January 2020

A Study Of Self-Evaluation And Estimator Error In Internal Medicine Residents And Its Influence On The Academic Achievement Of Medical Knowledge

William Morse
Wayne State University

A STUDY OF SELF-EVALUATION AND ESTIMATOR ERROR IN INTERNAL MEDICINE RESIDENTS AND ITS INFLUENCE ON THE ACADEMIC ACHIEVEMENT OF MEDICAL KNOWLEDGE

by

WILLIAM MORSE

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2020

MAJOR: LEARNING DESIGN AND

TECHNOLOGY

Approved By:

____
Advisor Date

____

____

____

____

ACKNOWLEDGMENTS

I would like to express my gratitude to my program advisor Ingrid Guerra-López, Ph.D. my departmental professors and dissertation committee which includes Sean Drake, MD, Monica Tracey, Ph.D, and Ke Zhang, Ph.D. It was an honor to work and study under your guidance. Your support, expertise and insightful comments have been invaluable to my education.

Thank you to Eric Scher, MD, Kimberly Baker-Genaw, MD, Odaliz Abreu Lanfranco, MD and Henry Ford Health System. Collectively, you create an environment that inspires higher learning and academic achievement through opportunities that foster personal growth. The support I have received as I furthered my education has been immeasurable.

I would also like to thank Mrs. Kathy Huber—the only person to have read my dissertation and provide feedback that wasn't part of Wayne State University committee. Thank you for your friendship and caring support.

Finally, I would like to thank my wife (Leia), children (Spencer, Madison and Gavin) and my other family members for their continued love, support, prayer, and understanding the many days and nights I have spent working tirelessly toward this degree. Your support has helped me get through this process in the most positive way possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: DEFINITION OF THE PROBLEM

Introduction to the Problem

The graduate medical education (GME) system is adapting in a variety of ways because of the increasing challenges presented in an ever-evolving health care environment. Residency training programs must produce effective educational strategies to motivate and prepare residents to achieve both their personal and professional goals accompanied by societal needs. Competency-based medical education is expected to support the increasing complexities of care through the delineation of fundamental skills and attributes "that the profession and the public expect of a physician" (Carraccio & Englander, 2013, p. 1067). It is trusted that the regular practice and attainment of these professional competencies by physicians-in-training through the guidance of graduate medical education will have a significant impact on the development of future physicians to deliver outstanding patient care.

Graduate medical education was established based on experiential learning, whereby physicians-in-training become competent in clinical practice by providing patient care initially under supervision with the aim of shifting toward direct patient care gradually through training experiences. In the United States, the Accreditation Council for Graduate Medical Education (ACGME) is the regulatory body responsible for the administration, monitoring, and enforcement of educational standards and requirements for training programs through the process of accreditation (ACGME, 2016a). In 1999, the ACGME, together with the American Board of Medical Specialties (ABMS), endorsed a set of six core competency domains intended to provide a framework for training programs to phase into teaching and assessment (Carraccio & Englander, 2013). The competency domains included: (1) patient care, (2) medical knowledge, (3) practice-based learning and improvement, (4) interpersonal and communication skills, (5) professionalism, and (6) systems-based practice.

One competency domain, in particular, is central to understanding the background of this research, and that is *practice-based learning and improvement* (PBLI). The ACGME provides this general description of the competency: “practice-based learning and improvement is one of the defining characteristics of being a physician. It is the ability to investigate and evaluate the care of patients, to appraise and assimilate scientific evidence, and to continuously improve patient care based on constant self-evaluation and lifelong learning” (p. 20). This competency is related to the broader concept of self-directed learning. It further describes complex processes that include self-evaluation, competency-related beliefs, or self-perceptions concerning their abilities, self-motivation, and other latent constructs that mediate continuous improvement in daily practice (ACGME, 2016a).

Additionally, the ACGME (2016a), states “the intention of this competency [PBLI] is to help a physician develop the habits of mind required to continuously pursue quality improvement, well past the completion of residency” (p. 20). Based on the available literature, self-directed learning skills are necessary, in a field that is vastly transformative, for effective lifelong learning in physicians as a means of continuous improvement in medical knowledge and the daily practice of medicine (Burke, Benson, Englander, Carraccio & Hicks, 2014; Eva & Regehr, 2007). The principles that define *self-directed learning* for this study are based on Malcolm Knowles (1975) description “individuals take the initiative, with or without the help of others, in diagnosing their learning needs, formulating goals, identifying human and material resources for learning, choosing and implementing appropriate learning strategies, and evaluating learning outcomes” (p. 18).

The research literature acknowledges, and to set the context for the remainder of this study, a critical element of self-directed learning is considered the ability of learners to recognize their deficits in performance in order to motivate a response to rectify any discrepancies (Burke, Benson, Englander, Carraccio & Hicks, 2014; Eva & Regehr, 2007). It is theorized in the literature that a person's insight and awareness into their inadequacies through the process of self-evaluation is an essential prerequisite to the development of skills and the progress toward any competence (Dunning, 2011; Kruger & Dunning, 1999). Thus, the main objective of this study is to address the research question that is unobserved in the literature—how do residents self-evaluate and what, if any, influence does this have on their academic achievement of medical knowledge?

Statement of the Problem

After reviewing the literature, the research found that graduate medical education has built up a considerable amount of conceptual structure for its competency-based approach over the past decades; nonetheless, several outstanding issues are prompting further investigation. The identification of the factors contributing to achievement motivation is a way of promoting capable lifelong self-directed learners. However, the literature reveals that medical education training programs lack awareness of what competency-related tasks, training activities, and experiences stimulate resident behaviors to learn. Numerous theories have been developed in general education to explain student motivation to learn that has led to the development of educational interventions to support learning, but this practice is scarcely applied in the GME context (Cook & Artino, 2016; Kusurkar, Ten Cate, Van Asperen, & Croiset, 2011). Motivation is a generally neglected aspect in designing the medical education curricula (Schutte et al., 2017).

What is motivation, and why is it essential to self-directed learning skills? Nevid (2013) defines motivation as "factors that activate, direct, and sustain goal-directed behavior" (p. 288).

Further, research has demonstrated that identifying these factors are particularly crucial to predicting three achievement modes of behavior often cited in the literature: (1) the intensity of effort, (2) the direction of attention (e.g., the source of gratification, academic choice), and (3) the persistence of effort over time (Eccles & Wigfield, 2002; Mitchell & Daniels, 2003; Schunk, 2012). Cook and Artino (2016) argue that the motivation of residents is often unobserved in academic medical research. Thus, the lack of evidence and research findings inhibits medical education's ability to connect theoretical foundations that enable educators to think about achievement motivation systematically.

How does motivation relate to self-evaluation? A resident must be capable of defining and resolving their own learning needs to self-regulate their behavior toward the achievement of goals. It is thought that a resident's ability to identify their deficiencies relates to their ability to self-evaluate accurately (Burke, Benson, Englander, Carraccio & Hicks, 2014). As a whole, people are often overconfident in the accuracy of their own abilities (Kornel & Bjork, 2009). The Dunning and Kruger effect in self-evaluation emphasized that without proper strategies to achieve success, incompetent people are unable to realize their lack of ability and continue to perform as if they are doing fine (Kruger & Dunning, 1999). Additionally, Kruger and Dunning (1999) stated, "in essence, we argue that the skills that engender competence in a particular domain are often the very same skills necessary to evaluate competence in that domain—one's own or anyone else's" (p. 1121). Moreover, they contend the reason for this lack of insight and self-awareness is related to metacognition, a person's knowledge of their own knowledge or abilities.

How might insight and self-awareness improve medical knowledge and skill development? The problem with improving medical education often appears to relate to the absence of feedback relative to performance (Ende, 1983; Graber, Gordon & Franklin, 2002; Graber, Gordon, &

Franklin, 2005; Mamede et al., 2010). Without feedback, learners may have insufficient self-awareness of performance deficits making it more challenging to modify behaviors effectively (Kruger & Dunning, 1999). As Ende (1983) stated, “Without feedback, mistakes go uncorrected, good performance is not reinforced, and clinical competence is achieved empirically or not at all” (p. 778). The research literature acknowledges that the consequences of inadequate medical knowledge and skill development are associated with several problems affecting patient care outcomes (Ende, 1983; Mamede et al., 2010).

Absent in the literature is the determination of whether the self-evaluation skills of residents or more critically, the measurement errors in self-evaluation motivate any difference in the academic achievement of medical knowledge? There are consistent suggestions in the literature that the application of theoretical foundations of motivation, self-evaluations skills, the development of metacognition, and feedback are considerable problems among residency training programs (Bing-You et al., 2009; Branch & Paranjape, 2002; Ende, 1983; Cook & Artino, 2016). Achievement motivation theory, self-evaluation constructs, and performance-related outcomes are well-established concepts that have been the subject of several publications in other populations but limited in the resident population (Kruger & Dunning, 1999; Eccles & Wigfield, 2002; Cook & Artino, 2016). The researcher of this study speculates that establishing relevant and meaningful empirical evidence related to self-evaluation outcomes through research is a critical step toward assisting residents in fostering lifelong self-directed learning skills and academic achievement of medical knowledge.

Purpose of the Study

The main objective of this study is to address the research question—how do residents self-evaluate and what, if any, influence does this have on their academic achievement of medical knowledge? The research has two primary overarching purposes. First, a review of the literature provided an analysis of gaps within the existing graduate medical education population and context. Many everyday problems in research materialize from an absence of information. Based on the available literature, the researcher believes an exploratory study is necessary to establish evidence where no data exists for the resident population, and little is known about their response patterns. Thus, the researcher aims to provide educators and program directors with relevant and meaningful information about issues associated explicitly with self-evaluation skills and medical knowledge appraisal.

Second, research is an integral part of professional education practices and a way of questioning observations while at the same time attempting to study and interpret what is observed. Without relevant and meaningful new evidence, the possibilities of building on or extending theoretical foundations based in psychology and other related fields into graduate medical education are challenging. Because factors contributing to self-evaluation skills will not be known in the study population, this investigation intends to present a synthesis of research literature to establish a framework that offers distinct definitions that enable medical educators to think systemically about theoretical foundations in residency training and its effects on competence. From the literature, a clear set of structural parameters based on academic research will guide the operationalization of the study to increase confidence in the construct validity.

Research Questions

This study will present an initial effort to assess the outcomes of self-evaluation empirically in a new population of internal medicine residents to determine what, if any, influence it has on the academic achievement of medical knowledge? The study was designed with the intent of comparing residents' subjective assessments of performance (i.e., competency-related beliefs) on a study survey relative to their actual return on an existing objective operational standard—the Internal Medicine In-Training Examination (IM-ITE). The IM-ITE is a well-established instrument for the measurement of academic achievement of medical knowledge. According to the American College of Physicians (ACP) website, the IM-ITE is designed to evaluate the expected medical knowledge, diagnostic reasoning, and clinical judgment skills of a certified internist (American College of Physicians, 2019).

Available evidence suggests that competency-related beliefs, estimations, or expectancies are thought to constitute an essential source of self-motivation (Wigfield & Eccles, 2000; Eccles & Wigfield, 2002). Based on this theoretical foundation in literature, the researcher determined that measuring predictions for performance, as a type of competency-related belief, reflects the ability of a resident to estimate or judge their perceived learning needs. The author of this paper uses the terms *prediction*, *estimation*, and *expectancy* throughout this report interchangeably to denote the same character of competency-related belief; the construct is mostly focused on predictive self-judgments for performance.

In this research, the focus will be on self-evaluation skills as a potential mediator for the academic achievement of medical knowledge. Therefore, operationally, when the forecasted value of predicted performance is relatively compared to the objective resource value, dimensions can be revealed as a kind of bias between internal (i.e., subjective) and external positions—a type of *estimator error*. This is especially crucial to recognizing the residents' predisposition orientation of learning needs for the target (Kluger & DeNisi, 1996; Eccles & Wigfield, 2002). Hereafter, these perceptional biases, or measurement error dimensions, will be broadly termed *estimation errors* for the remainder of this report when considering these dimensional aspects. However, the *estimator error-index* is a specific integer that represents the quantitative difference between the predicted performance and actual performance as an interval value. The estimator error-index value operates to indicate the direction and intensity dimensions as a distinct measure.

Finally, to better understand some underlying response patterns or population features of residents, two common resident characteristics will be considered as independent variables in exploration of this study: (1) training level (i.e., post-graduate year) and (2) gender. As a result, this study will provide evidence intended to answer these six research questions:

1. How well do residents predict their actual performance on the IM-ITE?
2. How does the actual performance on the IM-ITE differ by training level or gender of residents?
3. How does the predicted performance on the IM-ITE differ by the training level or gender of residents?
4. How does the estimator error-index of residents influence actual performance on the IM-ITE?
5. How does the estimator error-index on the IM-ITE differ by training level or gender of residents?
6. Does the frequency of overestimation or underestimation have any influence on actual performance?

Significance of the Study

This research is one of the first known empirical studies to apply learning theory in graduate medical education to clarify the relationship between self-evaluation and resident performance. The theoretical foundations provide a set of principles that bridge the gap between observations and interpretation of the outcomes. Information about the implications of competency-related beliefs, estimation errors gained through this study could prove beneficial to supporting residents through improved learning and performance design, and evaluation methods in education.

The lack of evidence and research findings related to competency-related beliefs inhibits medical education's capability to connect self-evaluation concepts to theoretical foundations of achievement motivation. The results that emerge from this study will contribute to the extension of theoretical foundations based on psychology and related fields into graduate medical education. Kusurkar et al. (2011) argue that motivation is well-researched in general education but rare in medical education. Cook and Artino (2016) urged research that "builds and extends motivation theory for education generally and health professions education specifically" (p. 1012). No known studies have investigated how resident competency-related beliefs develop or their impact on academic achievement of medical knowledge.

Primarily, the lack of suitable measurement instruments negates our ability to study the unobservable inherent resident characteristics that contribute to competence. Fortunately, for this study, the IM-ITE is well-established through methodological rigor that improves its validity and reliability as an objective resource standard. What is needed, as Rokeach (1979) explains, is "a measurement instrument that identifies, or attempts to identify, major end-states of human existence and the behavioral modes for achieving them" (p. 50). A competency-related belief, generally speaking, is what materializes from an individual's preferences, predispositions, and prior experiences. While there is no generally agreed-upon definition of self-evaluation, the term reflects typically the ability of a person to estimate, predict, or judge their position relative to the presence of a target—in this case, the IM-ITE. The estimation can be viewed as a kind of "expectancy" for performance, indicating a self-perceived end-state or predisposition orientation of a learning need when corresponding to the external reference standard.

The expectancy-value theory of Eccles and Wigfield (2002) provides a clear set of structural parameters for expectancies in achievement motivation based on theoretical research that considers goal-directed behaviors. Wigfield and Eccles (2000) describe expectancies as “beliefs about how well a student will do on upcoming tasks, either in the immediate or longer-term future” (p. 70). There have been no empirical investigations into the competency-related beliefs of residents or a comparison of their estimation errors related to the achievement of medical knowledge. Without a study, medical education lacks awareness of how self-evaluation of these competency-related beliefs develop or motivate resident self-directed learning behavior toward learning outcomes.

Finally, the research literature indicates that the cognitive bias related to the overestimation of one’s ability has exhibited resistance to change or modification without an evaluative process that increases awareness of the dissonance between expectancy sources (Salmoni, Schmidt & Walter, 1984; Kluger & DeNisi, 1996); therefore feedback appears critical. The findings of this study will benefit our understanding of competency-related beliefs, which could have implications for the achievement of medical knowledge by adjusting metacognitive accuracy in recognizing learning needs. The findings produced through the research will add to the existing body of knowledge. The study is intended to be exploratory and constructive as a cycle in the instructional design process by identifying aspects that may be carried forward in future learning interventions or research.

CHAPTER 2: REVIEW OF THE LITERATURE

Introduction

A literature review was conducted in order to create a broad understanding of the background, the problem, challenges, existing methodologies, and relevant theories. The purpose of the literature review in this study was to provide a well-grounded appropriate selection of methods and validity and to inform the reader of background material. The literature review was based upon an examination of textbooks and relevant scientific journals.

Historical Overview of Medical Practice Requirements

In 1910, Abraham Flexner provided a report on the condition of medical education in the United States and Canada that revealed an over-production of uneducated and ill-trained medical practitioners (Flexner, Pritchet & Henry, 1910). Since the *Flexner Report,* physician education has made significant contributions to knowledge and progress. Today, physicians enter practice with strong scientific foundations in the biological and physical sciences (Eden et al., 2014). In all 50 states, District of Columbia and Puerto Rico, to provide direct patient care, a medical practitioner must complete training in an accredited undergraduate medical education program and meet licensure requirements (American Board of Medical Specialties, 2018).

Formal physician education begins with undergraduate medical education in an allopathic or osteopathic medical school (Eden et al., 2014, p. 1-4; Mowery, 2015). Graduate medical education (GME) describes the period of residency and fellowship training that is provided to physicians after they receive the M.D. or D.O degree (ACGME, 2016a; American Board of Medical Specialties, 2018; Mowery, 2015). Eden, Berwick & Wilensky (2014), "Board certifications are not required to practice medicine in any state as medical licenses are not specialty-specific" (p.118). Although voluntary, most physicians engage in training beyond the minimum licensure requirement to become board certified (ACGME, 2017; American Board of Medical Specialties, 2018; Eden et al., 2014); moreover, hospitals and other healthcare organizations are requiring it as "a condition of employment or practice privileges and by health insurers as a condition of physician enrollment" (Eden et al., 2014, p. 4-9).

ACGME. In 1981, the Accreditation Council for Graduate Medical Education (ACGME) was founded as an independent, non-profit organization assigned with the responsibility for review and accreditation of graduate medical education programs (ACGME, 2016a; Holmboe, Edgar & Hamstra, 2016; Nasca, Philibert & Brigham, 2012). According to the ACGME (2017), "The mission of the ACGME is to improve health care and population health by assessing and advancing the quality of resident physicians' education through accreditation" (ACGME, 2017, p.5). The overarching goal of residency training is to prepare each resident in six board competency domains: (1) patient care, (2) medical knowledge, (3) practice-based learning and improvement, (4) interpersonal and communication skills, (5) professionalism, and (6) systems-based practice (ACGME, 2016a; Holmboe et al., 2016).

Upon the conclusion of residency training, a resident can decide whether or not to pursue certification. The passing of the certification is designed to ensure that the resident has met the expected standards. As an example, the American Board of Internal Medicine website states, "certification means residents have demonstrated clinical judgment, skills, and attitudes essential for the delivery of excellent patient care" (American Board of Internal Medicine, 2019).

Residency. The term *residency* refers to the initial period of physician training required for board eligibility (ACGME, 2017). "Graduates of GME programs become eligible for board certification through specialty and subspecialty boards" (Eden et al., 2014, p. 4-9). According to the process established by the American Board of Medical Specialties (2018), a physician must demonstrate competence by way of (a) finishing the requisite premedical education; (b) earning a medical degree; (c) meeting licensure and procedural requirements; (d) completing a residency training program; (e) and finally, passing a rigorous knowledge assessment exam. Board certification is a designation granted by one or more of the specialty boards (Eden et al., 2014; Shaw, Cassel, Black, & Levinson. 2009) and is intended to assure the public that certified physicians have "knowledge, experience, and skills to provide quality healthcare within a given specialty" (American Board of Medical Specialties, 2018).

Competency-Based Medical Education. This section gives an overview of the competency-based medical education (CBME) framework and presents some definitions of competency. Given the impetus to improve physician qualifications in practice, and to validate its mission and objective existence, the ACGME has embarked on an initiative to refine its model of education using a competency-based approach. According to Frank et al. (2010), "the definition of competency-based medical education is highly variable in the literature" (p. 641). From the literature, Englander et al. (2017) stated referencing Frank et al. (2010), that competency-based medical education refers to "an outcomes-based approach to the design, implementation, assessment, and evaluation of medical education programs, using an organizing framework of competencies" (p. 584).

In the context of academic medicine, the description of *competency* commonly applied is defined as follows:

> Frank et al. (2010), an observable ability of a health professional related to a specific activity that integrates knowledge, skills, values, and attitudes. Since competencies are observable, they can be measured and assessed to ensure their acquisition. Competencies can be assembled like building blocks to facilitate progressive development. (As cited in Englander et al., 2017, p. 584)

However, for this research, the definition of *competency* will require more explanation to interpret these guiding principles operationally.

Although other characterizations of competency exist, this study highlights five distinct definitions. First, a measurement of an "observable activity," as described by Frank et al. (2010), operates to reflect the term *performance* (i.e., what is done) in a current observable moment. Guerra-López (2008) emphasizes that performance should focus on "the accomplishments of behavior rather than the behavior itself" (p. 25). Second, the organizing framework of competencies suggests that the term *competency* has most often been conceived as a representation of expectation or a standard for performance (i.e., what is expected to be done under standard conditions) now (i.e., in the moment) or in the future. Third, *competence* indicates an evaluative process of the behavioral tasks to determine the degree of provisional achievement in a contextual domain or area of study. Fourth, *competence* is developmental, impermanent, and evolutionary; therefore, competence is momentary (Epstein & Hundert, 2002; Leach, 2002). Lastly, the process of competency can broadly be summarized as an individual who can perform the competency-related activity to expectation is assumed momentarily *competent*. However, not synthesized in this outline of competency is the question—does competence involve more than observable behavior? Further, are there any consequences to the academic achievement of competency if the unobserved is overlooked?

As noted earlier by Frank et al. (2010), the integrated processes that guide behavior such as knowledge, skills, values, and attitudes—mostly related to unobservable latent constructs. For instance, we cannot observe what a person has learned, the values they hold, or the motivation they have toward the achievement of particular skills. Inherently, without a comprehensive account of more attributes that motivate the resident performance, the present behavioral task approach to assessment limits the depth of the analysis of achievement (i.e., acquisition) and overall depiction of what is accomplished. Supporting this view, Feather (1992) states the observable, "actions are constrained by beliefs" (p. 113), and therefore, it may be reasonable to consider *competency-related beliefs* of residents as a theoretical component of their achievement. Furthermore, this idea suggests a concept overlooked in competency-based medical education—*self-competence*.

Self-competence. Researchers have conceptualized and defined self-competence as an individual's perceived ability in broad academic areas or domains (Harter, 1982). Tafarodi and Swann Jr, (1995), explain the concept of self-competence as "the overall sense of oneself as capable, effective, and in control" (p.325). Self-competence is related to the broader concepts of self-efficacy, and competency-related beliefs, which stem from self-perceptions, and has been studied for its motivating effects in persistent learning behavior. A void in academic medical literature exists about the conceptual definition of self-competence and its role in achievement; however, it can be compared with other related constructs in psychology and other related fields.

Based on previous literature in the field of psychology, self-competence has most often been regarded in connection with self-efficacy (Tafarodi & Swann Jr, 1995), and appears as a central concept in social learning theory, and theories related to motivation (Harter, 1982). Bandura (1986) defined *self-efficacy* as:

> People's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances. It is concerned not with the skills one has, but with the judgments of what one can do with whatever skills one possesses. (p. 391)

Consistent with the previous definition, another similar example can be drawn related to competency-related beliefs. Achievement motivation theorist argues that student judgments can partially be explained by their perception of needs concerning the contextual demands to achieve goals. Nevid (2013) describes achievement motivation as "the desire to achieve success" (p. 292). More specifically, in the expectancy-value theory, as investigated by Eccles and Wigfield (2002), competency-related beliefs are designated as individual perceptions of current competence (i.e., ability belief) and expectancies. The competency-related belief of this study intends to correspond with Eccles and Wigfield's (2002) definition of *expectancy* as “beliefs about how well students will do on upcoming tasks, either in the immediate or longer-term future” (p. 70).

Empirical studies of the expectancy-value theory—as noted by Cook and Artino (2016), “nearly all of them outside of medical education” (p. 1003)—show that competency-related beliefs and values directly influence both engagement in learning activities and learning achievement (Eccles & Wigfield, 2002). More than a half-century of research on achievement motivation has demonstrated that competency-related beliefs and goal-directed behaviors are fundamental factors in academic performance (Atkinson, 1957; Wigfield, Eccles, Schiefele, Roseser, & Davis-Kean, 2006; Conley, 2012).

The assessment and evaluation of a resident’s perceived competence are essential to program-evaluation research and graduate medical education. However, there is little, if any, empirical research available to answer questions about the competency-related beliefs or achievement motivation of residents. At the practical level, it is reasonable to ask— does the accuracy of self-evaluation in an internal medicine resident population factor into their academic achievement of medical knowledge? How do overestimations and underestimations of competency-related beliefs relate to learning outcomes? How do estimator errors develop in residents?

Prerequisites Components and Underlying Factors

The next step in understanding the ability to self-evaluate is to consider a few of the multiple factors that influence the self-regulation of learning behavior and performance outcomes. These components, which contribute to an individual’s academic performance in other studies, include metacognition, self-awareness, and feedback.

Metacognition. Based on the available literature, an underlying problem with self-evaluation may involve poorly developed metacognitive skills. Metacognition denotes a learner's knowledge of their own knowledge and abilities (Colbert et al., 2015; Kruger & Dunning, 1990). Tulving and Madigan (1970) refer to this capability as a "truly unique characteristic of human memory; its knowledge of its own knowledge" (p. 477).

After reviewing the literature, the researcher found that improving the capacity of residents to self-assess their competence is of particular interest in medical education (Bing-You et al., 2009; El Saadawi et al., 2010; Thurlings, Vermeulen, Bastiaens, & Stijnen, 2013). As suggested by Colbert et al. (2015), teaching metacognitive skills might enable cognitive correction necessary for practice-based learning and improvement of patient care outcomes. Furthermore, several publications advocate that metacognitive skills are essential to the acquisition of knowledge and reducing cognitive errors in the practice of medicine (Croskerry, 2003; Redelmeier et al., 2005; Graber et at., 2005; Mamede et al., 2010). Garrett, Alman, Gardner & Born (2007) proposed that the label of medical knowledge deficiency may relate to deficits in metacognitive skills, resulting from a failure to self-regulate information adequately and causing a breakdown to adapt and learn. Self-awareness in research is frequently explained in the same terms as metacognition.

Self-awareness. Stephen Covey (2013) states that "Self-awareness enables us to stand apart and examine even the way we 'see' ourselves — our self-paradigm, the most fundamental paradigm of effectiveness. It affects not only our attitudes and behaviors but also how we see other people. It becomes our map of the basic nature of mankind" (p.74). Self-evaluation and feedback are essential to initiating the self-awareness process.

In research, self-awareness is a prerequisite to the modification of knowledge or behavior (e.g., motivation, self-regulation) (Kruger & Dunning, 1999). Research into feedback and metacognition studied in combination suggests insight into the dissonance between information sources might provide the impetus to activate learning (Ende, 1983; Kluger & DeNisi, 1996). The definition of *learning* refers to an enduring change, resulting from practice or experience (Medin et al., 2001; Salmoni, Schmidt & Walter, 1984). Feedback can function as a relevant stimulus to a metacognitive knowledge resulting in memory change (Questienne, Van Opstal, van Dijck & Gevers 2016).

In particular, in conflict situations where feedback generates a reassessment of our prior understanding, a control component of metacognition is thought to enable the selection of stimulus features potentially leading to an ability to update our working memory and, therefore, our ability to adapt our knowledge (Shimamura, 2000). As cited in Kluger and DeNisi (1996), "in control theory, when discrepancy (in feedback) is noted, people are motived to deduce it" (p.259). However, without feedback, a problem can occur with motivating a metacognitive reaction and, therefore, our ability to adapt (Questienne et al., 2016; Shimamura, 2000).

Furthermore, using the cognitive psychology perspective, the conflict monitoring theory states that the presentation of a discrepancy in information is sufficient by itself to be detected by the response level directing attention to the adaptation (change) mechanisms in memory (Carter & Van Veen, 2007; Questienne et al., 2016). Current findings suggest a metacognitive element could mediate the process of informing the adapting system (Desender, Van Opstal & Van den Bussche 2014; Questienne et al., 2016). Feedback has been used effectively as a metacognitive scaffold intended to influence cognitive gains in medical education (El Saadawi, 2010). Hence, it appears from the literature that feedback may act as a stimulus to bring to mind metacognition, and together, metacognition appears to delineate our ability to adapt our knowledge.

Emotional Intelligence and Self-awareness. Goleman (1995) described that a person's cognitive intelligence accounts for only 20% of their career success, leaving a wide range of other factors to consider that correspond to emotional intelligence. Several studies have demonstrated that emotional intelligence can be increased through a blend of awareness, training and practice (Goleman, 1995; Nelis, Quoidbach, Mikolajczak & Hansenne, 2009; Webb, Young & Baumer, 2010) and others have found a significant and positive association with academic achievement (Mohzan, Hassan & Halil, 2013). Furthermore, Goleman (1995) refers to self-awareness as an essential component of emotional intelligence, the ability to assess one's own competence for strengths and limitations (Kreitner, 2009). Furthermore, self-awareness promotes self-regulation, which facilitates harmonious interactions with others once people learn to adapt to their surroundings (Gailliot, Mead, & Baumeister, 2008); such adaptation is essential to establish and maintain therapeutic relations with patients and work effectively as a member of the healthcare team.

Feedback. Feedback is a basic need frequently acknowledged as an area for improvement in medical education (Bing-You et al., 2009; Branch & Paranjape, 2002; Ende, 1983; Van de Ridder, Stokking, McGaghie & Ten Cate, 2008). Medical students and residents often complain about not receiving enough feedback (Bing-You et al., 2009). Van de Ridder et al. (2008) compared the definition of feedback found in 36 sources of literature and discovered three dominant concepts: feedback as information; feedback as a reaction where information is given; and feedback as a cycle between information and reaction (p. 191). The importance of feedback in the acquisition of knowledge is essential to the modification of our understanding. The consequences of inadequate feedback are associated with cognitive errors revealed in the missed diagnosis and diagnostic errors found in the practice of medicine (Graber at al., 2002; Croskerry, 2003).

Some contributing causes of the feedback problem include findings that suggest physicians, as a group of professionals, have only a modest capacity to self-assess their knowledge (Bing-You et al., 2009; Davis et al. 2006). In general, people are often overconfident in the accuracy of their own knowledge (Kornell & Bjork, 2009), but in medicine, this can have serious implications for patient care decisions, including a failure to consider alternative information. Also, research implies that at least 40 types of cognitive bias may affect clinical reasoning (Mamede et al. 2010). Other evidence indicates that the ability of faculty members to evaluate medical knowledge of their residents can be unreliable (Holmboe & Hawkins, 1998; Hawkins, Sumption, Gaglione & Holmboe, 1999; Jones, Panda & Desbien 2008). Additionally, Dine, Ruffolo, Lapin, Shea & Kogan (2014) indicate that implementing patient feedback as a mechanism can be a barrier because it requires large numbers of responses to make judgments.

Finally, the researcher believes what might be needed is beyond the everyday objectively measured performance, but a feedback mechanism that reveals the predisposition of the resident to the empirical standard. It is postulated that the use of a complementary subjective assessment, as the study survey, in combination with the objective resource, like the IM-ITE, will reveal cognitive bias and estimation errors. The problem is—without the external feedback stimuli—learners are not intrinsically motivated to critically examine or modify their knowledge. This researcher considers estimation errors as a mediating variable in the intervening process that affects knowledge performance and the ability to learn. Research suggests that insight into the dissonance between information sources might provide the impetus to activate learning (Ende, 1984; Kluger & DeNisi, 1996). Furthermore, research conducted by Salmoni (1984) indicates, "When no sources of information about errors are present, no learning or at least no change in performance is evident" (p. 361). Additionally, Garrett et al. (2007) propose that the label of medical knowledge deficiency may actually relate to deficits in metacognitive skills, resulting from a failure to self-regulate information adequately, causing an inability to adapt and learn.

In-Training Examination and Empirical Research

Making improvements to feedback can be difficult without a better understanding of contributing factors and relevant instrumentation. Many program directors do not receive specialized training in administering performance tests to assess the medical knowledge of resident progress and have difficulty recognizing training needs (Powell & Carraccio, 2018). The medical knowledge domain is well supported with appropriate subject matter experts, but program directors and other faculty often lack expertise in creating, identifying, and aligning academic performance metrics (Anderson, 2012). According to Guerra-Lopez (2008), the purpose of performance measurement is to: (1) compare the results with expectations, (2) discover the barriers to expected performance, and (3) use the information to inform decision-making that improves the achievement of performance (p. 6). A well-defined performance measurement instrument commonly used in internal medicine residency training is the American College of Physicians (ACP) Internal Medicine In-Training Examination (IM-ITE).

Internal Medicine In-Training Examination (IM-ITE). According to the American College of Physician's (2019) website, the self-evaluation is administered online annually and requiring 9 hours with 7 hours devoted to the exam. The exam is composed of roughly 300 multiple-choice questions in 12 medical content areas: cardiology, endocrinology, gastroenterology, general internal medicine, geriatric medicine, hematology/oncology, infectious diseases, nephrology, neurology, pulmonary and critical care medicine, rheumatology, and high-value care (American College of Physicians, 2019). The examination is intended to evaluate resident knowledge, diagnostic reasoning, ordering and interpreting of testing results, and recommendations for treatment in clinical scenarios applicable to patient care. According to the ACP website, “The blueprint for the IM-ITE is modeled after the blueprint for the American Board of Internal Medicine’s certification exam” (American College of Physicians, 2019).

The results of the IM-ITE are intended to provide objective formative feedback to assist both educators and residents with insight and awareness into areas of training need. However, this performance measurement single-handedly does not offer input into other aspects of human judgment like cognitive dissonance or predisposition orientations toward content. No materials exist that jointly study the competency-related ability beliefs and estimation errors of residents associated with this examination. The failure to appraise the development of resident competency-related beliefs negates our capacity to study the integral part of their self-identified training needs, accuracy in their judgment, and the impact of estimator errors along with achievement motivation and actual IM-ITE performance.

Although the Internal Medicine In-Training Examination has experienced developmental changes over the years, there are previous studies that emphasize the potential impact and importance of such an instrument. For example, Garibaldi, Subhiyah, Moore & Waxman (1994), a statistical analysis of the IM-ITE, demonstrated the assessment to be reliable, internally consistent, and discriminating. The study suggested the IM-ITE be a useful instrument to assess the knowledge base of the resident during internal medicine training. Further, Garibaldi et al. (2002) analyzed the longitudinal results from the first 12 years of administration (1998-2000) for a total of 13 exams and discovered that more than 80% of residents take the exam annually. Also, the researchers stated that exam scores increased approximately by 5% with each year of training, which is substantial in understanding growth rates over time. In addition, the study found that graduates from international medical schools have scored higher on average than graduates from U.S. medical schools on every examination at every post-graduate year (PGY) level.

Predictive Quality of IM-ITE. Research studies over the years have suggested a predictive relationship between the IM-ITE and the American Board of Internal Medicine Certification Exam (ABIM-CE). A study titled, "Validity of the In-Training Examination for Predicting American Board of Internal Medicine Certifying Examination Score," Grossman et al. (1992), looked at the IM-ITE scores of 109 residents from a combination of six internal medicine programs and the subsequent performance on the ABIM-CE. Grossman et al. (1992), concluded the IM-ITE could be used to predict performance on the ABIM-CE accurately. Most importantly, the researchers, empirically derived a cutoff score of the ≥35th percentile, using a receiver operating characteristic (ROC) curve, which amplified the ability of the IM-ITE to discriminate between residents that were expected to pass and those who were likely to fail the ABIM-CE. As a result of these findings, the data presented residency training programs with a way to pointedly detect struggling residents.

Sequentially, Waxman et al. (1994), conducted a study involving 223 residents from eight teaching hospitals, constructed a receiver operating characteristic (ROC) curve was constructed and presented findings that the IM-ITE highly correlated with ABIM-CE pass rates. The researchers recommended that the results serve residents as an essential measure of their preparedness during training. Besides, the additional validity of the predictive relationship, the research suggested the use of a ≥ 40th percentile cutoff score.

Prior research primarily focused on IM-ITE percentile scores and receiver operating characteristic curves. Rollins, Martindale, Edmond, Manser, and Scheld (1998) presented an alternative mode of analysis to predict pass rates on the ABIM-CE using the percentage correct scores from the IM-ITE, a sample size of 155 internal medicine residents were selected from three residency programs in Virginia. A logistic regression model was constructed to calculate a regression coefficient for the IM-ITE percentage correct scores to predict pass or fail outcomes on the ABIM-CE. The effects of the study further demonstrated high predictability between IM-ITE and ABIM-CE. Specific findings of the study suggested benchmarks that IM-ITE scores above 66% correctly predicted a specific pass, while scores below 49% correctly predicted inevitable failure.

Knowledge Assessment in Medical Residents. Hawkins, Sumption, Gaglione, and Holmboe (1999), investigated resident perceptions regarding IM-ITE in connection with the ability of faculty members to evaluate the knowledge of their residents. The methods of the study involved asking residents about the perceived utility of the IM-ITE and to predict their own performance into upper, middle, or lower percentile classes. Additionally, faculty predicted resident percentile classes of PGY 2 and 3 while residents (PGY 2 and 3) predicted the scores of PGY 1 interns. The study concluded that 97% of residents perceived the IM-ITE to be useful, with 91% stating that they modified their study habits or clinical rotations schedule based on its results.

Furthermore, the study found that nearly half of the residents accurately predicted their percentile class. However, faculty predicted resident performance, precisely 49% of the time. Resident PGY 2 and 3 predicted PGY 1 intern scores 38% accurately. Also, both faculty and residents were more likely to overestimate than underestimate percentile rankings. Most importantly, the study demonstrated a lack of correlation between the ability to assess medical knowledge by raters.

In a study conducted by Jones, Panda, and Desbien (2008), findings further confirmed that internal medicine residents do not accurately assess their own medical knowledge. In the study, residents predicted their overall percentile performance before and after taking the IM-ITE. According to the study, of 26 residents who participated in the survey, 31% had IM-ITE scores that were within 10 points of their predictions. Also, most were pessimistic about their overall performance, with 69% underestimating their performance. This finding differs from the results of Hawkins et al. (1999), in the earlier study, but this might be accounted for by the use of different predictive class intervals between reviews.

Holmboe and Hawkins (1998) reviewed methods for evaluating the clinical competence of residents and the relative effectiveness of tools to measure specific elements of clinical expertise. The article discussed significant instruments available to educators: (a) medical record audit; (b) IM-ITE; (c) ABIM rating form; (d) clinical evaluation exercise; and (e) standardize patients concerning their assessment framework (e.g., medical knowledge, clinical judgment, communications skills). Corresponding to the study, clinical competence is complex, and no single evaluation tool can adequately assess a resident's knowledge, skills, and attitudes. Most importantly, the review pointed out that successful completion of the ABIM-CE is not an adequate measure of overall clinical competence because it only measures medical knowledge.

Theoretical Background Relevant to the Study

Environmental shaping. B.F. Skinner is distinguished for his theory of operant conditioning; how the surrounding environment controls behavior. He maintained that behavior is determined by reinforcement, like feedback, where the reinforcer is designed to strengthen the connection between the desired response (Medin, Ross & Markman, 2001); he contends that behavior is a product of environmental shaping. He argued that the science of practice consists of connections between patterns of reinforcement and behavioral responses or systematic behavior modification. In this theory, learning is defined by a change in behavior that is shaped through reinforcement (Medin, Ross & Markman, 2001; Skinner, 1954; Skinner, 1958; Spector & Yuen, 2016). As stated by Skinner (1954), "Once we have arranged the particular type of consequence called a reinforcement, our techniques permit us to shape the behavior of an organism almost at will" (p. 87). Kreitner (2009) summarized that the theory "involves managing environmental factors to get people to do the right things more often and the wrong things less often" (p 416). Behavior modification is the way toward accomplishing and maintaining goals in a harmonious relationship between individuals and the environment.

Reinforcement and feedback is also an essential component of several learning theories where learning is defined as a process of adaption or change (Little & Erickson, 2015; Skinner, 1954). Research into feedback suggests that insight into the discrepancy between our understanding relative to environmental or external information might provide the impetus to activate the learning (Ende, 1983; Kluger & DeNisi, 1996). Many motivation theories revolve around the discrepancy of needs or goal orientation. Feedback can evoke corrective action needed to resolve cognitive errors through the reassessment of our knowledge comparatively to the environmental or external information. Therefore, feedback can result in an elevated awareness that encourages self-assessment, change, and appropriate direction for a change. In other words, metacognition appears to delineate our ability to adapt our knowledge.

Motivational orientation. The literature has shown that a wide range of different factors causes behavior. The majority of learning experiences in graduate medical education are experiential and often unplanned. Therefore, the amount of time devoted to mastery learning is self-directed, and studying the motivational orientations of individuals to learn is critical to medical educators and program-evaluation researchers.

Research describes that behavior usually results from a motivational process involving a need or desire, and this discrepancy urges the action toward goal attainment (Grant, 2008). Motivation is a central element of most human activity. Deci, Ryan & Williams (1996) state, “To be motivated means to behave with the intention of achieving some outcome” (p. 166). Motivation differs from behavior; motivation is a psychological state; the outcome or result of that state is behavior (Nevid, 2013). Pinder (1998; 2014) described motivation as “the energy a person expends in relation to work” (p.1), and understanding goal orientation is a mechanism to understanding the intensity and direction of motivation.

Goals are seen as significant drivers to attention and action. The locus for motivation can be internal (“self-motivated”) or provided by the external environment. Deci, Ryan & Williams (1996) defined intrinsically motivated behaviors as performed freely out of self-interest or inherent satisfaction. Extrinsic motivation is defined as the desire to perform actions with the intent to obtain outcomes external to the effort itself, such as rewards or recognition (Amabile, 1993; Deci, Ryan & Williams, 1996; Grant, 2008).

Goals are a significant factor in motivation research (Ryan & Deci, 2008). Self-determination theory considers the way people interpret internal or external stimulus feedback for meaning to direct the fulfillment of needs (Deci & Ryan 1985, 1991, 2008). The theory aligns levels of self-regulation with motivation as a form of action control. Goals lead to actions. The process by which a goal is realized is self-regulation; self-regulation is the process of reaching one’s goals (Credé, & Phillips, 2011; Ryan, & Deci, 2008).

Literature indicates that congruence of motivation orientation has influences on performance. Goal difficulty and importance are associated with intensity. The expectancy theory is based on the assumption that individual motivation strength is determined by whether or not a person believes they can be successful at a task (Atkinson, 1957; Bandura, 1986; Eccles & Wigfield, 2002; Vroom, 1964). People tend to work harder shaped by the apparent conceivable outcomes of accomplishment. Perception is an essential element of this theory. Research conducted found a link between expectations and achievement (Bandura, 1986).

Self-determination theory. According to self-determination theorists, the concept of motivation is guided by three basic psychological needs that we all share. First, people everywhere are engaged in the process of relating to or making sense of our environment—referred to as relatedness. Deci & Ryan (1985) state, "that a basic need for interpersonal relatedness explains why people turn external goals into internal goals through internalization" (Eccles & Wigfield, 2002, p. 113). Second, we are motivated by the inner desire to improve or be competent (i.e., competence). "Competence is a psychological motive that both organizes daily experience and shapes our self-concept" (Conroy, 2017, p. 25). Lastly, by way of competence, we seek to gain some control over our environment—referred to as autonomy (Connell & Wellborn, 1991; Ryan & Deci, 2000). As stated by the ACGME (2019a), "residency programs provide the clinical experience and education to gradually and progressively achieve autonomy to deliver the highest quality patient care without supervision." Thus, learning is an active process where people seek out favorable stimulation and engagement in challenging learning activities which indicate their desired end-states or goals because people need relatedness, competence, and autonomy (Wigfield & Eccles, 2000; Ryan, & Deci, 2008).

Theoretical Foundations for the Study

To accomplish the purposes of the study, the researcher has assembled a collection of suitable theoretical foundations from the literature that conceptualize self-evaluation and achievement motivation constructs. A substantial body of research establishes the importance of motivation to performance outcomes (Atkinson, 1957; Wigfield, Eccles, Schiefele, Roseser & Davis-Kean, 2006; Conley, 2012). Motivation is a behavioral component of most activities, and "motives are the whys of behavior, the needs or wants that drive behavior and explain what we do" (Nevid, 2013, p. 288). A study of the motivational processes is required to make inferences about the unobserved patterns in resident self-evaluations and competency-related beliefs that may explain how typical behavioral patterns are formed to reach academic goals.

Expectancy-value theory. The contemporary understanding of expectancy-value theory is primarily based on the seminal research conducted by Eccles and Wigfield (Eccles et al. 1983; Wigfield & Eccles, 2000). Within this framework of thinking, motivation is a product of competency-related beliefs and subjective task values (McCoach, Gable & Madura, 2013). The theory is one of the most influential models that has been used to examine achievement motivation. Previous research demonstrated that competency-related beliefs mediate learner motivation and competence (Eccles & Wigfield, 2002). These factors are particularly crucial to predicting these three specific behavioral modes: (1) the intensity of effort; (2) the direction of attention (e.g., the source of gratification, academic choice); and (3) the persistence of effort over time (Eccles & Wigfield, 2002). The most extensive data on expectancy-value theory comes from studies of children. As of yet, there is no known comparative research in this area related to a resident population or graduate medical education context.

Theoretical Constructs for the Study. After reviewing the literature, the researcher found that the relationship between self-evaluation, estimation errors, and academic achievement of medical knowledge is less clear and determined it was necessary to develop a framework based on previous research. Given the predictive nature of competence-related beliefs to indicate performance outcomes, the expectancy-value theory research by Wigfield and Eccles provides a clear set of structural parameters that will guide the operationalization of this study.

Competence-related beliefs. Wigfield and Eccles' (2000) expectancy-value theory provides two frames of reference related to the timing of competence-related beliefs: (1) ability beliefs and (2) expectancies for success. Wigfield and Eccles (2000) defined *ability beliefs* as self-evaluation of the "individual's perception of his or her current competence at a given activity" (p. 70). Wigfield and Eccles (2000), explained *expectancies for success* as "beliefs about how well students will do on upcoming tasks, either in the immediate or longer-term future" (p. 70). Wigfield and Eccles (2000) note, "ability beliefs thus are distinguished conceptually from expectancies for success, with ability beliefs focused on present ability, and expectancies focused on the future" (p. 70).

For this study, the prediction for performance, as a competency-related belief, is intended to function similarly to the expectancy construct of the expectancy-value theory. The author of this paper uses the terminologies of prediction and estimation interchangeably to represent an equivalent expectancy construct. However, the *expectancy* term is not used exclusively in this report, because of how it operates differently from Wigfield and Eccles (2000) definition; the prediction for performance will occur after the participant completes the task (i.e., takes the IM-ITE). Nonetheless, the resident will be unaware of their actual IM-ITE score at the time; therefore, the belief construct is focused on the future expected outcome.

Academic achievement of medical knowledge. The IM-ITE is a multiple-choice examination used to assess medical knowledge in internal medicine residents. Consequently, the actual performance on the IM-ITE embodies a level of achievement for medical knowledge.

Dimensions. Competency-related beliefs are an affective attribute of residents. Anderson and Bourke (2000) state that there are three critical dimensions to all affective characteristics: intensity, direction, and the target. Gable and Wolf (1993) provides these definitions for the dimensions: (a) the *intensity* dimension is "the degree or strength of the feeling;" (b) whereas, the *direction* attribute "reflects the positive, neutral, or negative aspect of the feeling;" and finally, (c) the *target* "identifies the object, behavior, or idea at which the feeling is being directed" (p. 4). Further, these dimensions appear comparable to some of the achievement modes of behavior: (a) the intensity of effort or (b) the direction of attention (Eccles & Wigfield, 2002; Mitchell & Daniels, 2003; Schunk, 2012).

*Estimation Errors***.** The literature offers fairly consistent definitions for measurement error in education and psychology. Although terms like *calibration* and *bias* appear in the literature, they broadly apply to the same intensity and direction dimensions previously reported. As a result, the researcher acknowledges and provides some definitions from the literature for reference below, but will denote these errors by their dimensions and as dimensions of estimation error throughout the remainder of this report.

Chen (2003), "*calibration* is a measure of the accuracy of metacognitive monitoring in terms of congruence between one's perceptions of competence about performing a particular task and one's actual performance" (p. 80). Moreover, Chen (2003), the accuracy of calibration is measured as the intensity or magnitude of the error. The term and measurement of *bias* in the literature are often determined by the difference of expected (i.e., estimated) value and the actual value to determine the direction of judgment error (Chen, 2003).

Summary

Based on the literature review, the ideal course of action is to provide medical educators, program-evaluation researchers, and, most of all, program directors with empirical evidence that is generalizable and applicable to their educational practice. As Anderson (2012) states, "… most teaching in the clinical setting is carried out by clinicians who are not trained educators; it is perhaps not surprising that the translation from theory and research to practice has not occurred on a large scale" (p. 154). There remain outstanding questions about the significance of self-evaluation, and estimation errors in a resident population; thus, a research study will seek to provide more insight. The findings from this study have the potential to benefit the alignment and identification of resident training needs and their motivational orientations toward the achievement of medical knowledge.

The research was implications for educational practice because the study intends to build on and extend the theoretical foundations of existing research through new evidence. The benefits of this study will provide more comprehension into the effects of self-evaluation, competency-related beliefs, and whether the impact of estimation errors requires further consideration in program evaluation methods and instructional design. An understanding of how estimation errors develop in residents will provide an opportunity to clarify research questions through the conveyance of study findings. The central task is to provide empirical findings into the status of estimation errors and its relationship to performance so that the learning environment can consider and, if necessary, provide the instructional support that might improve the academic performance of residents.

CHAPTER 3: METHODOLOGY

Introduction

This chapter outlines the research design and reviews the methodological criteria applied in the study. A description of the study population and sample, instrumentation and measures, data collection process, and data analysis is provided.

The type of study design is quantitative, cross-sectional survey research using a non-random sample of convenience. The survey is designed to identify competency-related beliefs of residents as a prediction or estimation of their performance in conjunction with actual performance on the Internal Medicine In-Training Examination (IM-ITE). The design is intended to gather sample data from participants that represent the graduate medical education internal medicine resident population. The primary sources for data collection will be a survey administered online (i.e., subjective assessment), and the Internal Medicine In-Training Examination (i.e., objective resource) results. Some independent variables will focus on specific demographic characteristics—population features of residents (e.g., post-graduate year and gender).

The main objective of this study was to develop a better understanding of how residents self-evaluate and to address the research question—what, if any, influence does this have on their academic achievement of medical knowledge? The face validity of this research will show that insight into self-evaluation, including estimation errors, is essential to understanding academic achievement. This study intends to explore the nature of predictions for performance, as a type of competency-related belief, in comparison to the actual IM-ITE results in order to provide empirical data for the analysis of self-evaluation and its effects. This study has two main goals: (1) to investigate self-evaluation, estimation errors of residents with the objective measures of the IM-ITE associated with medical knowledge achievement, and (2) to build on and extend the theoretical foundations of motivation theory research to enable educators to think systematically about achievement motivation in graduate medical education. This study will be an essential step toward recognizing resident competency-related beliefs, estimation errors, and their potential impact on achievement motivation and learning outcomes.

Population and Sample

The researcher is currently employed by Henry Ford Health System and is a doctorate candidate at Wayne State University in the Learning, Design, and Technology program. Before conducting this research, institutional review board (IRB) approval was obtained from both Henry Ford Hospital and Wayne State University.

Setting characteristics. The data collection will be conducted at Henry Ford Hospital, an 877-bed acute care hospital in Detroit, Michigan. Henry Ford Hospital, recognized by the Joint Commission and is considered a general/teaching hospital. The Center for Medicare and Medicaid Services (CMS, 2019) defines a *teaching hospital* as “hospitals that receive payment for Medicare direct graduate medical education (GME), IPPS indirect medical education (IME), or psychiatric hospital IME programs during the last calendar year for which such information is available.” As of this writing, Henry Ford Hospital was a sponsoring institute of 51 total active ACGME-accredited residency and fellowship training programs. Of the training programs, 37.3% (19) were residency programs, and 62.7% (32) were fellowship programs. Henry Ford Hospital is designated as a Level 1 trauma center. The hospital is part of the more extensive Henry Ford Health System that consists of an integrated network with five regional hospitals. Furthermore, Henry Ford Hospital is affiliated with the Wayne State University School of Medicine.

Target population. According to the ACGME (2018b), the number of ACGME-accredited internal medicine training programs in the United States and Canada totals 529 with a population of 27,647 residents. The mean number of residents per program by internal medicine specialty is, on average, 52.3 residents. This mean is an important number for this study if inferences are to be made in this study regarding a typical internal medicine program. Internal medicine represents the largest active resident population by training specialty (20.4%), and over five years (2013 to 2018) has increased by 15.7%. The mean age of post-graduate year (PGY) 1 residents within the internal medicine specialty is, on average, 29.5 years old. The reported genders include 11,042 (39.9%) females and 14,979 (54.2%) males with 1,626 (5.9%) not reporting (ACGME, 2018b).

Sample Size. In nonprobability sampling, the degree to which the sample differs from the population is typically unknown. Therefore, the aim is to make statistical inferences with an acceptable level of confidence. Given a target United States and Canada total population of 27,647 internal medicine residents (ACGME, 2018b), a confidence level of 95%, and confidence interval of 10.94 a sample size of 80 residents would be required representing a 72.1% (80 out of 111) response rate of Henry Ford Hospital internal medicine residents.

Sampling Procedures. In view of the sampling procedures, the study will use a non-random (nonprobability) convenience sample method to establish an approximation of the results typical to exploratory research. The sample intends to reflect the characteristics of the study population previously identified. Completion of the online survey indicates voluntary participation in this research project as a subject.

Data Collection Methods

The researcher has taken intentional and strategic steps to create a study to assess self-evaluation skills following an objective operational standard to fulfill the need to make the research relevant and meaningful to the internal medicine resident population. The IM-ITE is the first source of data for this study and was administer independent of this research, followed by an online survey, which hereafter will be termed the *Subjective Assessment Survey (SAS)*.

Instrumentation. Two instruments will be used in this study: (1) the Internal Medicine In-Training Examination (IM-ITE), and (2) the Subjective Assessment Survey (See Appendix A).

Internal Medicine In-Training Examination. The Internal Medicine In-Training Examination (IM-ITE) is a well-established instrument that has been the subject of several publications. The American College of Physicians (2019) describes the IM-ITE as a self-evaluation of roughly 300 medical knowledge questions within 12 content areas: cardiology, endocrinology, gastroenterology, general internal medicine, geriatric medicine, hematology/oncology, infectious diseases, nephrology, neurology, pulmonary and critical care medicine, rheumatology, and high-value care.

Subjective Assessment Survey (SAS). The study survey is a 12-item instrument designed to measure the competency-related beliefs of residents. The 12 items of measurement directly corresponded to the 12 medical content areas of the IM-ITE as previously listed. The scoring requests that each participant provides an estimate of their performance as a percentage correct scale for each of the 12 medical content areas (See Appendix A).

Data Collection Procedures. In 2019, the Internal Medicine In-Training Examination (IM-ITE) was scheduled to be administrated from Thursday, August 22, 2019, to Wednesday, September 11, 2019, except Labor Day (Monday, September 2, 2019). The sample frame for the Subjective Assessment Survey (SAS) was intended to begin data collection shortly after the conclusion of the IM-ITE. The rationale for the SAS following the IM-ITE was concerning the first post-graduate year (PGY 1) residents having no prior experience with the examination. Therefore, the idea was to allow each level of training an opportunity to take the exam and then reflect on their performance afterward. Further, the results of the IM-ITE were unknown at this point in time, according to the American College of Physicians (2019) website, the IM-ITE results are typically available “roughly 4 to 6 weeks after the exam window”. Thus, the data collection using the SAS takes place in this window between the exam and results while the resident is unaware of their actual performance.

The Internal Review Board (IRB) process was completed with both the Henry Ford Health System and Wayne State University and approval received on September 18, 2019. On the following day, Thursday, September 19, 2019, one-hundred and eleven internal medicine residents at Henry Ford Hospital, who had taken the IM-ITE, were invited to participate in the online survey.

An email invitation was sent to each resident informing them of the study purpose and asking for their consent to participate by completing an online 12-item survey asking them to predict their overall percentage correct for each medical content area of the IM-ITE. The deadline to complete the survey was October 18, 2019, which provided nearly 4-weeks to participate (See Appendix B).

The online Subjective Assessment Survey was conducted within the Henry Ford Health System intranet behind the firewall; the survey and data collected were not visible outside the network. The individual email addresses, resident names, training level (i.e., post-graduate year), and gender were not received by the survey but were known to the researcher and stored in an encrypted desktop database offline. A link enclosed within the email invitation contained a unique coded identifier that allows the researcher to merge individual survey responses to other data sources later; the known demographic information (e.g., training level and gender) and to the IM-ITE results. The online survey responses were collected, stored only with the predicted performance scores and the unique coded identifier. Additionally, the identifier functioned to prevent duplicate submissions from the same participant. After the results of the IM-ITE were made available to the researcher and the data merged into a single table for analysis, the individual participants coded identifiers were removed and permanently deleted, rendering the data anonymous.

Variables and Measures

Based on the available literature, researchers have conceptualized and defined the constructs of this research in various ways. The following descriptive information is provided to explain how each variable will be measured in this study.

Actual performance. The actual performance measurement is obtained from the exact empirical Internal Medicine In-Training Exam (IM-ITE) values. The IM-ITE functions as the objective operational standard for this study. There were two principal IM-ITE values used for analysis in this study: (1) the overall percentage correct and (2) the percentage correct for each of the 12 medical content areas. The actual performance will be a dependent variable when examined in consideration of the demographic characteristics of residents.

Academic achievement of medical knowledge. The IM-ITE is designed to evaluate the expected medical knowledge, diagnostic reasoning, and clinical judgment skills of a certified internist (American College of Physicians, 2019). Likewise, the actual performance measurements, as previously mentioned, also embody a level of achievement for medical knowledge.

Predicted performance. The predicted performance measurement is obtained from the self-reported evaluation of residents' competency-related beliefs or predictions for performance on the IM-ITE. The Subjective Assessment Survey (SAS) instrument functions to collect the resident responses as they self-rate their performance for each of the 12 medical content areas identified by the IM-ITE. These competency-related beliefs predominantly correspond with the expectancy construct provided by Eccles and Wigfield (2002). The scale for this variable is calibrated deliberately to match the units of measurement for the actual IM-ITE performance—a percentage correct. The predicted performance will be a dependent variable when examined in consideration of the demographic characteristics of residents.

Estimation errors. The measurement error characteristics that stem from the estimator's ability to self-evaluate are exhibited by the subsequent comparison of predicted and actual performance values. The aspects of estimation error in the self-evaluation process has three critical dimensions: (1) intensity, (2) direction, and (3) target.

Estimator error-index. The estimator error-index is an integer that represents the quantitative difference between the predicted performance and actual performance as an interval value. The estimator error-index value operates to indicate the direction and intensity dimensions as a distinct measure. The estimator error-index will be a dependent variable when examined in consideration of the demographic characteristics of residents.

Overestimation. An overestimation is a type of measurement error related to the directional dimension, where the prediction performance value is higher than the actual performance value.

Underestimation. An underestimation is a type of measurement error related to the directional dimension where the prediction performance value is lower than the actual performance value.

Demographic characteristics of residents. For this study, two common resident attributes will be considered as variables to examine for differences in response patterns: (1) training level (i.e., post-graduate year) and (2) gender. The demographic characteristics of residents will be independent variables.

Data Analysis Methods

The quantitative research involves the empirical investigation of measurable variables. The principal components of the analysis will be performed based on IM-ITE results, the Subjective Assessment Survey, and categorical demographic groups. Data analysis, and inferential testing, as described below, will be conducted using IBM SPSS Statistics Version 25.0 for Windows.

Research questions and data analysis methods. The measures of the study will be analyzed to provide answers to the following research questions using the described statistical methods:

1. How well do residents predict their actual performance on the IM-ITE? The research will start with an examination of how well or how accurate residents are at self-evaluation with a comparison for differences between the predicted and actual group means for overall percentage correct using a matched-pair T-Test at the two-tailed .05 level of significance. A secondary analysis will be conducted to examine—is there a relationship between the predicted and actual overall percentage correct scores? A bivariate analysis of linear correlation (i.e., Pearson product-moment correlation), with a .05 level of significance, will be used to measure the degree of this relationship.
2. How does the actual performance on the IM-ITE differ by training level or gender of residents? Based on the literature, there is an assumption that performance on the IM-ITE would improve over time with training. Therefore, are there any significant differences for the independent variables of training level (i.e., post-graduate year) or gender that the actual performance depends (i.e., dependent variable)? To analyze this question, a one-way ANOVA, with a .05 level of significance, will be used to compare between-group means of overall actual percentage correct values for each demographic characteristic separately.

3. How does the predicted performance on the IM-ITE differ by the training level or gender of residents? The predicted or expected response on the Subjective Assessment Survey (SAS), in effect, is functioning as a competency-related belief of the resident to self-evaluate their IM-ITE performance. It is crucial to examine this dependent variable for any differences in perception based on the independent training level (i.e., post-graduate year) and gender variables. To analyze, a one-way ANOVA, with a .05 level of significance, will be used to compare between-group means of overall predicted percentage correct values for each demographic characteristic separately.
4. How does the estimator error-index of residents influence actual performance on the IM-ITE? The rationale behind this research question is to establish an understanding of how the combined estimation error dimensions may relate to actual performance. A bivariate analysis of linear correlation (i.e., Pearson product-moment correlation), with a .05 level of significance, will be used to measure the degree of relationship between the estimator error-index value and the actual overall IM-ITE percentage correct.
5. How does the estimator error-index on the IM-ITE differ by training level or gender of residents? To analyze this question, a one-way ANOVA, with a .05 level of significance, will be used to compare between-group means of estimator error-index values (i.e., dependent variable) for each demographic characteristic (i.e., independent variables) separately.
6. Does the frequency of overestimation or underestimation have any influence on actual performance? This question designed to evaluate the directional dimension of estimation error through an approach that considers the tendency in the resident response pattern distinctly from the single overall direction dimension. The frequency count of how often a

resident either overestimates or underestimates their performance across 12 categorical medical knowledge content areas used in relation to the actual overall IM-ITE percentage correct. A bivariate analysis of linear correlation (i.e., Pearson product-moment correlation), with a .05 level of significance, will be used to measure the degree of relationship for each direction of estimation error separately.

The researcher created tables for each research question analyzed, including the descriptive statistics and summary of analysis results. A scatterplot figure for individual data points plotted in two-dimensional space will be used to illustrate the relationship for each Pearson product-moment correlation. A boxplot figure will be used as a graphical representation of sample dispersion for each demographic characteristic by group mean.

CHAPTER 4: ANALYSIS OF THE DATA

The analysis of the data is characterized by the need to assess how residents self-evaluate, and the effects of self-evaluation on their performance of the IM-ITE. To achieve this aim, a predictive method utilizing self-ratings for performance on a study survey was analyzed in reference to the external standard. To the researcher's experience, no known studies have investigated the personal consequences of self-evaluation and estimation errors related to the academic achievement of medical knowledge.

Demographic Summary

A total of 58 residents self-selected to contribute to the study representing a 52.3% response from the 111 internal medicine residents invited to participate. The distribution of post-graduate year (PGY) of training was characterized by 18 PGY 1 residents corresponding to 31.0% of the overall sample; 22 PGY 2 residents corresponding to 37.9% of the total sample; and 18 PGY 3 residents corresponding to 31.0% of the total sample. Finally, the distribution of gender was characterized by 26 female residents, represented 44.8% of the overall sample, and 32 males represented 55.2% of the total sample. Table 1 provides a cross-tabulation of the demographic distribution.

Table 1. *Cross-tabulation of demographic data for training level and gender*

	Female	Male	Training Level Totals
PGY 1	7	11	18 (31.0%)
PGY 2	10	12	22 (37.9%)
PGY 3	9	9	18 (31.0%)
Gender Totals	26 (44.8%)	32 (55.2%)	58

Research Question Analysis

Question 1

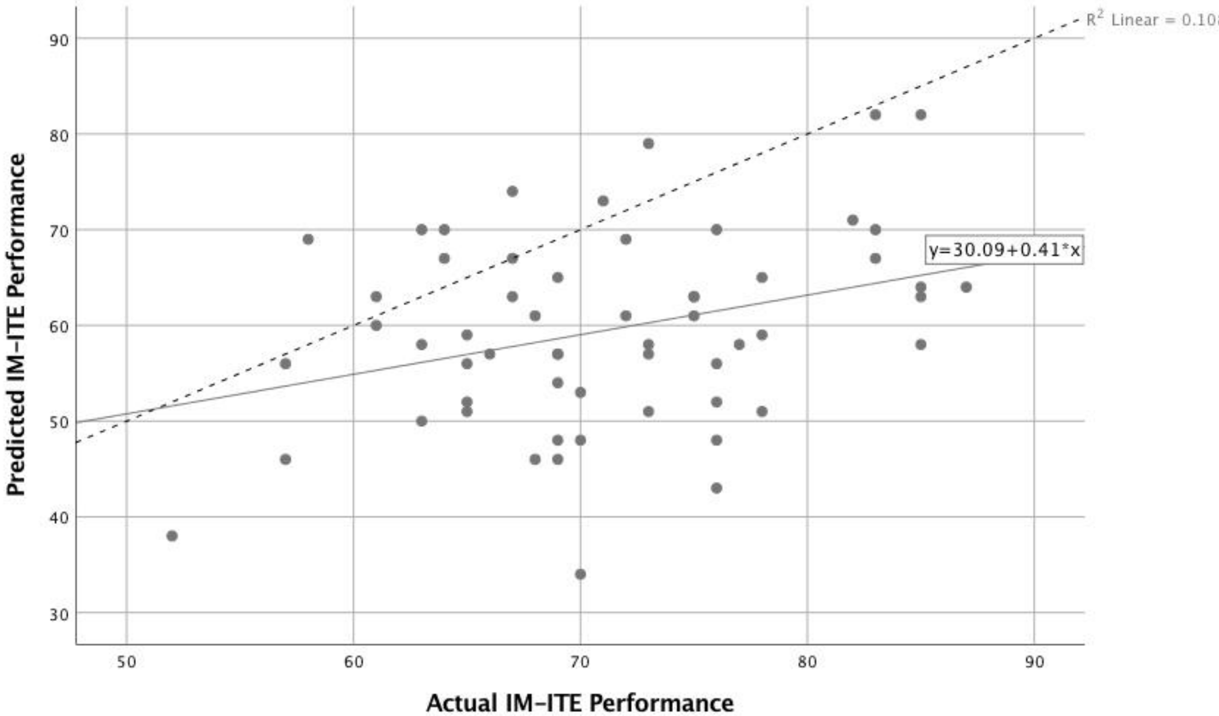
How well do residents predict their actual performance? To answer this research question, the question was evaluated with a focus on how well residents judge their performance (i.e., subjective assessment) using their overall prediction scores on the IM-ITE in comparison to the actual (i.e., objective resource) overall IM-ITE scores. The T-Test method for matched-pairs at the two-tailed .05 level of significance was used for data analysis (See Table 2). The 58 predicted overall IM-ITE scores (M = 59.534, SD = 10.060) and the 58 actual overall IM-ITE scores (M = 71.224, SD = 7.989) demonstrated a significant difference in scores: $t(57) = 8.403$, $p = 0.000$. As a result, there is sufficient evidence to support the conclusion that the residents of this study were not accurate with their predictions — further, they, as a group, tended to underestimate their performance significantly.

A secondary analysis of the same variables was conducted using Pearson's product-moment correlation coefficient to evaluate the degree of the relationship between the two variables and to provide a regression line. The correlation coefficient resulted in a significant, but weak positive relationship between the predicted overall IM-ITE scores and the actual overall IM-ITE scores: $r = 0.328$, $n = 58$, $p = 0.012$. A scatterplot summarizes the results in Figure 1.

In Figure 1, the regression line is represented by the solid line. A secondary dashed reference line was appended to the graph to illustrate what it would look like if there were a "perfect match" or accuracy between the two variables; moreover, the purpose of this reference line was to highlight something in the regression line that is not evident without the reference.

From this scatterplot, it can be clearly shown that the majority of scattered dots appear below the dashed reference line. In view of that, it is worth noting as the actual overall IM-ITE performance advances along the x-axis of the graph (i.e., where objective achievement scores increase), the solid regression line is progressively diverging from the reference line. This deviation can be interpreted as the predictions of overall IM-ITE performance are decreasing disproportionately to actual overall IM-ITE performance in intensity as the real performance increases. Thus, the error in accuracy is increasing as performance is increasing.

Figure 1. Scatterplot *of ability to predict actual IM-ITE performance*



Question 2

How does the actual performance on the IM-ITE differ by training level and gender of residents?

Training level characteristic. The first part of this question was evaluated by comparing actual (i.e., objective resource) overall scores on the IM-ITE and levels of training using a one-way ANOVA. Additionally, post-hoc Fisher's least significant difference (LSD) T-Tests between group means (values of p are for a two-tailed test) were performed for the training level independent variable. The descriptive statistics for actual IM-ITE performance by training level are presented in Table 3. The comparison of the 18 post-graduate year (PGY) 1 actual overall IM-ITE scores (M = 66.444, SD = 8.473) and the 22 PGY 2 actual overall IM-ITE scores (M = 75.773, SD = 7.374) demonstrated a significant difference in scores: t(38) = 4.143, p = 0.000. Additionally, the 22 PGY 2 actual overall IM-ITE scores were compared with the 18 PGY 3 actual overall IM-ITE scores, which also demonstrated a significant difference: $t(38) = 2.366$, $p = 0.023$. However, no statistical difference was found between PGY 1 and PGY 3: $t(36) = 1.739$, $p = 0.091$. The ANOVA summary for actual IM-ITE performance by training level is presented in Table 4.

Figure 2 shows the boxplots for the three levels of training and actual performance. Despite the significant differences in some scores at the levels of training that are obtained in this study, the more years in practice did not certainly result in a stepwise pattern of progressive development at each level of training, and for that reason, the researcher cannot conclude any meaningful findings from this part of the question without more evidence. Although these patterns have been observed in other studies, they are not fixed and will vary between training programs.

Gender characteristic. The second part of this question was evaluated based on actual (i.e., objective resource) overall score on the IM-ITE and the variable of gender. The descriptive statistics for actual IM-ITE performance by gender are presented in Table 5. Of the 26 female ($M = 69.308$, $SD = 8.019$) and 32 males ($M = 72.781$, $SD = 7.741$) participants the findings reveal no significant difference in performance by gender based on the evidence gathered: $t(56) = 1.672$, $p = 0.100$. The boxplot presented in Figure 3 graphically depicts the underlying statistical distribution for actual performance by gender. The ANOVA summary for actual IM-ITE performance by gender is presented in Table 6.

Question 3

How does the predicted overall performance on the IM-ITE differ by training level and gender of residents?

Training level characteristic. The first part of this question was evaluated by comparing predicted (i.e., subjective assessment) overall IM-ITE scores and levels of training using a one-way ANOVA. The descriptive statistics for predicted IM-ITE performance by training level are presented in Table 7. The 18 post-graduate year (PGY) 1 (M = 57.667. SD = 12.880), the 22 PGY 2 (M = 61.545, SD = 8.689), and the 18 PGY 3 (M = 58.944, SD = 8.419) demonstrate no significant differences in their predictions by training level: $F(2,55) = 0.775$, $p = 0.466$. The boxplot presented in Figure 4 graphically depicts the underlying statistical distribution for predicted performance by training level. The ANOVA summary for actual IM-ITE performance by training level is presented in Table 8. As a result, no post-hoc Fisher's least significant difference (LSD) T-Tests were conducted for this question.

Gender characteristic. The second part of this question evaluated the predicted (i.e., subjective assessment) overall IM-ITE score and the variable of gender using a one-way ANOVA. The descriptive statistics for predicted IM-ITE performance by gender are presented in Table 9. The 26 female (M = 54.615, SD = 8.841) and the 32 male (M = 63.531, SD = 9.287) participants demonstrated a significant effect for gender based on the responses: $F(1, 56) = 13.798$, $p = 0.000$. The boxplot presented in Figure 3 graphically depicts the underlying statistical distribution for predicted performance by gender. The ANOVA summary for actual IM-ITE performance by gender is presented in Table 10. The evidence reveals support to conclude that female residents notably underestimate their performance lower in comparison to their male colleagues.

To follow-up with this finding a closer examination using a one-way ANOVA to analyze the dependent variable of predicted overall IM-ITE scores in combination with isolated levels of training (i.e., each training level considered individually) and gender was conducted. The purpose of this subsequent analysis was to obtain a better understanding of the resultant effect present at each level of training. The descriptive statistics for predicted IM-ITE performance by each training level are presented in Tables 11 (PGY 1), 12 (PGY 2), and 13 (PGY 3). From the data and in this sample of residents, it can be concluded that the effect is significant in the first two years of training (i.e., PGY 1 and PGY 2), but is not significant by post-graduate year 3 (PGY 3):

1. Predicted overall IM-ITE performance for PGY 1 by gender: $F(1, 16) = 8.097$, $p = 0.012$.
2. Predicted overall IM-ITE performance for PGY 2 by gender: $F(1, 20) = 5.587$, $p = 0.028$.
3. Predicted overall IM-ITE performance for PGY 3 by gender: $F(1, 16) = 1.648$, $p = 0.218$.

On the basis of this subsequent analysis, the researcher concludes that there is enough evidence to support the idea that judgments in performance by gender change with time in-training.

Additionally, the analysis of this effect was carried out one-step further to isolate the genders and compare by the training levels, but no significant differences were detected from the within gender comparisons by training level:

a. Predicted IM-ITE for females by each PGY: $F(2, 23) = 2.532$, $p = 0.101$.
b. Predicted IM-ITE for males by each PGY: $F(2, 29) = 0.398$, $p = 0.676$.

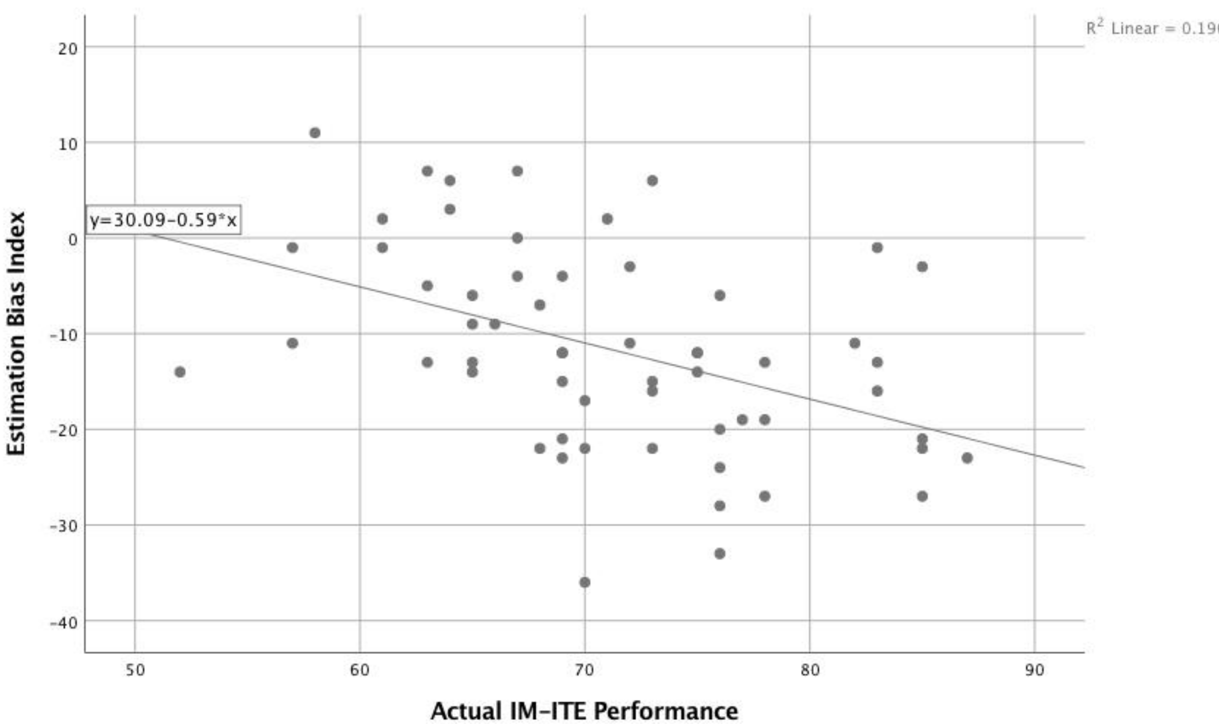
The descriptive statistics for predicted IM-ITE performance within each gender are presented in Table 14 (females) and Table 15 (males).

Question 4

How does the estimator error-index of residents influence actual overall performance on the IM-ITE? This question was evaluated to determine if a significant relationship exists between the estimation error-index (i.e., quantitative relationship properties of direction and magnitude rising from the comparison of subjective judgment and objective assessment) and actual (i.e., objective resource) overall IM-ITE score. The descriptive statistics for actual IM-ITE performance and estimator error-index are presented in Table 16. Pearson's product-moment correlation coefficient was conducted comparing the two variables. The 58 actual IM-ITE scores ($M = 71.224$, $SD = 7.989$) and the 58 estimator error-index scores ($M = -11.690$, $SD = 10.595$) demonstrated a significant negative correlation between variables: r = -0.442, n = 58, p = 0.001. Pearson's correlation matrix for actual IM-ITE performance and estimator error-index are presented in Table 17. A scatterplot summarizes the results in Figure 6.

In Figure 6, the scatterplot noticeably reveals that as actual performance on the IM-ITE increases along the x-axis, there is an effect related to both the direction and magnitude of the estimation error-index. From the results of this analysis, the researcher concludes there is enough evidence to support the idea that estimator error-index resulting from self-evaluation has a significant relationship between achievement and the acquisition of medical knowledge.

Figure 6. *Scatterplot for actual IM-ITE and estimator error-index*



Question 5

How does estimator error-index on the IM-ITE differ by training level or gender of residents?

Training level characteristic. The first part of this question was evaluated by comparing the estimator error-index (i.e., quantitative relationship properties of direction and magnitude rising from the comparison of subjective judgment and objective assessment) and levels of training using a one-way ANOVA. The descriptive statistics for the estimator error-index by training level are presented in Table 18. The 18 post-graduate year (PGY) 1 (M = -8.778, SD = 12.530), the 22 PGY 2 (M = -14.227, SD = 9.744), and the 18 PGY 3 (M = -11.500, SD = 9.205) demonstrate no significant differences in their estimator error-index by training level: $F(2,55) = 1.329$, $p = 0.273$. The boxplot presented in Figure 7 graphically depicts the underlying statistical distribution for the estimator error-index and training level. The data shows that each training level as a group underestimates themselves and that their ability to predict their actual scores does not inevitably get better over time. The ANOVA summary table for estimator error-index by training level are presented in Table 19.

Gender characteristic. The second part of this question evaluated the estimator error-index with the variable of gender. The descriptive statistics for estimator error-index by gender are presented in Table 20. The 26 female (M = -14.692, SD = 10.345) and the 32 male (M = -9.250, SD = 10.314) participants almost approached a sizable variance (i.e., $p = 0.051$), but evidence indicates there is not enough evidence to conclude a difference: $F(1,56) = 3.983$, $p = 0.051$. The boxplot presented in Figure 8 graphically depicts the underlying statistical distribution for the estimator error-index by gender. The ANOVA summary table for estimator error-index by gender are presented in Table 21.

Question 6

How does the frequency of overestimation or underestimation influence actual performance? With this research question, the researcher wanted to understand the degree of the relationship between how often a resident overestimates or underestimates their performance as determined in each of the 12 medical content categories of the IM-ITE as a measure of the tendency to actual IM-ITE performance? To evaluate this question, first, a frequency count was conducted for each resident in each of 12 content categories for each time they either overestimated or underestimated in a content category using the difference between the subjective assessment and objective resource to determine the direction. Then, these variables were then analyzed separately by frequency counts for overestimation and by the rate for underestimation in relationship to actual overall IM-ITE scores using Pearson's product-moment correlation coefficient.

Overestimation results. The descriptive statistics for overestimation frequency and actual IM-ITE scores are presented in Table 22. The results for frequency for overestimation reveal a moderate significant negative relationship between how often a resident overestimates and actual overall IM-ITE performance: $r = -0.464$, $n = 58$, $p = 0.000$. From this result, the research concludes there is enough evidence to support a significant negative relationship between the tendency to overestimate and achievement and acquisition of medical knowledge as measured by the actual overall IM-ITE performance. Pearson's correlation matrix for overestimation frequency and actual IM-ITE performance are presented in Table 23. The scatterplot shown in Figure 9 illustrates the relation between overestimation frequency and actual IM-ITE performance.

Underestimation results. The descriptive statistics for underestimation frequency and actual IM-ITE scores are presented in Table 24. In comparison, the results for the frequency of underestimation exhibit a moderate positive relationship between variables: $r = 0.498$, $n = 58$, $p = 0.000$. As a result, there appears to be enough evidence to support the conclusion that underestimation leads to better achievement and acquisition of medical knowledge as measured by the actual overall IM-ITE performance. Pearson's correlation matrix for underestimation frequency and actual IM-ITE performance are presented in Table 25. The scatterplot presented in

Figure 10 illustrates the relation between underestimation frequency and actual IM-ITE performance.

Analysis of Data Summary

In summary, the research clearly indicates six summative conclusions that correspond to each research question:

1. Residents are not very accurate with their self-evaluations. Residents, as a whole, tend to underestimate their performance in this sample.
2. The training level characteristic indicated some significant variation for actual IM-ITE performance at assorted ranks, but the researcher cannot conclude any meaningful revelations from the existing evidence. Additionally, the actual performance on the IM-ITE was uniform between genders.
3. The prediction of performance indicated no sizable dependence on the training level characteristic. However, an unexplained gender difference for predictions was detected. Females residents tend to underestimate their performance significantly more than their male colleagues. Furthermore, this gender difference appears to undergo substantial adjustment with time.
4. The estimation error-index has a significant relationship with actual performance. The directional dimension of estimation error indicates that overestimators are more likely to perform lower on the IM-ITE. Likewise, the underestimator is more likely to achieve higher on the IM-ITE. This finding is significantly crucial to understanding the relationship between self-evaluation and medical knowledge outcomes. Furthermore, the intensity dimension of estimation errors appears to influence performance outcomes in both directional aspects, but with contrary

effects. For instance, as inaccuracy grows in the overestimation direction, the intensity expands, the performance decreases. However, in the underestimation direction inaccuracy also grows, the intensity also expands, but performance increases.

5. The estimator error-index variable does not significantly depend on the demographic characteristics of residents (i.e., training and gender).
6. The actual performance on the IM-ITE is significantly influenced by the resident tendency for the directional dimension of estimation error. The frequency of overestimation was associated with lower performance on the IM-ITE. Likewise, the rate of underestimation was associated with higher performance on the IM-ITE.

CHAPTER 5: SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This chapter provides a summary of results from the current study, conclusions, and recommendations for future research. Furthermore, this chapter discusses the findings, which pertain to the self-evaluation skills of an internal medicine resident sample, the error of measurement in their self-evaluation, and its influence on the academic achievement of their medical knowledge as assessed by an objective standard. To the author's experience, this is the first study to investigate how these factors develop and contribute to resident achievement.

Summary

The present study was grounded in the concept that self-evaluation skills are central to the development of lifelong self-directed learning skills. Research literature indicates that a resident must be capable of defining and resolving their own learning needs in order to self-regulate their behavior toward the achievement of academic goals (Burke, Benson, Englander, Carraccio & Hicks, 2014; Eva & Regehr, 2007). It is theorized that a person's insight and awareness into their inadequacies through the process of self-evaluation is a fundamental prerequisite to the development of skills and the progress toward competency (Dunning, 2011; Kruger & Dunning, 1999). To determine which variables are influential is difficult without a proper study. To this end, this study was designed around a central question—how do residents self-evaluate their performance and what, if any, influence does this have on their academic achievement of medical knowledge?

This study aimed to understand how residents predict or estimate their performance on the Internal Medicine In-Training Examination (IM-ITE) using a voluntary self-reported survey. According to the American College of Physicians (2019), the IM-ITE is designed to evaluate the knowledge, diagnostic reasoning, and clinical judgment skills expected of a certified internist and functions as the objective resource for this study. The actual IM-ITE scores are offered as a percentage correct measurement in this study. The self-report survey required participants to assess their own internal beliefs by providing a self-rating also as a percentage correct as an expectancy for their performance for each of the 12 IM-ITE medical content areas. Measurement in this research aimed to analyze the properties exhibited in the assigned resident values between the two instruments. Likewise, the study intended to investigate the estimation errors that arise from the differences between the subjective assessment (i.e., survey) and objective resource (i.e., IM-ITE) for its influence on actual performance.

The resident evaluates the content of the survey against their relevant knowledge structures and experiences with the external (i.e., objective) resource in mind and provides a perceived reflection of their internalization. From the theoretical perspectives, self-evaluative responses are internal perceptions or beliefs, that is not directly observable, but encompass the intentions of an individual to act toward goal attainment (Eccles & Wigfield, 2002; Ryan & Deci, 2008). Predictions derived from this study are supported by empirical tests that examine individual predisposition orientations. In this report, the term "predisposition orientation" broadly applies to a self-evaluated learning need.

Conclusions

This study has evaluated the ability of residents to self-evaluate, the demographic characteristics, and the relationship between estimation errors and performance. This study was an exploratory study and limited by a lack of established evidence. The results of the survey nonetheless emphasize the potential influence of self-evaluation on performance in residents and support future research in this area. The following conclusions emerge from the study:

Ability to self-evaluate. Research from this study shows that residents are generally inaccurate with their self-evaluations of performance on the IM-ITE. Residents as a whole tended to underestimate their performance; on average, 11.7 percentage points lower than their actual performance. In total, 8 participants (representing 13.8% of the overall 58 total) overestimated their performance, 49 participants (84.5%) underestimated their performance, and 1 participant (1.7%) was accurate in the sample population. These estimation biases were studied further to determine if accuracy improves over the years of training; however, when a year-by-year comparison was conducted, there was no suggestion that residents were getting better with their accuracy. This finding was unexpected because residents further along in training years have had more encounters with the IM-ITE, including receiving results related to past performance than the residents taking it for the first time with no prior experience of results. Additionally, there was no indication for a gender difference linked to the accuracy in self-evaluations. These findings suggest that residents are not inherently different in their ability to self-evaluate from other student populations—more likely to be inaccurate than accurate (Dunning and Kruger, 1995).

Actual IM-ITE performance qualities. The actual IM-ITE performance scores were analyzed separately using the demographic characteristics of training level and gender, as independent variables, for a between-group analysis. Regarding the actual IM-ITE performance, the dependent variable, there were no unexpected findings to report. On the basis of evidence collected, the research shows there were some sizeable changes in actual performance on the IM-ITE related to training over time. The available evidence suggests that residents exhibited performance differences at the various levels of training, but the differences did not appear to follow a sequential step-wise pattern of growth. The results contrast with some previous studies that have detailed increased growth rates of approximately 5% for each year of training (Garibaldi et al., 2002). Nonetheless, the researcher acknowledges these discernible patterns have been observed in other studies, but are not fixed patterns, and in this sample, vary primarily amid the PGY 2 residents where, on average, they outperformed the PGY 3 residents. Lastly, an examination of gender revealed that the female and male participants' actual performance on the IM-ITE correspond to one another, and no gender differences were suggested from the data.

Predicted IM-ITE performance qualities. Likewise, the predicted IM-ITE performance scores were also analyzed separately using the demographic characteristics of training level and gender, as independent variables, for a between-group analysis. The available evidence suggests that there were no indications of a difference in predicted performance by training level. However, among genders, the study data indicates a significant difference between how female and male residents predict their performance. There is limited but consistent evidence from this study, that the female residents significantly underestimate their performance (M=54.6), on average, 8.9 percentage points lower than male residents (M=63.5); there is a high probability that this difference in means is not due to chance ($p = .000$). However, this study did not attempt to answer the subtle differences in the way these variables may manifest by demographic characteristics. Further, gender differences are poorly recognized in graduate medical education, and little is known about their existence or affects in training (Miller & Katz, 2018). Nonetheless, based on this evidence, the researcher determined a post-hoc examination of this gender difference should take into account the significance of this effect for each year of training.

It should be noted that the leading effect of the gender difference gap was most extensive in the first post-graduate year of training (PGY 1), where it measured, on average, 14.9 percentage points apart between the genders. Furthermore, the data suggest that this gender difference remains significant in the second year of training (PGY 2), but with the gap narrowing among genders for each successive year of training 14.9 (PGY 1) to 8 (PGY 2) percentage points. Finally, by the third year (PGY 3), the gender difference appears no longer significant, where it is shown at five percentage points away from each gender. The analysis was conducted one-step further to inspect within gender differences—females and males were compared independently—for each training year, but the data did not support any significant differences. The researcher concluded that the processes associated with the gender difference for the predictions are complex and not well understood; but, there is evidence that they are capable of adjusting with time.

A brief review of the literature following this conclusion reveals some possible explanations for this finding. Evidence provided in the literature about gender differences in the accuracy of self-evaluations of performance informs that gender-considered knowledge domains can exhibit these effects. Research by Beyer (1990, 1998, 2002) investigated the relation between self-perception biases and how expectancies affect post-task self-evaluations. Beyer (1990) details a history of findings where females have lower expectancies of success than males in many areas of achievement (see Beyer, 1990, p. 960). Furthermore, Beyer (1998) states, "gender differences in self-evaluations are unlikely to be caused by the mere fact of being male or female" (p. 105). Instead, there is some evidence to suggest that some knowledge domains, like sports or mathematics, are culturally gender-typed, and this perception mediates a difference in the accuracy of the self-evaluation. Characterized by Beyer's research, the difference in this study could be accounted for if the medical knowledge domain were judged as masculine.

Relationship between estimation errors and actual performance. Many factors influence the acquisition of medical knowledge, the differences among residents, concerning what they think and know, were examined for patterns using the combination of two dimensions of estimation error: (1) direction and (2) intensity. Evidence collected as part of this study suggests a significant negative correlation between the estimator error-index and actual performance.

Directional dimension. Data accessing the directional dimension of estimation errors in residents indicate that overestimations of performance are most often related to lower actual performance, which is contrasted by underestimations of performance that are associated with higher actual performance. While the process of this relationship is not explicitly identified in this study, theoretical models provide for some possible explanation. For example, Kruger and Dunning (1999) yielded comparable information for the directional dimension of estimation errors in their metacognitive research.

The Kruger and Dunning (1999) research focused on cognitive bias by examining the self-assessments of an undergraduate student population. They explained the reason for the estimation error as a deficiency in metacognitive self-awareness. For the overestimator, this cognitive bias, is a failure to recognize their shortcomings and allows people to assess their cognitive abilities as being better than they are. As a general rule, people who are lacking ability at something are unable to recognize their own inability, and thereby, tend to overestimate.

Further, corresponding to the findings of this current study, Kruger and Dunning (1999) also found that the higher-performing students tended to underestimate their performance. Their explanation for the underestimation direction in people appears to stem from an error in self-perceived ability relative to the perceived strength of others. In other words, the discrepancy is rationalized by a failure in assessing the proficiency of others as more exceptional without proper recognition of personal competence; the perception is directed toward oneself, which evokes placing oneself in a lower position. The findings from this study are generally consistent with previously reported results.

Throughout this report, the researcher broadly applies the expectancy-value theory framework of Eccles and Wigfield (2000, 2002) to examine several important aspects of competency-related beliefs, so that interpretations from theoretical foundations can be applied to findings that enable educators to think in connection to achievement motivation. The available evidence suggests based on achievement motivation theories and constructs, that competency-related beliefs cognitively engage resident motivational processes that direct the learning behavior manifested in the outcomes. The explanation provided in the literature for the directional dimension may be thought of as an orientation about internal latent states underlying the behavioral processes that direct the learning choices made. Hence, the underestimation leads attention and behavior toward increasing correspondence with the target; whereas, the overestimation directs attention and practice away from the goal. Therefore, the directional dimension affects the promotion of learning priorities consonant with their point of view, predicting the occurrence of specific learning behavior affecting the performance outcomes.

As a result, the directional dimension of self-evaluation has implications for educational practice that is self-directed. There is evidence to support that the academic achievement of medical knowledge is, to some extent, motivated by the competency-related beliefs of residents. It is hypothesized that residents may make distinctions among content areas that are essential to identify their areas of interest and motivation (Harter, 1982). The evidence provided in this study reveals potential forces for change that could be considered for improving the acquisition of medical knowledge by supporting the self-awareness of estimator error through interventional feedback mechanisms as applied in this research.

Intensity dimension. This study also found that the intensity dimension of estimation error increases as actual performance increases. From the existing data, Figure 1 illustrated the researcher's conceptual model, demonstrating that as performance increases, the predictions for performance grew inaccurate. In this one-directional model, there is a relation between the intensity dimensions of estimation error that corresponds to the highest levels of performance.

Is it better to be accurate or inaccurate? Although some studies suggest that the degree of accuracy played a role in performance outcomes, the conclusions in the literature are inconsistently reported. In psychology and related fields, there is no agreement about why the degree of accuracy contributes to higher performance in some students that may lead to diminished performance in others (Bol & Hacker, 2001; Kruger & Dunning, 1999; Chen 2003).

Figure 6 elaborates on this relationship considering multiple dimensions, the direction, and intensity. The regression line passes through the point of accuracy, where the y-axis is zero. Hence, accuracy is exhibited as a point without dimension. However, inaccuracy or intensity is an extensive dimension, and there is some evidence to support the idea that the degree of accuracy can be detrimental or beneficial depending on the direction dimension. This is a key distinction that must be made regarding the interpretation of this intensity dimension and may explain inconsistent findings in other studies. In this study, if the degree of accuracy or intensity dimension were considered as an absolute value, it may create distorted impressions and false conclusions as to its influence on performance. It would be misleading to report this intensity dimension without reference to the directional dimension because the amount is irrelevant without also knowing the orientation to outcomes. This is not necessarily the same condition with the directional dimension.

Overestimation and underestimation tendency. As reported earlier, the evidence indicates that there is a significant relationship for the direction of estimation errors, which constitute an essential source of information for anticipating performance outcomes. Another critical component of this same dimension was to consider if the frequency or tendency of the resident for a direction contributed to actual IM-ITE performance. The available evidence from this study confirms that the tendency to overestimate was significantly associated with reduced performance. Moreover, the tendency to underestimate was connected considerably with increased actual IM-ITE performance. The evidence further indicates that this directional dimension has a significant impact on the academic achievement of medical knowledge.

Study Limitations

The present research adds to our understanding of the self-evaluation process through a study of the moderating effects of competency-related beliefs and academic achievement of medical knowledge. However, the scope of this study acknowledges three main limitations: sample, methodology, and time.

First, by necessity, this was a non-probability sample of convenience where residents self-selected to participate in the study voluntarily. The sample size was restricted to an internal medicine residency program in Detroit, Michigan. Thus, there will be a limited number of participants in the study. As a result, the findings of the study may not represent internal medicine residents at other training institutes. Further, the internal medicine specialty may not represent other medical specialties, like surgery or pediatrics in graduate medical education.

Second, the methodology of the research design provides restrictions. The study intended to obtain information, which could be analyzed to describe, compare, and explore relations. The study design and data collection instruments may not identify all factors that influence participant behavior or outcomes; the study is limited to the research questions. A section of the research will be conducted over the Internet resulting in a lack of experimental control.

Third, there is always a restriction related to time. The review of the literature was not exhaustive but extensive. The data collection was conducted over a relatively short period of time, which could affect the rate of response. Common barriers include a lack of resident time to participate.

Recommendations for Future Research

This study is an essential step toward identifying some contributing factors that decrease and increase the academic achievement of medical knowledge among a sample of residents. The results, discussion, and implications of this study suggest the need for further research. This research has limitations and imperfections, just as any other research design does, but the researcher believes the findings of this study establish well-grounded evidence in an under-investigated population that future research might address. As documented throughout this report, much was unknown about the study population. These research findings help us to develop even better future research questions. Before the research, not enough was known to create hypotheses for the resident population, so concerns related to random sampling or experimental design were not appropriate. However, the study allowed the researcher the opportunity to make critical decisions, to determine differences, and report some distinct discoveries. These distinctions must be accounted for in subsequent studies before researchers can make sound generalizations about the strength of the competency-related belief and performance relationship in residents. Future research may focus on the control of potentially influential factors through experimental design to reduce any confounding variables.

For instance, the self-selecting, non-probability sampling might be suitable for investigating new populations; the findings may not reflect the general study population. Further, research is needed to identify additional mediating factors that relate to self-evaluation. The gender differences for the predictions of performance identified in this study are not understood. This study did not attempt to answer the subtle differences in the way these variables may manifest by demographic characteristics of residents. Further, it might be expected to see a similar gender difference also indicated in the estimation error-index, since predictions were an aspect of the index; and although the statistical determination was close to significant ($p = .051$), there was not enough evidence to conclude a distinction. Subsequent studies with larger sample sizes may someday account for a different finding. The results of this study have the potential to be transformative in the way residents recognize their cognitive bias and improve recognition of learning needs, thereby someday enhancing self-directed learning skills. Prior to the research, there was little known about the application of theoretical frameworks in GME to interpret findings to support learning interventions. The report reiterates several recommendations, but the most important might be using the methodology of this study to provide feedback.

Finally, feedback may be a promising opportunity for achieving performance change. The risk of overestimation can lead to insufficient efforts to learn because the student thinks of themselves better than they are (Elliot, Dweck, & Yeager, 2017). The potential to move the research agenda forward through increased feedback interventions, improved recognition of estimation errors in self-evaluation, and its influence on performance has the potential to support underachieving residents increase their medical knowledge resulting in better patient care. Although research suggests that self-evaluation can influence performance, little is known about how to best improve self-evaluation skills, which contributes to improvement in skills, and which advantages are derived from reducing estimation errors.

Recommendations for Future Practice

The recommendation for future practice is to consider the methods used in this study in the application of GME training to encourage self-awareness through self-evaluation. The results of this study underscore the feasibility and potential benefits of using self-evaluation of performance in combination with the objective IM-ITE standard. Self-evaluation and the feedback process are two areas that offer significant opportunities for improvement in the practice-based learning and improvement competency and the development of self-directed learning skills. The initial efforts of this study provide evidence to support the implications of competency-related beliefs as a mediator to self-motivation and academic achievement of medical knowledge. The next challenge is to establish the effectiveness of feedback in training residents to recognize their learning needs and commitment to self-regulation of learning behaviors. The potential to move the research agenda forward through increased feedback interventions and improved recognition of blind spots in self-evaluation has the promise to positively impact underachieving residents' increase medical knowledge, possibly resulting in better patient care.

TABLES

Table 1

Cross-tabulation of demographic data for training level and gender

	Female	Male	Training Level Totals
PGY 1	7	11	18 (31.0%)
PGY 2	10	12	22 (37.9%)
PGY 3	9	9	18 (31.0%)
Gender Totals	26 (44.8%)	32 (55.2%)	58

Table 2

T-Test comparing predicted and actual IM-ITE performance

	N	Mean	SD	t	DF	p
Actual IM-ITE	58	71.224	7.989	8.403	57	.000
Predicted	58	59.534	10.060			

Table 3

Descriptive statistics for actual IM-ITE by PGY

	N	Mean	SD
PGY 1	18	66.444	8.473
PGY 2	22	75.773	7.374
PGY 3	18	70.444	4.841

Table 4

ANOVA summary table for actual IM-ITE by PGY

Source of Variation	DF	Sum of Squares	Mean Squares	F-Ratio	Significance Level
A	2	877.334	438.667	8.739	0.001
Error	55	2760.753	50.196		
Total	57	3638.086			

Table 5

Descriptive statistics for actual IM-ITE by gender

	N	Mean	SD
Female	26	69.308	8.019
Male	32	72.781	7.741

Table 6

ANOVA summary table for actual IM-ITE by gender

Source of Variation	DF	Sum of Squares	Mean Squares	F-Ratio	Significance Level
A	1	173.079	173.079	2.797	0.100
Error	56	3465.007	61.875		
Total	57	3638.086			

Table 7

Descriptive statistics for predicted IM-ITE by PGY

	N	Mean	SD
PGY 1	18	57.667	12.880
PGY 2	22	61.545	8.689
PGY 3	18	58.944	8.419

Table 8

ANOVA summary table for predicted IM-ITE by PGY

Source of Variation	DF	Sum of Squares	Mean Squares	F-Ratio	Significance Level
A	2	158.032	79.016	0.775	0.466
Error	55	5610.399	102.007		
Total	57	5768.431			

Table 9

Descriptive statistics for predicted IM-ITE by gender

	N	Mean	SD
Female	26	54.615	8.841
Male	32	63.531	9.287

Table 10

ANOVA summary table for predicted IM-ITE by gender

Source of Variation	DF	Sum of Squares	Mean Squares	F-Ratio	Significance Level
A	1	1140.308	1140.308	13.798	0.000
Error	56	4628.123	82.645		
Total	57	5768.431			

Table 11

Descriptive statistics for predicted for PGY 1 by gender

	N	Mean	SD
Female	7	48.57	11.193
Male	11	63.45	10.586

Table 12

Descriptive statistics for predicted for PGY 2 by gender

	N	Mean	SD
Female	10	57.20	7.613
Male	12	65.17	8.077

Table 13

Descriptive statistics for predicted for PGY 3 by gender

	N	Mean	SD
Female	9	56.44	6.405
Male	9	61.44	9.774

Table 14

Descriptive statistics for predicted for females by PGY

	N	Mean	SD
PGY 1	7	48.57	11.193
PGY 2	10	57.20	7.613
PGY 3	9	56.44	6.405

Table 15

Descriptive statistics for predicted for males by PGY

	N	Mean	SD
PGY 1	11	63.45	10.586
PGY 2	12	65.17	8.077
PGY 3	9	61.44	9.774

Table 16

Descriptive statistics for actual IM-ITE and estimator error-index

	N	Mean	SD
Actual IM-ITE performance	58	71.224	7.989
Estimator Error-Index	58	-11.690	10.595

Table 17

Pearson's product-moment correlation matrix for actual IM-ITE and estimator error-index

		Actual IM-ITE
Estimator Error-Index	r	-.442
	N	58
	SE	0.106
	t	3.691
	p	.001

Table 18

Descriptive statistics for estimator error-index by PGY

	N	Mean	SD
PGY 1	18	-8.778	12.530
PGY 2	22	-14.227	9.744
PGY 3	18	-11.500	9.205

Table 19

ANOVA summary table for estimator error-index by PGY

Source of Variation	DF	Sum of Squares	Mean Squares	F-Ratio	Significance Level
A	2	294.939	147.470	1.329	0.273
Error	55	6103.475	110.972		
Total	57	6398.414			

Table 20

Descriptive statistics for estimator error-index by gender

	N	Mean	SD
Female	26	-14.692	10.345
Male	32	-9.250	10.314

Table 21

ANOVA summary table for estimator error-index by gender

Source of Variation	DF	Sum of Squares	Mean Squares	F-Ratio	Significance Level
A	1	424.875	424.875	3.983	0.051
Error	56	5973.538	106.670		
Total	57	6398.414			

Table 22

Descriptive statistics for overestimation frequency and actual IM-ITE

	N	Mean	SD
Actual IM-ITE performance	58	71.224	7.989
Overestimation Frequency	58	2.793	2.894

Table 23

Pearson's product-moment correlation matrix for overestimation frequency and actual IM-ITE

		Actual IM-ITE
Overestimation Frequency	r	-.464
	N	58
	SE	0.103
	t	3.918
	p	.000

Cronbach's Alpha = .634

Table 24

Descriptive statistics for underestimation frequency and actual IM-ITE

	N	Mean	SD
Actual IM-ITE performance	58	71.224	7.989
Underestimation Frequency	58	9.000	3.009

Table 25

Pearson's Product-Moment Correlation Matrix for underestimation frequency and actual IM-ITE

		Actual IM-ITE
Underestimation Frequency	r	.498
	N	58
	SE	0.099
	t	4.303
	p	.000

Cronbach's Alpha = .665

FIGURES

Figure 1

Scatterplot of ability to predict actual IM-ITE performance

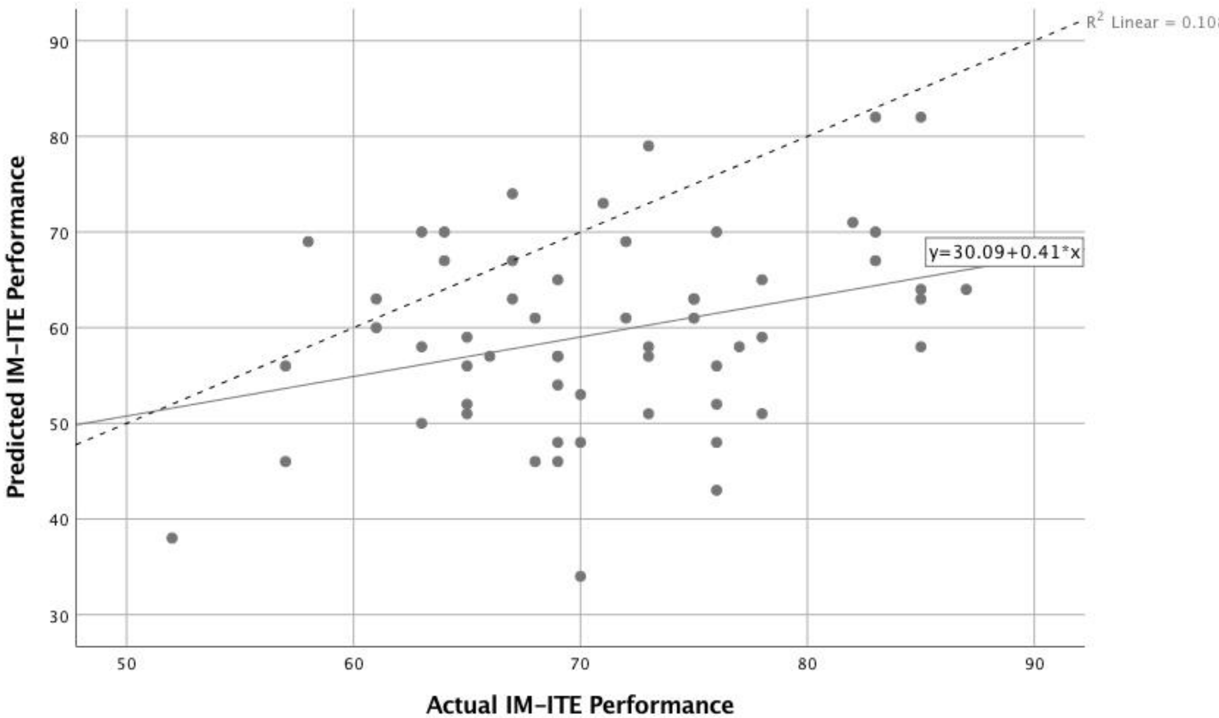


Figure 2

Boxplot of actual IM-ITE performance by PGY

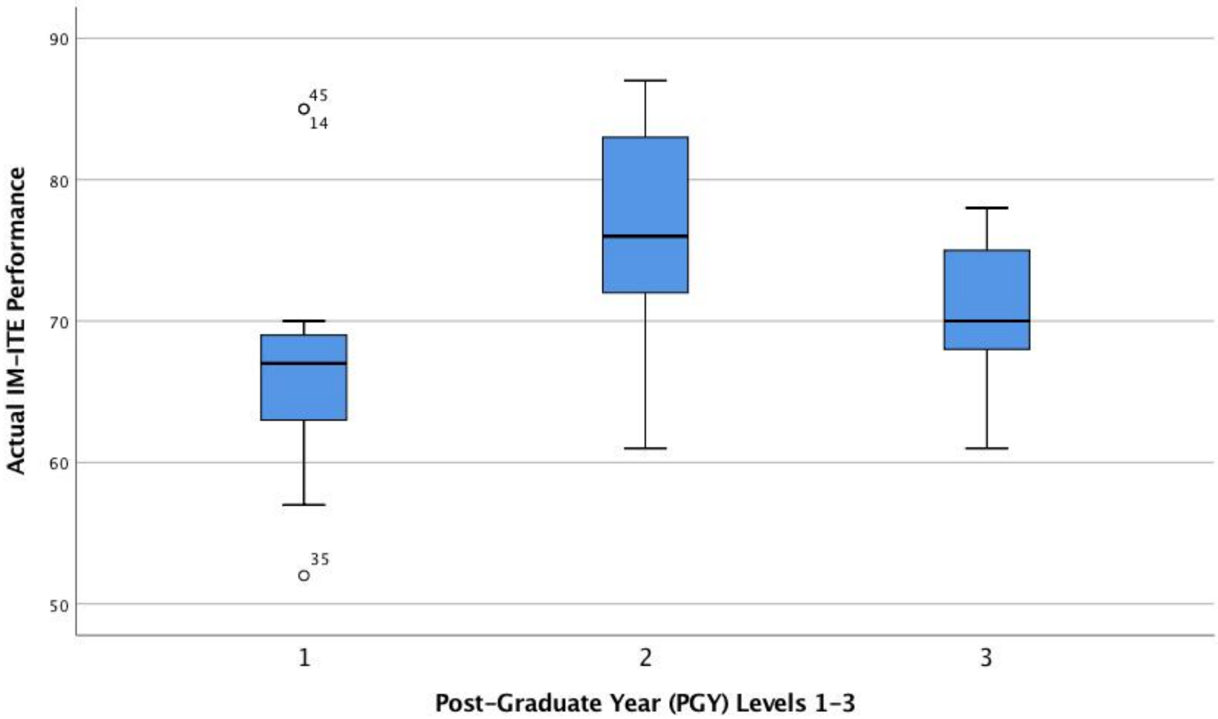


Figure 3

Boxplot for actual IM-ITE performance by gender

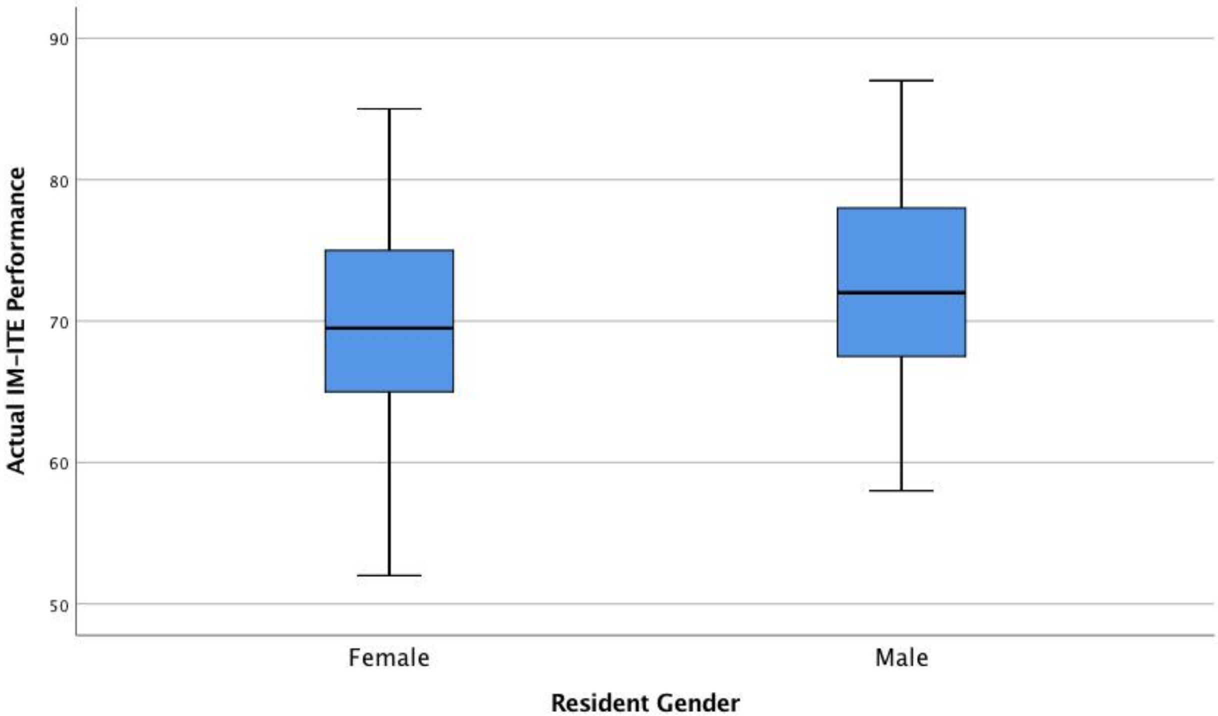


Figure 4

Boxplot for predicted IM-ITE by PGY

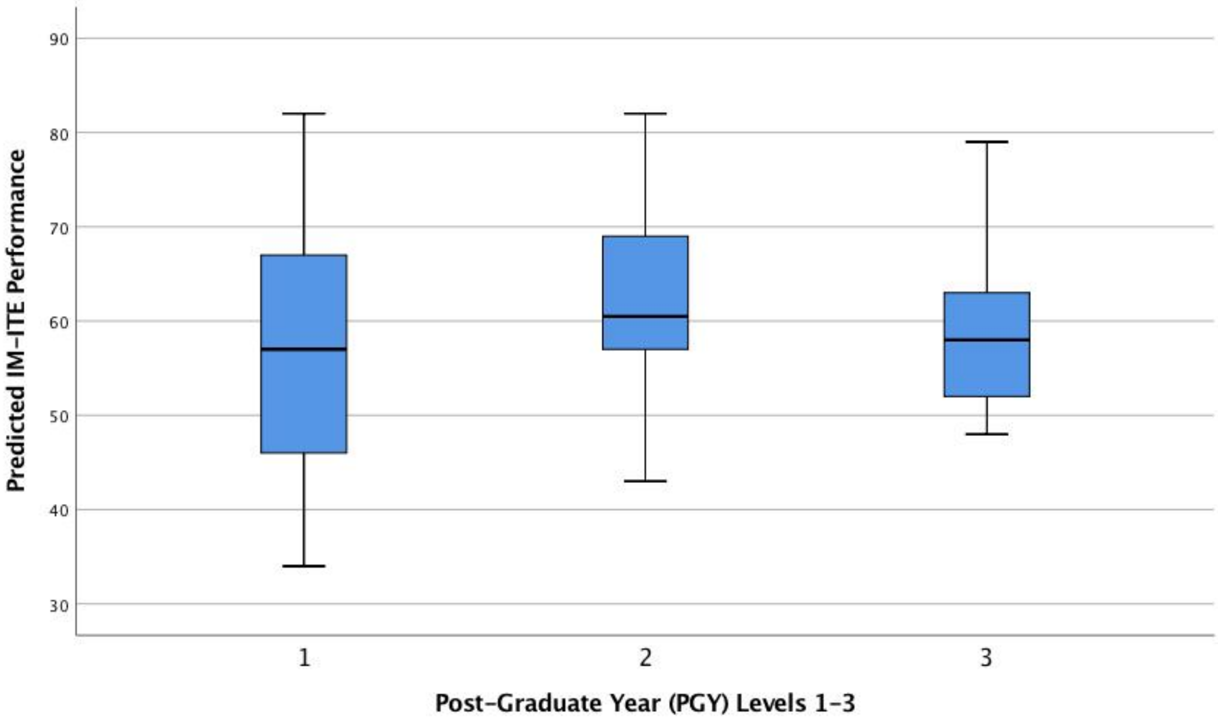


Figure 5

Boxplot for predicted IM-ITE by gender

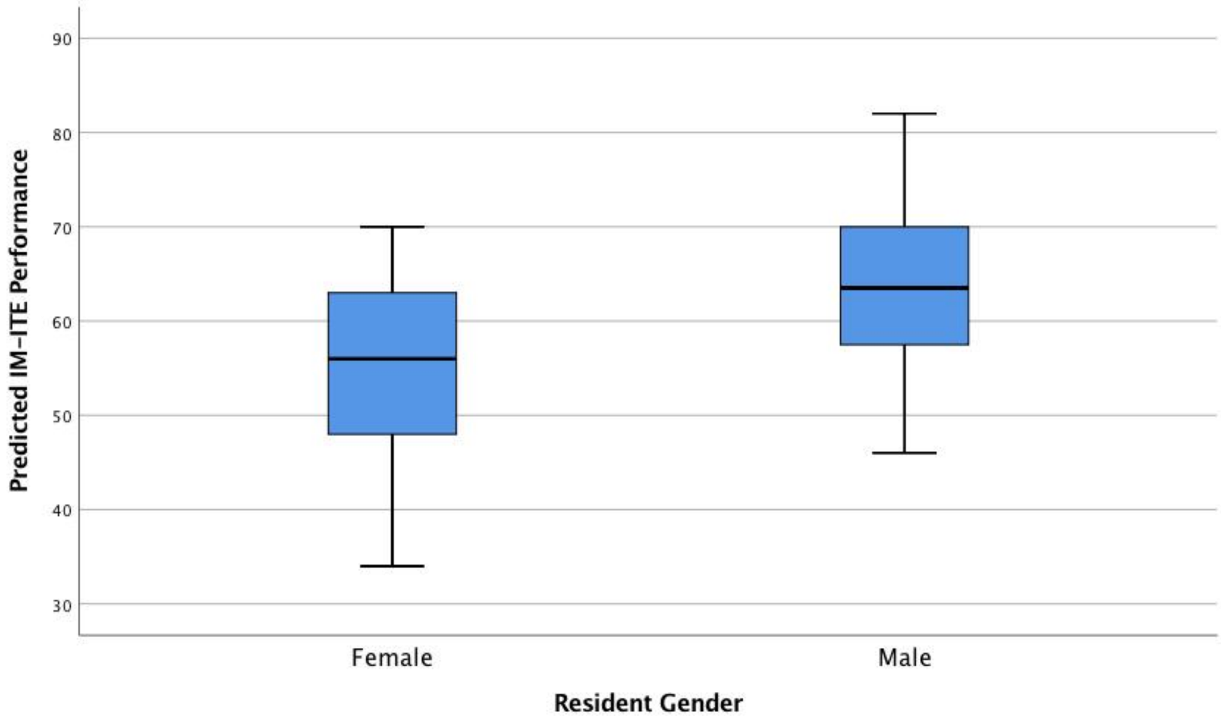


Figure 6

Scatterplot for actual IM-ITE and estimator error-index

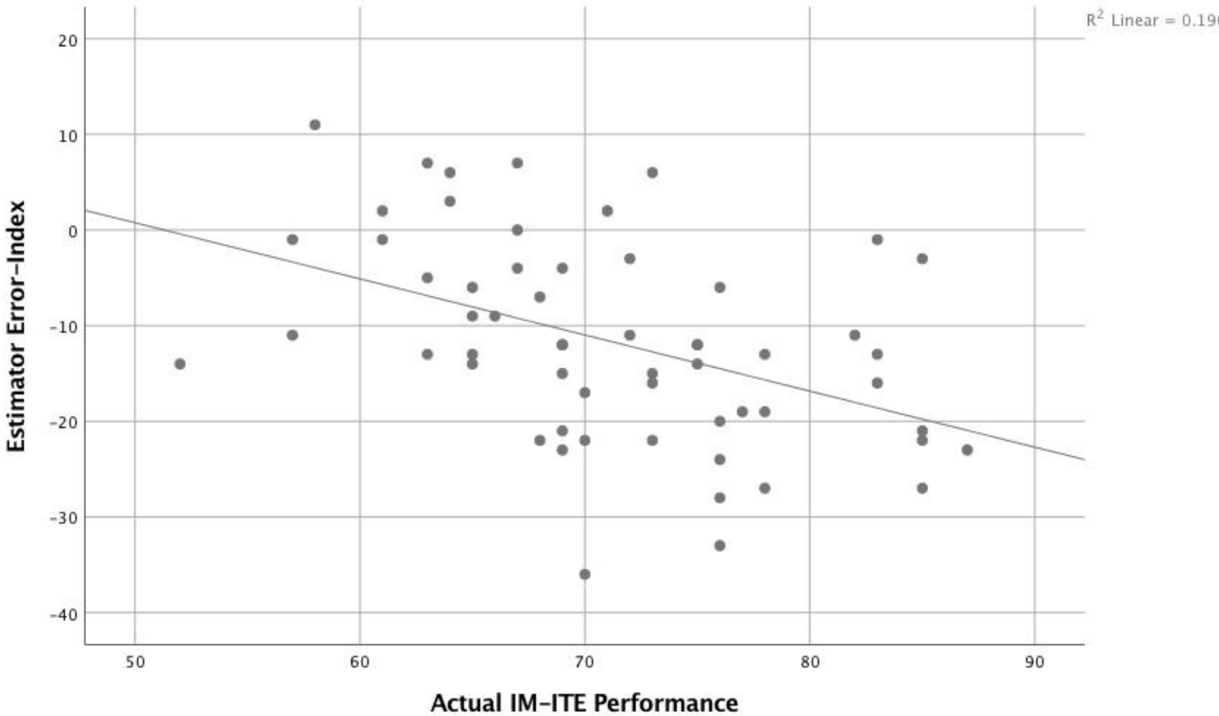


Figure 7

Boxplot for estimator error-index by PGY

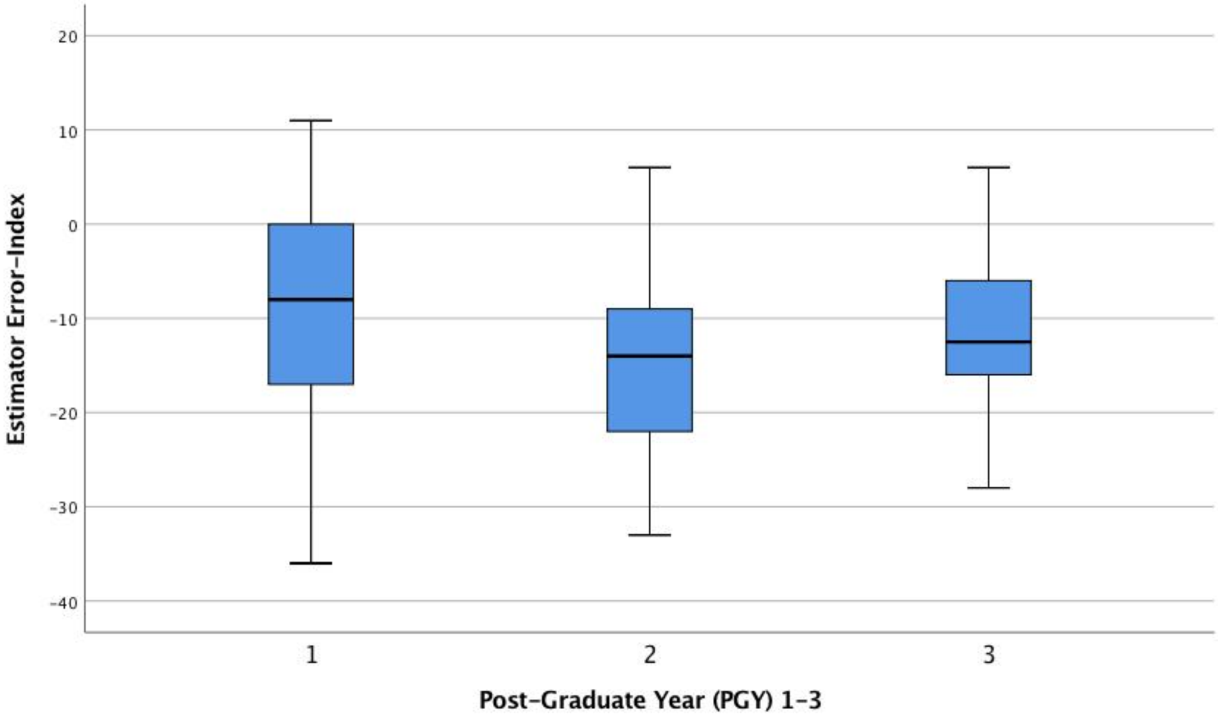


Figure 8

Boxplot for estimator error-index by gender

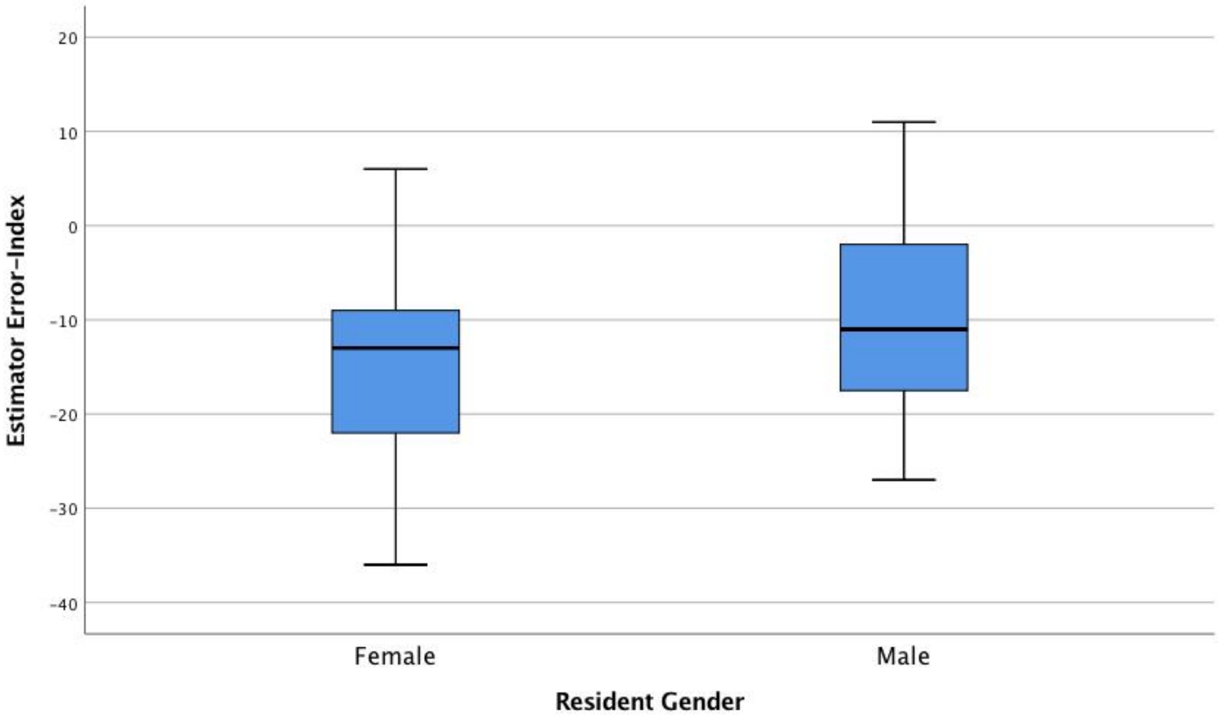


Figure 9

Scatterplot for overestimation frequency and actual IM-ITE

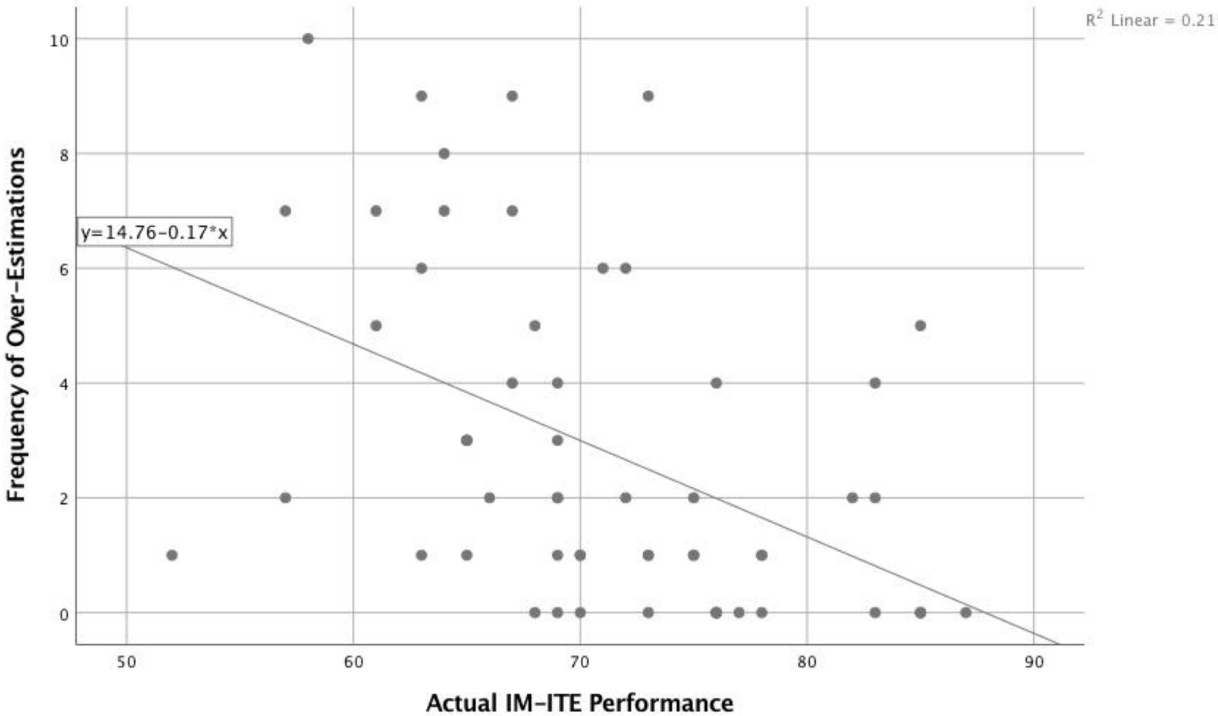
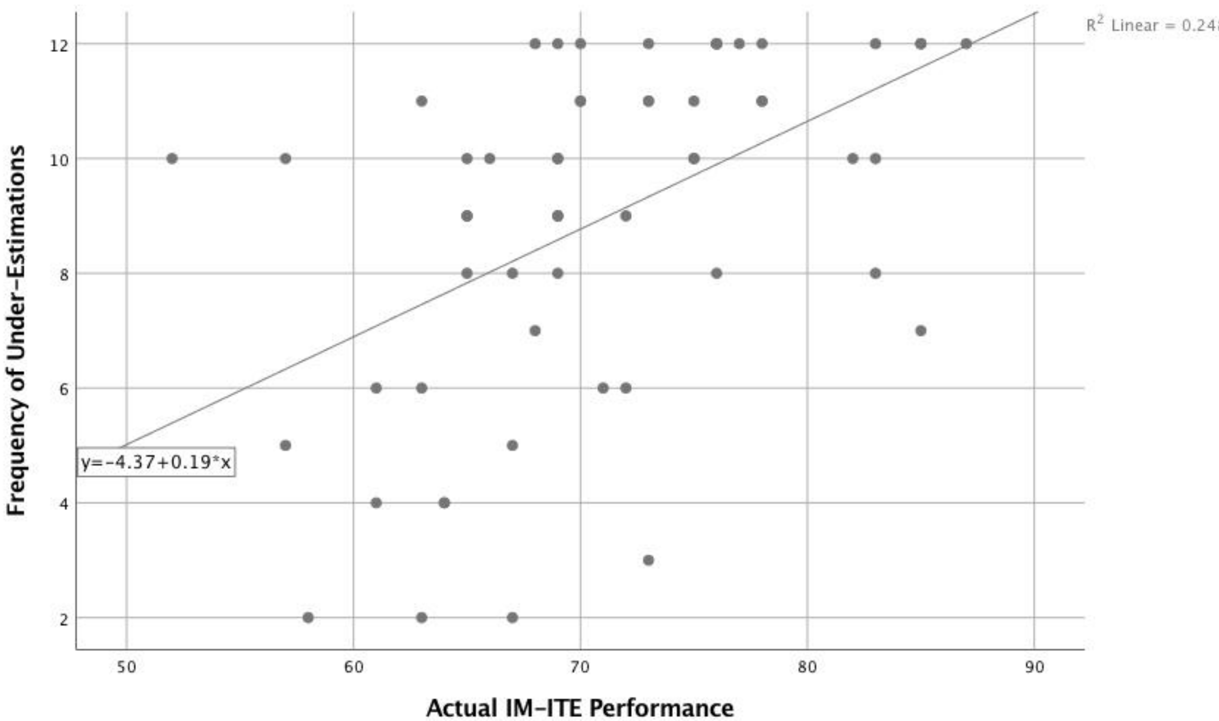
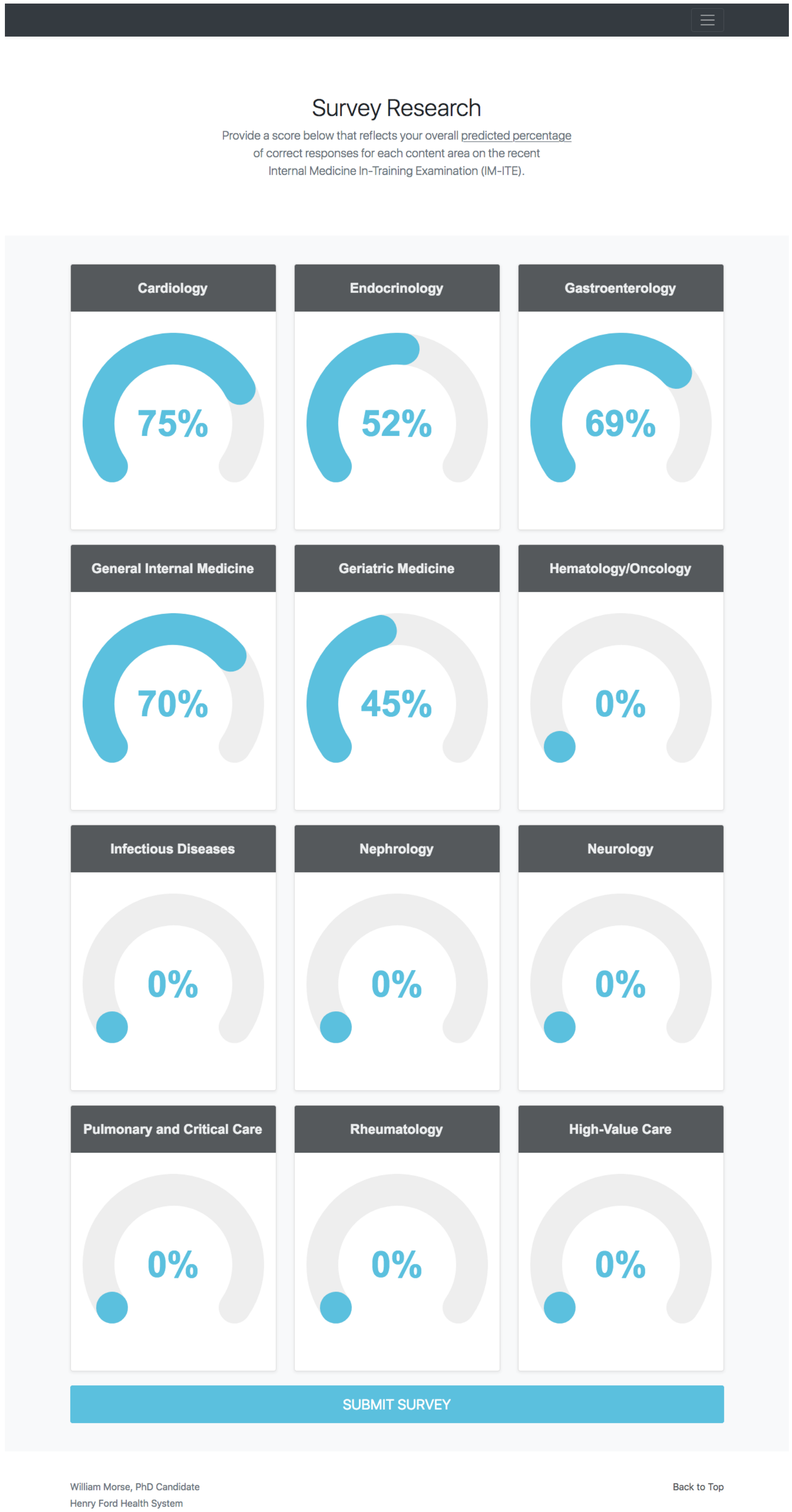


Figure 10

Scatterplot for underestimation frequency and actual IM-ITE



APPENDIX A. SUBJECTIVE ASSESSMENT SURVEY (SAS)


Survey Research
Provide a score below that reflects your overall predicted percentage
of correct responses for each content area on the recent
Internal Medicine In-Training Examination (IM-ITE).
Cardiology
75%
Endocrinology
52%
Gastroenterology
69%
General Internal Medicine
70%
Geriatric Medicine
45%
Hematology/Oncology
0%
Infectious Diseases
0%
Nephrology
0%
Neurology
0%
Pulmonary and Critical Care
0%
Rheumatology
0%
High-Value Care
0%
SUBMIT SURVEY
William Morse, PhD Candidate
Henry Ford Health System
Back to Top

APPENDIX B. EMAIL INVITATION TO PARTICIPATE IN SURVEY

Hello,
I am a doctorate candidate at Wayne State University and an employee of the Henry Ford Health System in graduate medical education. As part of my dissertation research, I am conducting a review of resident estimations of their current levels of knowledge in twelve content areas as measured by the In-Training Examination. A critical part of the review is an online survey to evaluate self-reported perceptions of your beliefs. The survey should take about 10-15 minutes to complete. It is critically important to obtain as many respondents as possible for a successful review process. The analysis of the data is group-focused, not individually focused. Participation is voluntary and if you self-select to contribute your responses, they will be confidential. The use of your email address and the coded identifier that appear in this invitation are necessary to manage the data collection process and will be removed before the analysis of data rendering your participation anonymous when complete. The completion of the survey implies informed consent. Taking part in this study is voluntary. You have the right to choose not to take part in this study. Your decisions will not change any present or future relationship with Wayne State University or Henry Ford Hospital or its affiliates, or other services you are entitled to receive.

Please complete the survey as soon as possible. The deadline to complete the survey is October 18, 2019.

SURVEY LINK: http://meded.hfhs.org/wm/index.php?ucode=abc123

Thank you for participating in this important review process. Additional research informed consent information is available below.

Sincerely,
William Morse

Principal Investigator (PI): William Morse wmorse1@hfhs.org (313) 916-0905 (HFHS OFFICE)
148 Clara Ford Pavilion
Co-Investigator : Odaliz Abreu Lanfranco Director of Internal Medicine Residency Program abreu2@hfhs.org

APPENDIX C. WAYNE STATE UNIVERSITY IRB APPROVAL LETTER

WAYNE STATE UNIVERSITY

IRB Administration Office
87 East Canfield, Second Floor
Detroit, Michigan 48201
Phone: (313) 577-1628
FAX: (313) 993-7122
http://irb.wayne.edu

NOTICE OF EXPEDITED APPROVAL

To: William Morse
Administration & Organization Stud
49241 Monte Road

From: For Dr. Scott Millis or designee D. Bielawski, Ph.D / BC
Chairperson, Behavioral Institutional Review Board (B3)

Date: September 17, 2019

RE: IRB #: 094719B3A

Protocol Title: A study of estimation bias in internal medicine residents and its influence on performance outcomes on the in-training examination

Funding Source:

Protocol #: 1909002518

Expiration Date: September 16, 2020

The above-referenced Administrative Application request for you to use another IRB as the IRB of Record was APPROVED on 09/17/2019. The IRB Administration Office is in receipt of 1) the IRB Approval Letter from the IRB of Record Institution indicating that Wayne State University has been approved as an additional site under their approved protocol 2) Wayne State University's IRB Authorization to Use Another IRB for Protocol Approval Agreement with appropriate signatures; and 3) an Administrative Application appropriately completed and signed by the principal investigator.

- Administrative Application dated 09/10/2019 received in the IRB Office 09/16/2019
- Authorization for WSU researcher to use another IRB of Record
- Name of IRB of Record: Henry Ford Health Systems
- Receipt of Wayne State University IRB Authorization to Use Another IRB for Protocol Approval Agreement signed and dated by PI on 09/13/2019 and signed by IRB Institutional Official 09/17/2019

- Please forward a copy of this approval memo and attached agreement to the collaborating institution's IRB upon receipt.
- All amendments and correspondence should be directed to the collaborating institution's IRB, unless instructed by that institution to send a copy to the WSU IRB Administration Office.
- Yearly Continuation approval from the collaborating institution must be submitted to the WSU IRB Administration Office along with an updated Administrative Application (check the continuation box at the top of the form). This submission should be received at least six weeks before the expiration date.
- Please reference the Protocol # (above) on all communication to the IRB Administration Office related to this research.
- Failure to receive approval for continuation before the expiration date will result in the automatic suspension of the approval of this protocol on the expiration date. Information collected following suspension is unapproved research and can never be reported or published as research data.

Enc: Copy of signed Agreement

Notify the IRB of any changes to the funding status of the above-referenced protocol.

APPENDIX D. IRB AUTHORIZATION BETWEEN INSTITUTES

WAYNE STATE UNIVERSITY

IRB Administration Office
87 E. Canfield, Second Floor
Detroit, MI 48201
Office (313) 577-1628
http://irb.wayne.edu/index.php

Wayne State University IRB Authorization
to Use Another IRB for Protocol Approval Agreement

- This form must be submitted along with the Administrative Application form to the IRB.
- The Officials signing below agree that the Wayne State University may rely on the Collaborating Institution's designated IRB for review and continuing oversight of its human participant research for the following protocol:

Name of Research Project:A study of estimation bias in internal medicine residents and its influence on performance outcomes on the in-training examination	
Name of Collaborating Institution Providing IRB Review	Henry Ford Health System, Henry Ford Hospital
Collaborating Institution's OHRP Federal Wide Assurance	00005846
Principal Investigator at Collaborating Institution	William Morse
Sponsor or Funding Agency for above named protocol	☒ N/A
Award Number for above named protocol	☒ N/A
WSU FWA Number	00002460
Principal Investigator at WSU	William Morse

The review and continuing oversight performed by the Collaborating Institution's designated IRB will meet the human participants' protection requirements of the Wayne State University's OHRP-approved FWA. The IRB at the Collaborating Institution will follow their written policies and procedures. Relevant minutes of IRB meetings will be made available to Wayne State University upon request. Wayne State University remains responsible for ensuring compliance with the IRB's determinations and with the terms of its OHRP- approved Assurance. This document must be kept on file at both institutions and provided to OHRP upon request.

Signature of Principal Investigator at Collaborating Institution — Date: 9/13/2019

Signature of Principal Investigator at Wayne State University — Date: 9/13/2019

Signature of IRB Institutional Official at Collaborating Institution — Date: 9/16/19

Printed Full Name: MARGOT LAPOINTE — Title: VP OF RESEARCH

Signature of IRB Institutional Official at Wayne State University — Date: 9/1?/19

[illegible] Asst VP [illegible]

Form Date: 04/2015

APPENDIX E. HENRY FORD HEALTH SYSTEM IRB APPROVAL LETTER

Henry Ford HEALTH SYSTEM

RESEARCH ADMINISTRATION

Research Administration
Henry Ford Health System
1 Ford Place – 2F
Detroit, MI 48202-2689
(313) 874-4464 Office
(313) 874-4288

To: Odaliz Abreu Lanfranco MD

From: Jonathan Ehrman, Ph.D.
IRB Chair

IRB No.: 13139

Title: A study of estimation bias in internal medicine residents and its influence on performance outcomes on the in training-examination

EXEMPT CATEGORY:	**2**
DETERMINATION SENT:	**August 29, 2019**

On August 28, the Henry Ford Health Systems (HFHS) Institutional Review Board (IRB) reviewed the initial submission for this research study.

The IRB determined that this research study is exempt from federal regulations, pursuant to 45 CFR 46.104 and if applicable, 21 CFR 56.104.

The Investigator reported that 116 subjects will be enrolled for this research study.

The IRB is expected to review all documents and activities that bear directly on the rights and welfare of participants of research. Any revisions to the protocol, including exceeding the number of reported subjects, must be reviewed by the IRB prior to implementation to ensure that the research study continues to be exempt.

This research study is exempt from the requirements of Continuing Review. However, the Investigator must notify the IRB Administration Office when this research is complete to close the study.

This protocol will be presented as an informational item at a subsequent IRB meeting.

Please contact the IRB Administration Office at 313-874-4464 if you have any questions or concerns.

APPENDIX F. PRINCIPLE INVESTIGATOR CHANGE FORM

RESEARCH ADMINISTRATION

Research Administration
Henry Ford Health System
1 Ford Place – 2F
Detroit, MI 48202-2689
(313) 874-4464 Office
(313) 874-4288

To: William Morse
Graduate Medical Education

From: Courtney A. Cloutier, JD
IRB Research Compliance Supervisor

IRB No.: 13139

Title: A study of estimation bias in internal medicine residents and its influence on performance outcomes on the in training-examination

DATE SENT:	**September 9, 2019**

On August 28, 2019, the Henry Ford Health Systems (HFHS) Institutional Review Board (IRB) reviewed the initial submission for this research study, determining that the study was Exempt from the Federal regulations. The initial submission listed Odaliz Abreu Lanfranco, MD, as the lead investigator and William Morse as a Co-Investigator on the submission.

On August 30, 2019, the IRB approved a change in personnel from Dr. Odaliz Abreu Lanfranco as the Principal Investigator to William Morse.

Please contact the IRB Administration Office at 313-874-4464 if you have any questions or concerns.

Courtney A. Cloutier, JD
IRB Research Compliance Supervisor

REFERENCES

ACGME (2016a) ACGME Common Program Requirements, available at: http://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRs_2017-07-01.pdf (accessed 7 March, 2019).

ACGME (2016b) ACGME What We Do, available at: http://www.acgme.org/What-We-Do/Overview (accessed 30 April, 2018).

ACGME (2017) ACGME Data Resource, available at: http://www.acgme.org/About-Us/Publications-and-Resources/Graduate-Medical-Education-Data-Resource-Book (accessed 30 April, 2018).

ACGME (2018a) ACGME Glossary of Terms, available at: https://www.acgme.org/Portals/0/PDFs/ab_ACGMEglossary.pdf (accessed 7 March, 2019).

ACGME (2018b) ACGME Data Resource, available at: https://www.acgme.org/About-Us/Publications-and-Resources/Graduate-Medical-Education-Data-Resource-Book (accessed 29 March, 2019).

ACGME (2019a) ACGME About Us, available at: https://www.acgme.org/About-Us/Overview (accessed 29 March, 2019).

Amabile, T. M. (1993). Motivational synergy: Toward new conceptualizations of intrinsic and extrinsic motivation in the workplace. *Human resource management review*, *3*(3), 185-201.

American Board of Internal Medicine. (2019). About ABIM. https://www.abim.org/about/default.aspx. Last accessed on April 30, 2019.

American Board of Medical Specialties [ABMS]. (2018). Based on Core Competencies. http://www.abms.org/board-certification/a-trusted-credential/based-on-core-competencies/. Last accessed on May 15, 2018.

American College of Physicians. (2019). https://www.acponline.org/featured-products/medical-educator-resources/im-ite#blueprint. Last accessed on April 30, 2019. Philadelphia, PA: American College of Physicians.

Anderson, L. W., & Bourke, S. F. (2000) Assessing affective characteristics in the schools.

Anderson, P. A. (2012). Giving feedback on clinical skills: are we starving our young?. *Journal of graduate medical education*, *4*(2), 154-158.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. Psychological review, 64(6p1), 359.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. Journal of social and clinical psychology, 4(3), 359-373.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of personality and social psychology*, *59*(5), 960.

Beyer, S. (1998). Gender differences in self-perception and negative recall biases. *Sex Roles*, *38*(1-2), 103-133.

Beyer, S. (2002). The effects of gender, dysphoria, and performance feedback on the accuracy of self-evaluations. *Sex Roles*, *47*(9-10), 453-464.

Bing-You, R. G., & Trowbridge, R. L. (2009). Why medical educators may be failing at feedback. *Jama*, *302*(12), 1330-1331.

Bloom, B. S. (1968). Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1. *Evaluation comment*, *1*(2), n2.

Bloom, B. S. (1976). Human characteristics and school learning. McGraw-Hill.

Branch Jr, W. T., & Paranjape, A. (2002). Feedback and reflection: teaching methods for clinical settings. Academic Medicine, 77(12), 1185-1188.

Burke, A. E., Benson, B., Englander, R., Carraccio, C., & Hicks, P. J. (2014). Domain of competence: practice-based learning and improvement. Academic pediatrics, 14(2), S38-S54.

CMS. (2019). CMS Definition of teaching hospital, available at: https://www.cms.gov/OpenPayments/About/Glossary-and-Acronyms.html (accessed 29 March, 2019).

Carroll, John, A model of school learning.Teachers College Record, 1963, 64: 723-733.

Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. Cognitive, Affective, & Behavioral Neuroscience, 7(4), 367-379.

Colbert, C. Y., Graham, L., West, C., White, B. A., Arroliga, A. C., Myers, J. D., ... & Clark, J. (2015). Teaching Metacognitive Skills: Helping Your Physician Trainees in the Quest to ‘Know What They Don't Know’. The American journal of medicine, 128(3), 318-324.

Conley, A. M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. Journal of educational psychology, 104(1), 32.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes.

Conroy, D. E. (2017). Achievement motives. Handbook of Competence and Motivation: Theory and Application, 25.

Cook, D. A., & Artino Jr, A. R. (2016). Motivation to learn: an overview of contemporary theories. Medical education, 50(10), 997-1014.

Covey, S. R. (2013). *The 7 habits of highly effective people: Powerful lessons in personal change*. Simon and Schuster.

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and individual differences*, *21*(4), 337-346.

Croskerry, P. (2003). The importance of cognitive errors in diagnosis and strategies to minimize them. Academic medicine, 78(8), 775-780.

Croskerry, P. (2009). Clinical cognition and diagnostic error: applications of a dual process model of reasoning. Advances in health sciences education, 14(1), 27-35.

Davis, D. A., Mazmanian, P. E., Fordis, M., Van Harrison, R. T. K. E., Thorpe, K. E., & Perrier, L. (2006). Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *Jama*, *296*(9), 1094-1102.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. Journal of research in personality, 19(2), 109-134.

Deci, E. L., Ryan, R. M., & Williams, G. C. (1996). Need satisfaction and the self-regulation of learning. *Learning and individual differences*, *8*(3), 165-183.

Desender, K., Van Opstal, F., & Van den Bussche, E. (2014). Feeling the conflict: The crucial role of conflict experience in adaptation. *Psychological science*, *25*(3), 675-683.

DeVellis, R. F. (2017). *Scale development: Theory and applications* (Fourth Edition). Sage publications.

Dine, C. J., Ruffolo, S., Lapin, J., Shea, J. A., & Kogan, J. R. (2014). Feasibility and validation of real-time patient evaluations of internal medicine interns' communication and professionalism skills. *Journal of graduate medical education*, *6*(1), 71-77.

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In Advances in experimental social psychology (Vol. 44, pp. 247-296). Academic Press.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. Annual review of psychology, 53(1), 109-132.

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

Elliot, A. J., Dweck, C. S., & Yeager, D. S. (Eds.). (2017). *Handbook of competence and motivation: Theory and application*. Guilford Publications.

El Saadawi, G. M., Azevedo, R., Castine, M., Payne, V., Medvedeva, O., Tseytlin, E., ... & Crowley, R. S. (2010). Factors affecting feeling-of-knowing in a medical intelligent tutoring system: the role of immediate feedback as a metacognitive scaffold. *Advances in Health Sciences Education*, *15*(1), 9-30.

Ende, J. (1983). Feedback in clinical medical education. Jama, 250(6), 777-781.

Englander, R., Frank, J. R., Carraccio, C., Sherbino, J., Ross, S., Snell, L., & ICBME Collaborators. (2017). Toward a shared language for competency-based medical education. Medical teacher, 39(6), 582-587.

Epstein, R. M., & Hundert, E. M. (2002). Defining and assessing professional competence. Jama, 287(2), 226-235.

Eva, K. W., & Regehr, G. (2007). Knowing when to look it up: a new conception of self-assessment ability. Academic Medicine, 82(10), S81-S84.

Feather, N. T. (1992). Values, valences, expectations, and actions. Journal of Social Issues, 48(2), 109-124.

Flexner, A., Pritchet, H., & Henry, S. (1910). Medical education in the United States and Canada bulletin number four (The Flexner Report). New York (NY): The Carnegie Foundation for the Advancement of Teaching.

Frank, J. R., Snell, L. S., Cate, O. T., Holmboe, E. S., Carraccio, C., Swing, S. R., ... & Harden, R. M. (2010). Competency-based medical education: theory to practice. Medical teacher, 32(8), 638-645.

Gable, R. K., & Wolf, M. B. (1993). Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings (2nd Edition). Springer Science & Business Media.

Gailliot, M. T., Mead, N. L., & Baumeister, R. F. (2008). Self-regulation. *Handbook of personality: Theory and research*, 472-491.

Garibaldi, R. A., Subhiyah, R., Moore, M. E., & Waxman, H. (2002). The in-training examination in internal medicine: an analysis of resident performance over time. Annals of internal medicine, 137(6), 505-510.

Garrett, J., Alman, M., Gardner, S., & Born, C. (2007). Assessing students' metacognitive skills. *American Journal of Pharmaceutical Education*, *71*(1), 14.

Goleman, D. P. (1995). Emotional intelligence: Why it can matter more than IQ for character, health and lifelong achievement.

Graber, M., Gordon, R., & Franklin, N. (2002). Reducing diagnostic errors in medicine: what's the goal? *Academic Medicine*, *77*(10), 981-992.

Graber, M. L., Franklin, N., & Gordon, R. (2005). Diagnostic error in internal medicine.

Archives of internal medicine, *165*(13), 1493-1499.

Grant, A. M. (2008). Does intrinsic motivation fuel the prosocial fire? Motivational synergy in predicting persistence, performance, and productivity. *Journal of applied psychology*, *93*(1), 48.

Grossman, R. S., Fincher, R. M. E., Layne, R. D., Seelig, C. B., Berkowitz, L. R., & Levine, M. A. (1992). Validity of the in-training examination for predicting American Board of Internal Medicine certifying examination scores. Journal of general internal medicine, 7(1), 63-67.

Guerra-López, I. J. (2008). *Performance evaluation: Proven approaches for improving program and organizational performance* (Vol. 21). John Wiley & Sons.

Harter, S. (1982). The perceived competence scale for children. *Child development*, 87-97.

Hawkins R.E., Sumption K.F., Gaglione M.M., Holmboe E.S. (1999). The in-training examination in internal medicine: resident perceptions and lack of correlation between resident scores and faculty predictions of resident performance. *Am J Med.* 1999;106(2):206–10.

Holmboe, E. S., Edgar, L., & Hamstra, S. (2016). The milestones guidebook. *American Council of Graduate Medical Education*. http://www.acgme.org/Portals/0/MilestonesGuidebook.pdf

Holmboe, E. S., & Hawkins, R. E. (1998). Methods for evaluating the clinical competence of residents in internal medicine: a review. *Annals of internal Medicine*, *129*, 42-48.

Jones R., Panda M., Desbiens N. (2008). Internal medicine residents do not accurately assess their medical knowledge. *Advances in Health Sciences Education: Theory and Practice*, 13(4), 463–468.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin*, *119*(2), 254.

Knowles, M. S. (1975) Self-Directed Learning. A guide for learners and teachers, Englewood Cliffs: Prentice Hall/Cambridge. 135 pages. Programmatic guide that is rather objective oriented. Sections on the learner, the teacher and learning resources.

Kornell, N., & Bjork, R. A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning. Journal of Experimental Psychology: General, 138(4), 449.

Kreitner, R. (2009). Principles of Management, 11th. *Sorth—Western, a part of Cengage Learning*.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. Journal of personality and social psychology, 77(6), 1121.

Kusurkar, R. A., Ten Cate, T. J., Van Asperen, M., & Croiset, G. (2011). Motivation as an independent and a dependent variable in medical education: a review of the literature. Medical teacher, 33(5), e242-e262.

Leach, D. C. (2002). Competence is a habit. Jama, 287(2), 243-244.

Little, D., & Erickson, G. (2015). Learner identity, learner agency, and the assessment of language proficiency: Some reflections prompted by the common european framework of reference for languages. *Annual Review of Applied Linguistics*, *35*, 120-139.

Mamede, S., van Gog, T., van den Berge, K., Rikers, R. M., van Saase, J. L., van Guldener, C., & Schmidt, H. G. (2010). Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. *Jama*, *304*(11), 1198-1203.

McCoach, D. B., Gable, R. K., & Madura, J. P. (2013). Instrument development in the affective domain. New York, NY: Springer. doi, 10, 978-1.

Medin. D., Ross, B., Markman, A. (2001). Cognitive Psychology 3rd Ed. Harcourt, Inc. Orlando, Fl.

Miller, J., & Katz, D. (2018). Gender differences in perception of workplace experience among anesthesiology residents. *The journal of education in perioperative medicine: JEPM*, *20*(1).

Mitchell, T. R., & Daniels, D. (2003). Observations and commentary on recent research in work motivation. Motivation and work behavior, 7(1), 225-254.

Mohzan, M. A. M., Hassan, N., & Halil, N. A. (2013). The influence of emotional intelligence on academic achievement. *Procedia-Social and Behavioral Sciences*, *90*, 303-312.

Mowery, Y. M. (2015). A primer on medical education in the United States through the lens of a current resident physician. Annals of translational medicine, 3(18).

Nasca, T. J., Philibert, I., & Brigham, T. (2012). The next GME accreditation system. *The New England Journal of Medicine*, *366*(24), 2325.

Nelis, D., Quoidbach, J., Mikolajczak, M., & Hansenne, M. (2009). Increasing emotional intelligence:(How) is it possible?. *Personality and individual differences*, *47*(1), 36-41.

Nevid, J. (2013). Psychology: Concepts and applications. Belmont, CA: Wadworth.

Nousiainen, M. T., Caverzagie, K. J., Ferguson, P. C., Frank, J. R., & ICBME Collaborators. (2017). Implementing competency-based medical education: What changes in curricular structure and processes are needed?. *Medical teacher*, *39*(6), 594-598.

Pinder, C. C. (2014). *Work motivation in organizational behavior*. Psychology Press.

Powell, D. E., & Carraccio, C. (2018). Toward competency-based medical education. New England Journal of Medicine, 378(1), 3-5.

Questienne, L., Van Opstal, F., van Dijck, J. P., & Gevers, W. (2016). Metacognition and cognitive control: behavioural adaptation requires conflict experience. *The Quarterly Journal of Experimental Psychology*, 1-15.

Redelmeier, D. A. (2005). The cognitive psychology of missed diagnoses. *Annals of internal medicine*, *142*(2), 115-120.

Rokeach, M. (1979). Understanding human values: Individual and Societal. Simon and Schuster.

Rollins, L. K., Martindale, J. R., Edmond, M., Manser, T., & Scheld, W. M. (1998). Predicting pass rates on the American Board of Internal Medicine certifying examination. Journal of general internal medicine, 13(6), 414-416.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American psychologist, 55(1), 68.

Ryan, R. M., & Deci, E. L. (2008). Self-determination theory and the role of basic psychological needs in personality and the organization of behavior. Handbook of personality: Theory and research, 3, 654-678.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. *Psychological bulletin*, *95*(3), 355.

Schunk, D. H. (2012). Learning theories an educational perspective sixth edition. Pearson.

Schutte, T., Tichelaar, J., Dekker, R. S., Thijs, A., De Vries, T. P., Kusurkar, R. A., ... & Van Agtmael, M. A. (2017). Motivation and competence of participants in a learner-centered student-run clinic: an exploratory pilot study. BMC medical education, 17(1), 23.

Shaw, K., C. Cassel, C. Black, and W. Levinson. 2009. Shared medical regulation in a time of increasing calls for accountability and transparency: Comparison of recertification in the United States, Canada, and the United Kingdom. *JAMA* 302(18):2008-2014.

Shimamura, A. P. (2000). Toward a cognitive neuroscience of metacognition.

Skinner, B. F. (1954). "The Science of Learning and the Art of Teaching," Harvard Educational Review, 24:2 (Spring 1954), pp. 86-97.

Skinner, B. F. (1958). Teaching machines. *Science*, *128*(3330), 969-977.

Spector, J. M., & Yuen, A. H. (2016). *Educational technology program and project evaluation*. Routledge.

Tafarodi, R. W., & Swann Jr, W. B. (1995). Self-linking and self-competence as dimensions of global self-esteem: initial validation of a measure. Journal of personality assessment, 65(2), 322-342.

Thorndike, E. L. (1913). *Educational psychology* (Vol. 2). Teachers college, Columbia university.

Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding feedback: A learning theory perspective. *Educational Research Review*, *9*, 1-15.

Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual review of psychology*, *21*(1), 437-484.

US Census Bureau. (2017). QuickFacts: Michigan; United States, available at: https://www.census.gov/quickfacts/fact/table/mi,US/PST045217 (accessed 29 March, 2019).

Van de Ridder, J. M., Stokking, K. M., McGaghie, W. C., & Ten Cate, O. T. J. (2008). What is feedback in clinical education?. *Medical education*, *42*(2), 189-197.

Van den, V., van Dulmen, S., van Dalen, J., Scherpbier, A. J., & van der Vleuten, C. P. (2013). Learning in context: Identifying gaps in research on the transfer of medical communication skills to the clinical workplace. *Patient education and counseling*, *90*(2), 184-192.

Vroom, V. H. (1964). *Work and motivation* (Vol. 54). New York: Wiley.

Waxman, H., Braunstein, G., Dantzker, D., Goldberg, S., Lefrak, S., Lichstein, E., ... & Schiffman, F. (1994). Performance on the internal medicine second-year residency in-training examination predicts the outcome of the ABIM certifying examination. Journal of general internal medicine, 9(12), 692-694.

Webb, A. R., Young, R. A., & Baumer, J. G. (2010). Emotional intelligence and the ACGME competencies. *Journal of Graduate Medical Education*, *2*(4), 508-512.

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. Educational psychology review, 6(1), 49-78.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary educational psychology*, *25*(1), 68-81.

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2006). Development of achievement motivation. Hoboken.

Wigfield, A., Rosenzweig, E. Q., & Eccles, J. S. (2017). Achievement Values. Handbook of competence and motivation: Theory and application, 116.

ABSTRACT

A STUDY OF SELF-EVALUATION AND ESTIMATOR ERROR IN INTERNAL MEDICINE RESIDENTS AND ITS INFLUENCE ON THE ACADEMIC ACHIEVEMENT OF MEDICAL KNOWLEDGE

by

WILLIAM MORSE

May 2020

Advisor: Dr. Ingrid Guerra-López

Major: Learning Design and Technology

Degree: Doctor of Philosophy

The main objective of this study is to address the research question—how does a sample of internal medicine residents self-evaluate and what, if any, influence does this have on their academic achievement of medical knowledge? The research builds on and extends our understanding of the self-evaluation process through the moderating effects between competency-related beliefs and academic achievement of medical knowledge. The study design is quantitative, cross-sectional survey research using a non-random sample of 58 internal medicine residents at Henry Ford Hospital in Detroit, Michigan. The primary sources for data collection included a study survey that measured competency-related beliefs as a subjective assessment of predicted performance in combination with the Internal Medicine In-Training Examination (IM-ITE), the objective resource. Findings indicate that residents are not very accurate with their self-evaluations. Residents, as a whole, tend to underestimate their performance. A gender difference was exhibited where female residents predicted their performance significantly lower than their male colleagues. Most significantly, actual performance on the IM-ITE was significantly influenced by the relationship with estimation error.

AUTOBIOGRAPHICAL STATEMENT

This student does not wish to disclose additional personal or professional information under the Family Educational Rights and Privacy Act (FERPA) of 1974 (20 U.S.C. § 1232g; 34 CFR Part 99).