

ABSTRACT

WHICKER, SHARI ANN. Medical Resident Ability to Self-Assess in Relation to their Competence within the Assessed Domain. (Under the direction of James E. Bartlett II, PhD).

The purpose of the study was to examine the relationship between resident communication skills competence and their ability to self-assess within that domain. Six research questions guided the study.

The first research question sought to explore the varied levels with which residents demonstrated competence within the domain of communication skills as measured by Standardized Patient Parents (SPPs) using the Breaking bad news Assessment Schedule (BAS) instrument. The second research question examined how residents' self-assessments corresponded with Standardized Patient Parent assessment of their communication skills performance. The third research question asked how resident communication skill competence related to their ability to self-assess within the domain. The fourth question explored how resident gender related to their ability to self-assess their communication skills performance. The fifth question explored how the number of years of residency training experience related to residents' ability to self-assess their communication skills. The sixth research question explored the amount of significant variance in resident ability to self-assess their communication skills that could be explained by gender, year in training, ethnicity, and theoretical element of communication skills competence.

The methodology implemented was a non-experimental, cross-sectional research design. A purposeful sample of Duke pediatric residents was used for this study. Each resident participated in a standardized patient parent (SPP) scenario in which they were asked to break bad news to a set (mother and father) of SPPs. Immediately subsequent to the

scenario, the residents were asked to self-assess their performance using the Breaking Bad News Assessment Schedule (BAS) and both SPPs were asked to assess each resident's performance as well.

Data analysis techniques included computation of means, standard deviations, frequencies, and percentages for the descriptive data. Chi Square tests, paired t-tests, and a multiple regression were also used to answer the six research questions.

Findings indicated that resident communication skill competence varied considerably from very low on the scale used to very high. Resident self-assessed scores also varied considerably, though by a narrower margin. When looked at overall, resident self-assessed scores were shown not to correlate with SPP scores.

Using tertiles, it was clearer that residents who were scored by the SPPs to have the least communication skills competence (bottom tertile) overestimated their performance, residents in the middle tertile were generally accurate in their self-assessments, and those in the top tertile underestimated their performance. Therefore, Kruger and Dunning's (1999) self-assessment framework was supported by the findings of this study. While residents in each tertile were shown to be somewhat inaccurate in their self-assessment, the greatest concern remains for those who overestimated, particularly those in the bottom tertile who overestimated to the greatest degree.

Of the variables gender, level of residency training, ethnicity, and four theoretical elements to effectively breaking bad news, none were shown to have an impact on residents' ability to self-assess their communication skills. One theoretical element, *Breaking the News*, was shown to explain 35% of the variance in residents' ability to self-assess.

General recommendations from this study include streamlining enhanced communication skills training into the framework of the existing residency curriculum. This enhancement should be implemented for all residents regardless of gender, training level, or ethnicity to improve their communication skills. Based on the literature and the corroborating results of this study, this should, in turn, enhance the residents' ability to more effectively self-assess their communication skills.

Research recommendations include efforts to determine the most effective and efficient means of streamlining communication skills training into the existing residency training curriculum. Furthermore, studies should be done for additional knowledge and skill domains to determine if Kruger and Dunning's self-assessment framework is generalizable to residency domains outside that of communication skills.

© Copyright 2012 by Shari Ann Whicker

All Rights Reserved

Medical Resident Ability to Self-Assess in Relation to their
Competence within the Assessed Domain

by
Shari Ann Whicker

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Education

Adult and Community College Education

Raleigh, North Carolina

2012

APPROVED BY:

James E. Bartlett II, PhD
Committee Chair

Diane D. Chapman, EdD

Colleen O'Connor Grochowski, PhD

James Kiwanuka-Tondo, PhD

DEDICATION

If the family were a boat, it would be a canoe that makes no progress unless everyone paddles. Letty Cottin Pogrebin

This dissertation is dedicated to my incredibly dedicated and supportive family who would never let me give up, even if I wanted to. I could never have gotten here without you all.

To my husband, Chris, who deserves an honorary doctorate for his unwavering support throughout our eleven years of marriage...and two graduate degrees. You have been my perfectly shaped complementary rock when I've been up against the hardest of places.

To my sweet children, Lynden and Leyton, who have been unbelievable troopers and "let mommy focus!" when they really wanted mommy to play. Lynden, I will forever treasure the nights you cuddled in the big yellow chair, reading your book close to me while I worked at my desk. Leyton, I will never forget your periodic kisses, hugs, and "just ONE more question(s)" you peppered me with nightly as I stared blankly at the computer, just when I needed them most.

Mom and dad, I cannot imagine more supportive parents. You both have inspired me to persevere beyond challenges. Thank you for the awesome family structure and hard work you have modeled year after year. No, dad, my degree does not mean that I can (or will!) assess/cure your ailments. You can always, however, take two and call me in the morning, noon, or night.

To my siblings, Carolyn, Dennis, and Liz who cheered me up through the thick of it. You all are so wonderful and accomplished in so many ways and continue to inspire me daily.

I pray I can make good on all of my promises to be a better wife, mom, daughter, sister, friend “after the craziness ends.” I love you all dearly and absolutely could not have done this without you.

BIOGRAPHY

Shari Whicker was born in Niagara Falls, NY and then moved to upstate New York, Ohio, and Michigan. After spending ten years in Michigan, she moved to North Carolina in 1997.

In North Carolina, Shari completed her undergraduate studies with a focus on Elementary Education and English at the University of North Carolina at Greensboro (UNCG). From 1999-2004, Shari developed an interest in teaching adults while working for Wake Forest University; she trained various health care providers and support staff on the medical computer systems and eventually moved into the role as the Training and Development Manager. Her focus on adults as learners led Shari to seek her Master's of Education (MEd) in Training and Development from North Carolina State University (NCSU).

In 2004, Shari joined the Pediatric staff at Duke University Medical Center, initially as the Pediatric Residency Training Program Coordinator and currently as the Education Administrator for the Office of Pediatric Education. In this role, Shari functions as the education specialist, working closely with the Vice Chair of Education to provide curricular guidance and oversight for all Duke pediatric educational programs, related to medical student, resident, and fellow training programs, along with faculty development.

Shari serves on several departmental, institutional, and multi-institutional committees. She is the Co-Chair for the Duke Residency Curriculum Committee and serves on mentoring committees for faculty and trainees interested in pursuing educator-related pathways. She

also provides instruction and counseling to residents, fellows, and faculty in many areas, including time management, learning styles, as well as study and teaching skills.

Furthermore, she is active in the area of medical education research and has co-authored multiple medical education articles and presented posters and workshops at national conferences.

Shari's passion for adult learning and desire to further develop her knowledge and skills in the field led her to pursue her doctoral studies in Adult Learning at NCSU beginning in 2008. Throughout her doctoral experience, Shari has developed a particular interest in pediatric resident communication skills and their ability to self-assess those skills.

Above all, Shari's greatest joy is her family. Strong, loving, generous Chris; smart, kind, angelic Lynden; and spirited, lovable, hilarious Leyton make every minute worth it all.

ACKNOWLEDGMENTS

I must first acknowledge my incredibly dedicated and supportive committee. James B., where would I be if I hadn't found you? Thank you dearly for taking me under your wing, providing clarity, and helping to sort through my ignorance. I know so much more because of you. I will remember fondly (and painfully) the many hours you sat with me and painstakingly sorted through (literally) thousands of data points.

Diane, I definitely would never be here without you. You were outstanding as an advisor throughout my master's program experience and saved me from drowning in my doctoral experience. You will forever hold a special place in my heart.

Colleen, your unrelenting dedication to advocating for others and practical, no-nonsense attitude have made you invaluable to me (and to so many others). You are an outstanding mentor and leader. Thank you so much for believing in me.

James K., my independent study with you was incredibly valuable in that you pushed me to dig deep into the literature to challenge myself and what I knew (and did not). You truly helped me to grow.

Deborah and Alisa, thank you so very much for all of your incredible guidance and support throughout the toughest of times. You knew I would get here (even when I had my doubts!).

Aditee, my partner/my committee! Your tireless dedication to medical education and research have encouraged me to do my very best to keep up with you. Thank you for your support, guidance, and patience along the way.

Kammy, I am still continually amazed at how you make it possible to dedicate 100% to each of the millions of things you do. Thank you for being a wonderful boss, advocate, friend, and confidant. Your flexibility and support have been truly remarkable.

I am humbled by you all. Thank you so very much.

TABLE OF CONTENTS

LIST OF TABLES	xiv
LIST OF FIGURES	xv
CHAPTER ONE: INTRODUCTION.....	1
Self-Regulated Learning for Professionals	4
Self-Assessment for Professionals	7
Physician Self-Assessment	10
Physician Communication Skills	12
Theoretical Framework	14
Statement of the Problem.....	16
Purpose Statement.....	18
Research Questions	19
Significance.....	20
Research & Theory	20
Practice & Policy	21
Delimitations.....	21
CHAPTER TWO: REVIEW OF THE LITERATURE	23
Introduction.....	23
Self-Regulated Learning for Professionals	26
Metacognition	29
Metacognitive Monitoring	30

Metacognitive Control	31
Self-Assessment for Professionals	32
Kruger & Dunning’s Theoretical Framework of Self-Assessment	35
Physician Self-Assessment	42
Context of Medical Education	44
Self-Assessment in Residency	48
Resident ability to self-assess as a static skill	51
Gender and Ethnicity	54
Explicit self-assessment interventions	55
Competence begets self-assessment ability	62
Communication Skills Assessment Tools.....	67
Summary	71
CHAPTER THREE: METHODOLOGY	74
Research Design.....	74
Research Questions	75
Variables Used in Study.....	76
Sample Site and Sample.....	77
Instrument	79
Validity and Reliability	81
Scale modification	83
Data Collection and Confidentiality	84

Preliminary Data Analysis85

 Univariate Distribution86

 Multivariate Assumptions86

Statistical Data Analysis87

Summary89

CHAPTER FOUR: DATA ANALYSIS.....90

 Research Questions90

 Instrument91

 Setting the Scene92

 Breaking the News92

 Eliciting Concerns93

 Information Giving93

 General Information93

 Demographics94

 Research Question One: Resident Communication Skills Domain Competence94

 Research Question Two: Resident Self-Assessment Ability96

 Research Question Three: Relationship between Resident Self-Assessment Ability and
their Communication Skills Domain Competence99

 Research Question Four: Relationship between Resident Gender and their Self-Assessment
Ability102

Research Question Five: Relationship between Resident Level of Training Experience and their Self-Assessment Ability	103
Research Question Six: Variance in Resident Ability to Self-Assess Explained by Demographic Variables, Gender, Year in Training, Ethnicity, and Theoretical Element of Communication Skills Competence.....	105
Summary	107
CHAPTER FIVE: SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	109
Chapter One	110
Chapter Two.....	110
Chapter Three.....	113
Chapter Four	114
Chapter Five	114
Discussion of Findings and Conclusions	114
Findings and Conclusions by Research Question.....	115
Research Question One.....	115
Conclusion One.....	115
General Recommendation One	116
Recommendation for Future Research One	116
Research Question Two	116
Conclusion Two.....	117
General Recommendation Two	118

Recommendation for Future Research Two	118
Recommendation for Future Research Three	118
Research Question Three	118
Conclusion Three	120
General Recommendation Three	121
Recommendation for Future Research Four	121
Research Question Four	121
Conclusion Four	122
General Recommendation Four	122
Recommendation for Future Research Five	122
Research Question Five	122
Conclusion Five	122
General Recommendation Five	123
Research Question Six	123
Conclusion Six	123
General Recommendation Six	123
Recommendation for Future Research Six	124
Limitations	124
Sample	124
Data Collection	125
Assessed Domain	125

Instrument Used	125
References	126
APPENDICES	144
Appendix A: Revised BAS for Self-Assessment.....	145
Appendix B: Duke IRB Form	150
Appendix C: Duke IRB Exemption Approval	151
Appendix D: North Carolina State University IRB Exemption Approval	152

LIST OF TABLES

Table 1	Independent and Dependent Variables Used for this Study	77
Table 2	Resident Participant Demographics	78
Table 3	SPP Descriptive Statistics	95
Table 4	Tertile Distribution of Resident Communication Skills Competence	96
Table 5	Resident Self-Assessment Descriptive Statistics	97
Table 6	Tertile Distribution of Resident Self-Assessment Scores	98
Table 7	Impact of Domain Competence Tertile on Degree of Self-Assessment Ability...101	
Table 8	Gender Impact on Self-Assessment Accuracy.....103	
Table 9	Level of Training Impact on Self-Assessment Accuracy	104
Table 10	Stepwise Multiple Regression Analysis to Explore if Level, Gender, Ethnicity, or Theoretical Breaking Bad News Sections Impacted Residents' Self-Assessment Ability106	

LIST OF FIGURES

Figure 1 Comparison between SPP (competence) and Self-Assessed Scores99

CHAPTER ONE: INTRODUCTION

Society depends on the expertise of professionals for the knowledge and skills that guide the activities of our daily lives. It has long been known that, “the professions have become essential to the very functioning of our society” (Schön, 1983, p.3). It is professionals to whom we turn “for the definition and solution of our problems, and it is through them that we strive for social progress” (p. 4). We rely on these professionals to maintain the highest levels of knowledge and skills in their areas of expertise.

For professionals to maintain a high level of expertise, they must be competent in their ability to recognize both the strengths of their knowledge and skills and the gaps where improvement and gains can still be made. Professionals who are unable to recognize points at which their knowledge or skills may be inadequate are ultimately underprepared to address those inadequacies (Schwartz & Perfect, 2002; Violato & Lockyer, 2006). Lack of effective self-assessment will only perpetuate their inability to improve their skills and knowledge base in the future. Otherwise preventable errors are likely to occur when professionals’ knowledge and skills have not been adequately developed (Graber, 2005; Mamede, Schmidt & Rikers, 2007).

Healthcare professionals, in particular, are granted extraordinary rights and privileges within our ever-evolving societal context (Schön, 1983). We depend upon physicians, for example, to maintain and continually update their own medical knowledge and expertise in order to provide safe and effective patient care (Eva, Cunningham, Reiter, Keane & Norman, 2004). Effective self-regulated learning is the means by which physicians are able to

maintain a high level of knowledge and skills within their area of expertise. Therefore, a main element of the preparatory professional training for physicians is ensuring that physicians are equipped with the skills to become effective self-regulated learners (Gordon et al., 2000).

Overconfident physicians are said to contribute significantly to the hundreds of thousands of serious injuries and deaths each year that result from preventable medical errors (Graber, 2005; Mamede et al., 2007). If physicians are unable to recognize their ineptitude in a given area, they are unlikely to seek assistance when it is needed, and the outcomes could be disastrous. As a result, they may fail to order appropriate tests or they might prescribe medications that counteract with other medications the patient is taking. Ineffective self-assessors might also incorrectly perform a procedure or provide inappropriate information regarding a patient's condition.

This chapter will provide a foundational discussion on the importance of self-assessment for professionals. An overview of the theory of self-regulated learning and metacognition will be discussed in relation to the development of professionals' ability to perform a self-assessment of their knowledge and skills. This chapter will provide the framework that emphasizes the integral nature of effective self-assessment within professional practice. An additional emphasis will be put on professionals, particularly physicians, developing the ability to effectively self-assess during their professional preparatory training.

As will be noted in the later discussion, there continues to be differing opinions about the key elements for developing self-assessment ability. Two different schools of thought regarding the development of physician self-assessment ability during their preparatory professional (residency) training will be discussed. Reasoning for the decision to apply Kruger and Dunning's theoretical framework to residents within the domain of communication skills (CS) will be explained. The importance of residents developing communication skills and the ability to self-assess within the domain of communication skills will also be discussed within this chapter. Each section within this chapter will lay the groundwork for the value of this study.

Self-regulated learners demonstrate “personal initiative, perseverance, and adaptive skill” in pursuing their learning (Zimmerman, 2001, p. 1). As a part of this process, self-regulated learners practice active metacognition (Pintrich & Zusho, 2002; Zimmerman, 1990). Active metacognition involves ongoing monitoring and control (Nelson & Narens, 1990) of one's own cognition and it is through these two elements of metacognition that self-regulated learners identify and address their learning needs.

Identifying one's own strengths and weaknesses through self-assessment during the monitoring process is what guides self-regulated learners to identify the learning needs which should be addressed within the control process. Physicians with effective self-assessment ability are able to recognize their own strengths and weaknesses to determine appropriate areas on which to focus their continued learning development. Without effective self-assessment practices, metacognitive control would be invalidated (Schwartz & Perfect, 2002;

Violato & Lockyer, 2006). Without effective self-assessment ability, physicians are unable to recognize which weaknesses need to be addressed. Physician self-assessment is a significant component of their professional medical practice because the incremental development of deficits in knowledge and skill is instrumental to quality patient care (Wooliscroft, TenHaken, Smith & Calhoun, 1993).

Self-Regulated Learning for Professionals

Self-regulated learning refers to the “self-directive process through which learners transform their mental abilities into task-related academic skills” (Zimmerman, 2001, p. 1). Pintrich and Zusho (2002) define self-regulated learning as an active process whereby learners monitor and control their own learning. This process has been emphasized throughout the literature as a critical component of professional development (Pintrich & Zusho, 2002; Zimmerman, 2001). The independent nature of professional practice requires practitioners to be self-regulated in the way in which they approach their learning (Boud, 1995; Eva et al., 2004). The structured guideline for the learning development for professionals ends once they complete their formal preparatory education and training. It is the professionals themselves who are predominantly responsible for their continued learning and professional development throughout their careers (Lajoie, 2008; Schein, 1972).

The monitoring and controlling of self-regulated learning are essential to professional learning development (Eva et al., 2004; Zimmerman, 2001) and are also the two interrelated components of metacognitively active learning (Nelson & Narens, 1990; Perfect & Schwartz, 2002). Throughout the literature, the concept of metacognition has been described as the

function of two distinct but influencing elements: metacognitive monitoring and metacognitive control. Together, these elements both directly influence comprehension and development for self-regulated learners (Hacker, Dunlosky & Graesser, 1998; Nelson & Narens, 1990; Perfect & Schwartz, 2002).

Metacognitive monitoring is considered “the process that allows individuals to observe, reflect on, or experience his or her own cognitive processes” (Perfect & Schwartz, 2002, p. 4). During the metacognitive monitoring process, self-regulated learners come to recognize the current state of their knowledge. The recognition often reveals the gap between what they do and do not know, and therefore allows them to discover what they should know. Self-monitoring is the point at which self-regulated learners identify their learning needs. It has been said that until learners recognize potential gaps in their knowledge, they are unlikely to appropriately address them by setting appropriate learning goals, making decisions, or taking action on those decisions (Hacker et al., 1998; Nelson & Narens, 1990; Perfect & Schwartz, 2002). Therefore, metacognitive monitoring plays an important role for self-regulated learners, as it is what allows them to determine when further learning development is necessary and, in turn, take action. Without metacognitive monitoring, learners are unable to effectively self-regulate their learning (Hacker et al., 1998).

Metacognitive control involves individuals setting goals, making decisions, and taking action on those decisions based on their self-monitoring (Nelson & Narens, 1990; Perfect & Schwartz, 2002). Metacognitive control is the point at which self-regulated

learning addresses learning deficits that they may have identified during self-monitoring. Once a self-regulated learner identifies their learning needs as a part of the metacognitive monitoring process, metacognitive control takes over. During the control process, self-regulated learners might outline specific learning tasks, choose appropriate learning venues, and complete the learning tasks. These elements of controlling one's learning are all based on appropriately identified learning needs through self-monitoring. Therefore, the control process is said to be inextricably dependent on effective metacognitive monitoring (Hacker et al., 1998; Nelson & Narens, 1990; Perfect & Schwartz, 2002; Violato & Lockyer, 2006).

Often discussed in tandem with metacognitive monitoring is the concept of learner self-assessment (Hacker, Bol & Bahbahani, 2008; Rivers, 2001). The emphasis throughout the literature on the importance of effective metacognitive monitoring for successful self-regulated learning has led researchers to focus on the level of effectiveness with which professionals are able to self-assess (Bredert & Fite, 2009; Gordon, 1991; Ross, 2007; Violato & Lockyer, 2006; Yammarino & Atwater, 1997; Yoo, Son, Kim & Park, 2009).

Self-assessment research has focused not only on the effectiveness with which the professionals are able to assess their own knowledge and skills, but also their performances associated with that knowledge and skills (Hacker et al., 1998; Nelson & Narens, 1990; Perfect & Schwartz, 2002). Parallel to the way in which effective metacognitive monitoring is a prerequisite to effective metacognitive control, effective self-assessment is a precursor to professionals taking appropriate steps to address the knowledge and skill deficits identified

by the self-assessment process. Effective self-assessment ability is essential to the ongoing development and success of effective professionals.

Self-Assessment for Professionals

Effective self-assessment has been defined as the “ability to accurately assess one’s strengths and weaknesses” (Ward, Gruppen & Regehr, p.63). Professionals are more likely to take steps to further develop their knowledge or skills in an area in which they see themselves as particularly weak (Hacker et al., 1998; Nelson & Narens, 1990; Perfect & Schwartz, 2002). If areas are not identified as weak, professionals feel that there is no reason to pursue learning endeavors to strengthen areas they have defined as being strong.

Professionals who see their knowledge and skills as being strong in a particular aspect of their expertise, for example, are less likely to set goals, make decisions, and take actions to pursue further development of knowledge and skills in that area.

Researchers have been primarily concerned with professionals who overestimate their knowledge and skills. However, those who underestimate their knowledge and skills have also been noted in the literature (Eva & Regehr, 2005). Professionals who view themselves as weak in an area in which they are actually strong may waste time and effort setting goals, making decisions, and taking action based on inaccurate information. These under-confident professionals may also be less likely to persevere with confidence in the face of difficult situations or set appropriately challenging learning goals when necessary (Eva & Regehr, 2005).

Because of the implications of professionals not effectively recognizing their strengths and weaknesses, self-assessment has been said to be meaningful only if it is accurate (Perfect & Schwartz, 2002). However, it has been suggested that there may be limits to some professionals' ability to accurately self-assess despite the fact that it is critical to their self-regulated learning process (Edwards, 2007; Falchikov & Boud, 1998; Hacker et al., 1998; Kreuter & Strecher, 1995; MacDonald et al., 2003; Ross, 2006; Spears, 1980; Taras, 2001; Violato & Lockyer, 2006; Ward et al., 2003; Westberg & Jason, 1994).

Throughout the literature, self-assessment effectiveness has been measured by professionals either predictively or postdictively (retroactively) judging their performance in comparison with an external measure. As suggested by Gordon (1991), quantitative measurement of accurate self-assessment involves "gaining reasonable concurrence between self-claimed and other [external] validated measures of performance" (Gordon, 1991, p.762). External measures have included objective tests (Hawkins, Sumption, Gaglione & Holmboe, 1999; Parker, Alford & Passmore's, 2004), customer (Spsychalski, Quiñones, Gaugler, & Pohley, 1997) or patient surveys (Lockyer & Violato, 2006), standardized patient assessments (Hodges, Regehr & Martin, 2001; Martin, et al. 1998), faculty or supervisor ratings (Fuqua, Johnson, Newman, Anderson & Gade, 1984), and peer assessments (Boud & Tyree, 1979).

Studies of professionals' self-assessment ability have involved professionals assessing their own participation in a demonstration of their knowledge, skills, or behaviors. These demonstrations might include performance on written tests (Hawkins et al., 1999;

Parker et al., 2004), specific interactions with customers or clients (Spsychalski et al., 1997), standardized role-modeling experiences (Hodges et al., 2001; Martin et al., 1998), or everyday practice (Boud & Tyree, 1979). At some point subsequent to these demonstrations, these professionals assessed how well they felt they performed. Their self-assessments were then compared with at least one form of external measurement in order to determine their accuracy. A consistent variability has been demonstrated in professionals' ability to self-assess across multiple professions.

Findings from a study of military professionals indicated variability from low to moderate in the subjects' ability to accurately assess their performance in comparison with the external assessments of their supervisors (Breidert & Fite, 2009). Falchikov and Boud (1989) found physical therapy students' scores of their own task performance to correlate moderately with the external scores of their instructors. Throughout 55 studies of professionals' ability to self-assess their performance, Mabe and West (1982) found a high variance across the board in a variety of professions. These professions included management, engineering, physician practice, and nurse practice. The variability in these results across the professions suggests a need for professionals to develop the ability to perform effective and accurate self-assessments if they are to be successful in their self-regulated learning throughout their careers.

Several authorities suggest that self-assessment ability should be developed as a part of formal professional preparation training (Curran et al., 2007; Gordon, 1991; Schunk & Zimmerman, 2008; Westberg & Jason, 1994). This assertion is a natural one, considering

that at the conclusion of most professional preparation training practitioners are left to rely on themselves (Lajoie, 2008; Schein, 1972) for monitoring and controlling of their own learning experiences.

When faced with a lack of knowledge and skills for effective performance, professionals in practice must first identify their appropriate strengths and weaknesses. Then, they must determine which learning activities are most appropriate to address those weaknesses. Therefore, educators and theoreticians are particularly interested in determining how to best develop learners' self-assessment ability during professional education and training, with the assumption that their self-assessment of necessary knowledge and skills will be vital for their careers (Edwards, 2007; Falchikov & Boud, 1998; Kreuter & Strecher, 1995; MacDonald et al., 2003; Ross, 2006; Spears, 1980; Taras, 2001; Violato & Lockyer, 2006; Ward et al., 2003; Westberg & Jason, 1994).

Physician Self-Assessment

While the development of self-assessment ability is important for all professionals, self-assessment ability is particularly relevant for high stakes professions such as medicine. Society at large and individual lives depend heavily on the accuracy of physicians' knowledge and skills expertise (Eva & Regehr, 2005; Violato & Lockyer, 2005; Ward et al., 2003; Wooliscroft et al., 1993). The fundamentals of medicine are predicated on the assumption that physicians are aware of their capabilities and limitations (Violato & Lockyer, 2005; Ward et al., 2003). Effective self-assessment ability is therefore not a luxury for physicians.

Physicians who recognize their weaknesses in deficient areas have the opportunity to seek out appropriate literature, consult a colleague, ask for assistance, or seek other methods of developing their knowledge, skills, or behavior. However, as has been demonstrated by the research literature on other professionals, physicians have varied ability to conduct these self-assessments of their knowledge and skills (Colthart et al., 2008; Dunning et al., 2004; Eva & Regehr, 2005; Fuhrmann & Weissburg, 1978; Gordon, 1991; Ward et al., 2003; Westburg & Jason, 1994;).

Several studies indicate strong correlations between physicians self-assessed scores externally assessed scores (Brewster et al., 2008; Mandel, Goff & Lentz, 2005; Moorthy, Munz, Adams, Pandey & Darzi, 2006), while others show moderate (Wood et al., 2004), low (Makoul, Krupat & Chang, 2007), or non-significant correlations (Barnsley et al., 2004; Brewster et al., 2008; Hawkins et al., 1999; Moorthy et al., 2006; Pandey et al., 2008; Wood et al., 2004). There is continued concern about physicians' capability to recognize their own problematic gap in the knowledge and skills necessary to be successful in their professional careers.

Further, it has been shown that a physician's capabilities in one knowledge or skill domain does not directly influence the capabilities of the same individual in another knowledge or skill domain (Brinkman et al., 2007; Dunning et al., 2011; Eva & Regehr, 2005; Lipsett, Harris, & Downing, 2011; Sargeant et al., 2011). These differences in a physician's capabilities are also present in their ability to self-assess across various knowledge and skill domains. As a result, physician self-assessment ability has been said to

be context-specific (Brinkman et al., 2007; Dunning et al., 2011; Eva & Regehr, 2005; Lipsett et al., 2011; Sargeant et al., 2011). Thus, even when strong self-assessment ability is demonstrated within a singular domain, a concern remains for physician's capabilities for self-assessment across the other key domains of physician practice.

Physician Communication Skills

Based upon current national standards, the major domains within physician practice and medical education mirror the core competencies outlined by the councils which accredit schools and programs at each level of physician education and training. These core competencies were first adopted by the American Board of Medical Specialties (ABMS) and the Accreditation Council for Graduate Medical Education (ACGME) in 1999. Each of the core competencies represents a key domain determined to be essential for quality patient care in medicine (American Board of Medical Specialties, 2012). The 6 core competencies are medical knowledge, patient care, interpersonal and communication skills, professionalism, systems-based practice, and problem-based learning and improvement (ACGME, 2011; LCME, 2010; ABMS, 2009).

Roberts and Bucksey (2007) have stated that, among the core competencies for a physician, communication skills are the most important skills for health care professionals to master. The domain of communication skills has been identified as one that is clearly significant to physician practice and central to effective patient care (Rider & Keefer, 2006, Simpson et al., 1991; Tulsky, 2005; Watling & Brown, 2007). Research has directly linked

physician to patient communication skills effectiveness to patient adherence and overall improvements in patient management (Eid, Pety, Hutchings, & Thompson, 2009).

Poor communication between physician and patient can have significant clinical consequences. Previous research has shown a relationship between poor physician/patient communication and poor patient satisfaction, poor patient compliance with medications, and other poor clinical outcomes (Bull et al., 2002; Ciechanowski, Katon, Russo & Walker, 2001; Stewart, 1995). Furthermore, poor physician/patient communication has been said to lead patients to misunderstand diagnoses, the severity of their illness, or potential risks. Poor communication has also presented a barrier for physicians' ability to appropriately assess and manage patient pain or other physical or mental symptoms that may change diagnoses (Franks, Fiscella & Shields, 2005). The inability to effectively communicate with patients has also been attributed to physician burnout (Krasner et al., 2009).

The conclusion to be drawn from the results described above is that effective communication skills remain a challenge throughout the field of medicine (Back et al., 2007; Rider & Keefer, 2006; Rosas-Blum, Shirsat, Leiner, 2007; Simpson et al., 1991; Watling & Brown, 2007). Physicians are not always effective at recognizing how well they interact and communicate with patients and their families. An inability for physicians to recognize deficits in their communication skills with patients and families precludes their tendency to seek to improve upon these skills (Gourgey, 2001; Westberg & Jason, 1994).

In the discussion of their findings, many studies have postulated that physicians should monitor their communication skills effectiveness as a part of their self-regulated

learning process throughout their careers (Brinkman et al., 2006; Gruppen et al., 1997; Miller, Hope & Talbot, 1999; Howells, Davies, Silverman, Archer & Mellon, 2010; Schirmer et al., 2005). The effective practice of self-assessing their communication skills can help physicians to seek methods for alleviating their communication skill weaknesses and approach areas of communication skill strengths with confidence. By doing this, physicians can enhance their overall communication skills. Enhanced communication skills presents the opportunity for improved patient satisfaction, patient compliance with medications, and clinical outcomes (Bull et al., 2002; Ciechanowski et al., 2001; Stewart, 1995). Additionally, patients will likely better understand their diagnoses, the severity of their illnesses, and potential risks. Furthermore, physicians may be better able to appropriately assess and manage patient concerns and thus make more appropriate diagnoses (Franks et al., 2005).

Theoretical Framework

This study was to explore the relationship between competence in a specific skill domain and the ability to self-assess within the domain. Specifically, this study explored the relationship between the residents' (physicians in training) communication skills domain competence with their ability to self-assess within the communication skills domain. To explore this relationship, Kruger and Dunning (1999) have suggested that when one develops competence within a domain, one also increases his/her self-assessment ability within it. Evidence has suggested that deficits in metacognitive skill, particularly the monitoring component, explain the lack of self-insight for those deemed incompetent in a given domain.

This theoretical framework presumes that “the skills needed to perform the cognitive, affective, or behavioral task (producing the response) are the same skills necessary for metacognitive tasks (judging the response)” (Dunning, 2005, p. 15). Therefore, incompetence within the given domain, in essence, robs individuals of the ability to recognize competence (or the lack thereof) when they see it. This relationship would suggest that when a physician is competent in a given domain, the physician has the ability to recognize their strengths and weaknesses within it. Under this theoretical framework, increasing competence in a given domain is the most direct means of increasing physicians’ ability to self-assess their performance within it.

The underlying assumptions and theory of this framework suggest its value for considering the training and professional practice of physicians. Findings that support the use of this framework are those that directly link medical resident self-assessment ability to competence within the assessed domain throughout research studies and the literature (Hodges et al., 2001; Parker et al., 2004). These studies report cases of the least competent residents within a given domain being the least able to effectively assess their own knowledge and skills within that domain. These residents rated themselves as having performed above (often well above) the mean. However, objective external assessments showed them to actually perform at a level within the least competent group (bottom quartile or tertile). Such dramatic overestimation of competence can be particularly dangerous for physicians (Graber, 2005; Mamede et al., 2007), heightening the need for researchers and theorists to examine the relationships between competence and self-assessment.

Statement of the Problem

It is often assumed that physicians in training and practice are able to effectively identify their own strengths and weaknesses (Violato & Lockyer, 2005; Ward et al., 2003). However, research has shown that this is not always the case (Barnsley et al., 2004; Brewster et al., 2008; Hawkins et al., 1999; Makoul et al., 2007; Moorthy et al., 2006; Pandey et al., 2008; Wood et al., 2004). Therefore, the ACGME (2011) has included the development of resident self-assessment ability as a requirement for residency training programs (ACGME, 2011). There is a clear need, supported by these accreditation requirements (ACGME 2011), for physicians to develop the ability to effectively self-assess within their preparatory (residency) training. Despite this need and requirements, the problem exists that it is unclear to the leaders of residency programs (program directors) and other resident educators how to best help their residents develop their self-assessment ability.

There are currently two main schools of thought expressed throughout the literature to guide medical educators to develop resident self-assessment ability throughout the course of residency training. One school of thought regarding the development of resident self-assessment ability involves explicit instructional interventions on self-assessment. The other school of thought applies Kruger and Dunning's framework to the development of resident self-assessment ability.

Historically, the literature has suggested that the accuracy of self-assessment is a skill that will remain underdeveloped until it is explicitly addressed by instructional interventions on self-assessment as its own domain (Colthart et al., 2008; Dunning et al., 2004; Fuhrmann

& Weissburg, 1978; Gordon, 1991; MacDonald et al., 2003; Ward et al., 2003; Westburg & Jason, 1994). While some studies of these explicit interventions have been shown to be successful in improving resident self-assessment ability, they may in fact be unnecessary for the development of resident self-assessment ability. If there were a more direct means of developing resident self-assessment ability, explicit interventions may be superfluous. Furthermore, self-assessment instructional interventions are typically conducted within a singular domain. Given the context-dependent nature of self-assessment (Brinkman et al., 2007; Dunning et al., 2011; Eva & Regehr, 2005; Lipsett et al., 2011; Sargeant et al., 2011), the instructional self-assessment interventions conducted within one or two domains may not be generalizable across other domains.

Alternatively, since Kruger and Dunning's 1999 study, medical education researchers and theoreticians have applied their framework to residents. The application of Kruger and Dunning's framework would suggest a focus on developing residents' skills within knowledge and skill domains in which they are the least competent. According to Kruger and Dunning's framework, increased competence within these domains would directly increase the residents' ability to assess their strengths and weaknesses within the domains. Increasing resident self-assessment ability via increased competence within deficient knowledge and skill domains would make explicit interventions on self-assessment as its own domain unnecessary.

Four extant studies have examined the relationship between levels of resident competence within a given domain and their ability to self-assess within it (Jones, Panda &

Desbiens, 2008; Mandel et al., 2005; Parker et al., 2004; Hodges et al., 2001). These studies will be discussed in greater depth in chapter two. Only one of these studies has included communication skills as an assessed domain (Hodges et al., 2001). Further research is needed to provide additional evidence to support the claim within the domain of communication skills for residents. Providing additional evidence of a relationship between increased self-assessment ability with increased competence within the domain of communication skills would be a significant contribution to practice and this body of research.

Purpose Statement

The purpose of this study was to provide additional evidence to support claims throughout the literature that Kruger and Dunning's (1999) theoretical framework is generalizable to residents' ability to self-assess. Kruger and Dunning's framework suggests that residents' ability to self-assess will directly correlate with the assessed domain. If this is the case, residents with higher levels of competence within a given domain will be better at self-assessing within that domain. Residents with less competence within the domain will self-assess with less accuracy when compared with scores of external assessors.

Research Questions

To determine the potential relationship between the development of resident competence within a given domain and resident self-assessment ability within the same domain, this study has sought to answer the following research questions:

1. What are the varied levels with which residents demonstrate competence within the domain of communication skills as measured by Standardized Patient Parents (SPPs) using the Breaking bad news Assessment Schedule (BAS) instrument?
2. Do residents' self-assessments correspond with Standardized Patient Parent assessment of their communication skills performance?
3. Does resident communication skill competence relate to their ability to self-assess within the domain?
4. Does resident gender relate to their ability to self-assess their communication skills performance?
5. Does the number of years of residency training experience relate to residents' ability to self-assess their communication skills?
6. Can a significant amount of variance in resident ability to self-assess their communication skills be explained by gender, year in training, ethnicity, and theoretical element of communication skills competence (setting the scene, breaking the news, eliciting concerns, information giving, and general information)?

Significance

Research & Theory

This study will add to the body of knowledge related to resident self-assessment ability. By demonstrating a relationship between higher levels domain competence and resident ability to self-assess within that domain, it supports the generalizability of Kruger and Dunning's (1999) theoretical framework to the training and professional practice of physicians. While other research has suggested explicit instructional interventions on self-assessment as its own domain are necessary, these results would contradict that claim.

Further evidence to support the application of Kruger and Dunning's theoretical framework within residency training could mean a shift in research focus for those interested in developing physician self-assessment ability in residency training. Researchers may choose to focus less on explicit self-assessment instructional interventions. Alternatively, researchers may choose to focus more on methods for developing resident knowledge and skills in domains within which they may be deficient, with the direct impact on development of their aligned self-assessment ability.

Additionally, the results of this study may lead to similar research in other resident knowledge and skill domains. Such studies might focus on the remaining five core competencies of physician education: medical knowledge, patient care, professionalism, systems-based learning, and practice-based learning and improvement.

Practice & Policy

Medical residency educators are charged with facilitating the development of the residents' self-assessment skills. However, to enable this charge, residency educators need to better understand self-assessment in residency. Determining if residents are better able to recognize their strengths and weaknesses within a domain if they are competent within that domain will help to guide training efforts to enhance self-assessment skills throughout residency training.

In particular, a relationship shown between residents with higher communication skills competence and higher self-assessment accuracy within that domain could lead to increased focus on enhancing resident skills within communication skills over the explicit development of self-assessment skills. Further evidence of this relationship might also lead accrediting bodies to reconsider their emphasis on medical educators creating explicit programs to develop resident self-assessment ability. Therefore, the results of this study have the potential to significantly impact research, theory, practice, and policy for graduate medical education.

Delimitations

The findings of this study are bound by the population and sample, which included residents from one institution (Duke University) and one specialty (pediatrics). Additionally, this study uses a cross-sectional approach for data collection. This approach examines Duke Pediatric residents who were in training in June 2011. This study was also delimited by the

domain assessed and the instrument used to assess the domain. The instrument measured resident competence within the domain of communication skills.

CHAPTER TWO: REVIEW OF THE LITERATURE

Introduction

This study focused upon residents' (physicians in training) ability to effectively self-assess their communication skills. This study seeks to answer the following research questions:

1. What are the varied levels with which residents demonstrate competence within the domain of communication skills as measured by Standardized Patient Parents (SPPs) using the Breaking bad news Assessment Schedule (BAS) instrument?
2. Do residents' self-assessments correspond with Standardized Patient Parent assessment of their communication skills performance?
3. Does resident communication skill competence relate to their ability to self-assess within the domain?
4. Does resident gender relate to their ability to self-assess their communication skills performance?
5. Does the number of years of residency training experience relate to residents' ability to self-assess their communication skills?
6. Can a significant amount of variance in resident ability to self-assess their communication skills be explained by gender, year in training, ethnicity, and theoretical element of communication skills competence (setting the scene, breaking the news, eliciting concerns, information giving, and general information)?

There are six main sections presented for discussion within this chapter. Each of the sections will contribute to a foundational description of the critical nature of effective self-assessment for professionals and, more specifically, physicians. The bulk of the discussion will focus on the importance of developing self-assessment ability within preparatory training and what is currently known about how this is done within residency.

The first section provides an overview of research on self-regulated learning for professionals. This section focuses on the commonalities among studies of professionals and their preparatory education and training. Furthermore, this section illuminates self-regulated learning as an integral component of the development of professionals throughout their careers, post-formal preparatory education and training.

A focus on metacognition, a critical element of self-regulated learning (Gourgey, 2001; Zimmerman, 1990), and a discussion of the two elements of metacognition (Nelson & Narens, 1990) are included within the second section. One of the two elements of metacognition, metacognitive monitoring, which is often discussed in tandem with self-assessment, leads into the third main area of discussion which presents an overview of self-assessment for professionals.

Subsequent to the overview of self-assessment for professionals, a theoretical framework of self-assessment (Kruger & Dunning, 1999) is presented. Kruger and Dunning's framework proposed that the ability to effectively self-assess may be directly related to the self-assessors' competence within the knowledge or skill domain being

assessed. Their framework was applied to medical professionals in training (physician residents) for examination within this study.

Because the current study applied Kruger and Dunning's theoretical framework to residents' ability to effectively self-assess, the next area of discussion will focus on the context of medical education. An overview is provided on each phase of education and training necessary for physician development in the United States. Self-assessment for physicians is then discussed in detail.

Furthermore, three subsections within this section illuminate studies that have been conducted to examine three areas of residents' ability to effectively self-assess. The first subsection discusses studies that have examined residents' ability to self-assess as a static skill. These studies examined residents' ability to self-assess at one point in time to determine if they either did or did not have the ability to self-assess effectively. As opposed to other studies throughout the literature (Butterworth, 2010; Hodges et al., 2001; Jones et al., 2008; Mandel et al., 2005; Parker et al., 2004; Regehr, Hodges & McNaughton, 1998; Ward et al., 2003), those examining self-assessment as a static ability did not consider the potential for a change in the residents' ability to self-assess.

The two remaining subsections review the two main schools of thought regarding the development of resident self-assessment ability. The first school of thought proposes that explicit instructional interventions specific to self-assessment are necessary to develop resident self-assessment ability (Colthart et al., 2008; Dunning et al., 2004; Fuhrmann & Weissburg, 1978; Gordon, 1991; Westburg & Jason, 1994). Studies discussed within this

subsection include instructional interventions primarily focused on explicit feedback including video and benchmarking (Hodges et al., 2001; Martin et al., 1998; Ward et al., 2003) or repeated self-assessment practice (Butterworth, 2010), each within the context of one or two knowledge and skill domains.

The second school of thought involves the application of Kruger and Dunning's (1999) theoretical framework to residents. The studies reviewed within this subsection examined the relationship between residents' ability to self-assess and their competence within the assessed domain. Four studies have examined this relationship (Jones et al., 2008; Mandel et al., 2005; Parker et al., 2004; Hodges et al., 2001). Only one of these four studies was within the domain of communication skills (Hodges et al., 2001).

The current study directly examined the relationship between resident self-assessment ability and their communication skills competence. Self-assessment ability has typically been studied by directly comparing the correlation between self-assessment scores with the scores of valid external raters. Therefore, it is appropriate for assessment instruments typically used for external examination of resident communication skills to be utilized for a comparison between self- and external assessment (Gordon, 1991). A final section in this chapter provides a discussion of a reliable, validated communication skill measurement instrument discussed in the literature (Miller et al., 1999).

Self-Regulated Learning for Professionals

This section provides an overview of self-regulated learning theory for professionals. Self-regulated learning refers to the "self-directive process through which learners transform

their mental ability into task-related academic skills” (Zimmerman, 2001, p. 1). The concept of self-regulated learning views learning as something for which learners are proactively responsible rather than an event of which they are passive recipients (Zimmerman, 2001). Zimmerman emphasizes that self-regulated learning does not necessarily involve asocial activities completed in isolation from others. Rather, the key issue to self-regulated learning lies in the learner displaying “personal initiative, perseverance, and adaptive skill” in pursuing learning (Zimmerman, 2001, p. 1).

Self-regulated learning has been emphasized throughout the literature as a critical component of professional development (Pintrich & Zusho, 2002; Zimmerman, 2001). The independent nature of professional practice and the professional development therein requires practitioners to be self-regulated in the way in which they continue to develop their knowledge and skills throughout the course of their careers.

Jarvis described the professional as “one who continually seeks the mastery of the branch of learning upon which his occupation is based, so that he may offer services to his clients” (1983, p. 27). Society depends on these professionals for their expertise (Schön, 1983). Professionals typically develop their expertise throughout a structured continuum of education and training (Schein, 1972). Such formal education and training is designed to equip them with the professional knowledge and skills to serve society throughout their careers.

As a key researcher in the field of formal professional education, Edgar Schein examined professional knowledge in the 1970s in relationship to the growth of professional

knowledge and the changing needs of students and society. Schein (1972) described formal professional education as a sum of three main elements:

1. An underlying discipline or basic science component upon which the practice rests or from which it is developed.
2. An applied science or “engineering” component from which many of the day-to-day diagnostic procedures and problem solutions are derived.
3. A skills and attitudinal component that concerns the actual performance of services to the client, using the underlying basic and applied knowledge. (p. 43)

Most professional education programs have the basic knowledge, practical, and attitudinal elements described by Schein (1972) formally built into their curricula. Schein emphasized, however, that upon completion of this formal education, professionals are predominantly left to their own devices to keep their knowledge up to date and on par with the changes both in the knowledge base and to meet society’s evolving needs. Schein’s assertion remains true today (Lajoie, 2008). Beyond their formal education, professionals need to determine their own learning needs and pursue learning activities based on those needs. From the point of completion of their formal training throughout the duration of their careers, self-regulated learning becomes integral to professional development.

Pintrich and Zusho (2002) further conceptualized self-regulated learning as an active process whereby learners monitor and control their learning. Monitoring and controlling one’s own learning are essential components to professional learning development

(Zimmerman, 2001). These components, discussed in further depth in the following section, are also the two salient features of metacognition as it relates to the self-directed learner (Nelson & Narens, 1990; Perfect & Schwartz, 2002). Therefore, active metacognition is an essential element of the self-regulated learning process for professionals.

Metacognition

Learners who are effective self-regulators have been identified as metacognitively active participants in their own learning process (Zimmerman, 1990). Throughout the literature, metacognition has generally been defined as experience and knowledge we have about our own cognitive processes (Flavell, 1979). A metacognitively active participant has been said to actively participate in their own learning by self-monitoring, self-assessing, planning, and setting goals (Zimmerman, 1990). It has been said that this process is critical to cognitive development, which is central to ongoing professional growth (Gourgey, 2001). The concept of metacognition present in the literature today has grown from roots in two fields of psychology research: cognitive psychology of the 1960s and post-Piagetian (developmental) psychology of the 1970s (Schwartz & Perfect, 2002).

Cognitive psychologists of the 1960s (Hart, 1965, 1967) who studied metacognition were interested in the accuracy with which individuals made judgments about their memory. These researchers did studies to determine if adults had accurate feelings-of-knowing (judgments) about their cognition. That is, they were interested in knowing if the participants had feelings of knowing which answers they would answer correctly on a test. They were

most interested in if those judgments were valid predictors of behavior (if they were correct in their judgments).

Developmentalists (Flavell, 1979) were more interested in determining if memory ability was “a function of greater conscious understanding of the rules that govern memory and cognition” (Schwartz & Perfect, 2002, p. 3). Therefore, they were interested in determining if learners’ memory abilities were enhanced by a greater understanding of cognitive processes. Until the early 1990s, these two schools of thought had long permeated the literature separately, both under the term “metacognition.”

In the early 1990s, Nelson and Narens (1990) presented a new metacognition theory that brought together the influences of both the cognitive and the developmental psychologists. Their theory presented two distinct but influential elements: metacognitive monitoring and metacognitive control. Nelson and Narens’ theory guides much of the current research on metacognition.

Metacognitive Monitoring

Metacognitive monitoring, as described by Nelson and Narens (1990) is considered “the process that allows the individual to observe, reflect on, or experience his or her own cognitive processes” (Perfect & Schwartz, 2002, p. 4). Metacognitive monitoring is the predecessor to metacognitive control. The monitoring process is the point at which individuals might identify deficits in their knowledge, skills, or behaviors to be addressed within the metacognitive control process. For metacognitive monitoring to effectively influence the metacognitive control process, professionals must be accurate in their

monitoring ability. Nelson and Narens (1990) contend that the main means of generating data about metacognitive monitoring is to gather subjects' reports about their own introspections.

Metacognitive Control

Metacognitive control, as conceptualized by Nelson and Narens, (1990), involves individuals consciously or unconsciously setting goals, making decisions, and taking action on those decisions based on what they discover within the self-monitoring process (Schwartz & Perfect, 2002). During the self-monitoring process, a professional might identify deficits in their own knowledge, skills, or behaviors. Based on this information, metacognitive controlling would help them to determine what should be done to address the deficits. Through the control process, the professional would then make plans to address the goals, and ultimately address them until they have been met and the issues have been resolved.

The concept of metacognitive control makes it possible for professionals to self-regulate their own learning development in a way that best meets their individual needs. However, for individuals to effectively develop their learning, it is critical for them to have first appropriately identified deficits in their own knowledge, skills, or behaviors. Therefore, for metacognitive controlling to be effective, it must be based on valid information acquired through effective metacognitive monitoring (Schwartz & Perfect, 2002; Violato & Lockyer, 2006). Metacognitive monitoring's reports of introspection, as defined by Nelson and Narens (1990), are often framed in terms of learners' self-assessment.

Self-Assessment for Professionals

Successful metacognitive monitoring is dependent upon effective self-assessment of knowledge, skills, and behaviors (Hacker et al., 2008). Effective self-assessment has been defined as the “ability to accurately assess one’s [own] strengths and weaknesses” (Ward et al., 2002, p.63). This section provides an overview of the theory on self-assessment for professionals.

To effectively metacognitively monitor their learning, professionals must be aware of their strengths and weaknesses. It is the knowledge of their strengths which guides self-regulated learners to set and accomplish appropriately challenging goals (Eva & Regehr, 2005). Even more importantly, it is knowledge of their weaknesses that guides them to appropriately address areas of deficit.

Professionals are more likely to take steps to further develop their knowledge or skills in an area in which they see themselves as particularly weak (Hacker et al., 1998; Nelson & Narens, 1990; Perfect & Schwartz, 2002). Professionals who see their knowledge and skills as being strong in a particular aspect of their expertise, for example, are less likely to set goals, make decisions, and take actions to pursue further development of knowledge and skills in that area.

If professionals are incorrect in their knowledge and skill assessment, learning development and thus professional growth is compromised (Gourgey, 2001; Westberg & Jason, 1994). These professionals would miss the opportunity to seek appropriate resources to develop those skills. Therefore, that knowledge or skill would remain underdeveloped,

thus compromising their level of expertise in their profession. Addressing these deficits may be critical to their professional success and development (Eva & Regehr, 2005). Therefore, researchers have sought to study the accuracy with which professionals are able to self-assess.

Throughout self-assessment research, accurate self-assessment has been defined as “gaining reasonable concurrence between self-claimed and other, validated measures of performance” (Gordon, 1991, p.762). Gordon defines valid self-assessment as “judging one’s performance against appropriate criteria” (p. 762). Within this body of research, self-assessment accuracy has been measured both by learners predicting how they will score on a test or perform a skill, as well as retroactively (postdictively) judging their performance (Gordon, 1991). The self-assessments are then compared against how the learners actually performed as determined by a valid external measurement (calibration).

Self-assessment ability has been studied and reported in the literature throughout many professional fields (Dunning et al., 2004) such as military (Breidert & Fite, 2009), management (Yammarino & Atwater, 1997), teaching (Ross, 2007), nursing (Yoo et al., 2009), physical therapy (Falchikov & Boud, 1989), and physician practice (Violato & Lockyer, 2006). Typically the results show modest to meager accuracy, however they have been found to vary across the spectrum. The following description of studies illustrates the variability throughout various professions.

Findings from a study of military professionals indicate variability from low to moderate in military men and women’s ability to accurately assess their performance in

comparison with the external assessments of their supervisors (Bredert & Fite, 2009). Similarly, Falchikov & Boud (1989) found physical therapy student's scores of their own task performance to correlate moderately with the external scores of their instructors. Throughout 55 studies of professionals' ability to self-assess, Mabe and West (1982) found a high variance across the board in a variety of professions including management, engineering, physician practice, and nurse practice. The variability in these results across the professions suggests a need for professionals to develop the ability to effectively self-assess if they are to be successful in their self-regulated learning throughout their careers. The variability of these results is also a cause for concern given the roles of professionals as experts in their respective fields (Schön, 1983).

Given the essential nature of the professions to the function of our society (Schön, 1983), professionals are relied upon to be aware of their strengths and weaknesses. Therefore, many researchers have sought to determine why such variability exists and what can be done to minimize it. To explore the variability of self-assessment's correlation to external assessment, it is important to examine relationships that exist among the varying levels of self-assessment ability. The application of Kruger and Dunning's framework to professionals allows researchers to examine the relationship between the self-assessment ability of professionals and their competence (as determined by external assessment) within the assessed domain.

Kruger and Dunning's Theoretical Framework of Self-assessment

The current study examined the application of Kruger and Dunning's (1999) theoretical framework on physician residents. Kruger and Dunning have referred to their framework as "unskilled and unaware." The framework focuses on the hypothesis that individuals who demonstrate the least competence ("unskilled") in assessed domains are also the least likely to be able to recognize their (in)competence ("unaware"). Kruger and Dunning's position has been that competence in a given domain is a natural predecessor to individuals being able to recognize their (or others') competence within it. The concept that those who are the least skilled in knowledge and skill domains are also the least able to recognize their deficits is of particular interest to the study of professionals and professional development.

Kruger and Dunning's (1999) seminal studies, which examined self-assessment as it relates to levels of domain competence, has spawned a great deal of scholarly discussion and research on self-assessment throughout various professional disciplines in the years to follow. Over the course of four distinct but interrelated studies, Kruger and Dunning examined undergraduate student self-assessment ability within three domains: humor, logical reasoning, and grammar. Levels of competence were determined by measurements of objective external assessments.

Kruger and Dunning's four studies focused on testing four predictions: 1) incompetent individuals would dramatically overestimate their ability and performance in comparison with objective external assessment; 2) incompetent individuals would be less

likely (in comparison with their more competent peers) to recognize competence in themselves or others; 3) incompetent individuals would be less likely than their more competent peers to recognize their level of incompetence, even when compared with the more competent performances of others; and that 4) incompetent individuals would gain insight into their weaknesses only after they had become more competent within the assessed domain.

During the initial study, Kruger and Dunning were interested in testing their first prediction. Using 65 Cornell University undergraduate students, Kruger and Dunning tested the students' ability to determine what others would find to be humorous. The students' scores were compared with scores of expert comedians. The students were also asked to rate their ability to "recognize what's funny" (p. 1123) in comparison with the average Cornell student by assigning themselves a percentile score. The results of this first study led Kruger and Dunning to conclude that their first prediction was accurate. The students who were the least competent in recognizing what was humorous as compared with expert comedians, dramatically overestimated their ability to do so in comparison to their peers.

Because humor is such a subjective domain, Kruger and Dunning wanted to replicate their findings from their first study in a different, more objective domain (logical reasoning). The researchers also wished to eliminate the question that the score miscalibration shown in the first study may have resulted from participants underestimating the skills of their peers rather than overestimating their own skills.

For their second study, Kruger and Dunning asked 45 Cornell University undergraduates from an introductory psychology course to complete a logical reasoning test and then make three estimates about their ability and performance on the test: 1) a percentile ranking of their general logical reasoning ability in comparison with their peers who had also taken the test, 2) a percentile ranking for how their test scores would compare with that of their peers, and 3) an estimate for how many test questions they believed they had answered correctly.

The results of the second study substantiated Kruger and Dunning's claims from the first study in a more objective domain. Those performing in the lowest percentile not only overestimated their own percentile scores, but their estimates also significantly exceeded the mean percentile. Additionally, those performing in the bottom quartile significantly overestimated the number of test questions they had answered correctly. The level of overestimation confirmed that the effect was not simply the result of the subjects underestimating the skills of their peers.

Additionally, an unanticipated outcome that was revealed within the first two studies was that students in the top quartile had underestimated their performance in both studies, though to a much lesser extent than those who had overestimated. While underestimation of performance is less concerning than overestimates since individuals are not overlooking knowledge or skill deficits, they still warranted further investigation. Therefore, the concern for overestimation of the top performing individuals was addressed as a part of Kruger & Dunning's third study.

In their third study, Kruger and Dunning sought to replicate the results of studies one and two in a third domain (grammar) due to its “clear and decisive rules and facts” (p. 1125). In addition, they wished to explore their second prediction; incompetent individuals would be less likely (in comparison with their more competent peers) to recognize competence in themselves or others. Furthermore, Kruger and Dunning wished to explore why those who had scored in the top quartile were also inaccurate in their self-assessment ability.

Eighty-four Cornell University undergraduate students were first asked to complete a test assessing their grammar knowledge. They were also asked to estimate their overall ability to recognize correct grammar, how their test performance would compare with their peers, and how many items they had answered correctly on the grammar test. The results of these tests further substantiated Kruger and Dunning’s conclusions from their first two studies within a different domain. Students scoring in the lowest percentile continued to substantially overestimate their performance. Those scoring in the top quartile also continued to underestimate their performance. The researchers’ concern for the inaccuracy of the more highly competent participants was addressed in the second phase in study three.

In the second phase of their third study, Kruger and Dunning sought to examine their second and third predictions. Therefore, they invited participants from both the bottom and top quartiles back 4 to 6 weeks following the completion of study three’s first phase. The researchers then asked each participant to grade a set of five de-identified grammar tests taken by five of their peers, indicating how many questions they thought each had answered correctly. The tests reflected a range of performances. Following this grading exercise, the

participants were all asked to re-estimate the number of questions that they had answered correctly on their own test.

As anticipated by prediction two, the bottom quartile participants were less able to recognize the correct answers on their peers' tests than were the top quartile. Furthermore, in support of prediction three, the bottom quartile participants did not gain any additional insight into their own poor performance, despite seeing the higher-level performances of their peers.

Following scoring the tests of their peers, the top quartile participants increased their self-assessment accuracy by increasing estimates of their own general grammar ability. These results led Kruger and Dunning to conclude that top performer overestimation was likely a result of a false-consensus effect (Ross, Greene & House, 1977) rather than true underestimation of their own performance. The false-consensus effect implies that, without the benefit of having observed others' performances, high performers often incorrectly assume that their peers are performing at higher levels than they actually are.

Kruger and Dunning's fourth study examined their fourth prediction; incompetent individuals would gain insight into their weaknesses only after they become more competent within the assessed domain. The participants in this study were 140 Cornell University undergraduates. The participants were first asked to take a test on logical reasoning. Following the test, participants were asked to provide a percentile estimate of their logical reasoning skills and performance on the test relative to their peers. They were also asked to estimate the number of problems they solved correctly.

A random selection of 70 participants were given a logical reasoning training packet that should have provided them with the knowledge that they needed to answer the questions on the test correctly. Following the training period, participants in both groups were asked to review their own tests and indicate which items they thought that they had answered correctly and which they had answered incorrectly. They also re-estimated the total number of problems they answered correctly and provided another estimate percentile of their performance relative to their peers.

Prior to the training exercise, the results of participant estimation were consistent with what was found in the first three studies. Participants, on average, overestimated their ability and performance, with those in the bottom quartile clearly overestimating to the greatest extent. Participants in the top quartile also underestimated their ability and performance. Following the training intervention, however, those who received the training packet graded their own tests more accurately than did participants who did not receive the packet. The change was even more evident when the bottom quartile group was looked at exclusively. The training intervention even brought the bottom quartile self-assessments to be just as accurate as those who had originally scored in the top quartile. Underestimates for those within the top quartile also improved significantly, in further support of the concept of the false consensus effect.

Through Kruger and Dunning's (1999) four studies, each of their four predictions were confirmed. Their first prediction was confirmed in three separate domains; one relatively subjective and the other two increasingly objective. Incompetent individuals in all

studies were found to dramatically overestimate their ability and performance in comparison with objective external assessment. It was also confirmed that incompetent individuals were less likely to recognize competence in themselves or others. Additionally, Kruger and Dunning revealed that incompetent individuals were less likely than their more competent peers to recognize their level of incompetence, even when compared with the more competent performances of others. They also proposed, due to their results, that the underestimation of the top performers was likely due to the false consensus effect. Furthermore, Kruger and Dunning provided evidence that self-assessment ability was increased by means of increasing competence within the assessed domain.

While Kruger and Dunning's self-assessment research was not focused on professional self-assessment, it contributed significantly to the foundation on which professional studies of self-assessment have been based. Researchers in professional fields recognized the potential implications for professionals if the least skilled among them are unable to recognize their deficiencies. The inability for professionals to effectively monitor their own strength and weaknesses can be exceptionally concerning, especially if Kruger and Dunning's framework translates to the professional realm.

Most continuing professional development is based on the assumption that professionals are skilled at identifying their own learning needs and seek to address them accordingly (Violato & Lockyer, 2005). When professionals recognize their weaknesses, they have the opportunity to seek resources to alleviate them. Potential errors could be avoided by addressing these weaknesses. Unskilled professionals who overestimate their

knowledge or skills are unlikely to seek resources to assess their deficits because they fail to recognize that they have any (Scwartz & Perfect, 2002; Violato & Lockyer, 2006).

Therefore, many incorrectly self-assessed “experts” may remain unskilled in deficient areas. These professionals might continue to practice their respective professions as if they are fully skilled when, in fact, they are not. As a result, errors are likely to be made and may not be recognized until it is too late (Graber, 2005). Because of this concern, self-assessment ability is of particular importance in the high stakes field of medicine.

Physician Self-Assessment

Preventable medical errors are known to be responsible for hundreds of thousands of serious injuries and deaths each year. The estimated 98,000 annual deaths due to error alone are said to cost U.S. citizens \$29 billion dollars (American Association for Justice, 2012). It has been said that a considerable amount of these errors can be attributed to physician overconfidence in practice (Graber, 2005; Mamede et al., 2007). Overconfident physicians are less likely to seek assistance to develop their knowledge or skills in areas of weakness because they do not recognize the weakness. If Kruger and Dunning’s framework can be applied to physician practice, it is the least skilled physicians who will be the least likely to seek assistance to develop areas of weakness. Medical decisions made by unskilled physicians are prime opportunities for preventable medical errors to occur (Graber, 2005; Mamede et al., 2007). Therefore, a research focus on improving the self-assessment ability of physicians is of paramount importance (Eva & Regehr, 2008; Westberg & Jason, 1994).

Self-assessment has been clearly identified as a major priority to physicians' professional practice and the education and training programs that prepare physicians for that practice (ABMS, 2009; ACGME, 2011; LCME, 2010). All levels of the continuum of physician education are guided by six core competencies of medical education, each of which depends on self-assessment to further their development.

The six core competencies: patient care, medical knowledge, interpersonal and communication skills, professionalism, problem-based learning and improvement, and systems-based practice (ACGME, 2011) have been outlined by the Accreditation Council for Graduate Medical Education (ACGME) and the American Board of Medical Specialties (ABMS) as aspects of learning to be infused throughout the continuum of medical education. Self-assessment permeates each of the six core competencies throughout the entirety of a physician's learning continuum (Caverzagie, Shea & Kogan, 2008). Self-assessment is also specifically noted as a primary element of the practice-based learning and improvement (PBLI) competency to be developed throughout the course of medical training.

Wooliscroft et al. (1993) have also identified self-assessment as "central to the function of the clinician" (p. 290). The fundamentals of medicine are predicated on the assumption that physicians are aware of their capabilities and limitations (Violato & Lockyer, 2005; Ward et al., 2003). Awareness of their strengths assists physicians in being able to persevere with confidence in the face of difficult situations and set appropriately challenging learning goals (Eva & Regehr, 2005). Even more importantly, self-awareness of

their weaknesses sets the stage for physicians to be able to seek assistance when appropriate (Eva & Regehr, 2005).

Individuals everywhere rely on medical providers' professional knowledge, skills, and behaviors in order to protect their health and well-being and that of their families. Dependence on medical professionals sets extraordinarily high expectations for professional competence (Sullivan, 2000). The same dependence also emphasizes how flawed self-assessment in medicine can be dangerous, particularly in the cases in which physicians overestimate their competence (Austin & Gregory, 2007). Therefore, development of self-assessment ability during training is critical to the success of physicians throughout their careers (Eva & Regehr, 2008; Westberg & Jason, 1994).

Context of Medical Education

Because the focus of this study was on the development of self-assessment ability for physicians in training, this section will present an overview of the professional education and the training physicians must undergo in preparation for their careers.

The continuum of phases of physician preparatory education shares the same basic structure of other professional education programs, such as law, military, nursing, and engineering. Elements of the shared structure include an underlying discipline or basic science, an applied component during which learners are able to put the basic science to practice under supervision, and a skills and attitudinal component as they relate to services provided to the "client" (Schein, 1972).

The initial educational process for becoming a physician is lengthy and involved. Starting with undergraduate studies, those planning on a professional life as a physician seek majors that will prepare them for the study of medicine; typically in a field of science. Once their undergraduate studies have been completed, the potential physician undergoes four years of medical school (also referred to as undergraduate medical education or UME) leading to a medical doctorate (MD) or Doctor of Osteopathic Medicine (D.O.) degree. The Liaison Committee of Medical Education (LCME) is the accrediting authority for all UME programs in the United States and Canada. Medical school is university-based and provides a strong platform of basic science education (as outlined by Schein, 1972) and direct clinical experience, primarily in major teaching hospitals (Petersdorf, 1994). This basic science platform and introduction to clinical experiences provides the foundation needed to prepare future physicians for the relative independence of residency training.

Following graduation from medical school, individuals enter the graduate medical education (GME) phase of their education, an intense period of clinical training referred to as “residency,” during which they gain practical experience in their chosen specialty (e.g. internal medicine, pediatrics, surgery). The Accreditation Council for Graduate Medical Education (ACGME) is the accrediting authority for all GME programs in the United States and Canada. This phase of medical education predominantly represents the applied science and skills and attitude components professional education described by Schein (1972). Residency training is the period of time during which physicians primarily learn

experientially. During this phase of training, residents have the benefit of guidance from and oversight of supervisors experienced in the specialty.

The format of medical residency provides residents with the structure to meet their learning expectations with increasing levels of autonomy as determined appropriate by board-certified, experienced supervisors. Supervising physicians are clinical teachers to the residents. They guide residents through clinically relevant rotations that provide them with the practical experience necessary to prepare them to be competent practitioners in their chosen field of medicine.

The practical elements are taught through strategies such as hands-on patient care, case discussions, guided direction on methods to increase topical resident fund of knowledge and improve patient care, chart reviews, and role modelling. Per the ACGME (2011) program requirements,

For the resident, the essential learning activity is interaction with patients under the guidance and supervision of faculty members who give value, context, and meaning to those interactions. As residents gain experience and demonstrate growth in their ability to care for patients, they assume roles that permit them to exercise those skills with greater independence. This concept—graded and progressive responsibility—is one of the core tenets of American graduate medical education. Supervision in the setting of graduate medical education has the goals of assuring the provision of safe and effective care to the individual patient; assuring each resident’s development of

the skills, knowledge, and attitudes required to enter the unsupervised practice of medicine; and establishing a foundation for continued professional growth. (p. 1)

Those who wish to pursue subspecialization in GME (e.g. internal medicine cardiology, pediatric endocrinology, orthopaedic surgery) enter fellowship training (Petersdorf, 1994). Fellowship training provides physicians with a higher degree of knowledge and experience to prepare them to practice independently in a subspecialized field. The Accreditation Council for Graduate Medical Education (ACGME) is the responsible body that evaluates and accredits residency and fellowship training programs throughout the United States.

Once physicians complete their GME training, they move into the longest phase of their medical education, Continuing Medical Education (CME). It is during this time that each physician becomes primarily responsible for enhancing their own learning. During CME, the accredited physicians themselves must keep their knowledge and skills current and appropriate for the evolving needs of their patients.

Certified physicians' CME requirements are governed by the appropriate specialty medical board under the auspices of the American Board of Medical Specialties (ABMS). Among other requirements, stipulations regarding CME are outlined in each board's Maintenance of Certification (MOC) guidelines. However, CME is primarily dependent on physicians self-assessing their strengths and weaknesses to appropriately determine areas on which they should focus their CME activities (Davis et al., 2006; Duffy & Holmboe, 2006). Physicians must develop the ability to identify these areas of strengths and weakness while

still in a training environment and still able to draw upon external assessment and training (Eva & Regehr, 2008; Westberg & Jason, 1994).

Self-Assessment in Residency

All levels of physician training warrant the development and use of self-assessment skills. However, the experiential nature of residency training makes it a prime opportunity to foster this development in a semi-structured environment still under the guidance of supervisors who can help to identify resident skill deficits and correct inaccurate introspections. It has been noted that if graduate medical trainees “are to become effective, competent physicians who grow throughout their careers, they must value and possess the skills required for pursuing ongoing learning” (Westberg & Jason, 1994, p. 278). Residency training is the optimal time during the continuum of physician medical education for developing and enhancing self-assessment skills (Westberg & Jason, 1994). It is the point at which student-physicians’ practice most closely mirrors the way in which they will practice medicine throughout the duration of their careers while still operating under the guidance of a formalized learning program and of experienced supervisors to guide their learning and the development of self-assessment ability.

Throughout the literature, studies on resident self-assessment have asked residents to rate their knowledge, skills, and behaviors either predictively or postdictively (Barnsley et al., 2004; Brewster et al., 2008; Hawkins et al., 1999; Mandel et al., 2005; Moorthy et al., 2006; Pandey et al., 2008; Wood et al., 2004). As with self-assessment studies throughout other fields, the residents’ self-assessed scores are then compared with some form of

objective measurement of their knowledge, skills, or behavior (Breidert & Fite, 2009; Falchikov & Boud's, 1989; Violato & Lockyer, 2006; Yammarino & Atwater, 1997; Yoo et al., 2009). These objective measurements may include physician faculty (Fuqua et al., 1984), peer (Boud & Tyree, 1979), other team member, patient, standardized patient (Hodges et al., 2001; Martin et al., 1998), or previously determined valid and objective written assessments (Hawkins et al., 1999; Parker et al., 2004). There is considerable discussion throughout the literature on residents' ability to effectively self-assess.

Methodological issues permeate the self-assessment literature both in general and more specifically throughout the medical education literature (Boud & Falchikov, 1989; Colthart et al., 2008; Davis et al., 2006; Gordon, 1991; Ward et al., 2002). Noted methodological issues include researchers not providing clear (or, in some cases, any) information on scale validity and/or reliability (Fox, Ingham Clark, Scotland, Dacre, 2000; Hoppe et al., 1990; Stern et al., 1995). Some studies have external assessors using different assessment criteria from the self-assessing physicians (Ward et al., 2002), which can produce additional variability not accounted for when looking at direct correlations. Oftentimes, validity data or information on statistical methods are missing or unclear (Biernat et al., 2003; Fox et al., 2000). Therefore, the literature purports that methodologically sound self-assessment studies on medical residents are scarce.

Though infrequent, some valid and reliable studies have been undertaken to examine resident ability to self-assess. Others have also examined the development of resident self-assessment ability in preparation for their professional careers. These studies will be

discussed in the upcoming subsections. As seen throughout the self-assessment literature on professionals in other fields, the studies discussed have illustrated a distinct and problematic variability in resident ability to self-assess.

The first subsection of the following review of literature covers studies which have examined resident self-assessment ability from the static perspective. Their focus has been on residents' capability of effectively self-assessing. Studies presented in the two subsequent subsections present studies that examined the development of resident self-assessment ability from two schools of thought throughout the literature.

The first discussion of the development of resident self-assessment ability reviews studies that have experimentally observed the change in resident self-assessment ability due to explicit self-assessment interventions. The next subsection presents a discussion of studies that have applied Kruger and Dunning's (1999) framework, discussed earlier, to residents. The studies that have incorporated Kruger and Dunning's framework seek to determine if a relationship exists between resident competence in a domain and their ability to self-assess within it.

The studies within the following three subsections illustrate continued variability in residents' ability to self-assess, despite the importance of developing these skills within residency training. High or low correlations are not consistently tied to specific domains. The variability of these studies validates suggestions throughout the literature that resident self-assessment should be developed throughout residency training (Eva & Regehr, 2008;

Westberg & Jason, 1994) in ways that are generalizable across domains (Brinkman et al., 2007; Dunning et al., 2011; Eva & Regehr, 2005; Lipsett et al., 2011; Sargeant et al., 2011).

Resident ability to self-assess as a static skill. The majority of the resident self-assessment literature has sought to determine if residents have the general capability to effectively self-assess their knowledge or skills. Typically, these studies are each conducted in one or two domains. Within this subsection, the discussion will focus on studies that have shown variability in residents' demonstrated ability to self-assess. These studies were conducted throughout a variety of domains. Presented first is a focus on three studies that yielded high correlations between resident self-assessed scores and externally assessed scores, suggesting that residents are able to self-assess well. Following that is a discussion of studies which demonstrated moderate, low, and no significant correlations.

Studies which have reported the strongest correlations involve surgical residents evaluating their technical skills. Residents in studies done by Moorthy, Munz, Adams, Pandey and Darzi (2006); Mandel, Goff and Lentz (2005); and Brewster et al. (2008) demonstrated strong correlations between self-assessment of their technical surgical skills and faculty expert scores of their performance on simulated surgical procedures.

Moorthy et al. and Brewster et al. also examined their residents' ability to self-assess non-technical skills. For the Moorthy et al. study, this included pre-operative preparation, communication, vigilance, and leadership skills. The Brewster et al. study included pre-, intra-, and post-surgical communication skills. Despite the high-correlations found for their

residents' ability to self-assess their technical skills, no significant correlations were found for their ability to self-assess their non-technical skills in either study.

As demonstrated in part by Moorthy et al.'s and Brewster et al.'s results, even residents who have been effective at self-assessing in some domains have not always been able to effectively self-assess their knowledge or skills in all domains. While each of these studies showed high self-assessment correlations in some domains, low correlations were demonstrated in other domains and in other studies. These studies further substantiated claims that self-assessment ability is context-dependent (Brinkman et al., 2007; Dunning et al., 2011; Eva & Regehr, 2005; Lipsett et al., 2011; Sargeant et al., 2011).

In examining additional studies of resident self-assessment ability in a variety of domains, moderate, low, and no correlations between resident self-assessments and external assessments were reported. Contrary to the studies discussed previously, residents in Pandey et al. (2008) and Barnsley et al. (2004) studies demonstrated a poor ability to self-assess their technical skills when compared with external assessors. Both of these studies resulted in no significant correlation between self and external assessment. Similarly, variability in resident self-assessment correlations have been demonstrated throughout other knowledge and skill domains as well.

Three other studies examined resident self-assessment skills as a static ability. These studies solely measured non-technical knowledge and skills, as opposed to those which have already been discussed. Two of these studies examined residents' ability to self-assess their communication skills (Makoul, et al., 2007; Wood et al., 2004), and one examined residents'

ability to predict their medical knowledge scores on a standardized practice exam (Hawkins et al., 1999).

In a study conducted by Wood et al. (2004), resident to faculty assessor correlations were found to be moderately significant, and no significant correlations were found between resident to patient assessors. Makoul, et al. (2007) conducted a similar study in which they developed and tested an instrument that would allow patients to validly and reliably assess the communication skills of residents. The inference from these study results suggests that the residents studied were not effective assessors of their own communication skills.

One final study of static resident self-assessment ability was reviewed. The domain of Hawkins et al.'s (1999) focus was medical knowledge. The study examined Internal Medicine residents' self-assessment ability in terms of their performance on an impending Internal Medicine In-Training Examination (IMITE). Thirty-five residents participated in the study.

Of the thirty-five residents in this study, just over half (18) of them accurately predicted their score in the appropriate tertile. Ten residents underestimated their performance, and 7 overestimated their performance. No significant trends were noted, likely due to the small n. However, this study requested residents to rank their own knowledge among what they perceived their peers' knowledge to be (without having witnessed their performance). This added another level of variability since comparison required the residents to be able to first recognize the competence of their peers. Residents were then asked to measure their own performance against that of their peers. The additional

level of variability may have impacted the residents' ability to accurately assess their own performance in this study.

Gender and Ethnicity. Some studies on self-assessment have looked at the impact of gender and ethnicity on the ability to effectively self-assess. Some have reported a tendency for men to overestimate and for women to underestimate themselves but the results from the medical education literature (Colthart et al. 2008) is inconsistent.

A small study by Minter et al. (2005) examined gender differences in surgical residents' ability to accurately self-assess. The authors reported that both male and female residents underestimated their abilities compared with faculty. In comparison female residents underestimated their abilities to a greater extent than their male counterparts, but the difference was not significant.

Zonia & Stommel's (2000) study compared seventy-three interns' self-assessments of their medical knowledge and skills against those of their faculty. They found gender to have had no influence on either the interns' or faculty's ratings. However, the authors did not present any data to support their conclusion. Parker et al. (2004) also found no difference in their family medicine residents' ability to self-assess based on their gender.

While some studies in the medical education literature (Woolliscroft et al., 1993; Fitzgerald et al., 2003) have looked at the impact of medical student ethnicity on their ability to self-assess, no sound studies were found that look at this for residents.

In summary, this section of literature review has discussed studies that demonstrate considerable variability in resident ability to self-assess their performance as a static skill.

Clearly, some residents have the ability to self-assess their knowledge and skills in some contexts. As noted by the studies of Moorthy et al. (2006), Mandel, et al.(2005), Brewster et al. (2008), Wood et al. (2004), Makoul, et al. (2007), and Hawkins et al. (1999), there is a strong likelihood that residents are better at self-assessing technical over non-technical skills. However, that is not true for all cases (Pandey et al., 2006; Barnsley et al., 2004). The results of studies that examined the impact of resident ability to self-assess based on gender are inconclusive and the literature does not provide any evidence regarding the impact of ethnicity on resident self-assessment ability.

Some of the authors concluded that, without explicit instructional interventions in self-assessment, residents were not reliable self-assessors. However, the researchers did not suggest specific interventions. The following two subsections suggest means of resident self-assessment development from two different schools of thought: by incorporating explicit interventions and by increasing resident competence within deficient domains.

Explicit self-assessment interventions. Evidence throughout the medical education literature has suggested that self-assessment is a skill that will remain underdeveloped until it is explicitly addressed by instructional interventions specific to self-assessment as its own domain (Colthart et al., 2008; Dunning, et a., 2004; Fuhrmann & Weissburg, 1978; Gordon, 1991; Westburg & Jason, 1994). Most of the studies in this subsection demonstrated at least marginal improvement in resident self-assessment ability subsequent to an explicit intervention. However, the interventions were isolated to instruction and practice within one or two domains. The context-dependent nature of self-assessment ability (Brinkman et al.,

2007; Dunning et al., 2011; Eva & Regehr, 2005; Lipsett et al., 2011; Sargeant et al., 2011) leads to questions of generalizability of explicit interventions throughout other domains.

Multiple studies have indicated that resident self-assessment ability can and must be enhanced by explicit interventions such as direct and ongoing feedback, particularly video or verbal, the use of benchmarking (Colthart et al., 2008; Eva & Regehr, 2008; Hodges et al., 2001; Martin et al., 1998; Ward et al., 2003), explicit ongoing practice (Butterworth, 2010; Eva & Regehr, 2005; MacDonald; et al., 2003), and other direct instructional interventions (Colthart et al., 2008).

The importance of direct, explicit feedback is one focus of the literature discussing resident self-assessment development (Colthart et al., 2008; Davis et al.; Eva & Regehr, 2008; Hodges et al., 2001; Ward et al., 2003). Davis et al. (2006) suggested that individual, unguided self-assessment is not sufficient for recognizing strengths and weaknesses. It has been contended that “external feedback is essential” (Eva & Regehr, 2008, p. 18) to effective resident self-assessment. Using Schön’s (1983) theory of reflection on practice as the foundation for their discussion, Eva and Regehr posited that ongoing and explicit feedback during residency helps physicians to develop habits of seeking external sources of feedback for developing a “coherent self-awareness of [their] strengths and weaknesses” (p. 18).

Direct feedback has been discussed in terms of residents watching themselves perform on video, either independently or with the guidance of faculty. There are several studies throughout the resident self-assessment literature that have reported on the impact of feedback, typically via video or benchmarking their ability to self-assess. In these studies,

residents have been asked to assess benchmark videos of others performing the same skills that they had previously been asked to perform and self-assess. They were then asked to re-score their own performance. The intent of the benchmarking process is, first, to determine if the residents are able to recognize varying levels of competence in others. Secondly, benchmarking videos have been used to provide residents with a standard against which to base measurement of their own performance. Researchers then have examined if the residents' ability to effectively recognize their own level of performance competence improved based on their review of the benchmarking video(s).

Three studies in this subsection involved a combination of interventions of both direct feedback and residents viewing benchmark videos in an effort to increase their level of self-assessment ability. In a small study, Ward et al. (2003) studied senior resident surgeons' ability to self-assess their operative performance. Twenty-six senior surgical residents were videotaped performing laparoscopic procedures on pigs. The residents were asked to self-assess their performance subsequent to performing the procedure. These scores were compared with external assessment of the videotaped performances. The residents were asked to again self-assess their performance after reviewing their own video and then again after reviewing four benchmark performances of others demonstrating varying degrees of procedural proficiency. The initial correlation between resident self- and expert assessment was found to be moderate. The correlation increased significantly after residents reviewed their own videoed performance, but remained constant after viewing the benchmark videos.

Similar to the findings from the Moorthy et al., Mandel et al., and Brewster et al. studies, the results of Ward et al.'s study led the researchers to conclude that senior surgical residents were fairly accurate at self-assessing their surgical skills. Furthermore, they concluded that reviewing themselves on video enhanced the residents' ability to self-assess, but reviewing benchmarking videos did not.

Martin, Regehr, Hodges and McNaughton (1998) conducted a similar, but somewhat larger study in which they used video and benchmarks to determine the level of improvement in resident self-assessment ability. In their study, Martin et al. asked fifty first and second year Family Medicine residents to assess their performance after conducting a complex standardized patient consultation about suspected child abuse and then again after viewing others performing the same consultation.

The initial resident self-assessment prior to the viewing of any videos showed low to moderate correlation with external rating. The resident scores of the benchmarking videos correlated relatively highly with the expert assessments. These results indicated that, when observing other residents performing at various levels of competence, the residents were generally able to distinguish between the gradations of performance. Residents were not offered the opportunity to review their own videotaped performance. However, following viewing and rating the benchmark videos, the correlation between resident re-self-assessment of their own performance and expert ratings increased significantly to a moderate correlation. This increased correlation suggests that the residents were able to more accurately recognize

their own level of competence, at least to some extent, subsequent to identifying the varying levels of competence in others.

In an additional small benchmarking study, Hodges, et al. (2001) examined the self-assessment ability of 24 family medicine residents. The results of this study emphasized the growing concern regarding ineffective self-assessment skills in medicine. The residents in the Hodges et al. study performed simulated communication sessions during which they were asked to break bad news to standardized patients. The residents in this study subsequently viewed 4 videotaped communication sessions of other individuals demonstrating a range of communication skill ability from highly incompetent to highly skilled. Residents were each asked to score their performance both immediately following their standardized session and then again following their review of the benchmarking video. These scores were both compared with external assessments of their performance.

Because the main intent of this study was to determine improvements pre- and post-benchmark video review, only comparison significance numbers were given rather than the actual pre- and post-video significance numbers themselves. Therefore, it is difficult to ascertain the baseline resident self-assessment ability. However, it was indicated that the majority of the residents were inaccurate in their ability to self-assess. These results improved to a moderate level of correlation following resident review of benchmarking videos.

These researchers also peripherally examined the relationship between resident self-assessment ability of their communication skills and their competence within the domain.

Therefore, additional results from this study will be discussed further in the next subsection related studies that have tested this theory.

Researchers have also considered other explicit interventions to enhance resident self-assessment ability. Some research has also noted explicit repeated self-assessment practice within a given domain as an interventional approach to increasing resident self-assessment ability. Butterworth's (2010) study, for example, examined the relationship between resident self-assessment ability and a combination of explicit faculty feedback and repeated self-assessment practice over time. Her results suggest that residents' self-assessment ability may be benefitted by faculty feedback and repeated self-assessment practice in some domains. Furthermore, contrary to other self-assessment intervention studies, Butterworth examined resident self-assessment ability across multiple domains. The residents in Butterworth's study varied in their ability to self-assess across domains. These results confirm assertions from other researchers that resident self-assessment ability is domain specific (Brinkman et al., 2007; Butterworth, 2010; Dunning et al., 2011; Eva & Regehr, 2005; Lipsett et al., 2011; Sargeant et al., 2011).

In her study, fifteen Emergency and General Practice residents were asked to complete a self-assessment form each month over the course of six months. In addition, for each month, an external observer scored each resident on the same performance. Pearson Correlations were calculated for each of twelve criteria. The twelve criteria reflected seven separate domains, including: history taking, physical examination skills, communication skills, clinical judgment, professionalism, organization and efficiency, and overall clinical

care. Self-assessments were compared with faculty assessments using Pearson Correlation Coefficients and changes over time were documented.

Over the six months, thirteen out of the fifteen residents showed consistent accuracy in self-assessment when compared with faculty scores in at least one of the seven individual domains. Twelve of the residents moved from inaccurate self-assessment of their performance to accuracy for at least one of the domains. The domains the residents demonstrated accurate self-assessment varied by resident. However, only two residents corrected under or overestimate of their performance for their summative scores across the domains. This study demonstrated that ongoing feedback and repeated practice in self-assessment can be effective for helping to improve resident self-assessment ability for individual domains. However, the improvement does not appear to universally translate across domains.

The results of Butterworth's study echo others throughout the literature that have demonstrated variability in resident ability to self-assess. The variability across domains within this study and throughout multiple domains across a variety of studies substantiates common assertions that self-assessment ability is context-dependent (Brinkman et al., 2007; Dunning et al., 2011; Eva & Regehr, 2005; Lipsett et al., 2011; Sargeant et al., 2011). Therefore, the effectiveness of explicit interventions to enhance resident self-assessment ability is called into question, particularly given that they are typically conducted within a singular domain. The concern over the generalizable impact of instructional self-assessment interventions begs the question, "how might self-assessment ability be enhanced for various

domains without explicit interventions for each?” Considering the development of resident self-assessment ability by means of developing resident competence within the deficient domain may be the answer.

Competence begets self-assessment ability. The results of Kruger and Dunning’s study revealed that those with the least skill in the assessed domain might be most at risk of inaccurately assessing their ability. Some would suggest, in accordance with Kruger and Dunning’s (1999) theoretical framework, that residents will remain unaware of their strengths and weaknesses within a domain unless and until they become competent within that domain. This subsection will focus on studies that have examined the relationship between resident ability to self-assess and their level of competence within the assessed domain.

Mandel et al.’s (2005) study of surgical resident self-assessment ability regarding their technical skills did not find their results to be consistent with the framework laid out by Kruger and Dunning. Rather, the correlations between self-assessed and external scores were consistently high throughout the range of competence levels. Given the lack of variability shown in resident self-assessment ability in this study, a demonstrated relationship between self-assessment ability and domain competence would not be expected.

To the contrary, in a relatively recent study conducted by Jones et al. (2008) to determine Internal Medicine resident ability to self-assess their medical knowledge, the results indicated that residents at all levels of medical knowledge competence were inaccurate in predicting their performance on the IMITE. Only eight of twenty-six assessed

residents had scores within ten points of their predictions, with the majority of residents underestimating their performance. Standing alone, the results of this small study suggest that there is not a direct relationship between domain competence in medical knowledge and residents' ability to self-assess this domain. However, the size of the study and methods used for self-assessment examination must also be taken under consideration.

There were only twenty-six participants in this study, making it considerably small for effective data analysis (Kumar, 1996). An additional level of variability was also present, similar to that in the Hawkins et al. (1999) study discussed in the subsection on static resident self-assessment ability. In both of these studies residents were asked to self-assess their own performance against what they perceived their peers' performance would be. This technique requires residents to not only be able to assess their own level of competence, but also to estimate that of their peers. Therefore, it becomes uncertain as to whether poor correlations between self-assessed scores and external scores were the result of residents inaccurately assessing their own performance or that of their peers. Given this, it would be difficult for the researchers to examine a true relationship between resident self-assessment ability and their competence within the assessed domain.

Parker et al. (2004) conducted a very similar but much larger study of resident ability to assess their own medical knowledge compared with their annual Family Medicine In-Training Exam (ITE). In their study, 311 residents from thirteen different Family Medicine programs throughout Texas and Oklahoma participated. For this study, the residents were asked to complete a self-assessment of their knowledge in nine medical knowledge content

areas represented on their annual Family Medicine In-Training Exam (ITE). Though the measurement tool also asked the residents to predict their performance in relative percentiles against that of their peers, the results of this study supported Kruger and Dunning's framework. Residents in the lowest quartile were the least able to recognize their level of (in)competence. Similar to what was shown in Kruger and Dunning's studies, the middle quartiles were the most accurate in their assessment ability and the top quartile underestimated their competence.

One final study was found in the literature examining the relationship between resident self-assessment ability and their level of competence within the assessed domain. Hodges, et al.'s (2001) study of the impact of benchmarking videos on resident ability to self-assess their communication skills also peripherally compared their residents' self-assessment ability against their level of competence within the domain. This study examined the relationship between resident ability to self-assess their communication skill and their competence within the domain only peripherally. The relationship was examined ancillary to the main intentions of the study. The main intentions of the study were to examine the impact of viewing benchmarking videos on resident self-assessment ability at various competence levels and to validate the use of z-scores to standardize the residents' self-assessment scores as an appropriate means of self-assessment ability analysis. Results would determine the impact of video benchmarking on resident ability to self-assess. Therefore, only inferences, not concrete data, were provided regarding the relationship between resident self-assessment ability and their communication skills competence.

Residents within the top and bottom tertiles demonstrated inaccurate self-assessment ability. As shown in other studies, the least competent residents (as determined by external assessment) overestimated their performance and the highest competent residents underestimated (Parker et al., 2004). The middle group demonstrated accurate self-assessment ability. Only the two inaccurate groups were re-assessed following benchmarking.

Every resident within the top tertile demonstrated improvement in their ability to self-assess after reviewing the benchmarking videos. However, the bottom tertile demonstrated considerable variability in their re-assessment. Only two out of the eight in this group demonstrated an appropriate shift toward accurate self-assessment. Others adjusted their scores too dramatically, showed no change, or scored themselves even more highly than they did originally. This study notes a growing concern regarding ineffective self-assessment skills in medicine, particularly for the least competent in the assessed domain.

It is clear that there is extreme variability in medical resident ability to self-assess. Given the importance of effective self-assessment ability to physician practice, this variability emphasizes the necessity for residents to develop this ability prior to graduation from formal training (Eva & Regehr, 2008; Westberg & Jason, 1994). However, the variability in effective development of resident self-assessment ability is also of concern. Even though some instructional interventions have shown an increase in resident ability to self-assess, the context-dependent nature of self-assessment ability makes interventions, typically implemented and practiced within one or two domains, less than ideal.

The goal of this study was to determine if increasing resident competence within the assessed domain is an effective means of developing resident self-assessment ability during their residency training. Providing additional evidence of this relationship proposed within Kruger and Dunning's framework may refocus researchers and practitioners on the development of resident knowledge and skill domain deficits. This approach would take the place of a focus on developing explicit self-assessment interventions to improve self-assessment ability. According to this framework, improving resident knowledge and skill domain deficits naturally increase their ability to recognize their strengths and weaknesses within those domains. However, further evidence of the relationship as it applies to residents is necessary (Hodges et al., 2001) to validate this claim.

The current study has focused on applying Kruger and Dunning's framework to residents' ability to self-assess their communication skills. Methodologically sound studies undertaken to examine the relationship between domain competence and resident ability to self-assess within that domain have not been plentiful, particularly in the domain of communication skills. Of the four studies applying Kruger and Dunning's framework to residents, one was related to resident technical skills (Mandel et al., 2005), two were related to resident medical knowledge (Jones et al., 2008; Parker et al., 2004), and only one was related to communication skills (Hodges et al., 2001).

Of the two medical knowledge studies, an additional level of variability to examining resident self-assessment ability existed in that residents first needed to be able to estimate the competence of their peers prior to assigning themselves relative self-assessment percentile

scores. Additionally, the sample size for the Jones et al. (2008) was considerably small (n=26), making it difficult to establish an accurate analysis of the data (Kumar, 1996). In the one study relating resident competence in communication skills to their ability to self-assess that domain, the sample size was also considerably small. Furthermore, the relationship was only studied peripherally to the main intent of the study, which was to assess the impact of benchmarking videos on resident self-assessment ability. Therefore, specifics on self-assessment correlations and significance were not reported. This study provides additional evidence to support claims that Kruger and Dunning's (1999) theoretical framework is applicable to resident self-assessment ability, particularly as it relates to assessment of their communication skills.

Communication Skills Assessment Tools

Because resident self-assessment ability has been shown to be context-specific, it is important to measure the ability separately within various domains. Otherwise, it cannot be assumed that the results of resident self-assessment ability within 1 domain will translate to others. This study has examined resident self-assessment ability within the domain of communication skills. Most of the self-assessment research on resident communication skills either has been focused on developing a tool that can be used for communication skill self-assessment or the use of self-assessment as an intervention (Brinkman et al., 2007). Therefore, valid, reliable tools for comparing residents' self-assessment of their communication skills with external measurements are not prevalent throughout the research and literature.

As illustrated within the previous section discussing resident self-assessment studies, there is limited research correlating resident self-assessment of communication skills with valid and reliable external measurement. Assertions have been made throughout the literature about the importance of using the same assessment tool for both external and self-assessment when examining resident self-assessment ability (Colthart, 2008; Gordon, 1991). Therefore, the literature related to externally assessing resident communication skills was reviewed for appropriate measurement tools for both external and self-assessment for this study.

Physician communication skills studies have been reported extensively throughout the literature. Literally thousands such studies have been conducted over the past ten years. One measurement tool that stood out among the rest for this study was the Breaking bad news Assessment Schedule (BAS) developed by Miller et al. (1999). The BAS addresses a broad range of elements related to physician communication skills. It also provides a descriptive set of anchors for what constitutes assigning each score for each item throughout the scale.

The BAS is a communication skills assessment scale that has been previously established in the literature. The BAS measures general communication skills, but it also provides some focus on skills specifically relevant to breaking bad news. The BAS was developed by Miller et al. (1999) in an effort to standardize a means of assessing residents' communication skills, particularly as they related to breaking bad news. This scale was

developed and tested with the intent of providing a valid, standardized way of assessing the communication skills necessary to effectively break bad news.

The content included in the BAS was determined by an extensive literature search. The literature search was performed to determine a range of key behaviors important to effectively communicating bad news to patients. The literature search initially identified eighty-one key behaviors, which were, in turn, grouped into the twenty-three main elements reflected in the BAS. The final BAS comprised twenty-three questions (based on the main elements identified) to be rated on a 5-point likert scale. Each item included clarifying descriptions of behavioral points indicative of the most desirable communication behavior.

Content validity was established by agreement of the scale content to that which has been established throughout the literature as key elements of effective communication skills for breaking bad news to patients. External validity is particularly difficult to establish, considering the lack of a “gold standard” of measurement throughout the communication skills literature. However, patients’ own views regarding an encounter have been postulated as a potential gold standard (Miller et al., 1999) and standardized patients which were used for this study are often utilized to approximate the patient perspective (Brewster, 2008; Epstein et al., 2004; Iramaneerat, Myford, Yudkowsky & Lowenstein, 2009; Lang et al., 2004; Miller et al., 1999; Pandey et al., 2008; Yudkowsky, Downing & Sandlow, 2006; Ward et al., 2003).

The researchers (1999) used Cronbach’s alpha to measure the Scale’s internal consistency or reliability. The Scale’s Cronbach’s alpha was found to be .93. Interrater

reliability was calculated in two ways: using a weighted kappa statistic (Landis & Koch, 1977) and an Analysis of Variance (ANOVA) was performed. Cohen's Kappa (commonly referred to as "Kappa") is often used as a measure of agreement between two individuals. Kappa measures the amount of agreement that could be expected due to chance alone. The range of kappa is -1 to 1. Landis and Koch (1977) have suggested the following to interpret Kappa coefficient scores: ≤ 0 = poor, 0.01–0.20=slight, 0.21–0.40=fair, 0.41–0.60=moderate, 0.61–0.80=substantial, and 0.81–1=almost perfect. Miller, et al. (1999) found the interrater reliability for the BAS to be "moderate to good (0.45-0.68) and considerably greater than that which would be expected by chance" (Miller et al., 1999).

An ANOVA was performed to analyze the Intraclass Correlation Coefficient (ICC), systematic bias, and random error (Brennan & Sillman, 1992) to determine areas of disagreement. As expected, the variation in the interview performances was shown to account for a moderate to high level (62%) of the amount of variance in the scoring between raters.

The BAS has been modified slightly, used, and further validated in two separate studies since 1999. In 2007, English, Naeve-Velguth, Rall, Uyehara-Isono and Pittman published the results of their study using a modified version of the BAS to assess audiology students' counseling ability when breaking bad news to parents about their child's hearing. The revised scale was titled the "Audiologic Counseling Evaluation." These researchers found the ACE's internal consistency to be quite high, with a cronbach alpha of 0.91 and interrater reliability moderate to good (0.57-0.67). A 2011 article (Schildmann, Kupfer,

Burchardi & Vollmann) reported on the use of a modified BAS used by standardized patients to assess medical student performance on a standardized patient visit during which the student was asked to break bad news to the patient. No updated internal consistency was reported for the instrument. However, interrater reliability was noted to be quite high at 0.86.

The BAS's high internal consistency and moderate to high interrater reliability as demonstrated in Miller et al.'s (1999) original study, and also in subsequent studies support its value as an effective tool for measuring resident communication skills. Suggestions throughout the literature noting the importance of using the same measurement tool for both external and self-assessment for meaningful evaluation of resident self-assessment ability (Colthart, 2008; Gordon, 1991) support this tool's use for measuring resident self-assessment ability of their communication skills.

Summary

Self-assessment is a primary component of metacognitive monitoring and is essential to effective self-regulated learning. The literature suggests that self-regulated learning skills, including self-assessment, can and should be developed and evaluated through experiences in professional education (Colthart et al., 2008; Curran et al., 2007; Dunning et al., 2004; Fuhrmann & Weissburg, 1978; Gordon, 1991; MacDonald et al., 2003; Schunk & Zimmerman, 2008; Ward et al., 2003; Westburg & Jason, 1994). The self-assessment skills developed throughout a physician's residency training will prepare them for the self-regulated learning they will face throughout their careers as independent healthcare providers.

The variability of resident self-assessment ability as a static skill is problematic, but additional studies focused on the potential for development of physician self-assessment skills during their professional training are promising. However, the studies are sparse, and no conclusive evidence has been presented in support of either of the two main schools of thought regarding resident self-assessment development.

Researchers have long suggested that explicit instructional interventions to teach residents self-assessment as its own domain are necessary. Several key studies discussed here have examined the impact of explicit interventions on resident ability to self-assess. However, the variability of the impact of explicit interventions on resident self-assessment ability persists. The application of the Kruger and Dunning (1999) framework, relating resident competence within a given domain to their ability to self-assess their strengths and weaknesses within it, presents an alternative approach to self-assessment development in residency training. A small subset of the literature has examined this relationship within residency training. However, the studies available are less than robust, and only one study even peripherally addresses this relationship within the domain of communication skills. The Breaking bad news Assessment Schedule (BAS) (Miller et al., 1999) was presented as an instrument to measure both external and self-assessment and compare the resulting scores accordingly to answer the research questions for this study.

Self-assessment has been said to be context-specific, and instructional interventions throughout the literature have only focused on one or two domains. If Kruger and Dunning's theoretical framework effectively translates to the field of medicine, it might suggest that

researchers should focus on methods to improve knowledge and skill domains identified as resident deficits to accordingly improve physician ability to determine their strengths and weaknesses within these areas. Such knowledge and skill-based intervention could be done in lieu of developing separate instructional interventions for individual domains. However, further studies need to be conducted in the domain of resident communication skills to substantiate these claims.

CHAPTER THREE: METHODOLOGY

The purpose of the chapter is to present the research design and methods used for the current study. The chapter begins with a discussion of the appropriateness of the explanatory, non-experimental research design and the six guiding research questions used for the study. Next, the chapter includes a presentation of the variables used in the study. Other sections of the chapter include descriptions of the sample site and sample and instrumentation. Given that this study has used secondary data for analysis, a brief description of the procedure for how the data was collected will be provided. Preliminary data analysis and statistical data analysis will also be discussed. The preliminary data analysis section covers assumptions and reliability for the instrument used in the study. The chapter concludes with an overall summary of the research design and methodology.

Research Design

This study has employed a descriptive, non-experimental research design (Johnson, 2001). Kerlinger and Lee (2000) stated, “Nonexperimental research is systematic empirical inquiry in which the scientist does not have direct control of independent variables because their manifestations have already occurred or because they are inherently not manipulable” (p. 558). Sproull (2002) defined non-experimental design as one that does not have an experimental variable, but does include a variable that can be measured. Additionally, Sproull (2002) stressed that the researcher “has control of who or what to measure, when the measurement takes place and what to ask or observe” (Sproull, 2002, p. 153). For the current study, the researcher chose to perform a cross-sectional examination of Duke University

pediatric residents' ability to self-assess their performance in breaking bad news in a standardized communication skills session. The focus of the study was on the relationship between resident competence in the domain of communication skills and their ability to self-assess within the domain. The researcher determined the variables to measure and the time to conduct the study.

This study used secondary data collected from a 2011 standardized communication skills session during which each resident had the opportunity to self-assess and be externally assessed by Standardized Patient Parents (SPPs). The SPP scores were used as a proxy to determine the resident competence. The residents' self-assessed scores self-assess were determined by their own perceptions of communication skills. Their ability to self-assess was calculated by subtracting the SPP score from the residents' self-assessed score.

Research Questions

The research questions for this study sought to examine resident ability to self-assess their communication skills competence when breaking bad news. To examine the relationship between the development of resident competence within the given domain and resident self-assessment ability within the same domain, this study sought to answer the following research questions:

1. What are the varied levels with which residents demonstrate competence within the domain of communication skills as measured by Standardized Patient Parents (SPPs) using the Breaking bad news Assessment Schedule (BAS) instrument?

2. Do residents' self-assessments correspond with Standardized Patient Parent assessment of their communication skills performance?
3. Does resident communication skill competence relate to their ability to self-assess within the domain?
4. Does resident gender relate to their ability to self-assess their communication skills performance?
5. Does the number of years of residency training experience relate to residents' ability to self-assess their communication skills?
6. Can a significant amount of variance in resident ability to self-assess their communication skills be explained by gender, year in training, ethnicity, and theoretical element of communication skills competence (setting the scene, breaking the news, eliciting concerns, information giving, and general information)?

Variables Used in the Study

The research variables for this study were organized as dependent or independent variables for each of the six research questions. The dependent variable has been defined as the residents' self-assessed ratings minus that of the SPPs ratings. The main independent variable has been defined as Duke Pediatric residents' level of communication skill competence as measured by Standardized Patient Parent (SPP) ratings of their performance in a simulated interaction. Scores were categorized into tertiles and relationships were determined accordingly. Additional independent variables, were also explored to determine if a relationship exists between these variables and residents' ability to self-assess.

Table 1

Independent and Dependent Variables Used for this Study

Variables	
Independent Variables	Resident Communication Skills Competence Year in Training Program Gender Ethnic Background
Dependent Variable	Resident Ability to Effectively Self-Assess Their Communication Skills

The dependent variable for this study, self-assessment ability, was determined by comparing mean self-assessed scores to mean standardized patient parent (SPP) scores. Mean sums of scores were tabulated by independently totaling and averaging self-assessed and SPP scores for each tertile of competence and each level of residency training.

Sample Site and Sample

The site for this investigation was the Duke Children's Hospital and Health Center located within the Duke University Medical Center in Durham, North Carolina. Duke Children's Hospital serves as both a primary care facility and a referral center for patients from North Carolina and throughout the Southeast. Each year, there are over 5,000 inpatient pediatric admissions, 2,500 deliveries, and 176,000 pediatric ambulatory visits.

While external validity is important, the researcher for this study is not trying to generalize to populations outside the specific sample. Therefore, a purposeful sample was used. All participants were Duke University Pediatric residents of varying levels of training and communication skills competence. Data was collected for all Duke Pediatric residents

and also residents in the combined Duke Internal Medicine/Pediatric (MP) residents. Select individuals may not have participated in this data assessment because they were unable to attend due to prior clinical commitments or pre-scheduled vacation. Residents in the data collection included those who were starting residency, as well as those ending their first through to their third and final year of training. A total of eighty-two residents participated in the study. However, due to missing data, only seventy-nine were included in the analysis. Resident demographic diversity is described in Table 2.

Table 2

Resident Participant Demographics

Variable	n	%
Level		
1	20	25.3
2	21	26.5
3	23	29.1
4	15	19.0
Gender		
Male	26	32.9
Female	53	67.1
Ethnicity		
White	60	75.9
Asian	5	6.3
Black	12	15.2
Hispanic	2	2.5

Instrument

The nature of self-assessment ability research does not lend itself to using a specific instrument to determine self-assessment ability across multiple fields or domains in the traditional sense (Falchikov & Boud, 1989). Rather, self-assessment ability has been determined by correlating self-assessed scores with the scores of external assessors (Gordon, 1991). These studies have most often been done within a domain selected by the researcher. Assertions have been made throughout the literature on the importance of using the same assessment tool for both external and self-assessment when examining resident self-assessment ability (Colthart, 2008; Gordon, 1991). Therefore, for resident self-assessment studies, an instrument designed for measuring residents' level of competence within the given domain have been used by both external assessors and the self-assessors themselves. The resulting scores have later been compared for analysis of self-assessment ability.

Communication skills is the domain of measurement for resident self-assessment ability in this study. Communication is a competency that residency-training programs are required to teach and assess by the Accreditation Council for Graduate Medical Education (ACGME) (ACGME, 2011). There is also evidence throughout the literature that identifies communication skills as a domain that is significant to physician practice and central to effective patient care (Rider & Keefer, 2006, Simpson et al., 1991; Tulskey, 2005; Watling & Brown, 2007). However, this domain has not been extensively studied in relation to residents' ability to assess their strengths and weaknesses within it, particularly related to their competence with the domain. The communication skills case utilized for this study

involves each resident breaking bad news (an element specific to the ACGME requirements) (ACGME, 2011) to standardized patient parents regarding their child's health.

A communication skills assessment scale that has been previously established in the literature, the Breaking Bad News Assessment Schedule (BAS) (Miller et al., 1999), was utilized for assessment in this study. The BAS measures general communication skills, but provided a focus on skills specifically relevant to breaking bad news. The BAS was modified slightly, as addressed in a subsequent section, for the purposes of this study. The final scale used for this study included eighteen items purported to reflect five separate components of communication skills important to breaking bad news to patients and/or families. These components included: setting the scene, breaking the news, eliciting concerns, information giving, and general information.

The BAS was utilized in this study both for resident self-assessment of their own performance as well as for the standardized patient parents (SPPs) to assess each resident's performance. Standardized patients have been established throughout the literature as valid and reliable external/ assessors of resident communication skills (Brewster, 2008; Epstein, Dannefer & Nofziger, 2004; Iramaneerat, Myford, Yudkowsky & Lowenstein, 2009; Lang, McCord, Harvill & Anderson, 2004; Miller et al., 1999; Pandey et al., 2008; Yudkowsky, Downing & Sandlow, 2006; Ward et al., 2003). As has been done throughout the literature, the residents' self-assessed scores were compared with the SPP scores to establish if they correlate and at what levels.

Validity and Reliability

The BAS was developed in an effort to standardize a means of assessing resident communication skills, particularly as they related to breaking bad news. This scale was developed and tested with the intent of providing a valid standardized way of assessing the communication skills necessary to effectively break bad news.

The content included in the rating scale was determined by an extensive literature search by Miller et al. (1999). The literature search was performed to determine a range of key behaviors important to effectively communicating bad news to patients. The literature search initially identified 81 key behaviors that were, in turn, grouped into the 23 main elements reflected in the BAS. The final BAS was comprised of 23 questions (based on the main elements identified) to be rated on a 5-point likert scale. Each item includes clarifying descriptions of behavioral points indicative of the most desirable communication behavior.

The researchers established content validity for the Scale by aligning the Scale content to the key content identified in the literature as significant elements of effective communication skills for breaking bad news to patients. External validity was particularly difficult to establish for a communication skills scale, given the lack of a “gold standard” of measurement throughout the communication skills literature. However, patients’ own views regarding an encounter have been postulated as a potential gold standard (Miller et al., 1999). Standardized patients, or patient parents, which were used for this study, have often been utilized to approximate the patient perspective (Brewster, 2008; Epstein et al., 2004; Iramaneerat et al., 2009; Lang et al., 2004; Miller et al., 1999; Pandey et al., 2008;

Yudkowsky et al., 2006; Ward et al., 2003). Standardized patient situations offer an opportunity for residents to practice and be assessed on their skill in a safe, non-threatening forum. Formal standardized patients are those whom have been appropriately trained to act out patient care scenarios in a consistent and realistic way. Consistency is important particularly for assessment scenarios so that all residents can be assessed in a standardized manner.

Items in a reliable assessment scale should correlate in a way that demonstrates that they are measuring the same thing (Choudhury, 2010). Miller et al. (1999) used Cronbach's alpha to measure the Scale's internal consistency or reliability. The Scale's Cronbach's alpha was found to be .93.

Interrater reliability was calculated in two ways: using a weighted kappa statistic (Landis & Koch, 1977) and an ANOVA. Cohen's Kappa (commonly referred to as "Kappa") is often used as a measure of agreement between two individuals. Kappa measures the amount of agreement that could be expected due to chance alone. The range of Kappa is -1 to 1. Landis and Koch (1977) have suggested the following to interpret Kappa coefficient scores: ≤ 0 = poor, 0.01–0.20=slight, 0.21–0.40=fair, 0.41–0.60=moderate, 0.61–0.80=substantial, and 0.81–1=almost perfect. Miller, Hope and Talbot (1999) found the interrater reliability for the BAS to be "moderate to good (0.45-0.68) and considerably greater than that which would be expected by chance" (Miller et al., 1999).

The original authors performed an ANOVA to analyze the Intraclass Correlation Coefficient (ICC), systematic bias, and random error (Brennan & Sillman, 1992) to

determine areas of disagreement. As expected, the variation in the interview performances was shown to account for a moderate to high level (62%) of the amount of variance in the scoring between raters.

Scale Modification

The researcher for this study made slight modification to the original BAS (Miller, Hope, Talbot, 1999) to make measurement more meaningful for the purpose of the study (Appendix A). The general instructions were first modified to remove the statement “The points below each question are for guidance only.” In addition, the statement, “When the doctor has delivered the diagnosis stop the tape and mark the first two sections before restarting and continuing to mark” was removed. Both were removed because they did not apply to this study.

Five scale items of the original twenty-three were removed from this scale. Item 1 was removed because it referred to the manner in which the resident may have arranged the environment to make it conducive to communicating this information. The standardized interaction for this study did not allow for the resident to manipulate aspects of the physical environment or setting.

Item 11, “Did the doctor explicitly check which areas were most important to the patient?” was removed due to its close nature to item 10 “Did the doctor explicitly attempt to obtain a complete list of the patient’s concerns?” and the limited time each resident has in the standardized session. The amount of time given for the interaction would not feasibly allow

for residents to thoroughly explore patient parent concerns much beyond that which is assessed with item 10.

Item 15 was removed because the standardized patient parent was not expected to assess the factual nature of the information provided. Item 16 and 17 were removed because they related directly to the breast cancer case utilized in the original study. Some of the examples from item 23 were removed to make it more meaningful to the standardized case utilized for this study. The term “patient” was replaced with the term “patient parents” or the acronym “pp” with related pronouns adjusted accordingly throughout the assessment. In total, 5 items were removed from the assessment, the instructions were modified slightly, and the terminology was changed to fit the case being evaluated. All changes to the assessment were made in an effort to make the instrument more meaningful for the intent of this study.

Data Collection and Confidentiality

Secondary data was used for this study. In the spring of 2011, each resident participated in one simulated communication session with a standardized patient mother and father. The standardized parents are trained actors hired through the Duke University School of Medicine Standardized Patient (SP) Program. The SP Program has over 100 actors trained to portray a patient or patient parent. For this study, a total of three trained Standardized Patient Parents (SPPs) were used to simulate patient parents during the various sessions using one SP breaking bad news scenario with one SPP being consistent throughout all sessions. The scenario used for this study involved each resident telling the parents of a newborn baby that the resident suspects their baby to have Down's Syndrome upon initial exam. Following

the simulated session, each resident was asked to self-assess their own performance using the BAS Scale (Appendix A). Also immediately following the interaction, the two SPP rated each resident's performance using the same scale.

This assessment of communication skills is an existing component of the pediatric curriculum. The program required participation in the communication sessions. However, each resident was given the option of not having his or her personal results included for data analysis. Each resident signed a Duke informed consent (Appendix B). This study was originally exempted by Duke University Medical Center IRB (Appendix C) and, later, also by North Carolina State University's IRB (Appendix D).

Results were collected and tabulated and will have no impact on formal resident evaluation or resident progress within the program. Results were coded using an anonymity code for comparison purposes. Results have been reported only in aggregate and have included no resident identifiers. Data is stored in this system and available only to those with password access.

Preliminary Data Analysis

The data collected for this study has been analyzed using the statistical package SPSS, version 20. Preliminary data analysis included addressing missing data and insuring assumptions have been met for appropriate data analysis. Data has been assessed for missing values in individual variables. Missing data was not greater than ten percent in any variable. Therefore, further analysis was not necessary to examine for non-random relationships.

Data was further examined using a variety of techniques to confirm assumptions were met for the selected statistical procedures and multivariate data analysis. According to Hair, Black, Babin, Anderson, and Tatham (2006), assumptions include: “1. Linearity of the phenomenon measured, 2. Constant variance of the error terms, 3. Independence of the error terms, 4. Normality of the error term distribution” (p. 204).

Univariate Distribution

Prior to examining assumptions of the multivariate model variate, the separate variables were tested for normality, kurtosis, and skewness. Osborne and Waters (2002) stated, “Non-normally distributed variables (highly skewed or kurtotic variables, or variables with substantial outliers) can distort relationships and significance tests.” Histograms were examined for all variables. Additionally, histograms were examined for variates. Further, skewness was used to examine the symmetry and kurtosis was used to examine the peakedness for the original and transformed variables and variates. Some skewness was anticipated and noted given the residents’ previous communication skills experience. However, skewness and kurtosis were found to otherwise be fine.

Multivariate Assumptions

Hair et al. (2006) advocated that it is essential to assess the assumptions for multiple regression. The four major assumptions include linearity of the phenomenon measured, constant variance of the error terms, independence of the error terms, and normality of the error term distribution (Hair et al, 2006). The researcher noted the importance to assess the assumptions “not only for each dependent and independent variable, but for the variate as

well” (p.208). According to Hair et al. (2005), “Perhaps the most frequently encountered assumption violation is non-normality of the independent or dependent variables or both” (p. 208). Statistical and graphical analyses were used to assess for the assumptions and transformations was completed where needed.

Statistical Data Analysis

For analysis, the residents were separately stratified according to their respective competence tertile (as determined by SPP assessment scores). Much of the literature that has examined self-assessment skills in relation to learner competence in the assessed domain has stratified residents according to competence quartile in the assessed domain (Austin & Gregory, 2007; Davis, Mazmanian, Fordis, Van Harrison, Thorpe & Perrier, 2006; Hodges et al., 2001; Regehr & Mylopoulos, 2008). However, the second and third quartiles have often shown to have such similar findings. These findings indicate that the quartiles can be collapsed into tertiles for effective analysis (Hodges et al., 2001).

The first research question sought to examine the varied levels with which residents demonstrate competence within the domain of communication skills as measured by SPPs using the Breaking bad news Assessment Schedule (BAS) instrument and the varied levels with which the residents assess their own competence using the same instrument. To examine these levels, the residents were first categorized into tertiles based on their SPP scores. Frequencies, percents, means, and standard deviations were then computed for the data analysis.

The second research question sought to describe the relationship between residents' self-assessed scores and those of the SPPs to determine the accuracy with which the residents are able to self-assess. These two variables were analyzed with a correlation and t test. The paired-samples t test is appropriate for comparing the means of two variables. It computes the difference between the two variables for each case, and tests to see if the average difference is significantly different from zero (O'Rourke, Hatcher, and Stepanski, 2005).

The third research question explored the relationship between resident competence within the communication skills domain and their ability to self-assess within the domain. This question employed a Chi Square test. According to Sharp (1979), a Chi Square test should be used to determine whether there is independence between the expected frequencies and the observed frequencies in one or more categories.

The fourth research question explored the relationship of gender on resident ability to self-assess their communication skills. This question was also analyzed using a Chi Square test.

The fifth research question explored the relationship of residency training program years of experience on resident ability to self-assess their communication skills. This question employed a Chi Square test as well.

The sixth research question explored the amount of significant variance in resident self-assessment ability of their communication skills explained by gender, year in training, ethnicity, and theoretical element of communication skills competence. A multiple regression analysis was performed to answer this question.

Osborne (2008) has indicated that “exploring relationships between multiple variables in a sample to shed light on a phenomenon, with a goal of generalizing the new understanding to a general or specific population” (p. 300) is the purpose of multiple regression analysis for explanation.

Summary

This chapter described the research design and statistical methods that were used to examine how accurate pediatric residents were able to self-assess their communication skills as compared with standardized patient parent judged competence. The data utilized in this study was gathered using a physician communication skills assessment tool for which evidence of their validity and reliability will be presented. The BAS instrument was utilized to compare resident self-assessed scores against standardized patient parent assessed scores. To answer the research question, descriptive statistics and a paired t-test were used. Descriptive statistics, Chi Square tests, and a multiple regression analysis were completed to meet additional, complementary questions.

CHAPTER FOUR: DATA ANALYSIS

The chapter provides the analysis of data and findings from collected data concerning medical residents' ability to self-assess their communication skills. The chapter is divided into nine main sections; one section discussing the instrument, one to provide an overview of the resident demographics, one section for each of the 6 research questions, and a final section for an overall data analysis summary.

Research Questions

The analysis of data and findings from the study were organized based on the following research questions guiding this study.:

1. What are the varied levels with which residents demonstrate competence within the domain of communication skills as measured by Standardized Patient Parents (SPPs) using the Breaking bad news Assessment Schedule (BAS) instrument?
2. Do residents' self-assessments correspond with Standardized Patient Parent assessment of their communication skills performance?
3. Does resident communication skill competence relate to their ability to self-assess within the domain?
4. Does resident gender relate to their ability to self-assess their communication skills performance?
5. Does the number of years of residency training experience relate to residents' ability to self-assess their communication skills?

6. Can a significant amount of variance in resident ability to self-assess their communication skills be explained by gender, year in training, ethnicity, and theoretical element of communication skills competence (setting the scene, breaking the news, eliciting concerns, information giving, and general information)?

Each resident in four pediatric program levels from incoming interns (soon to begin their first year of pediatric residency training) to impending graduates participated in a breaking bad news standardized scenario with a set (mother and father) of patient parent actors. Immediately subsequent to the scenario, the residents were asked to self-assess their performance using the BAS and both SPPs were asked to assess each resident's performance as well.

Instrument

Data were collected using the Breaking bad news Assessment Schedule (BAS) as described in the previous chapter, completed by each resident and two standardized patient parents (SPPs) ("mother" and "father") for each. The BAS included eighteen items divided into five sections of communication skills important to breaking bad news to patients and/or families. These components included: *Setting the Scene*, *Breaking the News*, *Eliciting Concerns*, *Information Giving*, and *General Information*. Each item was measured using a 5-point Likert scale to assess each resident's performance in the standardized communication skills session. The 5 points of measurement ranged from 1 (no agreement with the item's statement) to 5 (full agreement with the item's statement). The lowest possible score for the full assessment would be 18 and the highest would be 90.

Though originally considered for this study, a factor analysis was not conducted for this instrument since a larger sample size is recommended for a factor analysis to be accurate. Gorsuch (1983) recommended a sample size of at least 100. Hutcheson and Sofroniou (1999) recommended at least 150-300 cases and Hatcher (1994) suggested that the number of subjects should be either 100 or 5 times the number of variables (whichever is larger). Other researchers have recommended a considerably larger sample. Comrey and Lee (1992), for example, suggested greater than 500 subjects. Based on these recommendations and with a sample size of 79, it was determined that a factor analysis would not be meaningful for this study. However, reliability coefficients were calculated for each section for this study.

Setting the Scene

The two items in the *Setting the Scene* section of the BAS assessed residents' ability to begin the interaction with the SPPs. Anchors for items in this section included descriptions regarding appropriately introducing oneself and establishing the SPPs' knowledge of recent events, and their emotional baselines. The reliability for *Setting the Scene* was .49 when calculated using the residents' self-assessed scores. When using the SPP scores, the reliability was .80.

Breaking the News

The five items in the *Breaking the News* section assessed residents' ability to effectively break the news itself. This section's anchors described determining SPP base knowledge of the situation and expectations of the interaction, approaching with sensitivity and giving the SPPs time to respond. The reliability for *Breaking the News* was .70 when

calculated using the residents' self-assessed scores. When using the SPP scores, the reliability was .76.

Eliciting Concerns

The *Eliciting Concerns* section included two items with anchors describing inviting specific SPP questions and concerns. The reliability for *Eliciting Concerns* was .62 when calculated using the residents' self-assessed scores. When using the SPP scores, the reliability was .42.

Information Giving

The *Information Giving* section included three items. Anchors for these items described giving information specific to SPP concerns and in an ordered, logical manner, and also framing the given information in a positive way. The reliability for *Information Giving* was .57 when calculated using the residents' self-assessed scores. When using the SPP scores, the reliability was .73.

General Information

The *General Information* section included six items. These items and their descriptive anchors addressed psychosocial issues, appearing supportive, body language, and pace and timing. The reliability for *General Information* was .78 when calculated using the residents' self-assessed scores. When using the SPP scores, the reliability was .75.

Though several of the sections demonstrate low reliability for either the residents or SPPs (or both), the overall reliability for the instrument was considerably higher. For the residents' self-assessed scores, the reliability was estimated at .84 and .89 for the SPP scores.

Therefore, the remainder of the analysis will be discussed in terms of the overall scale rather than each separate section.

Demographics

The demographic characteristics for this study included gender, age, ethnicity and level of pediatric training. Demographic data were collected using MedHub, Duke University Medical Center's Graduate Medical Education trainee database. Table 2 illustrates the frequencies and percentages for each of these variables. Only the research assistant, not the researcher, was able to identify the participants for demographic linking purposes.

The participants were fairly evenly distributed among levels of training. There were twenty participants from both the incoming intern class as well as the rising second year class. Twenty-four rising third year residents participated as well as fifteen impending graduates. The ages of the participants ranged from 25 to 33 with a mean age of 28. The vast majority of participants were female which is typical for the female-dominated field of pediatrics (Goodman, 2005). The majority of participants were white (n=60, 76%), twelve (15%) were of Asian descent, and seven (9%) were of ethnicities considered to be underrepresented in the field of Medicine (Black, n=5, 6%; Hispanic, n=2, 3%) (AAMC, 2004).

Research Question One: Resident Communication Skills Domain Competence

The following section presents the findings and analysis of the data for the first research question. Research question one: What are the varied levels with which residents

demonstrate competence within the domain of communication skills as measured by Standardized Patient Parents (SPPs) using the Breaking bad news Assessment Schedule (BAS) instrument? The SPP scores were used as a proxy to determine the resident competence. Research question one was measured by examining the descriptive data collected from the SPP scores. Because the SPPs did not answer items on two assessments, those assessments were invalidated. Thus the combined SPP score represent an n of 156 rather than 158.

Each of the two SPPs scored each resident on all 18 items on the BAS using the 5-point Likert scale. “Mother” and “father” SPP assessments were each analyzed separately. Table 3 illustrates the descriptive statistics for the mother and father SPP ratings. Mothers had an average score of 60.95 (SD-12.84) and fathers had an average score of 58.97 (SD=8.87).

Table 3

SPP Descriptive Statistics

SPP	Minimum	Maximum	Mean	SD
Mother	29.00	68.00	60.95	12.84
Father	43.00	88.00	58.97	8.87

Note. N = 79

Table 4

Tertile Distribution of Resident Communication Skills Competence

Tertile	N	%
Bottom	53	33.5
Middle	50	31.6
Top	53	33.5
Total	156	98.7

Note. N = 156, 2 missing

The data analyzed showed the amount of variation with which the mother and father SPP rated the residents. For example, the lowest score the mother awarded was 21 points lower than the lowest awarded by the father. Additionally, the highest score the mother awarded was 68; 20 points lower than the father's top score. However, the means were very close at 60.95 (SD 12.84) for the mother and 58.97 (SD 8.87) for the father. Additionally, the mother and father scores correlated significantly (.79, $p \leq .05$). Therefore, further analysis will be discussed in terms of combined SPP scores (n=156). As can be seen in Table 4, when combined SPP scores were analyzed, fifty-three SPP scores fell into the bottom tertile, fifty fell into the middle tertile, and fifty-three fell into the top tertile.

Research Question Two: Resident Self-Assessment Ability

The following section presents the findings and analysis of the data for the research question two. Research question two: How do residents' self-assessments correspond with Standardized Patient Parent assessment of their communication skills performance? Resident self-assessment ability has been determined by comparing resident self-assessed scores with

those of the SPPs. Residents scored themselves on each of the 18 BAS items on the 5-point Likert scale. Table 5 shows the descriptive statistics for how the residents assessed themselves).

Table 5

Resident Self-Assessment Descriptive Statistics

	Minimum	Maximum	Mean	SD
Residents Self-Assessment Score	37.00	75.00	54.53	8.81

Note. N = 79

Total scores for the full scale ranged from 37 to 75 with a mean score of 54.53 (SD=8.81). Some residents rated themselves as low as 37 on the overall assessment. Other residents rated themselves as high as 75. As mentioned earlier, the lowest possible score for the full assessment would be 18 and the highest would be 90.

A paired samples t-test was done to determine if a correlation existed between the resident and the SPP scores. This analysis indicated no significant correlations ($r=.13$, $p=.11$) and that, in general, the residents did not correlate very well with the SPPs. However, this analysis was not very informative to this study. This analysis looked at the residents in aggregate and did not distinguish between residents who overestimated their performance versus those who underestimated their performance. Rather, it grouped all correlations together providing misleading information as to resident ability to accurately self-assess.

However, when looked at using tertiles, a clearer picture is presented. The researcher hypothesized that residents who demonstrated low competence within the domain of communication skills would overestimate their performance and that residents who demonstrated high competence in communication skills would underestimate their performance.

As shown in Table 6, twenty-seven residents self-assessed their performance within the bottom tertile, twenty-six residents self-assessed within the middle tertile, and twenty-six residents self-assessed within the top tertile. Figure 1 shows the comparison of how the residents in each tertile of communication skills competence self-assessed their performance. This figure illustrates that the majority of residents self-assessed their performance as approximately average. For residents at the middle level of competence, this was an accurate assessment. However, this scoring approach has residents in the bottom tertile of competence overestimating their performance and those within the top level of competence underestimating their performance.

Table 6

Tertile Distribution of Resident Self-Assessment Scores

Tertile	N	%
Bottom	27	34.2
Middle	26	32.9
Top	26	32.9
Total	79	100.0

Note. N = 79

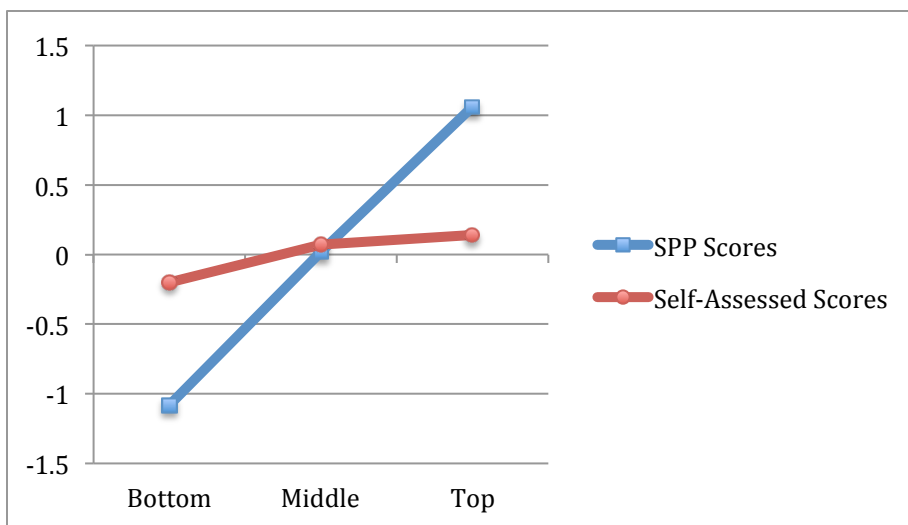


Figure 1

Comparison between SPP (competence) and Self-Assessed Scores

Research Question Three: Relationship between Resident Self-Assessment Ability and their Communication Skills Domain Competence

This section presents the findings and analysis of the data for the third research question. Research question three: How does resident communication skill competence relate to their ability to self-assess within the domain? This question was analyzed by examining the tertiles of resident communication skills competence in comparison with the tertiles of their self-assessment ability.

For question three's analysis, the researcher's intent was to compare tertiles of competence against tertiles of self-assessment ability. However, upon review, it was discovered that over 60% of the cells contained 0 values. Therefore, it was determined that

the analysis should be done by way of a Chi Square looking at the degree to which residents in each of the competence tertiles may have inaccurately or accurately assessed their performance. The researcher hypothesized that residents who demonstrated a low ability to self-assess within the domain of communication skills would overestimate their performance and that residents who demonstrated a high ability to self-assess their communication skills would underestimate their performance.

As shown in Table 7, residents who scored in the bottom competence tertile, but rated themselves within the top tertile of performance overestimated themselves by two degrees. Bottom competence tertile scoring residents who rated themselves within the middle tertile overestimated themselves by one degree. Bottom competence tertile scoring residents who rated their performance in the bottom tertile accurately assessed their performance. Middle competence tertile residents rating themselves in the top tertile overestimated themselves by one degree. Middle competence tertile self-ratings in the middle tertile accurately assessed their performance. Those who rated themselves within the bottom tertile underestimated themselves by one degree. Top performers who self-rated in the top tertile accurately assessed their performance. Top performers who self-rated in the middle tertile underestimated themselves by one degree. Top performers who rated themselves within the bottom tertile underestimated their performance by two degrees.

Table 7 illustrates a crosstabulation of the direction of the residents' self-assessment ability based on their level of communication skills competence. Twenty-one cases of residents in the bottom tertile were accurate in their self-assessment while thirty-two

overestimated, fifteen of which estimated to the highest degree (rated themselves within the top tertile). Within the middle tertile of resident competence, sixteen residents underestimated, thirteen were accurate, and twenty-one overestimated their performance. However, given their position within the middle tertile, it is only possible for this group to have under or overestimated their performance by one degree. Thirty-seven top tertile performers underestimated their performance and sixteen were accurate in their self-assessment. It was not possible for these top performers to have overestimated their performance.

Table 7

Impact of Domain Competence Tertile on Degree of Self-Assessment Ability

Degree of Accuracy	Bottom	Middle	Top	Total
Under by 2	0	0	16	16
Under by 1	0	16	21	37
Accurate	21	13	16	50
Over by 1	17	21	0	38
Over by 2	15	0	0	15

Note. $\chi^2 = 102.09$, $df = 8$, $p = .000$

This analysis shows a significant difference between the impact of domain competence and the degree to which the residents inaccurately self-assessed their performance. Those in the top and middle groups showed to be less likely to accurately assess themselves than do the bottom group. However, the top and middle groups'

inaccurate self-assessors also include those who underestimated their performance, which is not as concerning to this study as those who overestimated. Those who are in the bottom group and overestimated are of the greatest concern.

Research Question Four: Relationship between Resident Gender and their Self-Assessment Ability

This section presents the findings and analysis of the data for the fourth research question. Research question four: How does resident gender relate to their ability to self-assess their communication skills performance? Similar to with question three, degrees of self-assessment or accuracy were used rather than tertiles to determine the impact of resident gender on their self-assessment ability. A Chi Square, illustrated in Table 8, was used to determine if there was a relationship between resident gender and their self-assessment accuracy. The researcher hypothesized that there would be a significant difference of resident self-assessment ability based on resident gender. However, the analysis found no significant correlations between resident gender and their ability to effectively self-assess their communication skills performance.

Table 8

Gender Impact on Self-Assessment Accuracy

Degree of Accuracy	Female	Male	Total
Under by 2	14	2	16
<i>Expected</i>	<i>13.5</i>	<i>2.5</i>	
Under by 1	33	4	37
<i>Expected</i>	<i>31.3</i>	<i>5.7</i>	
Accurate	41	9	50
<i>Expected</i>	<i>42.3</i>	<i>7.7</i>	
Over by 1	33	5	38
<i>Expected</i>	<i>32.2</i>	<i>5.8</i>	
Over by 2	11	4	15
Total	132	24	156

Note. $\chi^2 = 2.571$, $df = 4$, $p = .632$

**Research Question Five: Relationship between Resident Level of Training
Experience and their Self-Assessment Ability**

This section presents the findings and analysis of the data for the fifth research question. Research question five: How does the number of years of residency training experience relate to residents' ability to self-assess their communication skills? Again similar to with questions three and four, degrees of self-assessment or accuracy were used rather than tertiles to determine the impact of resident level of residency training on their self-assessment ability. An additional Chi Square was done to determine if a relationship existed between resident level of training experience and their ability to self-assess their communication skills performance. The researcher hypothesized that a significant difference

would be found in resident ability to self-assess based on their level of residency training experience. However, the analysis found no significant relationship between level of training and resident self-assessment ability as shown in Table 9.

Table 9

Level of Training Impact on Self-Assessment Accuracy

Degree of Accuracy	Incoming Interns	Rising 2nd Years	Rising 3 rd Years	Impending Graduates	Total
Under by 2	1	2	1	3	7
<i>Expected</i>	<i>1.8</i>	<i>1.7</i>	<i>2.2</i>	<i>1.3</i>	<i>7.0</i>
Under by 1	3	3	7	4	17
<i>Expected</i>	<i>4.4</i>	<i>4.1</i>	<i>5.2</i>	<i>3.3</i>	<i>17.0</i>
Accurate	7	6	9	3	25
<i>Expected</i>	<i>6.4</i>	<i>6.1</i>	<i>7.7</i>	<i>4.8</i>	<i>25.0</i>
Over by 1	6	5	5	4	20
<i>Expected</i>	<i>5.1</i>	<i>4.9</i>	<i>6.2</i>	<i>3.8</i>	<i>20.0</i>
Over by 2	3	3	2	1	9
<i>Expected</i>	<i>2.3</i>	<i>2.2</i>	<i>2.8</i>	<i>1.7</i>	<i>9.0</i>

Note. $\chi^2 = 6.91, df = 12, p = .86$

This analysis showed that, despite their additional training experience, residents at higher levels of residency training were no different in their ability to self-assess than were residents at lower levels.

**Research Question Six: Variance in Resident Ability to Self-Assess Explained by
Demographic Variables, Year in Training, and Level of Communication Skill
Domain Competence**

This section presents the findings and analysis of the data for the sixth and final research question. Research question six: Can a significant amount of variance in resident ability to self-assess their communication skills be explained by gender, year in training, ethnicity, and theoretical element of communication skills competence (setting the scene, breaking the news, eliciting concerns, information giving, and general information)?

Findings for the sixth research question were determined by following the steps outlined by Mertler and Vannatta (2005) on conducting a regression analysis. As previous analysis showed, gender and level were shown to not impact resident self-assessment ability.

However, the researcher re-examined these factors along with ethnicity and the separate theoretical elements of breaking bad news as grouped by the BAS authors. Although the reliability was shown to be low for some of these theoretical elements, the researcher chose to examine the sections given that some of the sections showed to have higher reliability either by self-assessed or SPP scores. This was done to determine if they had any impact on the amount of self-assessment ability variance.

Table 10 provides the results to the stepwise regression analysis used to examine the amount of variance that each theoretical instrument section, level of training, gender, and ethnicity explained in the residents' ability to self-assess. As shown, neither ethnicity nor gender entered into the model.

Table 10

Stepwise Multiple Regression Analysis to Explore if Level, Gender, Ethnicity, or Theoretical Breaking Bad News Sections Impacted Residents' Self-Assessment Ability

Source of variation	SS	df	MS	F	p	
Regression	7633.04	1	7633.042	84.495	.000	
Residual	13911.88	154	90.337			
Total	21544.92	155				
	Overall R2	B	Beta	t	p	Collinearity Tolerance
Constant (-4.33)						
Breaking the News	35.4%	-6.99	-.595	-9.192	.000	
Excluded Variables						
Setting the Scene				1.15	.253	.545
Eliciting Concerns				1.085	.280	.968
Information Giving				-1.375	.171	.719
General Information				-1.293	.198	.502
Level of Training				.794	.429	.969
Gender				.310	.757	.993
Ethnicity				.305	.761	.980

The stepwise regression ($F = 84.495$, $df = 154$, $p = .000$) revealed that only one variable, the theoretical element of *Breaking the News* ($t = -9.192$, $p = .000$) entered the model. Overall, this element explained 35.4% of the variance in resident self-assessment

ability. Based upon the analysis of the excluded variables, no other variables showed significance.

Summary

This chapter has presented the findings of the analyzed data for this study. Descriptive statistics were provided for the resident participants to provide a framework of the population of interest. Descriptive statistics for SPP scores of resident communication skill performance were provided as the external assessment score and the gold standard upon which resident competence was determined. Resident self-assessed scores were compared with SPP scores to determine the level of accuracy with which residents self-assess their communication skills. Tertiles of resident self-assessment accuracy were compared with tertiles of their communication skills competence to determine if a relationship exists. Analyses were also done to determine the relationship between resident ability to self-assess and their gender, ethnicity, and level of training.

Findings from research question one found that the mother and father SPPs scored resident communication skill performance similarly to one another overall. Their combined scores demonstrated that the residents performed at considerable levels of variability; some scoring very low and others scoring near the top of the scale.

The findings from research question two indicated that the residents who performed poorly (as determined by SPP scores within the bottom tertile) overestimated their performance, middle tertile performers estimated fairly accurately, and high performers underestimated their performance.

Research question three clarified the findings from research question two. These findings showed the degree to which each tertile inaccurately self-assessed their performance. While each tertile inaccurately self-assessed somewhat and some in the top tertile underestimated their performance to a high degree, the residents in the bottom tertile who overestimated their performance were considerable and were of the greatest concern.

Research questions four and five found no correlations between either resident gender or their level of training on their ability to effectively self-assess their performance. The findings from research question six showed that only one theoretical element of *Breaking Bad News* explained a significant variance in the residents' ability to self-assess their performance. This element involved breaking the news itself. This analysis did not show gender, ethnicity, level of training, or any of the other theoretical elements to explain a significant variance.

CHAPTER FIVE: SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

A moment's insight is sometimes worth a life's experience.

[Oliver Wendell Holmes, Jr.](#)

The main purpose of this study was to provide additional evidence to determine if the theoretical framework proposed by Kruger and Dunning (1999) is generalizable to the field of graduate medical education and to examine other elements that may impact residents' ability to effectively self-assess. Does residents' ability to self-assess correlate with their competence within the assessed domain? This chapter presents an overall summary of the study including a review of the findings, conclusions, recommendations, and suggestions for future research, and limitations of the study. The study had six main research questions.

The first research question sought to explore the varied levels with which residents demonstrate competence within the domain of communication skills as measured by Standardized Patient Parents (SPPs) using the Breaking Bad News Assessment Schedule (BAS) instrument. The second research question examined how residents' self-assessments correspond with Standardized Patient Parent assessment of their communication skills performance. The third research question asked how resident communication skill competence relates to their ability to self-assess within the domain. The fourth question explored how resident gender relates to their ability to self-assess their communication skills performance. The fifth question explored how the number of years of residency training experience relates to residents' ability to self-assess their communication skills. The sixth

research question sought to explore the amount of significant variance in resident ability to self-assess their communication skills that can be explained gender, year in training, ethnicity, and theoretical element of communication skills competence.

Chapter One

Chapter one sought to provide a foundational discussion on the importance of self-assessment for professionals. An overview of the theory of self-regulated learning and metacognition was discussed in relation to the development of professionals' ability to perform a self-assessment of their knowledge and skills. This chapter provided the framework that emphasizes the integral nature of effective self-assessment within professional practice. An additional emphasis was put on professionals, particularly physicians, developing the ability to effectively self-assess during their professional preparatory training.

Two different schools of thought regarding the development of physician self-assessment ability during their preparatory professional (residency) training were also discussed. Reasoning for the decision to apply Kruger and Dunning's theoretical framework to residents within the domain of communication skills (CS) was explained. Chapter one also emphasized the importance of residents developing communication skills and the ability to self-assess within the domain of communication skills.

Chapter Two

Six main sections were presented for discussion within chapter two. Each of the sections contributed to a foundational description of the critical nature of effective self-

assessment for professionals and, more specifically, physicians. The bulk of the discussion focused on the importance of developing self-assessment ability within preparatory training and what is currently known about how this is done within residency.

The first section provided an overview of research on self-regulated learning for professionals. This section focused on the commonalities among studies of professionals and their preparatory education and training. Furthermore, this section illuminated self-regulated learning as an integral component of the development of professionals throughout their careers, post-formal preparatory education and training.

A focus on metacognition, a critical element of self-regulated learning (Gourgey, 2001; Zimmerman, 1990), and a discussion of the two elements of metacognition (Nelson & Narens, 1990) were included within the second section. One of the two elements of metacognition, metacognitive monitoring, which is often discussed in tandem with self-assessment, led into the third main area of discussion which presented an overview of self-assessment for professionals.

Subsequent to the overview of self-assessment for professionals, a theoretical framework of self-assessment (Kruger & Dunning, 1999) was presented. Kruger and Dunning's framework proposed that the ability to effectively self-assess might be directly related to the self-assessors' competence within the knowledge or skill domain being assessed. This framework was applied to medical professionals in training (physician residents) for examination within this study.

The next area of discussion focused on the context of medical education. An overview was provided on each phase of education and training necessary for physician development in the United States. Self-assessment for physicians was then discussed in detail.

Furthermore, three subsections within this section illuminated studies that have been conducted to examine three areas of residents' ability to effectively self-assess. The first subsection discussed studies that have examined residents' ability to self-assess as a static skill. These studies examined residents' ability to self-assess at one point in time to determine if they either did or did not have the ability to self-assess effectively. As opposed to other studies throughout the literature (Butterworth, 2010; Hodges et al., 2001; Jones et al., 2008; Mandel et al., 2005; Parker et al., 2004; Regehr, Hodges & McNaughton, 1998; Ward et al., 2003), those examining self-assessment as a static ability did not consider the potential for a change in the residents' ability to self-assess.

The two remaining subsections reviewed the two main schools of thought regarding the development of resident self-assessment ability. The first school of thought proposed that explicit instructional interventions specific to self-assessment are necessary to develop resident self-assessment ability (Colthart et al., 2008; Dunning et al., 2004; Fuhrmann & Weissburg, 1978; Gordon, 1991; Westburg & Jason, 1994). Studies discussed within this subsection included instructional interventions primarily focus on explicit feedback including video and benchmarking (Hodges et al., 2001; Martin et al., 1998; Ward et al., 2003) or

repeated self-assessment practice (Butterworth, 2010), each within the context of one or two knowledge or skill domains.

The second school of thought involved the application of Kruger and Dunning's (1999) theoretical framework to residents. The studies reviewed within this subsection examined the relationship between residents' ability to self-assess and their competence within the assessed domain. Four studies have examined this relationship (Jones et al., 2008; Mandel et al., 2005; Parker et al., 2004; Hodges et al., 2001). Only one of these four studies was done within the domain of communication skills (Hodges et al., 2001).

The current study directly examined the relationship between resident self-assessment ability and their communication skills competence. Self-assessment ability has been studied by directly comparing the correlation between self-assessment scores with the scores of valid external raters. Therefore, it is appropriate for assessment instruments typically used for external examination of resident communication skills to be utilized for a comparison between self- and external assessment (Gordon, 1991). A final section in chapter two provided a discussion of the BAS (Miller et al., 1999), which was the instrument of measurement used for the current study.

Chapter Three

The purpose of chapter three was to present the research design and methods used for the current study. The chapter began with a discussion of the appropriateness of the explanatory, non-experimental research design and the six guiding research questions used for the study. Next, the chapter included a presentation of the variables used in the study.

Other sections of the chapter included descriptions of the sample site and sample and instrumentation. Given that this study used secondary for analysis, a brief description of the procedure for how the data was collected was provided. Preliminary data analysis and statistical data analysis was also discussed. The preliminary data analysis section covered assumptions, confirmatory factor analysis, and reliability for the instrument used in the study. The chapter concluded with an overall summary of the research design and methodology.

Chapter Four

Chapter four provided a discussion of the analysis of data and findings from collected data concerning medical residents' ability to self-assess their communication skills. The chapter was divided into nine main sections; one section discussing the instrument, one to provide an overview of the resident demographics, one section for each of the 6 research questions, and a final section for an overall data analysis summary.

Chapter Five

This chapter presents an overall summary of the study including a review of the findings, conclusions, recommendations, and suggestions for future research, and limitations of the study. The study had six main research questions.

Discussion of Findings and Conclusions

As mentioned previously, the main purpose of this study was to provide additional evidence to determine if the theoretical framework proposed by Kruger and Dunning (1999) is generalizable to the field of graduate medical education and to examine other elements that may impact residents' ability to effectively self-assess.

This study examined residents' ability to self-assess their communication skills in relationship to their competence within the communication skills domain. Data was collected from a purposeful sample of Duke Pediatric residents at four levels of training: incoming first years, rising second and third years, and impending graduates.

Findings and Conclusions by Research Question.

Descriptive statistics were used to analyze resident communication skills competence based on Standardized Patient Parent (SPP) scores for the first research question.

Research Question One. The first research question sought to explore the varied levels with which residents demonstrate competence within the domain of communication skills as measured by Standardized Patient Parents (SPPs) using the Breaking bad news Assessment Schedule (BAS) instrument.

Conclusion One. Residents demonstrated a high level of variability in their performance breaking bad news to the SPPs. The SPPs scored some residents very low (29.00) on the competence spectrum (lowest possible score = 18, highest possible score = 90) and other residents near the top of the competence spectrum (88.00) indicating that the residents ranged from very low levels of communication skills competence to very high levels.

As exemplified by the data analysis for question one, many residents are poor communicators. Research has shown that effective communication skills are of utmost importance to physicians in practice (ACGME, 2011; LCME, 2010; ABMS, 2009; Eid, Pety, Hutchings, & Thompson, 2009; Rider & Keefer, 2006, Roberts & Bucksey, 2007; Simpson

et al., 1991; Tulskey, 2005; Watling & Brown, 2007). Poor communication between physicians and patients are said to lead to a variety of poor clinical consequences (Bull et al., 2002; Ciechanowski, Franks, Fiscella & Shields, 2005; Ciechanowski, Katon, Russo & Walker, 2001; Stewart, 1995) and have been attributed to physician burnout (Krasner et al., 2009). However, this study, as well as others (Back et al., 2007; Rider & Keefer, 2006; Rosas-Blum, Shirsat, Leiner, 2007; Simpson et al., 1991; Watling & Brown, 2007), has demonstrated that effective communication skills remain a challenge throughout the field of medicine.

General Recommendation One. Residency Training Program Directors and other residency educators should increase the emphasis on effective communication skills throughout residency training. Because of the current restrictions on the number of hours during which residents can participate in clinical and/or academic activities related to the program (ACGME, 2011), it is encouraged that additional instruction on communication skills be streamlined into other aspects of the current curriculum.

Recommendation for Future Research One. Further research should be conducted to determine the most effective and efficient means of incorporating communication skills training into residency training programs curricula.

Research Question Two. The second research question examined how residents' self-assessments correspond with Standardized Patient Parent assessment of their communication skills performance.

Physicians are not always effective at recognizing how well they interact and communicate with patients and their families. An inability for physicians to recognize deficits in their communication skills with patients and families precludes their tendency to seek to improve upon these skills (Gourgey, 2001; Westberg & Jason, 1994).

In the discussion of their findings, many studies have postulated that physicians should monitor their communication skills effectiveness as a part of their self-regulated learning process throughout their careers (Brinkman et al., 2006; Gruppen et al., 1997; Miller, Hope & Talbot, 1999; Howells, Davies, Silverman, Archer & Mellon, 2010; Schirmer et al., 2005). The effective practice of self-assessing their communication skills can help physicians to seek methods for alleviating their communication skill weaknesses and approach areas of communication skill strengths with confidence. By doing this, physicians can enhance their overall communication skills. Enhanced communication skills presents the opportunity for improved patient satisfaction, patient compliance with medications, and clinical outcomes (Bull et al., 2002; Ciechanowski et al., 2001; Stewart, 1995). Additionally, patients will likely better understand their diagnoses, the severity of their illnesses, and potential risks. Furthermore, physicians may be better able to appropriately assess and manage patient concerns and thus make more appropriate diagnoses (Franks et al., 2005).

Conclusion Two. The findings from research question two support the generalization of Kruger and Dunning's (1999) theoretical framework to resident ability to self-assess their communication skills. These findings show that, in general, the residents who performed poorly (as determined by SPP scores within the bottom tertile) overestimated their

performance, middle tertile performers estimated fairly accurately, and high performers underestimated their performance.

General Recommendation Two. Kruger and Dunning's (1999) framework is supported by the findings of research question two. This indicates, therefore, that residents improve in their ability to self-assess their communication skills performance with increased competence within the domain of communication skills. Therefore, it is recommended that a focus of residency training programs should be on enhanced communication skills (and other domain) competence over an explicit focus on the development of self-assessment skills as a separate domain.

Recommendation for Future Research Two. As with the first recommendation for future research, the analysis of research question two emphasizes the need for further research on effective and efficient residency training within the domain of communication skills.

Recommendation for Future Research Three. Furthermore, it is recommended that additional research be conducted within other domains of resident competence. This will help to determine if Kruger and Dunning's (1999) framework is generalizable for residents beyond the domain of communication skills.

Research Question Three. The third research question asked how resident communication skill competence relates to their ability to self-assess within the domain. Specifically, this research question describes the main goal of this study; to explore the relationship between the residents' communication skills domain competence with their

ability to self-assess within the communication skills domain. To explore this relationship, Kruger and Dunning (1999) have suggested that when one develops competence within a domain, one also increases his/her self-assessment ability within it. Evidence has suggested that deficits in metacognitive skill, particularly the monitoring component, explain the lack of self-insight for those deemed incompetent in a given domain.

This theoretical framework presumes that “the skills needed to perform the cognitive, affective, or behavioral task (producing the response) are the same exact ones necessary for metacognitive tasks (judging the response)” (Dunning, 2005, p. 15). Therefore, incompetence within the given domain, in essence, robs individuals of the ability to recognize competence (or the lack thereof) when they see it. This relationship would suggest that when a physician is competent in a given domain, the physician has the ability to recognize their strengths and weaknesses within it. Under this theoretical framework, increasing competence in a given domain is the most direct means of increasing physicians’ ability to self-assess their performance within it.

The underlying assumptions and theory of this framework suggest its value for considering the training and professional practice of physicians. Findings that support the use of this framework are those that have directly linked medical resident self-assessment ability to competence within the assessed domain throughout research studies and the literature (Hodges et al., 2001; Parker et al., 2004). These studies have reported cases of the least competent residents within a given domain being the least able to effectively assess their own knowledge and skills within that domain. These residents rated themselves as

having performed above (often well above) the mean. However, their objective external assessments showed them to actually perform at a level within the least competent group (bottom quartile or tertile). Such dramatic overestimation of competence can be particularly dangerous for physicians (Graber, 2005; Mamede et al., 2007), heightening the need for researchers and theorists to examine the relationships between competence and self-assessment.

Kruger and Dunning's position is that competence in a given domain is a natural predecessor to individuals being able to recognize their (or others') competence within it. The concept that those who are the least skilled in knowledge and skill domains are also the least able to recognize their deficits is of particular interest to the study of professionals and professional development. The goal of this study was to determine if increasing resident competence within the assessed domain is an effective means of developing resident self-assessment ability during their residency training.

Conclusion Three. As described in conclusion two, the findings from research question three further support the generalization of Kruger and Dunning's (1999) framework. These findings show that, as resident competence within the domain of communication skills increases, their ability to self-assess their competence improves. However, this study's findings also showed that highly competent residents tended to underestimate their performance. As described by Kruger and Dunning (1999), these findings were likely the result of a false-consensus effect (Ross, Greene & House, 1977) rather than true underestimation of their own performance. The false-consensus effect implies that, without

the benefit of having observed others' performances, high performers often incorrectly assume that their peers are performing at higher levels than they actually are, thus leading them to underestimate their own performance. Furthermore, as discussed, it is the least competent who overestimate their performance that are the most concerning to this study and throughout practice.

General Recommendation Three. Providing additional evidence of the relationship proposed within Kruger and Dunning's framework should refocus practitioners on the development of resident deficits within the communication skills domain. This approach would take the place of a focus on developing explicit self-assessment interventions to improve self-assessment ability.

Recommendation for Future Research Four. As discussed regarding research questions one and two, further research should be conducted on the most effective and efficient means of streamlining communication skills training into the residency-training curriculum.

Research Question Four. The fourth question explored how resident gender relates to their ability to self-assess their communication skills performance. Some studies on self-assessment have looked at the impact of gender on residents' ability to effectively self-assess. Some have reported a tendency for men to overestimate and for women to underestimate themselves but the results from the medical education literature (Colthart et al. 2008) has been inconsistent. Studies by Minter et al. (2005), Zonia and Stommel (2000), and Parker et

al. (2004) found no significant difference in residents' ability to self-assess their knowledge or performance based on their gender.

Conclusion Four. This study found no distinction between male and female residents' ability to effectively self-assess their performance.

General Recommendation Four. Residency training efforts to improve communication skills and self-assessment of those skills should be made without consideration of gender distinctions.

Recommendation for Future Research Five. Future research to improve communication skills and self-assessment of those skills should be conducted without consideration of gender distinctions.

Research Question Five. The fifth research question explored how the number of years of residency training experience relates to residents' ability to self-assess their communication skills. Medical residency provides residents with the structure to meet their learning expectations with increasing levels of autonomy as determined appropriate by board-certified, experienced supervisors. Given this format and the assumption that residents have had greater opportunities to practice self-assessment throughout increased years of residency training, it might be assumed that increased residency training experience would lead to improved self-assessment ability.

Conclusion Five. This study found residency-training level to have no significant impact on residents' ability to effectively self-assess. Despite residents' increased experience

with communication skills and self-assessment in advanced years of training, their ability to effectively self-assess their communication skills does not improve.

General Recommendation Five. Recommendations for enhanced communication skills training to improve resident self-assessment abilities within the domain are encouraged for all levels of residency training. This training should be initiated within the first year of residency training and continue until graduation from residency.

Research Question Six. The sixth research question sought to explore the amount of significant variance in resident ability to self-assess their communication skills that can be explained by gender, year in training, ethnicity, and theoretical element of communication skills competence.

While some studies in the medical education literature (Wooliscroft et al., 1993; Fitzgerald et al., 2003) have looked at the impact of medical student ethnicity on their ability to self-assess, no sound studies were found that look at this for residents.

Conclusion Six. Findings for research question six indicated that the only variable examined in addition to resident communication skills competence that had an impact on residents' ability to self-assess their communication skill performance was one theoretical element to breaking bad news: *Breaking the News*.

General Recommendation Six. Items included within the *Breaking the News* element of the BAS should be emphasized in communication skills training. Items within this element address the residents' ability to determine SPP base knowledge of the situation

and expectations of the interaction, approaching with sensitivity, and giving the SPPs adequate time to respond.

Recommendation for Future Research Six. Researchers should further explore the link between the element of *Breaking the News* over the other elements to improved ability for residents to self-assess their communication skills.

The results of this study have supported the generalization of Kruger & Dunning's (1999) theoretical framework to residency training. These results have suggested that increased competence within a domain (breaking bad news) improves a resident's ability to self-assess their performance within it. General suggestions for practice include increasing an emphasis on effective communication skills (breaking bad news) throughout residency training with a particular emphasis on elements related to breaking the news itself. The results of this study have also suggested that medical education researchers should conduct research to determine the most effective means of streamlining this training into the existing residency curriculum. Additional research should also be conducted within other domains of residency training to determine if the framework is equally generalizable to other domains.

Limitations

The limitations will address the sample, data collection, assessed domain and instrumentation.

Sample. The findings of this study are bound by the population and sample, which included residents from one institution (Duke University) and one specialty (pediatrics). Additionally,

the sample was relatively small, precluding the meaningfulness of an instrument factor analysis.

Data Collection. Additionally, this study used a cross-sectional approach for data collection.

This approach examines Duke Pediatric residents who were in training in June 2011.

Assessed Domain. This study assessed residents within one domain, that of communication skills. Therefore, it is not known if the results are generalizable to other residency knowledge and skill domains.

Instrument Used. While the BAS was found to have overall good reliability, the five theoretical elements were inconsistent in their reliability. Additionally, given the relatively small sample size, a factor analysis was not conducted.

References

- Accreditation Council for Graduate Medical Education (ACGME). (2011). *ACGME Program Requirements for Graduate Medical Education*. Retrieved from http://www.acgme.org/acWebsite/downloads/RRC_progReq/320_pediatrics_07012007.pdf.
- American Board of Medical Specialties (ABMS). (2012). Maintenance of Certification (MOC) competencies and criteria. Retrieved from http://www.abms.org/Maintenance_of_Certification/MOC_competencies.aspx.
- American Association for Justice. (2012). Preventable medical errors: The sixth biggest killer in America. Retrieved from <http://www.justice.org/cps/rde/justice/hs.xsl/8677.htm>.
- Association of American Medical Colleges. (2004). Underrepresented in Medicine Definition. Retrieved from <https://www.aamc.org/initiatives/urm/>.
- Austin, Z. & Gregory, P. (2007). Evaluating the accuracy of pharmacy students' self-assessment skills. *American Journal of Pharmaceutical Education*, 71(5), 1-8.
- Back A., Arnold R., Baile W., Fryer-Edwards K., Alexander S., Barley G.,...& Tulsky J. (2007). Efficacy of communication skills training for giving bad news and discussing transitions to palliative care. *Archives of Internal Medicine*, 167(5), 453-460. doi: 10.1001/archinte.167.5.453
- Barnsley, L., Lyon, P.M., Ralston, S.J., Hibbert, E.J., Cunningham, I., Gordon, F.C., Field, M.J. (2004). Clinical skills in junior medical officers: a comparison of self-reported

- confidence and observed competence. *Medical Education*; 38, 358-367. doi: 10.1046/j.1365-2923.2004.01773.
- Biernat, K., Simpson, D., Duthie, E. Jr., Bragg, D., London, R. (2003). Primary care residents self assessment skills in dementia. *Advances in Health Sciences Education*, 8, 105-110. doi: 10.1023/A:1024961618669
- Boud, D. (1995). *Enhancing Learning through Self-assessment*. London and Philadelphia: Kogan Page.
- Boud, D.J. & Falchikov, N. (1989). Quantitative studies of student self assessment in higher education: A critical analysis of findings. *Higher Education* 18, 529–549. doi: 10.1007/BF00138746
- Boud, D.J. and Tyree, A.L. (1979) Self and peer assessment in professional education: a preliminary study in law, *Journal of the Society of Public Teachers of Law*, 15(1), 65-74.
- Breidert, J. T., & Fite, J.E. (2009). Self-assessment: Review and implications for training. (ARI Research Report 1900). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Brewster, L.P., Risucci, D.A., Joehl, R.J., Littooy, F.N., Temeck, B.K., Blair, P.G. & Sachdeva, A.K. (2008). Comparison of resident self-assessments with trained faculty and standardized patient assessments of clinical and technical skills in a structured educational module. *The American Journal of Surgery*, 195(1), 1-4. doi: 10.1016/j.amjsurg.2007.08.048

- Breidert, J. T., & Fite, J. E. (2009). *Self-assessment: Review and implications for training*. (ARI Research Report 1900). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Brennan, P., Silman, A., Black, C., Bernstien, R., Coppock, J., Maddison, P., Sheeran, T., Stevens, C. & Wollheim, F. (1992). Reliability of skin involvement measures in scleroderma. *British Journal of Rheumatology*, 31, 457–460.
- Brinkman, W.B., Geraghty, S.R., Lanphear, B.P., Khoury, J.C., Gonzalez del Rey, J.A. DeWitt, T.G. & Britto, M.T. (2007). Effect of multisource feedback on resident communication skills and professionalism. *Archives of Pediatric and Adolescent Medicine*, 161(1), 44-49. doi:10.1001/archpedi.161.1.44
- Bull SA, Hu XH, Hunkeler EM, et al. (2002). Discontinuation of use and switching of antidepressants: influence of patient-physician communication. *Journal of the American Medical Association*, 288, 1403-1409. doi:10.1001/jama.288.11.1403
- Butterworth, K. (2010). Developing self-assessment skills amongst doctors in nepal. *Medical Teacher*, 32(2), e85-e95. doi: 10.3109/01421590903206372
- Caverzagie, K.J., Shea, J.A. & Kogan, J.R. (2008) Resident identification of learning objectives after performing self-assessment based upon the ACGME core competencies. *Journal of General Internal Medicine* 23(7), 1024-7. doi: 10.1007/s11606-008-0571-7
- Choudhury, A. (2010). Correlation and Regression. Retrieved March 11, 2012 from Experiment Resources: <http://www.experiment-resources.com/cronbachs-alpha.html>.

- Ciechanowski, P.S., Katon, W.J., Russo, J.E. & Walker, E.A. (2001) The patient-provider relationship: Attachment theory and adherence to treatment in diabetes. *American Journal of Psychiatry*, 158, 29–35. doi: 10.1176/appi.ajp.158.1.29
- Cochran, S.B. & Spears, M.C. (1980) Student self-assessment and instructors' ratings: a comparison, *Journal of the American Dietetic Association*, 76, 253-257.
- Colthart, I., Bagnall, G., Evans, A., Allbutt, H., Haig, A., Illing, J. et al. (2008). The effectiveness of self-assessment on the identification of learner needs, learner activity, and impact on clinical practice: BEME Guide no. 10. *Medical Teacher*, 30(2), 124–145. doi:10.1080/01421590701881699
- Comrey, A. L., & Lee, H. B. (1992). *A first Course in Factor Analysis*. Hillsdale, NJ: Erlbaum.
- Curran, V.R., Butler R., Duke P., Eaton W.H., Moffatt, S.M., Sherman, G.P. & Pottle, M. (2007). Evaluation of the usefulness of simulated clinical examination in family-medicine residency program. *Medical Teacher*, 29(4), 406-407. doi: 10.1080/01421590701477464
- Davis, D.A., Mazmanian P.E., Fordis M., Van Harrison R., Thorpe K.E., Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *Journal of the American Medical Association*, 296(9), 1094-1102. doi:10.1001/jama.296.9.1094

- Duffy, F.D., & Holmboe, E.S. (2006). Self-assessment in lifelong learning and improving performance in practice: Physician know thyself. *Journal of the American Medical Association*, 296, 1137–9. doi: 10.1001/jama.296.9.1137
- Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Psychology Press.
- Dunning, D., Heath, C. & Suls, J.M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69-106. doi: 10.1111/j.1529-1006.2004.00018
- Edwards, S. (2007). Critical thinking: A two-phase framework. *Nurse Education in Practice*, 7(5), 303-314.
- Eid, A, Petty, M., Hutchins, L. & Thompson, R. (2009). Breaking bad news: Standardized patient intervention communication skills for hematology-oncology fellows and advanced practice nurses. *Journal of Cancer Education*, 24(2), 154-159.
- English, K., Naeve-Velguth, S., Rall, E., Uyehara-Isono, J. & Pitman, A. (2007). Development of an instrument to evaluate audiologic counseling skills. *Journal of the American Academy of Audiology*, 18, 677-789.
- Epstein, R.M., Dannefer, E.F., Nofziger, A.C., Hansen, J.T., Schultz, S.H., Jospe, N., Connard, L.W., Meldrum, S.C., & Henson, L.C. (2004). Comprehensive Assessment of Professional Competence: The Rochester Experiment. *Teaching and Learning in Medicine*, 16(2), 186-196. doi: 10.1207/s15328015t1m1602_12

- Eva, K. W., Cunnington, J. P., Reiter, H. I., Keane, D. R., & Norman, G. R. (2004). How can I know what I don't know? Poor self assessment in a well-defined domain. *Advances in Health Sciences Education, Theory and Practice*, 9(3), 211-224. doi: 10.1023/B:AHSE.0000038209.65714.d4
- Eva K.W. & Regehr, G. (2005). Self-assessment in the health professions: a reformulation and research agenda. *Academic Medicine*, 80, S46–53.
- Eva, K. W. & Regehr, G. (2008). “I’ll never play professional football” and other fallacies of self-assessment. *Journal of Continuing Education in the Health Professions*, 28(1), 14–19. doi: 10.1002/chp.150
- Falchikov, N. & Boud, D. (1989). Student self-assessment in higher education: a meta-analysis. *Review of Educational Research*, 59, 395–430. doi: 10.2307/1170205
- Fitzgerald, J., White, C. & Gruppen, L. (2003). A longitudinal study of self-assessment accuracy. *Medical Education*, 37, 645–649. 10.1046/j.1365-2923.2003.01567.x
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist*, 34, 906–911. doi: 10.1037//0003-066X.34.10.906
- Franks, P., Fiscella, K., Shields, C.G., Meldrum, S.C., Duberstein, P., Jerant, A.F., Tancredi, D.J. & Epstein, R.M. (2005). Are patients’ ratings of their physicians related to health outcomes? *Annals of Family Medicine*, 3(3), 229-234. doi: 10.1370/afm.267

- Fox, R.A., Ingham Clark, C.L., Scotland, A.D. & Dacre, J.E. (2000). A study of pre-registration house officers' clinical skills. *Medical Education*, 34, 1007-1012. doi: 10.1111/j.1365-2923.2000.00729.
- Fuhrmann, B.S. & Weissburg, M.J. (1978). *Self-assessment. Evaluation of clinical competence in the health professions*. C. V. Mosby, St. Louis.
- Fuqua, D.R., Johnson, A.W., Newman, J.L., Anderson, M.W., and Gade, E.M. (1984) Variability across sources of performance ratings. *Journal of Counselling Psychology*, 31(2), 249-252. doi: 10.1037/0022-0167.33.3.353
- Goodman, D.C. (2005). The pediatrician workforce: Current status and future prospects and the committee on pediatric workforce. *Pediatrics*, 16(1), e156-e173. doi:10.1542/peds.2005-0874
- Gordon, M.J. (1991). A review of the validity and accuracy of self-assessments in health professions training. *Academic Medicine*, 66, 762-769.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Gourgey, A. (2001). Metacognition in basic skills instruction. In H. Hartman (Ed.), *Metacognition in learning and instruction: Theory, research and practice* (pp. 17-32). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Graber, M.L., Franklin, N. & Gordon, R. (2005). Diagnostic error in internal medicine. *Archives of Internal Medicine*, 165(13), 1493-9. doi:10.1001/archinte.165.13.1493

- Gruppen, L. D., Garcia, J., Grum, C. M., Fitzberald, J. T., White, C. A., Dicken, L., Sisson, J. C., & Zweifler, A. (1997). Medical students' self-assessment accuracy in communication skills. *Academic Medicine*, 72(1), S57-S59.
- Hair, J. F., Black, W. C., Babin, B., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Englewood Cliffs: Prentice Hall.
- Hacker, D.J., Dunlosky, J. & Graesser, A.C. (Eds.) (1998). *Metacognition in Educational Theory and Practice*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hacker D.J., Bol L., Bahbahani K. (2008). *Explaining calibration accuracy in classroom contexts: The effects of incentives, reflection, and explanatory style*. *Metacognition & Learning*, 3, 101-121. doi:10.1007/s11409-008-9021-5
- Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, 56, 208–216. doi: 10.1037/h0022263
- Hatcher, L. (1994). *A Step-by-Step Approach to Using the SAS® System for Factor Analysis and Structural Equation Modeling*. Cary, NC: SAS Institute, Inc.
- 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. *American Journal of Medicine*, 106(2), 206-10.
- Hodges, B., Regehr, G., & Martin, D. (2001). Difficulties in recognizing one's own incompetence: Novice physicians who are unskilled and unaware of it. *Academic Medicine*, 76, 87-89.

- Hoppe, R.B., Farquhar, L.J., Stoffelmayr, H.R. (1990). Residents' attitudes towards and skills in counseling: using undetected standardized patients. *Journal of General Internal Medicine*, , 5, 415-420.
- Howells, R.J. , Davies, H.A., Silverman, J. D., Archer, J.C., & Mellon, A.F. (2010). Assessment of doctors' consultation skills in the paediatric setting: the paediatric consultation assessment tool. *Archives of Disease in Childhood*, 95, 323-329. doi: 10.1136/adc.2008.146191
- Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Thousand Oaks, CA: Sage Publications.
- Iramaneerat, C., Myford, C.M., Yudkowsky, R. & Lowenstein, T. (2009). Evaluating the effectiveness of rating instruments for a communication skills assessment of medical residents. *Advances in Health Sciences Education*, 14(4), 575-594. doi: 10.1007/s10459-008-9142-2
- Jarvis, P. (1983). *Professional Education*. London, England: Croom-Helm.
- Jones, R., Panda, M., & Desbiens, N. (2008). Internal medicine residents do not accurately assess their medical knowledge. *Advances in Health Sciences Education*, 13, 463-468. doi: 10.1007/s10459-007-9058-2
- Kerlinger, F. N. & Lee, H. B. (2000). *Foundations of Behavioral Research*. Fort Worth, TX: Harcourt College Publishers.
- Krasner MS, Epstein RM, Beckman H, et al. Association of an educational program in mindful communication with burnout, empathy, and attitudes among primary care

- physicians. *Journal of the American Medical Association*, 302(12), 1284-1293.
doi:10.1001/jama.2009.1384
- Kreuter, M., and Strecher, V. (1995). Changing inaccurate perceptions of health risk: Results from a randomized trial. *Health Psychology*, 14, 56–63. doi: 10.1037/0278-6133.14.1.56
- 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), 1121-1134. doi: 10.1037/0022-3514.77.6.1121
- Kumar, R. (1999). *Research methodology: A step-by-step guide for beginners*, New Delhi: Sage Publications.
- Lang, F., McCord, R., Harvill, L. & Anderson, D.S. (2004). Communication assessment using the common ground instrument: psychometric properties. *Family Medicine*, 36(3), 189-198.
- Lajoie, S.P. (2008) Metacognition, self-regulation, and self-regulated learning: A rose by any other name? *Educational Psychology Review*, (20), 469–475. doi: 10.1007/s10648-008-9088-1
- Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.

- Liaison Committee for Medical Education (LCME). (June, 2010). *Standards for accreditation of medical education programs leading to the M.D. degree*. Retrieved from <http://www.lcme.org/functions2010jun.pdf>.
- Lipsett, P.A., Harris, I. & Downing, S. (2011). Resident self-other assessor agreement. *Archives of Surgery, 146*(8), 901-906. doi: 10.1001/archsurg.2011.172
- Mabe, P.A. & West, S.G. (1982). Validity of self-evaluation of ability: a review and meta-analysis. *Journal of Applied Psychology, 67*(3), 280-296. doi: 10.1037/0021-9010.67.3.280
- MacDonald, J, Williams, R.G & Rogers, D.A. (2003). Self-assessment in simulation-based surgical skills training. *The American Journal of Surgery, 185*(4), 316-318. doi: 10.1016/S0002-9610(02)01420-4
- Mamede, S., Schmidt, H.G., Rikers, R. (2007). Diagnostic errors and reflective practice in medicine. *Journal of Evaluation of Clinical Practice, 13*(1), 138-145. doi: 10.1111/j.1365-2753.2006.00638.
- Mandel LS, Goff BA, Lentz GM. 2005. Self-assessment of resident surgical skills: Is it feasible? *American Journal of Obstetrics and Gynecology, 193*, 1817–1822. doi: 10.1016/j.ajog.2005.07.080
- Makoul G., Krupat, E. & Chang, C.H. (2007). Measuring patient views of physician communication skills: development and testing of the communication assessment tool. *Patient Education and Counseling, 67*(3), 333-342. doi: 10.1016/j.pec.2007.05.005

- Martin, D. Regehr, G., Hodges, B. & McNaughton, N. (1998). Using videotaped benchmarks to improve the self-assessment ability of family practice residents. *Academic Medicine*, 73(11), 1201-6. doi: 10.1016/S0002-9610(03)00069-2
- Martin, J. A., Regehr, G., Reznick, R., Macrae, H., Murnaghan, J., Hutchison, C. & Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *British Journal of Surgery*, 84(2), 273-278. doi: 10.1046/j.1365-2168.1997.02502
- Miller, S.J., Hope, T. & Talbot D.C. (1999). The development of a structured rating schedule (the BAS) to assess skills in breaking bad news. *British Journal of Cancer*, 80(5/6), 792–800. doi: 10.1038/sj.bjc.6690423
- Minter, R.M., Gruppen, L.D., Napolitano, K.S., Gauger, P.G. (2005). Gender differences in the self-assessment of surgical residents. *American Journal of Surgery*, 189, 647–650. doi: 10.1016/j.amjsurg.2004.11.035
- Moorthy, K., Munz, Y., Adams, S., Pandey, V., & Darzi, A. (2005). A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre. *Annals of Surgery*, 242, 631–639. doi: 10.1097/01.sla.0000186298.79308.a8
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (125–173). New York: Academic Press.

- O'Rourke, N., Hatcher, L., & Stepanski, E.J. (2005). A step-by-step approach to using SAS for univariate and multivariate statistics, Second Edition. Cary, NC: SAS Institute Inc.
- Pandey, V.A., Wolfe, J.H.N., Lindahl, A.K., Rauwerda, J.A., & Bergqvist, D. (2004). Validity of an exam assessment in surgical skill: EBSQ-VASC pilot study. *European Journal of Vascular & Endovascular Surgery*, 27(4), 341-348.
10.1177/153100350501700225
- Parker R.W., Alford C., Passmore C. (2004). Can family medicine residents predict their performance on the in-training examination? *Family Medicine*, 36, 705-709.
- Perfect, T. J., & Schwartz, B. L. (2002). *Applied metacognition*. Cambridge, England: Cambridge University Press.
- Petersdorf, R G. (1994). Medical curriculum, training, and the continuum of medical education. *Journal of the Royal Society of Medicine*, 87(22), 47-49.
- Pintrich, P. R., & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In A. Wigfield, & J.S. Eccles (Eds.), *Development of achievement motivation* (250-284). San Diego: Academic Press.
- Rider E., Keefer C. (2006). Communication skills competencies: definitions and a teaching toolbox. *Medical Education*, 40(7), 624-629. doi: 10.1111/j.1365-2929.2006.02500
- Roberts, L. & Bucksey, S. (2007). Communicating with patients: What happens in practice? *Physical Therapy*, 87(5), 586-594. doi: 10.2522/ptj.20060077

- Rosas-Blum, E., Shirsat, P. & Leiner, M. (2007). Communicating genetic information: a difficult challenge for future pediatricians. *BMC Medical Education*, 7(17). doi: 10.1186/1472-6920-7-17
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment Research & Evaluation*, 11(10), 1–13.
- Schein, E. H. (1972). In Carnegie Commission on Higher Education. (Ed.), *Professional education: Some new directions*, New York, McGraw-Hill [c1972].
- Schön, D.A. (1983) *The Reflective Practitioner: How Professionals Think in Action*, New York: Basic Books.
- Schildmann, J., Kupfer, S., Burchardi, N. & Vollmann, J. (2011). Teaching and evaluating breaking bad news: A pre–post evaluation study of a teaching intervention for medical students and a comparative analysis of different measurement instruments and raters. *Patient Education and Counseling* 86, 210–219.
10.1016/j.pec.2011.04.022
- Schirmer, J.M., Mauksch, L., Lang, F., Marvel, M.K., Zoppi, K., Epstein, R.M., Brock, D. & Pryzbylski, M. (2005). Assessing communication competence: a review of current tools. *Family Medicine*, 37(3), 184-92.
- Schwartz, B. L., & Perfect, T. J. (2002). Introduction: toward an applied metacognition. In Perfect, T. J., & Schwartz, B. L. (2002). *Applied Metacognition*. Cambridge University Press. pp. 1-14.

- Simpson, M., Buckman, R., Stewart, M., Maguire, P., Lipkin, M., Novack, D. & Till, J. (1991). Doctor-patient communication: the Toronto consensus statement. *British Medical Journal*, 303(6814), 1385-1387.
- Sproull, N. L. (2002). Handbook of research methods: A guide for practitioners and students in the social sciences (2nd ed.). MD: The Scarecrow Press, Inc.
- Spychalski, A. C., Quinones, M. A., Gaugler, B. B., & Pohley, K. (1997). A survey of assessment center practices in organizations in the United States. *Personnel Psychology*, 50, 71-90.
- Stern, D.T., Linzer, M., O'Sullivan, P.S., Weld, L. (1995). Evaluating medical residents' literature-appraisal skills. *Academic Medicine*, 70, 152-154.
- Stewart MA. (1995). Effective physician-patient communication and health outcomes: a review. *Canadian Medical Association Journal*, 152, 1423-33.
- Sullivan, W. (2000). Medicine under threat: Professionalism and professional identity. *Canadian Medical Association Journal* 162, 673-75.
- Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment tasks: towards transparency for students and for tutors. *Assessment & Evaluation in Higher Education*, 26(6), 605-614. doi: 10.1080/02602930120093922
- Tulsky, J. (2005). Interventions to enhance communication among patients, providers, and families. *Journal of Palliative Medicine*, 8(1), s95-s101. doi:10.1089/jpm.2005.8.s-95

- Violato, C., Lockyer, J. (2006). Self and peer assessment of pediatricians, psychiatrists and medicine specialists: Implications for self-directed learning. *Advances in Health Science Education, 11*, 235-244. doi: 10.1007/s10459-005-5639-0
- Ward, M., MacRae, H., Schlachta, C., Mamazza, J., Poulin, E., Reznick, R. & Regehr, G. (2003). Resident self-assessment of operative performance. *The American Journal of Surgery, 185*(6), 521-524. doi: 10.1016/S0002-9610(03)00069-2
- Watling, C.J. & Brown, J.B. (2007). Communication skills for neurology residents: Structured teaching and reflective practice. *Neurology, 69*(22), e20-26. doi: 10.1212/01.wnl.0000280461.96059.44
- Ward, M., Gruppen, L. & Regehr, G. (2002). Research in self-assessment: Current state of the art. *Advances in Health Science Education, 7*, 63–80.
- Westberg, J. & Jason, H. (1994). Fostering learners' reflection and self-assessment. *Family Medicine, 26*(5), 278-282. doi:10.1186/1756-0500-2-152
- Wood, J., Collins, J., Burnside, E.S., Albanese, M.A., Propeck, P.A., Kelcz, F., Spilde, J.M., Schmaltz, L.M. (2004). Patient, faculty, and self-assessment of radiology resident performance: a 360-degree method of measuring professionalism and interpersonal/communication skills. *Academic Radiology, 11*, 931-939. doi:10.1016/j.acra.2004.04.016
- Woods, R., McCarthy, T., Barry, M.A. & Mahon, B. (2004). Diagnosing smallpox: would you know it if you saw it? *Biosecurity Bioterrorism, 2*(3), 157-163. doi:10.1089/bsp.2004.2.157

Wooliscroft, J.O., TenHaken, J., Smith, J. & Calhoun, J.G. (1993). Medical students' clinical self-assessments: comparisons with external measures of performance and the students' self-assessments of overall performance and effort. *Academic Medicine*, 68, 285–294.

Yammarino, F.J. & Atwater, L.E. (1997). Do managers see themselves as others see them? Implications of self-other rating agreement for human resources management. *Organizational Dynamics*, 25(4), 35-44. doi: 10.1016/S0090-2616(97)90035-8

F143

Yoo M.S., Son Y.S., Kim Y.S., Park J.H. (2009). Video-based self-assessment: Implementation and evaluation in an undergraduate nursing course. *Nurse Education Today* 29, 585–589. doi: 10.1016/j.nedt.2008.12.008

Yudkowsky, R., Downing, S.M. & Sandlow, L.J. (2006). Developing an institution-based assessment of resident communication and interpersonal skills. *Academic Medicine*, 81(12), 1115-1122. Doi: 10.1097/01.ACM.0000246752.00689.bf

Zimmerman, B.J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1), 13-17.

Zimmerman, B.J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B.J. Zimmerman & D.H. Schunk (Eds.), *Self-regulated learning and academic achievement: theoretical perspectives* (1-36). New York: Lawrence Erlbaum Associates.

Zonia SL, Stommel M. (2000). Interns self-evaluations compared with their faculty's evaluations. *Academic Medicine*, 75, 742.

APPENDICES

Appendix A

Revised BAS for Self-Assessment

Breaking bad news Assessment Schedule (BAS)

When marking please place a circle around the number which reflects the score you wish to assign yourself based on your performance on each stated item. Statements below each question reflect desirable behaviors that would result in higher scores for the applicable question.

A. Setting the scene: this section looks at whether you facilitated an initial rapport before breaking the bad news. This can be done by providing an environment which allows private and comfortable communication, by you introducing yourself, and by you showing an interest in the patient parents as individuals.

1. Did you use an appropriate greeting and introduction? **definitely 5 _ 4 _ 3 _ 2 _ 1 not at all**

You may have

- stood up to greet the patient's parents?
- established the patient's parents name?
- introduced yourself using your own name?
- given a brief description of your occupation?

2. Did you show interest in the patient's parents' current state of well-being and personal circumstances at the beginning of the interview? **definitely 5 _ 4 _ 3 _ 2 _ 1 not at all**

You may have

- used open questions?
- established recent events for the patient's parents?
- established their physical state?
- asked how they felt emotionally?
- inquired into their social circumstances?
- given them time to finish their statements?

B. Breaking the news: this section specifically focuses on whether you were sensitive to the patient's parents' perspective when you delivered the news.

Individual patients may vary in the amount of information they wish to receive during this interview, and in the rate at which they assimilate the news.

3. Before breaking the news did you check what they knew already?

definitely 5 _ 4 _ 3 _ 2 _ 1 not at all

Did you

- ask the patient's parents what they believed was the nature of the problem?
- inquire into what they thought the purpose of this meeting was?
- check if they had thoughts about the possible outcomes from this consultation?
- ensure that they understood their perspective at this stage of the interview?

4. **Before** breaking the news did you introduce it with sensitivity?

definitely 5 _ 4 _ 3 _ 2 _ 1 not at all

Did you

- gently alert the patient's parents to the fact that what followed was going to be important, before using any specific terms?
- take the lead from the patient's parents as to whether to speak or listen after introducing the news?

5. When delivering the news did you allow you to decide the detail and language used?

definitely 5 _ 4 _ 3 _ 2 _ 1 not at all

Did you

- begin by using non-specific lay terminology?
- respond to the patient's parents' cues, or ask them if they wanted more detail, before becoming more specific?
- check that they were satisfied with their own understanding of the terms used?

6. Did you allow the patient's parents to set the pace for the delivery of the news?

definitely 5 _ 4 _ 3 _ 2 _ 1 not at all

Did you

- deliver appropriate information when it was asked for?
- give the news at a rate which gave the patient's parents time to think and respond?
- check that the patient's parents had understood and assimilated what had been said before giving more information?

<p>7. Did you use an appropriate pause after giving the news? definitely 5 _ 4 _ 3 _ 2 _ 1 not at all</p> <p>Did you</p> <ul style="list-style-type: none"> • allow the news about the diagnosis and its implications to sink in? • give the patient’s parents time to respond? • appropriately break the silence if the pause was too long?
<p><i>C. Eliciting concerns:</i> this section focuses on whether you actively attempted to gain a clear idea of the personal implications and meaning of the news to you, and the concerns that it generated.</p>
<p>8. Did you specifically invite questions? definitely 5 _ 4 _ 3 _ 2 _ 1 not at all</p> <p>You may have needed to invite questions repeatedly.</p>
<p>9. Did you explicitly attempt to obtain a complete list the patient’s parents’ concerns? definitely 5 _ 4 _ 3 _ 2 _ 1 not at all</p> <p>Did you explore</p> <ul style="list-style-type: none"> • the patient’s parents’ feelings and emotions about the news just given? • the patient’s parents’ concerns about treatment? • the patient’s parents’ concerns about prognosis? • the concerns arising from family and relationship issues? • the patient’s parents’ concerns about the effect on their social setting, for example their employment?
<p><i>D. Information giving:</i> this section looks at aspects other than giving the news itself.</p>
<p>10. Did you give information tailored to the patient’s parents’ expressed concerns? definitely 5 _ 4 _ 3 _ 2 _ 1 not at all</p> <p>Did you</p> <ul style="list-style-type: none"> • give information in a manner which related to the patient’s parents’ expressed concerns? • answer the patient’s parents’ questions?

11. Did you clearly explain any information given so that the patient's parents understood?
definitely 5 _ 4 _ 3 _ 2 _ 1 not at all

Did you

- give information in an ordered and logical manner?
- use terms appropriate to the patient's parents using plain English and avoiding jargon?
- check that the patient's parents understood and offer clarification?
- summarize points for the patient's parents?

12. Did you manage to focus on any positive aspects? **definitely 5 _ 4 _ 3 _ 2 _ 1 not at all**

Did you

- frame treatment options in a positive way?
- achieve a good balance between explaining benefits and side-effects?
- manage to give correct information about the prognosis without extinguishing hope?

E. General Information: the following points relate to the interview as a whole.

13. Were the psychosocial issues which the patient's parents flagged up during the interview **explored**?
definitely 5 _ 4 _ 3 _ 2 _ 1 not at all

Did you

- acknowledge: the patient's parents' feelings and emotions; and the effects on family and relationships, and social circumstances?
- allow the patient's parents to talk about these issues?
- enter into a dialogue?

14. Did you manage to appear supportive during the interview?
definitely 5 _ 4 _ 3 _ 2 _ 1 not at all

Did you

- show warmth?
- show emotional supportiveness?
- convey a sense that this really mattered to you?
- convey a personal sense of strength and resourcefulness that was available to help the patient's parents?

15. Did you use appropriate body language during the interview?

definitely 5 _ 4 _ 3 _ 2 _ 1 not at all

Did you

- maintain an appropriate level of eye contact?
- look interested and alert to the patient's parents' needs?
- show a competent and caring professional manner?

16. Did you avoid appearing clumsy during the interview?

never clumsy 5 _ 4 _ 3 _ 2 _ 1 often clumsy

Did you

- introduce difficult topics gently?
- deal with painful issues sensitively?
- show flexibility and sensitivity to the patient's parents' needs?
- avoid non sequitur?
- avoid using phrases that were inappropriate?

17. Did you tailor the pace of the interview to suit the patient's parents?

definitely 5 _ 4 _ 3 _ 2 _ 1 not at all.

Did you

- let the patient's parents speak without interruption?
- respond to their cues regarding timing and delivery?
- deliver appropriate information when it was asked for?
- use pauses where appropriate to give them time to think and respond?
- check that they had finished with a topic before moving on to another?

18. Did you manage the time available? **very well 5 _ 4 _ 3 _ 2 _ 1 poorly**

Did you

- cover the important issues in this session?
- make a plan for future action?
- bring the interview to a conclusion?

Appendix B

Duke University IRB Form

Dear Residents,

As you know, the Duke Pediatric Residency Training Program is committed to providing you with excellent education to best train you for independent practice. Part of independent practice is the ability to incorporate self-assessment within your medical practice. In addition, it is a requirement of the ACGME that residency training programs teach residents about self-assessment and conduct ongoing program assessments.

To fulfill these requirements, we will have residents complete a self-assessment exercise as part of their residency program curriculum. This is an expectation for all residents in the program. Residents will complete the standardized patient parent Communications session of the Clinical Skills Fair (as is currently being done). At the end of their session, residents will complete a short self-assessment of their performance. The standardized patient parents will also complete an assessment of your performance which you will be provided with at a later date for feedback on your performance. After a period of 2-4 weeks, residents will return to the residency office for a review of their recorded communication sessions. Each resident will then complete the same short self-assessment based on the video review of their performance. Each resident will review their individual recorded session on a handheld device which they will be allowed to keep for future sessions.

We will compare the difference in resident self-assessment scores with the standardized patient parent assessment scores. Additionally, pre- and post- video review resident self-assessment scores will be compared. This data will be compared in aggregate form. Individual resident scores will NOT be able to be traced to the resident. Scores will NOT be used in individual resident promotion or evaluation.

Your participation in the self-assessment is an expectation of the residency training program curriculum. We are requesting that we include your scores (in aggregate and without identifying information) in further analysis. This portion of your participation is voluntary. If you choose to participate, you will be allowed further use of the handheld device for future sessions. If you decline, you will not be adversely affected.

The major benefit of participation is to help us improve our program and allow us to ensure you are receiving the best training available. Additionally, participation will allow you to assess your own strengths and weaknesses for your further study.

Please contact Shari Whicker, MEd or the Duke IRB with any questions or concerns.

Appendix C

Duke IRB Exemption Approval

DECLARATION OF EXEMPTION FROM IRB REVIEW

The DUHS IRB has determined that the following protocol meets the criteria for a declaration of exemption from further IRB review as described in 45 CFR 46.101(b)(1), 45 CFR 46.102 (f), or 45 CFR 46.102 (d), satisfies the Privacy Rule as described in 45 CFR 164.512(i), and satisfies Food and Drug Administration regulations as described in 21 CFR 56.104, where applicable.

Protocol ID: Pro00030337
Protocol Title: Pediatric Resident Self-Assessment of Communication Skills
Principal Investigator: Shari Whicker
Review Date: 5/1/2011
Expiration Date: **Does not expire*
Exempt Category: Exempt from IRB Review (45CFR46.101(b)(1))

*This Declaration of Exemption from further IRB Review is in effect from 5/1/2011 and does not expire. However, changes to the proposed research will require an amendment requesting re-review for exemption. Reportable serious adverse events and unanticipated problems related to the research that place subjects or others at risk of physical, psychological, economic, or social harm must be promptly reported to the IRB and will result in reconsideration of the activity's exempt status.



DUHS Institutional Review Board
2424 Erwin Rd | Suite 405 | Durham, NC | 919.668.5111
Federalwide Assurance No: FWA 00009025



DukeMedicine
Institutional Review Board for Clinical Investigations

Appendix D

North Carolina State University IRB Exemption Approval

North Carolina State University is a land-grant university and a constituent institution of the University of North Carolina

Office of Research and Innovation
Division of Research Administration

NC STATE UNIVERSITY

Campus Box 7514
Raleigh, North Carolina 27695-7514

919.515.2444 (phone)
919.515.7721 (fax)

From: Carol Mickelson, IRB Coordinator
North Carolina State University
Institutional Review Board

Date: April 17, 2012

Title: Medical Resident Ability to Self-Assess

IRB#: 2618

Dear Shari Whicker

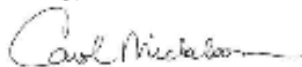
The research proposal named above has received administrative review and has been approved as exempt from the policy as outlined in the Code of Federal Regulations (Exemption: 46.101. b.4). Provided that the only participation of the subjects is as described in the proposal narrative, this project is exempt from further review.

NOTE:

1. This committee complies with requirements found in Title 45 part 46 of The Code of Federal Regulations. For NCSU projects, the Assurance Number is: FWA00003429.
2. Any changes to the research must be submitted and approved by the IRB prior to implementation.
3. If any unanticipated problems occur, they must be reported to the IRB office within 5 business days.

Please forward a copy of this letter to your faculty sponsor, if applicable. Thank you.

Sincerely,



Carol Mickelson
NC State IRB