science GPA and MCAT Percentile



APPENDIX J

SUNY Upstate Medical University IRB Approval

UPSTATE
MEDICAL UNIVERSITY

Institutional Review Board for the
Protection of Human Subjects
750 East Adams Street
Syracuse, NY 13210

DATE:	March 18, 2019
TO:	Jennifer Welch, MS
FROM:	SUNY Upstate IRB
SUBMISSION TYPE:	New Project
PROJECT TITLE:	[1410005-1] Noncognitive Traits and Success on Step 2
ACTION:	DETERMINATION OF EXEMPT STATUS
DECISION DATE:	March 17, 2019
REVIEW CATEGORY:	Exemption category # 2 (ii) & 4 (ii)

Thank you for your submission of New Project materials for this project. The SUNY Upstate IRB has determined this project is EXEMPT FROM IRB REVIEW according to federal regulations.

Any changes to this study will necessitate further review by this office.

If you have any questions, please contact Michelle Saya at (315) 464-4349 or sayam@upstate.edu. Please include your project title and reference number in all correspondence with this committee.

APPENDIX K

Syracuse University IRB Approval

SYRACUSE UNIVERSITY

INSTITUTIONAL REVIEW BOARD
MEMORANDUM

TO:	Catherine Engstrom
DATE:	April 12, 2019
SUBJECT:	**Determination of Exemption from Regulations**
IRB #:	19-105
TITLE:	*Non-Cognitive Traits and Medical Student Success on Step 2*

The above referenced application, submitted for consideration as exempt from federal regulations as defined in 45 C.F.R. 46, has been evaluated by the Institutional Review Board (IRB) for the following:

1. determination that it falls within the one or more of the eight exempt categories allowed by the organization;
2. determination that the research meets the organization's ethical standards.

It has been determined by the IRB this protocol qualifies for exemption and has been assigned to category 2 & 4. This authorization will remain active for a period of five years from **April 11, 2019** until **April 10, 2024**.

CHANGES TO PROTOCOL: Proposed changes to this protocol during the period for which IRB authorization has already been given, cannot be initiated without additional IRB review. If there is a change in your research, you should notify the IRB immediately to determine whether your research protocol continues to qualify for exemption or if submission of an expedited or full board IRB protocol is required. Information about the University's human participants protection program can be found at: http://researchintegrity.syr.edu/human-research/ Protocol changes are requested on an amendment application available on the IRB web site; please reference your IRB number and attach any documents that are being amended.

STUDY COMPLETION: Study completion is when all research activities are complete or when a study is closed to enrollment and only data analysis remains on data that have been de-identified. A Study Closure Form should be completed and submitted to the IRB for review (Study Closure Form).

Thank you for your cooperation in our shared efforts to assure that the rights and welfare of people participating in research are protected.

Tracy J. Cromp

Tracy Cromp, M.S.W.
Director

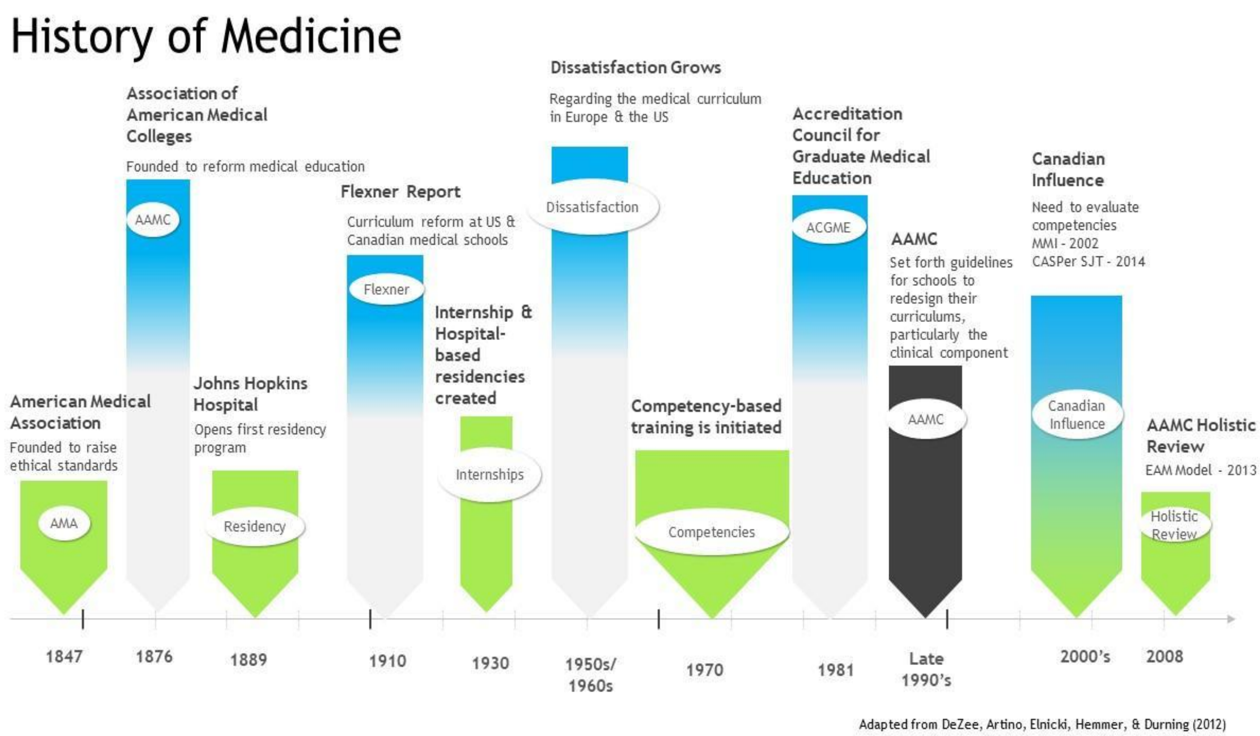
DEPT: Higher Education, 350 Huntington Hall

STUDENT: Jennifer Welch

Research Integrity and Protections | 214 Lyman Hall | Syracuse, NY 13244-1200 | 315.443.3013 | orip.syr.edu

APPENDIX L

Historical timeline of medical education in the United States.



REFERENCES

Abrams, A., & Ducharme, J. (2020, April 13). *Meet the medical students becoming doctors in the middle of a pandemic*. Time. https://time.com/5820046/medical-students-covid-19/

Agrawal, J. R., Vlaicu, S., & Carrasquillo, O. (2005). Progress and pitfalls in underrepresented minority recruitment: Perspectives from the medical schools. *Journal of the National Medical Association*, *97*(9), 1226.

Aiken, L. (2011). Nurses for the future. *The New England Journal of Medicine*, *364*(3), 196-198. doi: 10.1056/NEJMp1011639

Akos, P., & Kretchmar, J. (2017). Investigating grit at a non-cognitive predictor of college success. *The Review of Higher Education*, *40*(2), 163-186.

Alameida, M. D., Prive, A., Davis, H. C., Landry, L., Renwanz-Boyle, A., & Dunham, M. (2011). Predicting NCLEX-RN success in a diverse student population. *Journal of Nursing Education 50*(5), 261-267. doi: 10.3928/01484834-20110228-01

Albanese, M. A., Snow, M. H., Skochelak, S. E., Huggett, K. N., & Farrell, P. M. (2003). Assessing personal qualities in medical school admissions. *Academic Medicine*, *78*(3), 313-321.

Altus Assessments. (2020). *CASPer: How it works*. Altus Assessments. https://altusassessments.com/how-it-works.

Alvin, M. D. (2016). The USMLE Step 2 CS: Time for a change. *Medical Teacher*, *38*(8), 854-856.

Alvarado, L., Capozza, K., Jackson, K., & Russell, A. (2008). HEALTH PROFESSIONS ANNOTATED BIBLIOGRAPHY.

Ambady, N., LaPlante, D., Nguyen, T., Rosenthal, R., Chaumeton, N., & Levinson, W. (2002). Surgeons' tone of voice: A clue to malpractice history. *Surgery*, *132*(1), 5-9.

American Medical Association. (2019, June 12). New policy aimed at increasing diversity in physician workforce. [Press release]. https://www.ama-assn.org/press-center/press-releases/new-policy-aimed-increasing-diversity-physician-workforce.

Anderson, C., Turner, A. C., Heath, R. D., & Payne, C. M. (2016). On the meaning of grit… and hope… and fate control… and alienation… and locus of control… and… self-efficacy… and… effort optimism… and…. *The Urban Review*, *48*(2), 198-219.

Andre, B. K. (2020). An Analysis of Inclusion Resources, As a Pathway to Retention of Minority Students in the Medical Profession.

Andriole, D. A., & Jeffe, D. B. (2012). A national cohort study of US medical school students who initially failed Step 1 of the United States Medical Licensing

Arredondo, A. (2015). *How incoming student experiences, attributes and metrics are related to academic performance in a combined-degree medical program: A comparison of holistic and non-holistic admissions practices.* [Doctoral dissertation, University of Kansas]. KU Scholar Works. https://kuscholarworks.ku.edu/handle/1808/19501

Association of American Medical Colleges (AAMC). (2015). *Assessing Financial Need Using the AAMC Socioeconomic Status Indicators (E01, E02). AAMC.* https://www.aamc.org/professional-development/affinity-groups/gsa/webinars/assessing-financial-need

Association of American Medical Colleges (AAMC). (2016). *Holistic review*. AAMC. https://www.aamc.org/services/member-capacity-building/holistic-review.

Association of American Medical Colleges (AAMC). (2018). *Graduation Rates and Attrition Rates of US Medical Students.* AAMC.

https://www.aamc.org/system/files/reports/1/graduationratesandattritionratesofu.s.medicalstudents.pdf

Association of American Medical Colleges. (2019). *New findings confirm predictions on physician shortage.* AAMC. https://www.aamc.org/news-insights/press-releases/new-findings-confirm-predictions-physician-shortage

Association of American Medical Colleges. (2020a). *2019 FACTS: Applicants and matriculants data*. AAMC. https://www.aamc.org/data-reports/students-residents/interactive-data/2019-facts-applicants-and-matriculants-data

Association of American Medical Colleges. (2020b). *2020 AMCAS applicant guide*. AAMC. https://aamc-orange.global.ssl.fastly.net/production/media/filer_public/b2/23/b223c482-8ba3-44dd-bb1c-8835ac84f3e6/2020amcasapplicantguide-060419.pdf

Association of American Medical Colleges. (2020c). *Age of applicants to US medical schools*. AAMC. https://www.aamc.org/system/files/d/1/321468-factstablea6.pdf.

Association of American Medical Colleges (2020d). *Diversity in medicine: Facts and figures 2019*. AAMC. https://www.aamc.org/data/

Association of American Medical Colleges. (2020e). *Underrepresented in medicine definition*. AAMC. https://www.aamc.org/what-we-do/mission-areas/diversity-inclusion/underrepresented-in-medicine

Association of American Medical Colleges. (2020f). *Postbaccalaureate premedical programs*. AAMC. https://apps.aamc.org/postbac/#/index

Association of American Medical Colleges. (2020g). *Research to Evaluate the Fairness, Use, and Predictive Validity of the MCAT® Exam Introduced in 2015.* AAMC.

https://www.aamc.org/system/files/2020-03/services-mcat-article-collection-academic-medicine-03212020.pdf

Association of American Medical Colleges. (2020h). *Who we are*. AAMC. https://www.aamc.org/who-we-are

Austin, E. J., Evans, P., Goldwater, R., & Potter, V. (2005). A preliminary study of emotional intelligence, empathy and exam performance in first year medical students. *Personality and Individual Differences*, *39*(8), 1395-1405.

Axelson, R., Kreiter, C., Ferguson, K., Solow, C., & Huebner, K. (2010). Medical school preadmission interviews: Are structured interviews more reliable than unstructured interviews? *Teaching and Learning in Medicine*, *22*(4), 241-245.

Babaria, P., Abedin, S., & Nunez-Smith, M. (2009). The effect of gender on the clinical clerkship experiences of female medical students: Results from a qualitative study. *Academic Medicine*, *84*(7), 859-866.

Bach, P. B., Pham, H. H., Schrag, D., Tate, R. C., & Hargraves, J. L. (2004). Primary care physicians who treat blacks and whites. *New England Journal of Medicine*, *351*(6), 575-584.

Baill, I. C., Khallouq, B. B., Joledo, O., Jacobs, A., & Larkin, R. (2019). How Postbaccalaureate Career Changer and Traditional Medical Students Differ Academically. *Southern Medical Journal*, *112*(12), 610-616.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*(1-2), 9-30.

Basco, Jr., W. T., Way, D. P., Gilbert, G. E., & Hudson, A. (2002). Undergraduate institutional MCAT scores as predictors of USMLE Step 1 performance. *Academic Medicine*, *77*(10), S13-S16.

Bassett, A. M., Brosnan, C., Southgate, E., & Lempp, H. (2018). Transitional journeys into, and through medical education for First-in-Family (FiF) students: a qualitative interview study. *BMC Medical Education*, *18*(1), 102.

Baum, S., & Steele, P. (2017). Who goes to graduate school and who succeeds? *AccessLex Institute Research Paper*, (17-01).

Beacham, T., Askew, R. W., & Williams, P. R. (2009). Strategies to increase racial/ethnic student participation in the nursing profession. *ABNF Journal*, *20*(3), 69-72.

Behling, O. (1998). Employee selection: Will intelligence and conscientiousness do the job? *Academy of Management Perspectives*, *12*(1), 77-86.

Benner, P., Sutphen, M., Leonard, V., & Day, L. (2010). *Educating nurses: A call for radical transformation*. Jossey-Bass.

Bickel, J. (2001). Gender equity in undergraduate medical education: A status report. *Journal of Women's Health & Gender-Based Medicine*, *10*(3), 261-270.

Bickel, J., Wara, D., Atkinson, B. F., Cohen, L. S., Dunn, M., Hostler, S., Johnson, T. R. B., Morahan, P., Rubenstein, A. H., Sheldon, G. F., & Stokes, E. (2002). Increasing women's leadership in academic medicine: Report of the AAMC Project Implementation Committee. *Academic Medicine*, *77*(10), 1043-1061.

Biggs, J. S. G., Najman, J. M., Schulz, E. B., & Williams, G. (1991). Parental problems influencing the academic achievement of medical students: A prospective study. *Medical Education*. *25*(5), 374-382. doi:10.1111/j.1365-2923.1991.tb00084.x.

Bills, S. E., Karst, G., & Meyer, K. (2019, October 20). *Chart a course for holistic admissions transformation using an interprofessional model*. [Poster/presentation]. University of Nebraska Medical Center Educational Leadership Conference, Omaha. https://digitalcommons.unmc.edu/cahp_pres/1/

Blumenthal, D. (2020, April 13). *Professionalism drives health care workers to treat patients with Covid-19. We mustn't jeopardize it.* STAT News. https://www.statnews.com/2020/04/13/professionalism-health-care-workers-treat-covid-19/

Boatright, D., Ross, D., O'Connor, P., Moore, E., & Nunez-Smith, M. (2017). Racial disparities in medical student membership in the Alpha Omega Alpha honor society. *JAMA Internal Medicine*. *177*(5), 659. doi:10.1001/jamainternmed.2016.9623.

Bogdan, R. C., & Biklen, S. K. (2006). *Qualitative research for education: An introduction to theory and methods* (5th edition). Pearson.

Bore, M., Munro, D., & Powis, D. (2009). A comprehensive model for the selection of medical students. *Medical Teacher*, *31*(12), 1066-1072.

Bourdieu, P. (1986). The forms of capital. In J. C. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241-258). Greenwood Press.

Bowen, W. G., & Bok, D. (2016). *The shape of the river: Long-term consequences of considering race in college and university admissions* (Vol. 96). Princeton University Press.

Bowen, W. G., & Rudenstine, N. L. (2003, February 7). Race-sensitive admissions: Back to basics. *Chronicle of Higher Education*, B7-11.

Bowman, N. A., Hill, P. L., Denson, N., & Bronkema, R. (2015). Keep on truckin'or stay the course? Exploring grit dimensions as differential predictors of educational achievement, satisfaction, and intentions. *Social Psychological and Personality Science*, *6*(6), 639-645.

Boyle, G. J., Matthews, G., & Saklofske, D. H. (Eds.). (2008). *The SAGE handbook of personality theory and assessment: Personality measurement and testing* (Vol. 2). SAGE.

Bradby, H., Lindenmeyer, A., Phillimore, J., Padilla, B., & Brand, T. (2020). 'If there were doctors who could understand our problems, I would already be better': Dissatisfactory health care and marginalization in superdiverse neighborhoods. *Sociology of Health & Illness*, *42*(4), 739-757.

Brady, L. M., Germano, A. L., & Fryberg, S. A. (2017). Leveraging cultural differences to promote educational equality. *Current Opinion in Psychology*, *18*, 79-83.

Brandt, J. S., Kramer, S. E., & Chasen, S. T. (2013). Academic differences among male and female candidates applying for obstetrics and gynecology residency: The experience of one program. *Journal of Surgical Education*, *70*(4), 439-442. doi:10.1016/j.jsurg.2013.03.012.

Brosnan, C., Southgate, E., Outram, S., Lempp, H., Wright, S., Saxby, T., Harris, G., Bennett, A., & Kelly, B. (2016). Experiences of medical students who are first in family to attend university. *Medical Education*, *50*(8), 842-851.

Brummer S., Reyes I., Martin M. L., Walker L. U., & Heron S. L. (2016) Racial/ethnic health care disparities and inequities: Historical perspectives. In M. Martin, S. Heron, L. Moreno-Walton, & A. Jones (Eds.), *Diversity and inclusion in quality patient care* (pp. 11-21). Springer.

Buerhaus, P. I., Staiger, D. O., & Auerbach, D. I. (2009). *The future of nursing workforce in the United States: Data, trends, and implications*. Jones and Bartlett, LLC.

Burk-Rafel, J., Pulido, R. W., Elfanagely, Y., & Kolars, J. C. (2019). Institutional differences in USMLE Step 1 and 2 CK performance: Cross-sectional study of 89 US allopathic medical schools. *PloS one*, *14*(11).

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences*, *29*(6), 1057-1068.

Cammarata, L. (2010). Foreign language teachers' struggle to learn content-based instruction. *L2 Journal*, *2*(1), 89-118.

Camp, C. L., Wang, D., Turner, N. S., Grawe, B. M., Kogan, M., & Kelly, A. M. (2019). Objective predictors of grit, self-control, and conscientiousness in orthopaedic surgery residency applicants. *Journal of the American Academy of Orthopaedic Surgeons*, *27*(5), e227-e234.

Cantwell, B., Canche, M., Milem, J., & Sutton, F. (2010). Do the data support the discourse? Assessing holistic review as an admissions process to promote diversity at a US medical school. A*nnual Meeting of the Association for the Study of Higher Education.* Paper presented at the meeting of the Association for the Study of Higher Education, Indianapolis, Indiana.

Carnevale, A. P., & Strohl, J. (2013). *Separate & unequal: How higher education reinforces the intergenerational reproduction of white racial privilege*. Georgetown Public Policy Institute.

Carter, P. L., & Welner, K. G. (Eds.). (2013). *Closing the opportunity gap: What America must do to give every child an even chance*. Oxford University Press.

Caruso, J. C. (2000). Reliability generalization of the NEO personality scales. *Educational and Psychological Measurement*, *60*(2), 236-254.

Casas, R. S., Xuan, Z., Jackson, A. H., Stanfield, L. E., Harvey, N. C., & Chen, D. C. (2017). Associations of medical student empathy with clinical competence. *Patient Education and Counseling*, *100*(4), 742-747. doi:10.1016/j.pec.2016.11.006.

Case, S. M., Ripkey, D. R., & Swanson, D. B. (1996). The relationship between clinical science performance in 20 medical schools and performance on Step 2 of the USMLE licensing examination. 1994-95 validity study group for USMLE Step 1 and 2 pass/fail standards. *Academic Medicine, 71*(1), S31-3.

Casey, P. M., Palmer, B. A., Thompson, G. B., Laack, T. A., Thomas, M. R., Hartz, M. F., Jensen, J. R., Sandefur, B. J., Hammack, J. E., Swanson, J. W., Sheeler, R. D., & Grande, J. P. (2016). Predictors of medical school clerkship performance: A multispecialty longitudinal analysis of standardized examination scores and clinical assessments. *BMC Medical Education, 16*(1), 128.

Cestone, C., Russo, D., Berardi-Demo, L., Weber, V., White, M., Shumsky, J., Baird, D., Walker, L., & Núñez, A. (2018). Broadening diversity in medicine: Outcomes of the Drexel Pathway to Medical School postbaccalaureate medical science program. *Journal of Best Practices in Health Professions Diversity*, *11*(1), 1-12.

Chang, A., Boscardin, C., Chou, C. L., Loeser, H., & Hauer, K. E. (2009). Predicting failing performance on a standardized patient clinical performance examination: The importance of communication and professionalism skills deficits. *Academic Medicine, 84*(10), S101-S104.

Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*(2), 268. doi: 10.1037/0022-0663.97.2.268.

Chaudhry, H. J., Katsufrakis, P. J., & Tallia, A. F. (2020). The USMLE Step 1 decision: An opportunity for medical education and training. *Journal of American Medical Association*, *323*(20), 2022-3.

Chaytor, A. T., Spence, J., Armstrong, A., & McLachlan, J. C. (2012). Do students learn to be more conscientious at medical school? *BMC Medical Education*, *12*(1), 54

Christopoulou, M., Lakioti, A., Pezirkianidis, C., Karakasidou, E., & Stalikas, A. (2018). The role of grit in education: A systematic review. *Psychology*, *9*(15), 2951-2971.

Chur-Hansen, A. (2003). Mature-aged medical students: A qualitative study. *Learning in Health and Social Care*, *2*(3), 159-168. doi:10.1046/j.1473-6861.2003.00048.x.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

Cohen, J. J., Gabriel, B. A., & Terrell, C. (2002). The case for diversity in the health care workforce. *Health Affairs*, *21*(5), 90-102.

Cohen, J. J., & Steinecke, A. (2006). Building a diverse physician workforce. *Journal of the American Medical Association*, *296*(9), 1135-1137.

Cohen, M. J., Kay, A., Youakim, J. M., & Balaicuis, J. M. (2009). Identity transformation in medical students. *The American Journal of Psychoanalysis*, *69*(1), 43-52.

Colbert-Getz, J. M., Fleishman, C., Jung, J., & Shilkofski, N. (2013). How do gender and anxiety affect students' self-assessment and actual performance on a high-stakes clinical skills examination? *Academic Medicine*, *88*(1), 44-48.

College Board and Educational Testing Service. (2016). 2015 College-bound seniors: Total group profile report, 2015. College Board. https://reports.collegeboard.org/pdf/total-group-2016.pdf

Conard, M. A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality*, *40*(3), 339-346.

Conrad, N., & Patry, M. W. (2012). Conscientiousness and academic performance: A mediational analysis. *International Journal for the Scholarship of Teaching and Learning*, *6*(1), n1.

Conrad, S. S., Addams, A. N., & Young, G. H. (2016). Holistic review in medical school admissions and selection: A strategic, mission-driven response to shifting societal needs. *Academic Medicine*, *91*(11), 1472-1474.

Cooper, R. A. (1994). Seeking a balanced physician workforce for the 21st century. *Journal of the American Medical Association*, *272*(9), 680-687.

Cooter, R., Erdmann, J. B., Gonnella, J. S., Callahan, C. A., Hojat, M., Xu, G. (2004). Economic diversity in medical education. *Evaluation & the Health Professions*, *27*(3), 252-264. doi:10.1177/0163278704267041.

Cortez, A. R., Winer, L. K., Kassam, A. F., Hanseman, D. J., Kuethe, J. W., Sussman, J. J., Quillin, III, R. C. (2020). Exploring the relationship between burnout and grit during general surgery residency: A longitudinal, single-institution analysis. *The American Journal of Surgery, 219*(2), 322-327.

Costa, P. T., & McCrae, R. R. (1989). *NEO five-factor inventory (NEO-FFI)*. Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, *4*(1), 5.

Costa, P. T., & McCrae, R. R. (2008). The revised NEO Personality Inventory (NEO-PI-R). *The SAGE Handbook of Personality Theory and Assessment*, *2*(2), 179-198.

Costa, Jr., P. T., & McCrae, R. R. (2010). Bridging the gap with the five-factor model. *Personal Discord*, *1*(2), 127-30. doi: 10.1037/a0020264.

Cuddy, M. M., Swanson, D. B., & Clauser, B. E. (2007). A multilevel analysis of the relationships between examinee gender and United States Medical Licensing Exam (USMLE) Step 2 CK content area performance. *Academic Medicine*, *82*(10), S89-S93. doi:10.1097/acm.0b013e3181400379.

Cuddy, M. M., Swanson, D. B., & Clauser, B. E. (2008). A multilevel analysis of examinee gender and USMLE Step 1 performance. *Academic Medicine*, *83*(10), S58-S62. doi:10.1097/acm.0b013e318183cd65.

Cuddy, M. M., Swygert, K. A., Swanson, D. B., & Jobe, A. C. (2011). A multilevel analysis of examinee gender, standardized patient gender, and United States medical licensing examination Step 2 clinical skills communication and interpersonal skills scores. *Academic Medicine*, *86*(10), S17-S20. doi:10.1097/acm.0b013e31822a6c05.

Cuddy, M. M., Winward, M. L., Johnston, M. M., Lipner, R. S., & Clauser, B. E. (2016). Evaluating validity evidence for USMLE Step 2 clinical skills data gathering and data interpretation scores: Does performance predict history-taking and physical examination ratings for first-year internal medicine residents? *Academic Medicine*, *91*(1), 133-139.

Dahl, M. (2016, May 9). *Don't believe the hype about grit, pleads the scientist behind the concept*. New York Magazine. https://www.thecut.com/2016/05/dont-believe-the-hype-about-grit-pleads-the-scientist-behind-the-concept.html#_ga=2.204071120.1101794668.1595190212-1317128504.1595190212

Dall, T., West, T., Chakrabarti, R., & Iacobucci, W. (2017). The complexities of physician supply and demand 2016 update: Projections from 2014 to 2025. IHS, Inc.

Damgaard, M. T., & Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review*, *64*, 313-342.

Dapremont, J. A. (2011). Success in nursing school: Black nursing students' perception of peers, family, and faculty. *Journal of Nursing Education, 50*(5), 254-260.

Datu, J. A. D., & McInerney, D. M. M. (2017). Does culture matter for grit? Mapping cross-cultural directions in grit research programs. In F. Guay, H. W. Marsh, D. M. McInerney, & R. G. Craven (Eds.), *Self: Driving positive psychology and well-being* (p. 113-133). IAP Information Age Publishing.

Datu, J. A. D., Yuen, M., & Chen, G. (2017). Development and validation of the Triarchic Model of Grit Scale (TMGS): Evidence from Filipino undergraduate students. *Personality and Individual Differences*, *114*, 198-205.

Davis, D., Dorsey, J. K., Franks, R. D., Sackett, P. R., Searcy, C. A., & Zhao, X. (2013). Do racial and ethnic group differences in performance on the MCAT exam reflect test bias? *Academic Medicine*, *88*(5), 593-602.

Davis, K., Schoenbaum, S. C., & Audet, A. M. (2005). A 2020 vision of patient-centered primary care. *Journal of General Internal Medicine*, *20*(10), 953-957.

Davis, D., Dorsey, J. K., Franks, R. D., Sackett, P. R., Searcy, C. A., & Zhao, X. (2013). Do racial and ethnic group differences in performance on the MCAT exam reflect test bias?. *Academic Medicine*, *88*(5), 593-602.

Dee, T. S., & West, M. R. (2011). The non-cognitive returns to class size. *Educational Evaluation and Policy Analysis*, *33*(1), 23-46.

Deil-Amen, R., & Tevis, T. L. (2010). Circumscribed agency: The relevance of standardized college entrance exams for low SES high school students. *The Review of Higher Education*, *33*(2), 141-175.

Del Canale, S., Louis, D. Z., Maio, V., Wang, X., Rossi, G., Hojat, M., & Gonnella, J. S. (2012). Physicians' empathy and disease complications: An empirical study of primary care physicians and their diabetic patients in Parma, Italy. *Academic Medicine, 87*, 1243-1249.

DeZee, K. J., Artino, A. R., Elnicki, D. M., Hemmer, P. A., & Durning, S. J. (2012). Medical education in the United States of America. *Medical teacher*, *34*(7), 521-525.

Doherty, E. M., & Nugent, E. (2011). Personality factors and medical training: A review of the literature. *Medical Education*, *45*(2), 132-140.

Donaldson, M. S., Corrigan, J. M., & Kohn, L. T. (Eds.). (2000). *To err is human: Building a safer health system* (Vol. 6). National Academies Press.

Dong, T., Swygert, K. A., Durning, S. J., Saguil, A., Gilliland, W. R., Cruess, D., DeZee, K. J., LaRochelle, J., & Artino Jr, A. R. (2014). Validity evidence for medical school OSCEs: Associations with USMLE® Step assessments. *Teaching and Learning in Medicine*, *26*(4), 379-386.

Donnon, T., Paolucci, E. O., & Violato, C. (2007). The predictive validity of the MCAT for medical school performance and medical board licensing examinations: A meta-analysis of the published research. *Academic Medicine*, *82*(1), 100-106.

Dore, K. L., Reiter, H. I., Eva, K. W., Krueger, S., Scriven, E., Siu, E., ... & Norman, G. R. (2009). Extending the interview to all medical school candidates—Computer-Based Multiple Sample Evaluation of Noncognitive Skills (CMSENS). *Academic Medicine*, *84*(10), S9-S12.

Duckworth, A. (2016). *Grit: The power of passion and perseverance*. Scribner.

Duckworth, A. (2018, November 3). *Don't quit on a bad day*. [Video]. Association of American Medical Colleges. https://www.aamc.org/news-insights/dont-quit-bad-day

Duckworth, A. L., Kirby, T., Tsukayama, E., Berstein, H., & Ericsson, K. (2010). Deliberate practice spells success: Why grittier competitors triumph at the National Spelling Bee. *Social Psychological and Personality Science, 2,* 174-181.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality & Social Psychology, 92*(6), 1087-1101.

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment, 91*(2), 166-174. doi: 10.1080/00223890802634290

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44*(4), 237-251.

Dudkin, E., Bodek, M., Niiler, T., Geveke, K., Finneran, D., & Tierney, J. (2015, August). Delivering a postbaccalaureate medical sciences certificate program to working adults. [Poster]. American Chemical Society. Boston.

Dunleavy, D., Sondheimer, H., Castillo-Page, L., & Bletzinger, R. B. (2011). Medical school admissions: More than grades and test scores. *Association of American Medical Colleges Analysis Brief*, (11), 1-2.

Dunleavy, D. M., Kroopnick, M. H., Dowd, K. W., Searcy, C. A., & Zhao, X. (2013). The predictive validity of the MCAT exam in relation to academic performance through medical school: A national cohort study of 2001–2004 matriculants. *Academic Medicine*, *88*(5), 666-671.

Dyrbye, L. N., Thomas, M. R., Huntington, J. L., Lawson, K. L., Novotny, P. J., Sloan, J. A., & Shanafelt, T. D. (2006). Personal life events and medical student burnout: A multicenter study. *Academic Medicine*, *81*(4), 374-384.

Eaglen, R. H. (2017). *Academic quality and public accountability in academic medicine: the 75-year history of the LCME*. Association of American Medical Colleges.

Ecker, D. J., Milan, F. B., Cassese, T., Farnan, J. M., Madigosky, W. S., Massie, S. F., Mendez, P., Obadia, S., Ovitsh, R. K., Silvestri, R., Uchida, T., & Daniel, M., (2018). Step up—not on—the Step 2 clinical skills exam: Directors of clinical skills courses (DOCS) oppose ending Step 2 CS. *Academic Medicine, 93*(5), 693-698.

Eden, J., Berwick, D. M., & Wilensky, G. R. (Eds.). (2014). *Graduate medical education that meets the nation's health needs*. National Academies Press.

Educational Testing Services (2008). Graduate record examination: Factors that can influence performance on the GRE general test 2006-2007. ETS. https://www.ets.org/Media/Tests/GRE/pdf/gre_0809_factors_2006-07.pdf

Eikeland, H. L., Ørnes, K., Finset, A., & Pedersen, R. (2014). The physician's role and empathy–A qualitative study of third year medical students. *BMC Medical Education, 14*(1), *165.*

Elam, C., Stratton, T., & Gibson, D. D. (2007). Welcoming a new generation to college: The millennial students. *Journal of College Admission*, *195*, 20-25.

Enns, M. W., Cox, B. J., Sareen, J., & Freeman, P. (2001). Adaptive and maladaptive perfectionism in medical students: a longitudinal investigation. *Medical Education*, *35*(11), 1034-1042.

Epps, A. C. (2015). The strategic impact of a post baccalaureate pre-medicine intervention program on medical school academic performance. *Journal of Health Care for the Poor and Underserved*, *26*(1), 8-20.

Ehrenfeld, M., & Tabak, N. (2000). Value of admission interviews in selecting of undergraduate nursing students. *Journal of Nursing Management*, *8*(2), 101-106.

Eskreis-Winkler, L., Duckworth, A. L., Shulman, E. P., & Beal, S. (2014). The grit effect: Predicting retention in the military, the workplace, school and marriage. *Frontiers in Psychology, 5*, 36.

Espenshade, T. J., & Chung, C. Y. (2010). *Standardized admission tests, college performance, and campus diversity*. [Unpublished manuscript]. Office of Population Research, Princeton University.

Eva, K., Reiter, H., Trinh, K., Wasi, P., Rosenfeld, J., & Norman, G. (2009). Predictive validity of the multiple mini-interview for selecting medical trainees. *Medical Education*, *43*(8), 767-775.

Evans. M. (2020, April 8). *New York City feels full force of hospital crisis*. The Wall Street Journal. A1, A6.

Farooqi, Y. N., Ghani, R., & Spielberger, C. D. (2012). Gender differences in test anxiety and academic performance of medical students. *International Journal of Psychology and Behavioral Sciences*, *2*(2), 38-43. doi:10.5923/j.ijpbs.20120202.06.

Faustinella, F., & Jacobs, R. J. (2018). The decline of clinical skills: A challenge for medical schools. *International Journal of Medical Education*, *9*, 195.

Fayard, J. V., Roberts, B. W., Robins, R. W., & Watson, D. (2012). Uncovering the affective core of conscientiousness: The role of self-conscious emotions. *Journal of Personality*, *80*(1), 1-32. doi:10.1111/j.1467-6494.2011.00720.x

Federation of State Medical Boards of the United States, Inc. and the National Board of Medical Examiners. (2020). USMLE content outline. https://www.usmle.org/pdfs/usmlecontentoutline.pdf

Ferguson, E., James, D., & Madeley, L. (2002). Factors associated with success in medical school: Systematic review of the literature. *BMJ Journal*, *324*(7343), 952-957.

Ferguson, E., McManus, I. C., James, D., O'Hehir, F., & Sanders, A. (2003). Pilot study of the roles of personality, references, and personal statements in relation to performance over the five years of a medical degree. Commentary: How to derive causes from correlations in educational studies. *BMJ Journal*, *326*(7386), 429-432.

Fields, S. A., Morris, C., Toffler, W. L., & Keenan, E. J. (2000). Early identification of students at risk for poor academic performance in clinical clerkships. *Academic Medicine*, *75*(10), S78-S80.

Figueroa O. (2014). The significance of recruiting underrepresented minorities in medicine: An examination of the need for effective approaches used in admissions by higher education institutions. *Medical Education Online*, *19*, 24891. https://doi.org/10.3402/meo.v19.24891

Filut, A., Alvarez, M., & Carnes, M. (2020). Discrimination toward physicians of color: A systematic review. *Journal of the National Medical Association*, 112(2), 117-140).

Finn, G., Sawdon, M., Clipsham, L., & McLachlan, J. (2009). Peer estimation of lack of professionalism correlates with low Conscientiousness Index scores. *Medical Education*, *43*(10), 960-967.

First, L. R., Chaudhry, H. J., & Melnick, D. E. (2013). Quality, cost, and value of clinical skills assessment. *New England Journal of Medicine*, *368*(10), 963-964.

Flier, L. A., Henderson, C. R., & Treasure, C. L. (2016). Time to eliminate the Step 2 Clinical Skills examination for US medical graduates. *Journal of American Medical Association Internal Medicine*, *176*(9), 1245-1246.

Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap?. *The High School Journal*, *91*(1), 29-42.

Fortin, Y., Kealey, L., Slade, S., & Hanson, M. D. (2016). Investigating Canadian medical school attrition metrics to inform socially accountable admissions planning. *Medical Teacher*, *38*(3), 286-290. doi:10.3109/0142159x.2015.1045847.

Fowler Jr, F. J. (2013). *Survey research methods*. SAGE Publications.

Freedle, R. (2003). Correcting the SAT's ethnic and social-class bias: A method for reestimating SAT scores. *Harvard Educational Review*, *73*(1), 1-43.

Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences*, *14*(1), 47-64.

Garcia, S. B., & Guerra, P. L. (2004). Deconstructing deficit thinking: Working with educators to create more equitable learning environments. *Education and Urban Society*, *36*(2), 150-168.

Gardner, A. K., Cavanaugh, K. J., Willis, R. E., & Dunkin, B. J. (2019, August 24-28). *Is reliance on the USMLE for residency screening at odds with efforts to enhance diversity among the surgical workforce?* [Podium presentation]. International Association for Medical Education (AMEE), Chicago, IL.

Gauer, J. L., & Jackson, J. B. (2017). The association of USMLE Step 1 and Step 2 CK scores with residency match specialty and location. *Medical Education Online*, *22*(1), doi:10.1080/10872981.2017.1358579

Gauer, J. L., & Jackson, J. B. (2018a). Relationships of demographic variables to USMLE physician licensing exam scores: A statistical analysis on five years of medical student data. *Advances in Medical Education and Practice*, *9*, 39. doi:10.2147/amep.s152684.

Gauer, J. L., & Jackson, J. B. (2018b). The association between United States Medical Licensing Examination scores and clinical performance in medical students. *Advances in Medical Education and Practice*, *10*, 209. doi:10.2147/amep.s192011.

Gauer, J. L., Wolff, J. M., & Jackson, J. B. (2016). Do MCAT scores predict USMLE scores? An analysis on 5 years of medical student data. *Medical Education Online*, *21*(1), 31795.

Ghaffari-Rafi, A., Lee, R. E., Fang, R., & Miles, J. D. (2019). Multivariable analysis of factors associated with USMLE scores across US medical schools. *BMC Medical Education*, *19*(1), 154.

Gilbert, G. E., Basco, W. T., Blue, A. V., & O'Sullivan, P. S. (2002). Predictive validity of the Medical College Admissions Test writing sample for the United States medical licensing examination Steps 1 and 2. *Advances in Health Sciences Education*, *7*(3), 191-200.

Gilliland, W. R., Rochelle, J. L., Hawkins, R., Dillon, G. F., Mechaber, A. J., Dyrbye, L., Papp, K. K., & Durning, S. J. (2008). Changes in clinical skills education resulting from the introduction of the USMLE™ Step 2 clinical skills (CS) examination. *Medical Teacher*, *30*(3), 325-327.

Giordani, B., Edwards, A. S., Segal, S. S., Gillum, L. H., Lindsay, A., & Johnson, N. (2001). Effectiveness of a formal post-baccalaureate pre-medicine program for underrepresented minority students. *Academic Medicine, 76*(8), 844-848.

Givens, J. L., Frederick, M., Silverman, L., Anderson, S., Senville, J., Silver, M., Sebastiani, P., Terry, D. F., Costa, P. T., & Perls, T. T. (2009). Personality traits of centenarians' offspring. *Journal of the American Geriatrics Society*, *57*(4), 683-685.

Goff, M., & Ackerman, P. L. (1992). Personality-intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology*, *84*(4), 537.

Gohara, S., Shapiro, J. I., Jacob, A. N., Khuder, S. A., Gandy, R. A., Metting, P. J., Gold, J., & Kleshinski, J. (2011). Joining the conversation: Predictors of success on the United States Medical Licensing Examinations (USMLE). *Learning Assistance Review, 16*(1), 11-20.

Gonzalez, P., & Stoll, B. (2002). *The color of medicine: Strategies for increasing diversity in the US physician workforce*. Community Catalyst.

Goode, C. A., & Talbot, C. (2016). A special master's program as a pipeline to increase diversity at an osteopathic medical college. *Journal of Best Practices in Health Professions Diversity: Education, Research & Policy*, *9*(2), 1198-126.

Gough, H. G., Bradley, P., & McDonald, J. S. (1991). Performance of residents in anesthesiology as related to measures of personality and interests. *Psychological Reports*, *68*(3), 979-994.

Grabowski, C. J. (2018). Impact of holistic review on student interview pool diversity. *Advances in Health Sciences Education*, *23*(3), 487-498.

Grbic, D., Jones, D. J., & Case, S. T. (2013, March). *Effective practices for using the AAMC socioeconomic status indicators in medical school admissions*. Association of American Medical Colleges. https://aamc.org/download/330166/data/seseffectivepractices.pdf

Grbic, D., Jones, D. J., & Case, S. T. (2015). The role of socioeconomic status in medical school admissions: Validation of a socioeconomic indicator for use in medical school admissions. *Academic Medicine, 90*(7), 953-960.

Grbic, D., Morrison, E., Sondheimer, H. M., Conrad, S. S., & Milem, J. F. (2019). The association between a holistic review in admissions workshop and the diversity of accepted applicants and students matriculating to medical school. *Academic Medicine*, *94*(3), 396-403.

Green, M., Jones, P., & Thomas, Jr., J. X. (2009). Selection criteria for residency: Results of a national program directors survey. *Academic Medicine*, *84*(3), 362-367.

Grehan, P. M., Flanagan, R., & Malgady, R. G. (2011). Successful graduate students: The roles of personality traits and emotional intelligence. *Psychology in the Schools*, *48*(4), 317-331.

Griffin, B., & Wilson, I. (2012). Associations between the Big Five personality factors and multiple mini-interviews. *Advances in Health Sciences Education*, *17*(3), 377-388. doi:10.1007/s10459-011-9316-1

Grimm, P. (2010). Social desirability bias. *Wiley International Encyclopedia of Marketing*.

Grumbach, K., & Chen, E. (2006). Effectiveness of University of California postbaccalaureate premedical programs in increasing medical school matriculation for minority and disadvantaged students. *Journal of American Medical Association*, *296*(9), 1079-1085.

Grumbach, K., & Mendoza, R. (2008). Disparities in human resources: addressing the lack of diversity in the health professions. *Health Affairs*, *27*(2), 413-422.

Hadinger, M. A. (2017). Underrepresented minorities in medical school admissions: A qualitative study. *Teaching and Learning in Medicine*, *29*(1), 31-41.

Hahn, E., Gottschling, J., & Spinath, F. M. (2012). Short measurements of personality–Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality*, *46*(3), 355-359.

Haight, S. J., Chibnall, J. T., Schindler, D. L., & Slavin, S. J. (2012). Associations of medical student personality and health/wellness characteristics with their medical school performance across the curriculum. *Academic Medicine*, *87*(4), 476-485.

Haist, S. A., Witzke, D. B., Quinlivan, S., Murphy-Spencer, A., & Wilson, J. F. (2003). Clinical skills as demonstrated by a comprehensive clinical performance examination: Who performs better–men or women? *Advances in Health Sciences Education*, *8*(3),189-199. doi:10.1023/A:1026072102739.

Hall, J. A., Roter, D. L., Blanch, D. C., & Frankel, R. M. (2009). Nonverbal sensitivity in medical students: Implications for clinical interactions. *Journal of General Internal Medicine*, *24*(11), 1217. doi:10.1007/s11606-009-1107-5.

Han, B. (2018). *Characterizing grit in medical students: Low grit associated with higher rates of dissatisfaction in medical school and more consideration for leaving medical school* [Capstone project]. University of Michigan Medical School.

Han, B., & Salles, A. (n.d.). *How the personal characteristic of grit and resilience relate to surgeon well-being*. American College of Surgeons. https://www.facs.org/education/division-of-education/publications/rise/articles/grit-resilience

Harfmann, K. L., & Zirwas, M. J. (2011). Can performance in medical school predict performance in residency? A compilation and review of correlative studies. *Journal of the American Academy of Dermatology*, *65*(5), 1010-1022.

Harris, T. B., Thomson, W. A., Moreno, N. P., Conrad, S., White, S. E., Young, G. H., Malmberg, E. D., Weisman, B., & Monroe, A. D. (2018). Advancing holistic review for faculty recruitment and advancement. *Academic Medicine, 93*(11), 1658-1662.

Harrison, L. E. (2019). Using holistic review to form a diverse interview pool for selection to medical school. *Baylor University Medical Center Proceedings*, *32*(2), 218-221.

Hartocollis, A. (2018, October 25). Is an Extroverted Applicant Better Suited for Harvard Than an Introvert? New York Times. p. A11. Retrieved from https://www.nytimes.com/2018/10/25/us/harvard-admissions-trial-asian-americans.html

Hawkins, R. E., Swanson, D. B., Dillon, G. F., Clauser, B. E., King, A. M., Scoles, P. V., Whelan, G. P., Burdick, W. P., Boulet, J. R., & Homan, A. G. (2005). The introduction of clinical skills assessment into the United States medical licensing examination (USMLE): A description of USMLE Step 2 clinical skills (CS). *Journal of Medical Licensure and Discipline*, *91*, 22-25.

Health Resources and Services Administration [HRSA]. (2006). The rationale for diversity in the health professions: A review of the evidence. Bureau of Health Professions. http://bhpr.hrsa.gov/healthworkforce/reports/diversityreviewevidence.pdf

Henning, K., Ey, S., & Shaw, D. (1998). Perfectionism, the impostor phenomenon and psychological adjustment in medical, dental, nursing and pharmacy students. *Medical Education*, *32*(5), 456-464.

Hill, P. L., Burrow, A. L., & Bronk, K. C. (2016). Persevering with positivity and purpose: An examination of purpose commitment and positive affect as predictors of grit. *Journal of Happiness Studies*, *17*(1), 257-269.

Hill, A., Jones, D., & Woodworth, L. (2018). A doctor like me: Physician-patient race-match and patient outcomes. SSRN. https://ssrn.com/abstract=3211276

Hirshfield, L. E., Yudkowsky, R., & Park, Y. S. (2019). Pre-medical majors in the humanities and social sciences: Impact on communication skills and specialty choice. *Medical Education*, *53*(4), 408-416.

Hojat, M., Erdmann, J. B., & Gonnella, J. S. (2013). Personality assessments and outcomes in medical education and the practice of medicine: AMEE Guide No. 79. *Medical Teacher*, *35*(7), e1267-e1301.

Hojat, M., Gonnella, J. S., Erdmann, J. B., & Vogel, W. H. (2003). Medical students' cognitive appraisal of stressful life events as related to personality, physical well-being, and academic performance: A longitudinal study. *Personality and Individual Differences*, *35*(1), 219-235.

Hojat, M., Gonnella, J. S., Mangione, S., Nasca, T. J., Veloski, J. J., Erdmann, J. B., Callahan, C. A., & Magee, M. (2002a). Empathy in medical students as related to academic performance, clinical competence and gender. *Medical Education*, *36*(6), 522-527. doi:10.1046/j.1365-2923.2002.01234.x.

Hojat, M., Gonnella, J. S., Nasca, T. J., Mangione, S., Veloksi, J. J., & Magee, M. (2002b). The Jefferson scale of physician empathy: Further psychometric data and differences by gender and specialty at item level. *Academic Medicine*, *77*(10), S58-S60. doi:10.1097/00001888-20021001-0019.

Hong, M., & Lee, S. (2019). Effects of grit and conscientiousness on academic performance: The mediation effects of self-determination motivation. *Journal of the Korea Convergence Society*, *10*(10), 143-151.

Hopkins, T. H. (2008). Early identification of at-risk nursing students: A student support model. *Journal of Nursing Education*, *47*(6), 254-259.

Hough, L. M. (1992). The 'Big Five' personality variables--construct confusion: Description versus prediction. *Human Performance*, *5*(1-2), 139-155.

Huff, K. L., Koenig, J. A., Treptau, M. M., & Sireci, S. G. (1999). Validity of MCAT scores for predicting clerkship performance of medical students grouped by sex and ethnicity. *Academic Medicine*, *74*(10), S41-4.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, *85*(6), 869.

Imlach, A. R., Ward, D. D., Stuart, K. E., Summers, M. J., Valenzuela, M. J., King, A. E., Saunders, N. L., Summers, J., Srikanth, V. K, Robinson, A., & Vickers, J. C. (2017). Age is no barrier: Predictors of academic success in older learners. *NPJ Science of Learning*, *2*(1), 13. doi:10.1038/s41539-017-0014-5.

Institute of Medicine, IOM. (2004). *In the nation's compelling interest: Ensuring diversity in the health-care workforce.* National Academies Press.

Isenberg, G., Brown, A., DeSantis, J., Veloski, J., & Hojat, M. (2020). The relationship between grit and selected personality measures in medical students. *International Journal of Medical Education*, *11*, 25-30.

IsHak, W., Nikravesh, R., Lederer, S., Perry, R., Ogunyemi, D., & Bernstein, C. (2013). Burnout in medical students: a systematic review. *The Clinical Teacher*, *10*(4), 242-245.

Isik, U., Wouters, A., Ter Wee, M. M., Croiset, G., & Kusurkar, R. A. (2017). Motivation and academic performance of medical students from ethnic minorities and majority: A comparative study. *BMC Medical Education*, *17*(1), 233.

Ivcevic, Z., & Brackett, M. (2014). Predicting school success: Comparing conscientiousness, grit, and emotion regulation ability. *Journal of Research in Personality*, *52*, 29-36.

Jackson, C. K. (2012). *Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina* (No. w18624). National Bureau of Economic Research.

Jackson, J. J., Bogg, T., Walton, K. E., Wood, D., Harms, P. D., Lodi-Smith, J., ... & Roberts, B. W. (2009). Not all conscientiousness scales change alike: A multimethod, multisample study of age differences in the facets of conscientiousness. *Journal of Personality and Social Psychology*, *96*(2), 446.

Jauhar, S. (2008). From all walks of life—Nontraditional medical students and the future of medicine. *New England Journal of Medicine*, *359*(3), 224-227.

Jayakumar, K. L. (2018). The limited value of USMLE Step 2 CS. *Academic Medicine*, *93*(3), 345.

Jerant, A., Griffin, E., Rainwater, J., Henderson, M., Sousa, F., Bertakis, K. D., Fenton, J. J., & Franks, P. (2012). Does applicant personality influence multiple mini-interview performance and medical school acceptance offers? *Academic Medicine*, *87*(9), 1250-1259. doi:10.1097/acm.0b013e31826102ad

Jerant, A., Henderson, M. C., Griffin, E., Talamantes, E., Fancher, T., Sousa, F., & Franks, P. (2018). Medical school performance of socioeconomically disadvantaged and underrepresented minority students matriculating after a multiple mini-interview. *Journal of Health Care for the Poor and Underserved*, *29*(1), 303-320.

Jerant, A., Sciolla, A. F., Henderson, M. C., Griffin, E., Talamantes, E., Fancher, T., & Franks, P. (2019). Medical student socioeconomic disadvantage, self-designated disadvantage, and subsequent academic performance. *Journal of Health Care for the Poor and Underserved*, *30*(4), 1419-1432.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big-five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins & L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (pp. 114-158). Guilford Press.

Johnson, B., Flemer, M., Khuder, S., & Puri, N. (2017). Premedical special master's programs increase USMLE Step 1 scores and improve residency placements. *PloS One, 12*(11), e0188036.

Johnson, H. (2003). A critical review of standardized patient examinations as part of the USMLE. *American Medical Association Journal of Ethics, 5*(12), 426-429.

Johnson, D. G., & Hutchins, E. B. (1966). Doctor or dropout? A study of medical student attrition. *Academic Medicine, 41*(12), 1099-274.

Johnson, D. G., Smith, V. C., Tarnoff, S. L. (1975). Recruitment and progress of minority medical school entrants, 1970-1972. Journal of Medical Education, 50, 731-755

Julian, E. R. (2005). Validity of the Medical College Admission Test for predicting medical school performance. *Academic Medicine*, *80*(10), 910-917.

Jurich, D., Santen, S. A., Paniagua, M., Fleming, A., Harnik, V., Pock, A., Swan-Sein, A., Barone, M. A., & Daniel, M. (2020). Effects of moving the United States medical licensing examination Step 1 after core clerkships on Step 2 clinical knowledge performance. *Academic Medicine*, *95*(1), 111-121.

Jurjus, R. A., Butera, G., Abdelnabi, M., & Krapf, J. M. (2017). Comparing the experience of mature-aged and traditional medical students in the clinical setting: A qualitative approach. *Journal of Academic Development and Education*, (7), 8-28. doi: 10.21252/keele-0000005.

Kahlenberg, R. D. (Ed.). (2004). *America's untapped resource: Low-income students in higher education*. Century Foundation.

Kaiser Family Foundation. (2008). Eliminating racial/ethnic disparities in health care: What are the options? KFF. https://www.kff.org/disparities-policy/issue-brief/eliminating-racialethnic-disparities-in-health-care-what/

Kalesnikava, V. A., Ekey, G. P., Ko, T. M., Shackelford, D. T., & Mezuk, B. (2019). Grit, growth mindset and participation in competitive policy debate: Evidence from the Chicago Debate League. *Educational Research and Reviews*, *14*(10), 358-371.

Kalter, L. (2019, February 5). *Trading places: When doctors become patients.* Association of American Medical Colleges. https://news.aamc.org/patient-care/article/trading-places/

Kamenetz, A. (2015, May 28). Nonacademic skills are key to success. But what should we call them? National Public Radio. http://www.npr.org/sections/ed/2015/05/28/404684712/non-academicskills-are-key-to-success-but-what-should-we-call-them

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.

Kataoka, H. U., Koide, N., Ochi, K., Hojat, M., & Gonnella, J. S. (2009). Measurement of empathy among Japanese medical students: Psychometrics and score differences by gender and level of medical education. *Academic Medicine*, *84*(9), 1192-1197.

Kaufman, D. M., Mann, K. V., Muijtjens, A. M., & van der Vleuten, C. P. (2000). A comparison of standard-setting procedures for an OSCE in undergraduate medical education. *Academic Medicine*, *75*(3), 267-271.

Keiser, H. N., Sackett, P. R., Kuncel, N. R., & Brothen, T. (2016). Why women perform better in college than admission scores would predict: Exploring the roles of conscientiousness and course-taking patterns. *Journal of Applied Psychology*, *101*(4), 569.

Keith, T. Z. (2006). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*. Routledge.

Kellow, J. T., & Jones, B. D. (2008). The effects of stereotypes on the achievement gap: Reexamining the academic performance of African American high school students. *Journal of Black Psychology*, *34*(1), 94-120.

Kelly, P. J. (2005). *As America becomes more diverse: The impact of state higher education inequality*. National Center for Higher Education Management. https://vtechworks.lib.vt.edu/handle/10919/83132

Kelly-Blake, K., Garrison, N. A., Fletcher, F. E., Ajegba, B., Smith, N., Brafford, M., & Lovis, E. B. (2018). Medical education in review rationales for expanding minority physician representation in the workforce: A scoping review. *Medical Education*, *52*(9), 925-935.

Kendi, I. X. (2016). Why Standardized Tests Have Standardized Postracial Ideology. *Academe, 102*(6), 26-29. Accessed September 20, 2020.

Kendi, I. X. (2017). *Stamped from the beginning: The definitive history of racist ideas in America*. Random House.

Kenny, N. P., Mann, K. V., & MacLeod, H. (2003). Role modeling in physicians' professional formation: Reconsidering an essential but untapped educational strategy. *Academic Medicine*, *78*(12), 1203-1210.

Kerrigan, N., Akabas, M. H., Betzler, T. F., Castaldi, M., Kelly, M. S., Levy, A. S., ... & Dolan, S. M. (2016). Implementing competency-based admissions at the Albert Einstein College of Medicine. *Medical Education Online, 21*(1), 30000.

Kessler, M. (2003). Conscientiousness, mastery, conflictual independence, and self-efficacy in relationship to student adaptation and academic success during the first semester of college: A comparative and predictive study of first-and second-generation college students.

Kick, S., Adams, L., & O'Brien-Gonzales, A. (2000). Unique issues of older medical students. *Teaching and Learning in Medicine*, *12*(3), 150-155. doi:10.1207/s15328015tlm1203_6.

Kim, S., & George, P. (2018). The relationship between preclinical grading and USMLE scores in US allopathic medical schools. *Family Medicine*, *50*(2), 128-131.

Kirch, D. G. (2012). Transforming admissions: The gateway to medicine. *Journal of American Medical Association*, *308*(21), 2250-2251.

Kirch, D. G., Henderson, M. K., & Dill, M. J. (2012). Physician workforce projections in an era of health care reform. *Annual Review of Medicine*, *63*, 435-445.

Kirch, D. G., & Petelle, K. (2017). Addressing the physician shortage: The peril of ignoring demography. *Journal of American Medical Association*, *317*(19), 1947-1948.

Kleshinski, J., Khuder, S. A., Shapiro, J. I., & Gold, J. P. (2009). Impact of preadmission variables on USMLE Step 1 and Step 2 performance. *Advances in Health Sciences Education*, *14*(1), 69-78.

Klostermann, B. K., Slap, G. B., Nebrig, D. M., Tivorsak, T. L., & Britto, M. T. (2005). Earning trust and losing it: Adolescents' views on trusting physicians. *Journal of Family Practice*, *54*(8), 679-688.

Koenig, T. W., Parrish, S. K., Terregino, C. A., Williams, J. P., Dunleavy, D. M., & Volsch, J. M. (2013). Core personal competencies important to entering students’ success in medical school: What are they and how could they be assessed early in the admission process? *Academic Medicine*, *88*(5), 603-613.

Komarraju, M., Karau, S. J., & Schmeck, R. R. (2009). Role of the Big Five personality traits in predicting college students' academic motivation and achievement. *Learning and Individual Differences*, *19*(1), 47-52.

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, *51*(4), 472-477:

Komarraju, M., Ramsey, A., & Rinella, V. (2013). Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline. *Learning and Individual Differences*, *24*, 103-109.

Körner, A., Czajkowska, Z., Albani, C., Drapeau, M., Geyer, M., & Braehler, E. (2015). Efficient and valid assessment of personality traits: Population norms of a brief version of the NEO Five-Factor Inventory (NEO-FFI). *Archives of Psychiatry & Psychotherapy*, *17*(1), 21-32.

Kreiter, C. D., Yin, P., Solow, C., & Brennan, R. L. (2004). Investigating the reliability of the medical school admissions interview. *Advances in Health Sciences Education*, *9*(2), 147-159.

Kruzicevic, S. M., Barisic, K. J., Banozic, A., Esteban, C. D., Sapunar, D., & Puljak, L. (2012). Predictors of attrition and academic success of medical students: a 30-year retrospective study. *PloS One*, *7*(6), e39144.

Kulasegaram, K., Reiter, H. I., Wiesner, W., Hackett, R. D., & Norman, G. R. (2010). Non-association between NEO-5 personality tests and multiple mini-interview. *Advances in Health Sciences Education*, *15*(3), 415-423. doi:10.1007/s10459-009-9209-8

Kulatunga-Moruzi, C., & Norman, G. R. (2002). Validity of admissions measures in predicting performance outcomes: The contribution of cognitive and non-cognitive dimensions. *Teaching and Learning in Medicine*, *14*(1), 34-42.

Kulesza, R. J., Hu, W., & Bateman, M. (2015). Longitudinal assessment and perceptions of medical post baccalaureate students. *Medical Science Educator*, *25*(1), 19-26.

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science, 315*(5815), 1080-1081.

Kushnir, T., Cohen, A. H., & Kitai, E. (2000). Continuing medical education and primary physicians' job stress, burnout and dissatisfaction. *Medical Education*, *34*(6), 430-436.

Kyllonen, P. C. (2005). The case for noncognitive assessments. *R&D Connections.* Educational Test Service Research and Development.

Kyllonen, P. C. (2012). The importance of higher education and the role of noncognitive attributes in college success. *PEL: Pensamiento Educativo*, *49*(2), 84-100. doi: 10.7764/PEL.49.2.2012.7

Kyllonen, P. C., Walters, A. M., & Kaufman, J. C. (2011). The role of noncognitive constructs and other background variables in graduate education. Educational Testing Service. Princeton, NJ.

Lakhan, S. E., & Laird, C. (2009). Addressing the primary care physician shortage in an evolving medical workforce. *International Archives of Medicine*, *2*(1), 14.

Lancaster, R. J., Baseman, E., & Smolinski, L. (2020). A quality improvement project: Defining and operationalizing a holistic admission selection policy. *Journal of Professional Nursing*. https://doi.org/10.1016/j.profnurs.2020.01.003

Lee, K. B., Vaishnavi, S. N., Lau, S. K., Andriole, D. A., & Jeffe, D. B. (2007). "Making the grade:" Noncognitive predictors of medical students' clinical clerkship grades. *Journal of the National Medical Association, 99*(10), 1138-1150.

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2011). *IBM SPSS for intermediate statistics: Use and interpretation* (4th ed.) Taylor and Francis Group.

Lehman, IV, E. P., & Guercio, J. R. (2013). The Step 2 Clinical Skills exam—A poor value proposition. *New England Journal of Medicine*, *368*(10), 889-891.

Leigh, J. P., Kravitz, R. L., Schembri, M., Samuels, S. J., & Mobley, S. (2002). Physician career satisfaction across specialties. *Archives of Internal Medicine*, *162*(14), 1577-1584.

Lemann, N. (2000). *The big test: The secret history of the American meritocracy*. Macmillan.

Lemay, J. F., Lockyer, J. M., Collin, V. T., & Brownell, A. K. W. (2007). Assessment of non-cognitive traits through the admissions multiple mini-interview. *Medical Education*, *41*(6), 573-579.

Levine, R. B., Levy, A. P., Lubin, R., Halevi, S., Rios, R., & Cayea, D. (2017). Evaluation of a course to prepare international students for the United States Medical Licensing Examination Step 2 clinical skills exam. *Journal of Educational Evaluation for Health Professions*, *14(25)*.

Leyerzapf, H., Abma, T. A., Steenwijk, R. R., Croiset, G., & Verdonk, P. (2015). Standing out and moving up: Performance appraisal of cultural minority physicians. *Advances in Health Sciences Education*, *20*(4), 995-1010.

Lieberman, S. A., Stroup-Benham, C. A., & Peel, J. L. (1998). Predictors of intellectual satisfaction in medical school: Sociodemographic, cognitive, and personality factors. *Academic Medicine*, *73*(10 SUPPL.), S44-S46.

Lievens, F., Coetsier, P., De Fruyt, F., & De Maeseneer, J. (2002). Medical students' personality characteristics and academic performance: A five-factor model perspective. *Medical Education*, *36*(11), 1050-1056.

Lievens, F., Ones, D. S., & Dilchert, S. (2009). Personality scale validities increase throughout medical school. *Journal of Applied Psychology*, *94*(6), 1514.

Lim, H., Heckman, S., Montalto, C. P., & Letkiewicz, J. (2014). Financial stress, self-efficacy, and financial help-seeking behavior of college students. *Journal of Financial Counseling and Planning*, *25*(2), 148-160.

Lipscomb, W. D., Mavis, B., Fowler, L. V., Green, W. D., & Brooks, G. L. (2009). The effectiveness of a postbaccalaureate program for students from disadvantaged backgrounds. *Academic Medicine*, *84*(10), S42-S45.

Lohfink, M. M., & Paulsen, M. B. (2005). Comparing the determinants of persistence for first-generation and continuing-generation students. *Journal of College Student Development*, *46*(4), 409-428.

Love, B. L. (2019). *We want to do more than survive: Abolitionist teaching and the pursuit of educational freedom*. Beacon Press.

Low, D., Pollack, S. W., Liao, Z. C., Maestas, R., Kirven, L. E., Eacker, A. M., Morales, L. S. (2019). Racial/ethnic disparities in clinical grading in medical school. *Teaching and Learning in Medicine*, *31*(5), 487-496.

Lowrance, A. M. (2017). *The disadvantaged status in the American Medical College Application Service: Medical students reflect on their experiences* [Doctoral dissertation, University of Northern Colorado]. https://digscholarship.unco.edu/dissertations/417

Lumsden, M. A., Bore, M., Millar, K., Jack, R., & Powis, D. (2005). Assessment of personal qualities in relation to admission to medical school. *Medical Education*, *39*(3), 258-265.

Magee, M., & Hojat, M. (1998). Personality profiles of male and female positive role models in medicine. *Psychological Reports*, *82*(2), 547-559.

Magzan, M. (2012). Mental models for leadership effectiveness: Building future different than the past. *Journal of Engineering Management and Competitiveness (JEMC)*, *2*(2), 57-63.

Maher, B. M., Hynes, H., Sweeney, C., Khashan, A. S., O'Rourke, M., Doran, K., ... & O'Flynn, S. (2013). Medical school attrition-beyond the statistics a ten-year retrospective study. *BMC Medical Education*, *13*(1), 13.

Majeed, U., Sardar, Z., Kiran, N., Suqrat, H., Sardar, H., & Adil, S. (2019). Association of grit with depression among medical students. *Annals of Punjab Medical College (APMC)*, *13*(4), 260-262.

Malau-Aduli, B. S., O'Connor, T., Ray, R. A., van der Kruk, Y., Bellingan, M., & Teague, P. A. (2017). Risk factors associated with academic difficulty in an Australian regionally located medical school. *BMC Medical Education, 17*(1), 266. doi:10.1186/s12909-018-1170-x.

Mann, K. V. (1999). Motivation in medical education: How theory can inform our practice. *Academic Medicine, 74*(3), 237-9.

Manusov, E. G., Livingston, H., Wang, A., Berne-Anderson, T., Alston, S., Foster, E., & Hurt, M. (2011). Student's perspective of success in a postbaccalaureate program. *Journal of the National Medical Association*, *103*(9-10), 822-831.

Marrah, A. K. (2012). The perceptions of standardized tests, academic self-efficacy, and academic performance of African American graduate students: A correlational and comparative analysis. [Doctoral dissertation, Old Dominion University].

Marrast, L. M., Zallman, L., Woolhandler, S., Bor, D. H., & McCormick, D. (2014). Minority physicians' role in the care of underserved patients: Diversifying the physician workforce may be key in addressing health disparities. *JAMA Internal Medicine*, *174*(2), 289-291.

Martini, S., Arfken, C. L., Churchill, A., & Balon, R. (2004). Burnout comparison among residents in different medical specialties. *Academic Psychiatry*, *28*(3), 240-242.

Mavis, B. (2001). Self-efficacy and OSCE performance among second year medical students. *Advances in Health Sciences Education*, *6*(2), 93-102.

McCrae, R. R. (1996). Social consequences of experiential openness. *Psychological Bulletin*, *120*(3), 323.

McCrae, R. R., & Costa, Jr., P. T. (1983). Joint factors in self-reports and ratings: Neuroticism, extraversion and openness to experience. *Personality and Individual Differences*, *4*(3), 245-255.

McCrae, R. R., & Costa, Jr., P. T. (1989). More reasons to adopt the five-factor model. *American Psychologist, 44*(2), 451-452.

McCrae, R. R., & Costa, Jr, P. T. (2007). Brief versions of the NEO-PI-3. *Journal of Individual Differences*, *28*(3), 116-128.

McCrae, R. R., & Costa, P. T. (2010a). *NEO inventories for the NEO Personality Inventory-3 (NEO-PI-3), NEO Five-Factor Inventory-3 (NEO-FFI-3), NEO Personality Inventory-Revised (NEO PI-R): Professional Manual*. Psychological Assessment Resources, Inc.

McCrae, R. R., & Costa Jr, P. T. (2010b). *NEO Inventories: Professional manual*. Psychological Assessment Resources, Inc.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, *15*(1), 28-50.

McCredie, M. N., & Kurtz, J. E. (2020). Prospective prediction of academic performance in college using self-and-informant-rated personality traits. *Journal of Research in Personality*, *85*, 103911.

McDade, W. A. (2019) Increasing graduate medical education diversity and inclusion. *Journal of Graduate Medical Education*, *11*(6), 736-738.

McGaghie, W. C. (2002). Assessing readiness for medical education: Evolution of the medical college admission test. *Journal of American Medical Association*, *288*(9), 1085-1090.

McGaghie, W. C., Cohen, E. R., & Wayne, D. B. (2011). Are United States Medical Licensing Exam Step 1 and 2 scores valid measures for postgraduate medical residency selection decisions? *Academic Medicine*, *86*(1), 48-52.

McLachlan, J. C., Finn, G., & Macnaughton, J. (2009). The conscientiousness index: A novel tool to explore students' professionalism. *Academic Medicine*, *84*(5), 559-565.

MedPage Today. https://www.medpagetoday.com/publichealthpolicy/generalprofessionalissues/73617

Mehta, J. (2015). The problem with grit. *Education Week*.

Mehta, N. P., & Kramer, D. B. (2005). A critique of the USMLE clinical skills examination. *Medscape General Medicine, 7*(2), 76.

Meredith, D. M. (2008). *Carolina Covenant: Low-SES, first generation college student's navigation of higher education* [Doctoral dissertation, University of Cincinnati].

Milem, J. F., O'Brien, C., Miner, D., Bryan, W. P., Sutton, F., Castillo-Page, L., & Schoolcraft, S. (2012). *The important role that diverse students play in shaping the medical school curriculum*. Arizona Medical Education Research Institute. https://www.coe.arizona.edu/sites/default/files/Milem,O%27Brien,Miner,Bryan,Castillo-Page,Schoolcraft(2012)-The_Important_Role_that_Diverse_Students_Play_in_Shaping_the_Medical_School_Curriculum.pdf

Miller-Matero, L. R., Martinez, S., MacLean, L., Yaremchuk, K., & Ko, A. B. (2018). Grit: A predictor of medical student performance. *Education for Health, 31*(2), 109. doi:10.4103/efh.efh_152_16

Monroe, A., Quinn, E., Samuelson, W., Dunleavy, D. M., & Dowd, K. W. (2013). An overview of the medical school admission process and use of applicant data in decision making: What has changed since the 1980s? *Academic Medicine, 88*(5), 672-681.

Monteiro, K. A., George, P., Dollase, R., & Dumenco, L. (2017). Predicting United States Medical Licensure Examination Step 2: Clinical Knowledge scores from previous academic indicators. *Advances in Medical Education and Practice, 8*, 385-391. doi:10.2147/AMEP.S138557

Moore, P. J., Adler, N. E., & Robertson, P. A. (2000). Medical malpractice: The effect of doctor-patient relations on medical patient perceptions and malpractice intentions. *Western Journal of Medicine*, *173*(4), 244.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, *60*(3), 683-729.

Morris, J. G. (1999). The value and role of the interview in the student admissions process: A review. *Medical Teacher*, *21*(5), 473-481.

Mullangi, S., & Jagsi, R. (2019). Imposter syndrome: Treat the cause, not the symptom. *Journal of American Medical Association*, *322*(5), 403-404.

Mullen, A. L., Goyette, K. A., & Soares, J. A. (2003). Who goes to graduate school? Social and academic correlates of educational continuation after college. *Sociology of Education*, 76, 143-169.

Muller, D., & Kase, N. (2010). Challenging traditional premedical requirements as predictors of success in medical school: The Mount Sinai School of Medicine Humanities and Medicine Program. *Academic Medicine*, *85*(8), 1378-1383.

Muller, D. (2014). FlexMed: a nontraditional admissions program at Icahn School of Medicine at Mount Sinai. *AMA Journal of Ethics*, *16*(8), 614-617.

Nakae, S. T. (2014). *The backgrounds and outcomes of allopathic medical school applicants: Exploring stratification and inequality* [Doctoral dissertation, Loyola University Chicago]. https://ecommons.luc.edu/cgi/viewcontent.cgi?article=2293&context=luc_diss

National Board of Medical Examiners (NBME). (2010). Assessment of communication skills within clinical encounters. *Examiner, 57*(2), 5-6.

National Board of Medical Examiners (NBME). (2015). *National Board of Medical Examiners: 100 Years in service to the public and health professions*. NBME. https://nbme100.org

National Center for Education Statistics. (2019). *IPEDS 2019-2020 Survey Materials Glossary.* NCES. https://surveys.nces.ed.gov/ipeds/Downloads/Forms/IPEDSGlossary.pdf

National Center for Fair and Open Testing. (2012). How standardized testing damages education. FairTest. https://fairtest.org/how-standardized-testing-damages-education-pdf.

Nivet, M. A. (2010). Minorities in academic medicine: Review of the literature. *Journal of Vascular Surgery, 51*(4), S53-S58.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*(1), 116.

Nolen, LaShyra. (2020, June 8). *This is what I want to tell my White professors when they ask 'How are you today?'*. Huffpost. https://www.huffpost.com/entry/black-medical-student-wants-white-professors-to-know_n_5ed91238c5b6e0feefc26315

Nora, L. M., McLaughlin, M. A., Fosson, S. E., Stratton, T. D., Murphy-Spencer, A., Fincher, R. M. E., German, D. C., Seiden, D. & Witzke, D. B. (2002). Gender discrimination and sexual harassment in medical education: Perspectives gained by a 14-school study. *Academic Medicine*, *77*(12 Part 1), 1226-1234.

Norcini, J. J., Boulet, J. R., Opalek, A., & Dauphinee, W. D. (2014). The relationship between licensing examination performance and the outcomes of care by international medical school graduates. *Academic Medicine*, *89*(8), 1157-1162.

Norcini, J. J., van Zanten, M., & Boulet, J. R. (2008). The contribution of international medical graduates to diversity in the US physician workforce: Graduate medical education. *Journal of Health Care for The Poor and Underserved*, *19*(2), 493-499.

Ntiri, D. W. (2001). Access to higher education for nontraditional students and minorities in a technology-focused society. *Urban Education*, *36*(1), 129-144.

Ogunyemi, D., & Taylor-Harris, D. (2005). Factors that correlate with the US Medical Licensure Examination: Step-2 scores in a diverse medical student population. *Journal of the National Medical Association*, *97*(9), 1258-1262.

Oliver, T., Hecker, K., Hausdorf, P. A., & Conlon, P. (2014). Validating MMI scores: are we measuring multiple attributes?. *Advances in Health Sciences Education*, *19*(3), 379-392.

Orfield, G., & Wald, J. (2000). Testing, testing. *The Nation*, *270*(22), 38-40.

Orozco Cortes, M. E. (2019). The Relationship Between Advanced Degrees and Performance on USMLE Step 1. [Unpublished manuscript]. Wright State University.

Orzeck, T., & Lung, E. (2005). Big-five personality differences of cheaters and non-cheaters. *Current Psychology*, *24*(4), 274-286.

Pascarella, E. T., Pierson, C. T., Wolniak, G. C., & Terenzini, P. T. (2004). First-generation college students: Additional evidence on college experiences and outcomes. *The Journal of Higher Education*, *75*(3), 249-284.

Pascarella, E. T., Wolniak, G. C., Pierson, C. T., & Terenzini, P. T. (2003). Experiences and outcomes of first-generation students in community colleges. *Journal of College Student Development*, *44*(3), 420-429.

Pau, A., Jeevaratnam, K., Chen, Y. S., Fall, A. A., Khoo, C., & Nadarajah, V. D. (2013). The multiple mini-interview (MMI) for student selection in health professions training–a systematic review. *Medical Teacher*, *35*(12), 1027-1041.

Pereira-Lima, K., & Loureiro, S. R. (2015). Burnout, anxiety, depression, and social skills in medical residents. *Psychology, Health & Medicine*, *20*(3), 353-362.

Perry, A. (2016). Black and brown boys don't need to learn grit; They need schools to stop being racist. *The Root*.

Petersdorf, R. G., Turner, K. T., Nickens, H. W., Ready, T. (1989). Minorities in medicine: Past, present, and future. *Academic Medicine*, *65*, 663–70.

Petersen, S., Hutchings, P., Shrader, G., & Brake, K. (2011). Integrating health care: The clear advantage for underserved diverse populations. *Psychological Services*, *8*(2), 69.

Petterson, S. M., Liaw, W. R., Phillips, R. L., Rabin, D. L., Meyers, D. S., & Bazemore, A. W. (2012). Projecting US primary care physician workforce needs: 2010-2025. *The Annals of Family Medicine*, *10*(6), 503-509.

Pieterse, A. L. (2007). Book review: Beyond the big test: Noncognitive variables in higher education. *Management and Evaluation in Counseling and Development*, *40*(3), 181-183.

Pike, G. R., & Kuh, G. D. (2005). First-and second-generation college students: A comparison of their engagement and intellectual development. *The Journal of Higher Education*, *76*(3), 276-300.

Poole, Jr., K. G. (2020). 'I want a doctor who looks like me': The dilemma of race-based requests. *Cleveland Clinic Journal of Medicine*, *87*(5), 268.

Poropat, A. E. (2009). A meta-analysis of the Five-Factor Model of Personality and academic performance. *Psychological Bulletin*, *135*(2), 322-338. doi: 10.1037/a0014996

Powell, K. (2013). Higher education: On the lookout for true grit. *Nature*, *504*(7480), 471-473.

Price, P. B., Lewis, E. G., Loughmiller, G. C., Nelson, D. E., Murray, S. L., & Taylor, C. W. (1971). Attributes of a good practicing physician. *Academic Medicine*, *46*(3), 229-37.

Prideaux, D., Roberts, C., Eva, K., Centeno, A., McCrorie, P., Mcmanus, C., Patterson, F., Powis, D., Tekian, A. & Wilkinson, D. (2011). Assessment for selection for the health care professions and specialty training: Consensus statement and recommendations from the Ottawa 2010 Conference. *Medical Teacher*, *33*(3), 215-223.

Prober, C. G., Kolars, J. C., First, L. R., & Melnick, D. E. (2016). A plea to reassess the role of United States Medical Licensing Examination Step 1 scores in residency selection. *Academic Medicine*, *91*(1), 12-15.

Rainey, M. (2001). How do we retain minority health professions students? (pp. 328-360) In *the Right Thing to Do, The Smart Thing to Do: Enhancing Diversity in Health Professions*: Summary of the Symposium on Diversity in Health Professions in Honor of Herbert W. Nickens, M. D, Washington, D. C.

Ramsey, J. H. (1997, August 14). Test scores aren't 100% of the picture. *San Jose Mercury News*, 10B.

Ray, R., & Brown, J. (2015). Reassessing student potential for medical school success: Distance traveled, grit, and hardiness. *Military Medicine*, *180*(4 Suppl), 138-141. doi:10.7205/MILMED-D-14-00578

Razack, S., Faremo, S., Drolet, F., Snell, L., Wiseman, J., & Pickering, J. (2009). Multiple mini-interviews versus traditional interviews: Stakeholder acceptability comparison. *Medical Education*, *43*(10), 993-1000.

Redford, G. (2020, April 3). "Itching to get back in": Medical students graduate early to join the fight. Association of American Medical Colleges. https://www.aamc.org/news-insights/itching-get-back-medical-students-graduate-early-join-fight

Reed, A. J., Schmitz, D., Baker, E., Nukui, A., & Epperly, T. (2012). Association of grit and satisfaction in rural and nonrural doctors. *Journal of American Board of Family Medicine, 25*(6), 832-839.

Reed, L., & Jeremiah, J. (2017). Student grit as an important ingredient for academic and personal success. In *Developments in Business Simulation and Experiential Learning: Proceedings of the Annual ABSEL conference, 44*, 252-256.

Reede, J. Y. (2003). A recurring theme: The need for minority physicians. *Health Affairs*, *22*(4), 91-93.

Rees, E. L., Hawarden, A. W., Dent, G., Hays, R., Bates, J., Hassell, A. B. (2016). Evidence regarding the utility of multiple mini-interview (MMI) for selection to undergraduate health programs: A BEME systematic review: BEME Guide No. 31. *Med Teach*, *38*(5), 443-455. doi:10.310910142159X.2016.1158799.

Reeves, R. E., Vishwanatha, J. K., Yorio, T., Budd, M., & Sheedlo, H. J. (2008). The post-baccalaureate premedical certification program at the University of North Texas Health Science Center strengthens admission qualifications for entrance into medical school. *Academic Medicine*, *83*(1), 45-51.

Reiter, H. I., Eva, K. W., Rosenfeld, J., & Norman, G. R. (2007). Multiple mini-interviews predict clerkship and licensing examination performance. *Medical Education*, *41*(4), 378-384.

Reiter, H. I., Lockyer, J., Ziola, B., Courneya, C. A., Eva, K., & Canadian Multiple Mini-Interview Research Alliance (CaMMIRA). (2012). Should efforts in favor of medical student diversity be focused during admissions or farther upstream? *Academic Medicine*, *87*(4), 443-448.

Relman, A. S. (1989). The changing demography of the medical profession. *New England Journal of Medicine, 321*(22), 1540-2.

Richardson, E. D. (2017). *An Assessment of Factors That Influence the Academic Achievement of Black Male Physicians* (Doctoral dissertation, Northeastern University).

Richman, J. A., Flaherty, J. A., Rospenda, K. M., & Christensen, M. L. (1992). Mental health consequences and correlates of reported medical student abuse. *Journal of American Medical Association*, *267*(5), 692-694.

Ris, E. (2016). The problem with teaching “grit” to poor kids? They already have it. Here’s what they really need.’ *The Washington Post*, 2016-2018.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261.

Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, *58*(1), 103-139.

Roberts, B. W., Hill, P. L., & Davis, J. P. (2017). How to change conscientiousness: The sociogenomic trait intervention model. *Personality Disorders: Theory, Research, and Treatment*, *8*(3), 199.

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In M. R. Leary & R. H. Hoyle (Eds.) *Handbook of individual differences in social behavior* (p.369-381). The Guilford Press.

Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M., & Hill, P. L. (2014). What is conscientiousness and how can it be assessed? *Developmental Psychology*, *50*(5), 1315.

Robertson-Kraft, C., & Duckworth, A. L. (2014). True grit: Trait-level perseverance and passion for long-term goals predict effectiveness and retention among novice teachers. *Teachers College Record*, *116*, 1-27.

Rojas, J. P., Reser, J. A., Usher, E. L., & Toland, M. D. (2012). Psychometric properties of the academic grit scale. [Research poster]. University of Kentucky College of Education. https://sites.education.uky.edu/motivation/files/2013/08/PojasPeserTolandUsher.pdf

Rolfe, I. E., Pearson, S., Powis, D. A., & Smith, A. J. (1995). Time for a review of admission to medical school?. *The Lancet*, *346*(8986), 1329-1333.

Roman, Jr., S. A., Sorenson, J. R., Davis, W. I., & Erickson, R. (1979). Predicting minority student performance in the first medical school year. *Journal of the National Medical Association*, *71*(7), 661.

Rosenthal, E. (2012, April 13). Pre-med's new priorities: Heart and soul and social science. *The New York Times.* https://www.nytimes.com/2012/04/15/education/edlife/pre-meds-new-priorities-heart-and-soul-and-social-science.html

Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research*, *21*(4), 37-59.

Roser, M. A. (1998, February 15) Test firm lauds A&M de-emphasis of MCAT. *Austin American-Statesman*, B1

Rosner, J. (2001). Disparate outcomes by design: University admission test. *Berkeley La Raza Law Journal*, 12, 377-386.

Roter, D. L., Hall, J. A., & Aoki, Y. (2002). Physician gender effects in medical communication: A meta-analytic review. *Journal of American Medical Association*, *288*(6), 756-764. doi:10.1001/jama.288.6.756.

Roth, K. S., Riley, W. T., Brandt, R. B., & Seibel, H. R. (1996). Prediction of students' USMLE Step 2 performances based on premedical credentials related to verbal skills. *Academic Medicine, 71*(2), 176-180.

Royal College of Physicians and Surgeons of Canada. (2020). About CanMEDS. Royal College. http:www.royalcollege.ca/rcsite/canmeds/about-canmeds-e.

Rubright, J. D., Jodoin, M., & Barone, M. A. (2019). Examining demographics, prior academic performance, and United States Medical Licensing Examination scores. *Academic Medicine*, *94*(3), 364-370. doi:10.1097/acm.0000000000002366.

Rumrill, Jr., P. D., Bellini, J. L., & Webb, J. M. (1999). Perspectives on scientific inquiry: Survey research in rehabilitation. *Journal of Vocational Rehabilitation*, *13*(3), 195-198.

Sacks, P. (2007). *Tearing down the gates.* University of California Press.

Sade, R. M., Stroud, M. R., Levine, J. H., & Fleming, G. A. (1985). Criteria for selection of future physicians. *Annals of Surgery*, *201*(2), 225.

Sadik, A., Woldemariam, B., & Wang, X. (2017). Master students' cGPA is a good predictor of success in medical school. *Medical Science Educator*, *27*(2), 193-199.

Saguil, A., Dong, T., Gingerich, R. J., Swygert, K., LaRochelle, J. S., Artino Jr, A. R., Coruess, D. F., & Durning, S. J. (2015). Does the MCAT predict medical school and PGY-1 performance? *Military Medicine*, *180*(suppl_4), 4-11.

Saipanish, R. (2003). Stress among medical students in a Thai medical school. *Medical Teacher*, *25*(5), 502-506.

Salsberg, E. S., & Forte, G. J. (2002). Trends in the physician workforce, 1980–2000. *Health Affairs*, *21*(5), 165-173.

Salsberg, E., & Grover, A. (2006). Physician workforce shortages: Implications and issues for academic health centers and policymakers. *Academic Medicine*, *81*(9), 782-787.

Salvatori, P. (2001). Reliability and validity of admissions tools used to select students for the health professions. *Advances in Health Sciences Education*, *6*(2), 159-175.

Samuelsen, B. T., Desai, V. S., Turner, N. S., Kelly, A. M., Grawe, B., & Camp, C. L. (2019). Generational differences in grit, self-control, and conscientiousness among orthopaedic surgeons: From Millennials to Baby Boomers. *Journal of Bone and Joint Surgery*, *101*(14), e71.

Santen, S. A., Holt, D. B., Kemp, J. D., & Hemphill, R. R. (2010). Burnout in medical students: examining the prevalence and associated factors. *Southern Medical Journal*, *103*(8), 758-763.

Schmidt, F. T., Lechner, C. M., & Danner, D. (2020). New wine in an old bottle? A facet-level perspective on the added value of Grit over BFI–2 Conscientiousness. *PloS One*, *15*(2), e0228969.

Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T. J., Billington, A. Q., Sinha, R., & Zorzie, M. (2009). Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of admitted students. *Journal of Applied Psychology*, *94*(6), 1479-1497. doi: 10.1037/a0016810

Sedlacek, W. E. (1977). Test bias and the elimination of racism. *Journal of College Student Personnel*, *18*, 16-20.

Sedlacek, W. E. (2001). *Why we should use noncognitive variables with graduate and professional students* [Unpublished doctoral dissertation, University of Maryland].

Sedlacek, W. E. (2004). *Beyond the big test: Noncognitive assessment in higher education*. Jossey-Bass.

Sedlacek, W. E. (2011). Using noncognitive variables in assessing readiness for higher education. *Readings on Equal Education*, *25*(13), 187-205.

Shah, M., Hasan, S., Malik, S., & Sreeramareddy, C. T. (2010). Perceived stress, sources and severity of stress among medical undergraduates in a Pakistani medical school. *BMC Medical Education*, *10*(1), 2.

Shanafelt, T. D., Boone, S., Tan, L., Dyrbye, L. N., Sotile, W., Satele, D., West, C. P., Sloan, J., & Oreskovich, M. R. (2012). Burnout and satisfaction with work-life balance among US physicians relative to the general US population. *Archives of Internal Medicine*, *172*(18), 1377-1385.

Shechtman, N., DeBarger, A. H., Dornsife, C., Rosier, S., & Yarnell, L. (2013). *Promoting grit, tenacity, and perseverance: Critical factors for success in the 21st century*. SRI International. Menlo Park, CA.

Shen, H., & Comrey, A. L. (1997). Predicting medical students' academic performances by their cognitive abilities and personality characteristics. *Academic Medicine*, *72*(9), 781-6.

Sherina, M. S., Rampal, L., & Kaneson, N. (2004). Psychological stress among undergraduate medical students. *Medical Journal of Malaysia*, *59*(2), 207-211.

Sherry, S. B., Hewitt, P. L., Flett, G. L., Lee-Baggley, D. L., & Hall, P. A. (2007). Trait perfectionism and perfectionistic self-presentation in personality pathology. *Personality and Individual Differences, 42*(3), 477-490.

Sheskin, D., & Sheskin, D. J. (2011). Test 29: Spearman's rank-order correlation coefficient. *Handbook of parametric and nonparametric statistical procedures* (5th ed). Chapman & Hall/CRC.

Silverman, B. D. (2012, January). Physician behavior and bedside manners: The influence of William Osler and The Johns Hopkins School of Medicine. In *Baylor University Medical Center Proceeding, 25*(1), 58-61.

Silvia, P. J., Eddington, K. M., Beaty, R. E., Nusbaum, E. C., & Kwapil, T. R. (2013). Gritty people try harder: Grit and effort-related cardiac autonomic activity during an active coping challenge. *International Journal of Psychophysiology*, *88*(2), 200-205. doi: http://dx.doi.org/10.1016/j.ijpsycho.2013.04.007

Simon, S. R., Bui, A., Day, S., Berti, D., & Volkan, K. (2007). The relationship between second-year medical students' OSCE scores and USMLE Step 2 scores. *Journal of Evaluation in Clinical Practice, 13*(6), 901-905. doi:10.1111/j.1365-2753.2006.00768.x

Slavin, R., & Smith, D. (2009). The relationship between sample sizes and effect sizes in systematic reviews in education. *Educational Evaluation and Policy Analysis*, *31*(4), 500-506.

Smedley, B. D., Stith, A. Y., Colburn, L., & Evans, C. H. (2001, August). The right thing to do, the smart thing to do. Enhancing diversity in the health professions. In *In: the right thing to do, the smart thing to do: enhancing diversity in the health professions–summary of the symposium on diversity in health professions in honor of Herbert W. Nickens, MD* (Vol. 1, p. 35).

Smith, C. K., Peterson, D. F., Degenhardt, B. F., & Johnson, J. C. (2007). Depression, anxiety, and perceived hassles among entering medical students. *Psychology, Health & Medicine*, *12*(1), 31-39.

Sobowale, K., Ham, S. A., Curlin, F. A., & Yoon, J. D. (2018). Personality traits are associated with academic achievement in medical school: A nationally representative study. *Academic Psychiatry*, 42(3), 338-345. doi:10.1007/s40596-017-0766-5

Socol, I. D. (2014). TAKING A CLOSER AT THE GRIT NARRATIVES. *Knowledge Quest*, *43*(1), 8.

Solórzano, R. W. (2008). High stakes testing: Issues, implications, and remedies for English language learners. *Review of Educational Research*, *78*(2), 260-329.

Sommerfeld, A. (2011). Recasting non-cognitive factors in college readiness as what they truly are: Non-academic factors. *Journal of College Admission*, *213*, 18-22.

Southgate, E., Brosnan, C., Lempp, H., Kelly, B., Wright, S., Outram, S., & Bennett, A. (2017). Travels in extreme social mobility: How first-in-family students find their way into and through medical education. *Critical Studies in Education*, *58*(2), 242-260.

Sprinthall, R. C. (2012). *Basic statistical analysi*s. Pearson.

Steele, C. (1999). Thin ice: "Stereotype threat" and black college students. *The Atlantic Monthly, 284* (2), 50-54.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613.

Stegers-Jager, K. M., Steyerberg, E. W., Lucieer, S. M., & Themmen, A. P. (2015). Ethnic and social disparities in performance on medical school selection criteria. *Medical Education*, *49*(1), 124-133.

Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology*, *3*(4), 292-316.

Sternberg, R. J. (2005). Accomplishing the goals of affirmative action - with or without affirmative action. *Change, Jan-Feb,* 6-13.

Sternberg, R. J. (2008). Increasing academic excellence and enhancing diversity are compatible goals. *Educational Policy, 22*(4), 487-514. doi: 12.1177/0895904807310037

Stoffel, J. M., & Cain, J. (2018). Review of grit and resilience literature within health professions education. *American Journal of Pharmaceutical Education*, *82*(2), 6150.

Stratton, T. D., McLaughlin, M. A., Witte, F. M., Fosson, S. E., & Nora, L. M. (2005). Does students' exposure to gender discrimination and sexual harassment in medical school affect specialty choice and residency program selection? *Academic Medicine*, *80*(4), 400-408.

Strayhorn, T. (2013). What role does grit play in the academic success of black male collegians at predominantly white institutions? *Journal of African American Studies*, *18*, 1-10. doi: 10.1007/s12111-012-9243-0

Sullivan, L. W. (2004). Missing persons: Minorities in the health professions, A report of the Sullivan Commission on Diversity in the Healthcare Workforce. The Sullivan Commission. http://hdl.handle.net/1903/22267

Sutherland, J. A., Hamilton, M. J., & Goodman, N. (2012). Affirming at risk minorities for success: ARMS retention graduation and success on the NCLEX-RN. *Journal of Nursing Education*, *46*(8), 347-353.

Suzuki, Y., Tamesue, D., Asahi, K., & Ishikawa, Y. (2015). Grit and work engagement: A cross-sectional study. *PloS ONE*, *10*(9), e0137501.

Swail, W. S. (2003). Retaining minority students in higher education: A framework for success. *ASHE-ERIC higher education report.* Jossey-Bass Higher and Adult Education Series. Washington, D. C.

Swails, J. L., Aibana, O., & Stoll, B. J. (2019). The conundrum of the United States Medical Licensing Examination score reporting structure. *Journal of the American Medical Association*, *322*(7), 605-606.

Swygert, K. A., Cuddy, M. M., van Zanten, M., Haist, S. A., & Jobe, A. C. (2012). Gender differences in examinee performance on the Step 2 Clinical Skills® data gathering (DG) and patient note (PN) components. *Advances in Health Sciences Education*, *17*(4), 557-571. doi:10.1007/s10459-011-9333-0.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5). Pearson.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5, pp. 481-498). Boston, MA: Pearson.

Tamblyn, R. (1998). Association between licensing examination scores and resource use and quality of care in primary care practice. *Journal of the American Medical Association, 280*(11), 989. doi:10.1001/jama.280.11.989

Taylor, M. L., Blue, A. V., Mainous, III, A. G., Geesey, M. E., & Basco, W. T., Jr. (2005). The relationship between the National Board of Medical Examiners' prototype of the Step 2 clinical skills exam and interns' performance. *Academic Medicine, 80*(5), 496-501.

Terenzini, P. T., Cabrera, A. F., & Bernal, E. M. (2001). *Swimming against the tide: The poor in American higher education.* College Entrance Examination Board. https://files.eric.ed.gov/fulltext/ED562879.pdf

Tett, R. R., Jackson, D. N., Rothstein, M., & Reddon, J. R. (1994). Meta-analysis of personality-job performance relations: A reply to Ones, Mount, Barrick, and Hunter (1994). *Personnel Psychology*, *47*(1), 157-172.

Tewell, E. (2020). The Problem with Grit: Dismantling Deficit Thinking in Library Instruction. *portal: Libraries and the Academy*, *20*(1), 137-159.

Thomas, N. K. (2004). Resident burnout. *Journal of American Medical Association*, *292*(23), 2880-2889.

Thundiyil, J. G., Modica, R. F., Silvestri, S., & Papa, L. (2010). Do United States Medical Licensing Examination (USMLE) scores predict in-training test performance for emergency medicine residents? *The Journal of Emergency Medicine*, *38*(1), 65-69. doi:10.1016/j.jemermed.2008.04.010

Ting, S. R. (2009). Impact of noncognitive factors on first-year academic performance and persistence of NCAA Division I student-athletes. *The Journal of Humanistic Counseling, Education and Development*, *48*(2), 215-228.

Tough, P. (2012). *How children succeed: Grit, curiosity, and the hidden power of character*. Houghton Mifflin Harcourt.

Towne, L., & Shavelson, R. J. (2002). *Scientific research in education*. National Research Council.

Tracey, T. J., & Sedlacek, W. E. (1984). Noncognitive variables in predicting academic success by race. *Measurement and Evaluation in Guidance*, *16*(4), 171-178.

Trapmann, S., Hell, B., Hirn, J. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Journal of Psychology*, *215*, 132-151.

Turner, B. O., Paul, E. J., Miller, M. B., & Barbey, A. K. (2018). Small sample sizes reduce the replicability of task-based fMRI studies. *Communications Biology*, *1*(1), 1-10.

Underdahl, L., Jones-Meineke, T., & Duthely, L. M. (2018). Reframing physician engagement: An analysis of physician resilience, grit, and retention. *International Journal of Healthcare Management*, *11*(3), 243-250.

United States Census Bureau. (2017). Hispanic or Latino, and not Hispanic or Latino by race for the population 18 years and over. [Data set]. https://data.census.gov/cedsci/table?q=race%20combinations%20of%20two%20races%20and%20not%20hispanic%202010&hidePreview=false&tid=DECENNIALSF12010.P11&y=2010&vintage=2010

United States Census Bureau. (2018). *United States quick facts*. Census. https://www.census.gov/quickfacts/fact/table/US/RHI225218

United States Center for Labor Statistics. (2019, September 4). Employment projections: 2018-2028. [Press release]. https://www.bls.gove/news.release/pdf/ecopro.pdf

United States Medical Licensing Examination. (2019). *USMLE score interpretation guidelines*. USMLE. https://www.usmle.org/pdfs/transcripts/USMLE_Step_Examination_Score_Interpretation_Guidelines.pdf

United States Medical Licensing Examination. (2020a). *Step 1*. USMLE. https://www.usmle.org/step-1/

United States Medical Licensing Examination. (2020b). *Step 2 CK*. USMLE. https://www.usmle.org/step-2-ck/

United States Medical Licensing Examination. (2020c). Step 2 Clinical Knowledge (CK): Content description and general information. USMLE. https://www.usmle.org/pdfs/step-2-ck/Step2CK_Content.pdf

United States Medical Licensing Examination. (2020d). *Step 2 CS*. USMLE. https://www.usmle.org/step-2-cs/

United States Medical Licensing Examination. (2020e). Step 2 CK. Performance Data. https://www.usmle.org/performance-data/

van Zanten, M., Boulet, J. R., McKinley, D. W., DeChamplain, A., & Jobe, A. C. (2007). Assessing the communication and interpersonal skills of graduates of international medical schools as part of the United States Medical Licensing Exam (USMLE) Step 2 Clinical Skills (CS) Exam. *Academic Medicine*, *82*(10), S65-S68. doi:10.1097/acm.0b013e318141f40a.

Veloski, J. J., Callahan, C. A., Xu, G., Hojat, M., & Nash, D. B. (2000). Prediction of students' performances on licensing examinations using age, race, sex, undergraduate GPAs, and MCAT scores. *Academic Medicine*, *75*(10), S28-S30. doi:10.1097/00001888-200010001-00009.

Verrier, D. (2010). Holistic review and medical school admissions: Implications for the health professions advisor. *The Advisor, 30*(1).

Verstegen, D. A. (2015). On doing an analysis of equity and closing the opportunity gap. *Education Policy Analysis Archives*, *23*, 41.

Villwock, J. A., Sobin, L. B., Koester, L. A., & Harris, T. M. (2016). Impostor syndrome and burnout among American medical students: A pilot study. *International Journal of Medical Education*, *7*, 364.

Wagerman, S. A., & Funder, D. C. (2007). Acquaintance reports of personality and academic achievement: A case for conscientiousness. *Journal of Research in Personality*, *41*(1), 221-229.

Wald, H. S. (2015). Professional identity (trans) formation in medical education: reflection, relationship, resilience. *Academic Medicine, 90*(6), 701-706.

Waldman, J. D., & Spector, R. A. (2003). Malpractice claims analysis yields widely applicable principles. *Pediatric Cardiology*, *24*(2), 109-117.

Walker, K. O., Moreno, G., & Grumbach, K. (2012). The association among specialty, race, ethnicity, and practice location among California physicians in diverse specialties. *Journal of the National Medical Association*, *104*(1-2), 46-52.

Wall, A. L., Aljets, A., Ellis, S. C., Hansen, D. J., Moore, W. M., Petrelli, H. M., Speedie, M. K., TenHoeve, T., Watchmaker, C., Winnike, J. S., & Wurth, S. D. (2015). White paper on pharmacy admissions: Developing a diverse workforce to meet the health-care needs of an increasingly diverse society: Recommendations of the American Association of Colleges of Pharmacy special committee on admissions. *American Journal of Pharmaceutical Education*, *79*(7), S7.

Walsh, M. J. (2020). Online doctoral student grade point average, conscientiousness, and grit: A moderation analysis. *Journal of Educators Online*, *17*(1), n1.

Warburton, E. C., Bugarin, R., Nuñez, A. M., & Carroll, C. D. (2001, May). Bridging the gap: Academic preparation and postsecondary success of first-generation students. [Statistical analysis report]. National Center for Education Statistics.

Wear, D., Aultman, J. M., & Borges, N. J. (2007). Retheorizing sexual harassment in medical education: Women students' perceptions at five US medical schools. *Teaching and Learning in Medicine*, *19*(1), 20-29.

Wendland, C. L. (2010). *A Heart for the Work: journeys through an African medical school*. University of Chicago Press.

Weiss, K. (2001, February 18). UC faculty chief backs dropping SAT as unfair. *Los Angeles Times*. http://articles.latimes.com/2001/feb/18/local/me-27050.

Werner, K. M., Milyavskaya, M., Klimo, R., & Levine, S. L. (2019). Examining the unique and combined effects of grit, trait self-control, and conscientiousness in predicting motivation for academic goals: A commonality analysis. *Journal of Research in Personality*, *81*, 168-175.

White, C. B., Dey, E. L., & Fantone, J. C. (2009). Analysis of factors that predict clinical performance in medical school. *Advances in Health Sciences Education*, *14*(4), 455-464.

Wiggins, N., Blackburn, M., & Hackman, R. J. (1969). Prediction of first-year graduate success in psychology: Peer ratings. *The Journal of Educational Research*, *63*(2), 81-85.

Wijesekera, T. P., Kim, M., Moore, E. Z., Sorenson, O., & Ross, D. A. (2019). All other things being equal: Exploring racial and gender disparities in medical school honor society induction. *Academic Medicine*, *94*(4), 562-569. doi:10.1097/acm.0000000000002463.

Williams, C., Kwan, B., Pereira, A., Moody, E., Angus, S., & El-Bayoumi, J. (2019). A call to improve conditions for conducting holistic review in graduate medical education recruitment. *MedEdPublish, 8, 1-10.*

Williams, D., Tricomi, G., Gupta, J., & Janise, A. (2015). Efficacy of burnout interventions in the medical education pipeline. *Academic Psychiatry*, *39*(1), 47-54.

Wilson, T. D., & Buttrick, N. R. (2016). New directions in social psychological interventions to improve academic achievement. *Journal of Educational Psychology*, *108*(3), 392.

Winseman, J., Roselande Marcellon MS, B. A., Huamantla, M., & Ruiz, C. (2018). First-generation medical student wellness in the United States: A cross-sectional snapshot. *Journal of Best Practices in Health Professions Diversity*, *11*(2), 96-106.

Wise, F. S. (2020). *Applying to Medical School: Persistence in the Nontraditional Postbaccalaureate Adult Learner* (Doctoral dissertation, Capella University).

Witzburg, R. A., & Sondheimer, H. M. (2013). Holistic review—shaping the medical profession one applicant at a time. *New England Journal of Medicine*, *368*(17), 1565-1567.

Wolfe, R. N., & Johnson, S. D. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement*, *55*(2), 177-185.

Woolf, K., McManus, I. C., Potts, H. W., & Dacre, J. (2013). The mediators of minority ethnic underperformance in final medical school examinations. *British Journal of Educational Psychology*, *83*(1), 135-159.

Wros, P., & Noone, J. (2018). Holistic admissions in undergraduate nursing: One school's journey and lessons learned. *Journal of Professional Nursing, 34*(3), 211-6.

Yedidia, M. J., Gillespie, C. C., Kachur, E., Schwartz, M. D., Ockene, J., Chepaitis, A. E., Snyder, C. W., Lazare, A., & Lipkin, Jr., M. (2003). Effect of communications training on medical student performance. *Journal of American Medical Association*, *290*(9), 1157-1165. doi:10.1001/jama.290.9.1157.

Yosso, T. J. (2005). Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethnicity and Education*, *8*(1), 69-91.

Yosso, T. J. (2013). *Critical race counterstories along the Chicana/Chicano educational pipeline*. Routledge.

Zivot, J. (2020, March 3). Doctors volunteer for COVID-19 duty: Who has their backs? Medpage Today. https://www.medpagetoday.com/infectiousdisease/covid19/85191

Zuger, A. (2004). Dissatisfaction with medical practice. *New England Journal of Medicine*, *350*(1), 69-75.

Zyphur, M. J., Bradley, J. C., Landis, R. S., & Thoresen, C. J. (2007). The effects of cognitive ability and conscientiousness on performance over time: A censored latent growth model. *Human Performance*, *21*(1), 1-27.

Vita

NAME OF AUTHOR: Jennifer C. Welch

PLACE OF BIRTH: Utica, New York, USA

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:
Syracuse University, Syracuse, NY
State University of New York at Oswego, Oswego, NY
State University of New York at Potsdam, Potsdam, NY
Mohawk Valley Community College, Utica, NY

DEGREES AWARDED:
Master of Science in Counselor Education, 1998, Syracuse University
Bachelor of Arts in Economics, 1992

AWARDS:
Presidential Excellence Award, Campus Leader of the Year, 2010

PROFESSIONAL EXPERIENCE:
Chief Enrollment Management Officer and Associate Dean, Admissions and Financial Aid, 2017-present, SUNY Upstate Medical University, Syracuse, NY
Associate Dean, Admissions and Financial Aid, 2031-2017, SUNY Upstate Medical University, Syracuse, NY
Director of Admissions, 2001-2013, SUNY Upstate Medical University, Syracuse, NY
District Registrar, School Counselor, and Department Chairperson, 2000-2001, Liverpool Central School District, Liverpool, NY
Assistant Director of Admissions, 1997-2000, SUNY Upstate Medical University, Syracuse, NY
Admissions Advisor, 1994-1997, SUNY Upstate Medical University, Syracuse, NY
Admissions Counselor, 1993-1994, Elmira College, Elmira, NY