# Colonoscopy Performance Correlates with Scores on the FES Manual Skills Test

Department of Education and Counselling Psychology
McGill University
MEd (Educational Psychology) Special Activity Project

Carmen Mueller, BSc(H), MD, FRCSC

May 5, 2014

## Preface

This work was performed in collaboration with multiple authors (see page 20). My role in creating this document included study design, ethics board application, data collection and analysis, and final manuscript creation. PK assisted instrumentally with data management and analysis. LF, GM and MV assisted with thoughtful study design and editing of the condensed version of this manuscript, which has been accepted for publication in *Surgical Endoscopy*. I gratefully acknowledge the salary support received from the Steinberg-Bernstein Centre for Minimally Invasive Surgery during the completion of this work.

## Abstract

Achieving proficiency in flexible endoscopy is a major priority for general surgery training programs. The Fundamentals of Endoscopic Surgery (FES) is a high-stakes examination of the knowledge and skills required to perform flexible endoscopy. The objective of this study was to establish additional evidence for the validity of the FES hands-on test as a measure of flexible endoscopy skills by correlating clinical colonoscopy performance with FES score. Participants included FES-naïve general surgery residents, gastroenterology fellows at all levels of training, and attending physicians who regularly perform colonoscopy. Each participant completed a live colonoscopy and the FES hands-on test within two weeks. Performance on live colonoscopy was measured using the Global Assessment of Gastrointestinal Endoscopic Skills–Colonoscopy (GAGES-C, maximum score 20), and performance on the FES hands-on test was assessed by the simulator's computerized scoring system. The clinical assessor was blinded to simulator performance. Scores were compared using Pearson's correlation coefficient. A total of 24 participants were enrolled (mean age: 30; 54% male) with a broad range of endoscopy experience; 17% reported no experience, 54% had fewer than 25 previous colonoscopies and 21% had more than 100. The FES and GAGES scores reflected the broad range of endoscopy experience of the study group (FES™ score range: 32-105; GAGES score range: 5-20). Pearson's correlation coefficient between GAGES-C scores and FES hands-on test scores was 0. 78 (0. 54-0. 90, $p < 0.0001$) All eight participants with GAGES-C score > 15/20 achieved a passing score on the FES hands-on test. There is a strong correlation between clinical colonoscopy performance and scores achieved on the FES hands-on test. These data support the validity of FES as a measure of colonoscopy skills.

# Colonoscopy Performance Correlates with Scores on the FES Manual Skills Test

## Importance of Endoscopy to General Surgery Practice

Provision of diagnostic endoscopy services (colonoscopy and upper endoscopy) is an important part of health-care delivery, and a core practice component for gastroenterologists and general surgeons in North America (Heneghan et al., 2005; Hilsden & Tepper, 2007). Several studies have tried to determine to what extent endoscopy services in North America are provided by general surgeons, in order for programs to understand and meet training requirements.

A recent review by a Canadian group was able to document the provision of flexible endoscopy services in Canada very well (Hilsden & Tepper, 2007). Using data from the National Physician Database and the Canadian Institute of Health Information, this group documented the practices of all physicians in Canada who performed at least 100 endoscopic procedures in 2002. They found that 1444 physicians qualified for inclusion in their study. Of those, surgeons provided 44% of all colonoscopies (lower endoscopies) and 28% of all gastroscopies (upper endoscopies). In contrast, gastroenterologists provided 53% and 59% respectively. The remainder were done by other specialists, mostly family physicians. In urban areas, gastroenterologists provided the bulk of services (71% upper; 65% lower) whereas, in nonmetropolitan areas, surgeons provided the majority of both upper and lower endoscopies (51% and 67% respectively).

A similar study documenting endoscopy volume based on billing data and central health records has not been carried out in the United States, most likely because the multiple payer health system makes it difficult to track data in this manner. However, a survey of 390 rural and 145 urban surgeons did document endoscopy practice patterns, among other practice statistics

(Heneghan et al., 2005). That study aimed to document differences in practice patterns, case volumes and types among surgeons in dense metropolitan areas in comparison to those living in communities of fewer than 50,000 people. As in the Canadian study of endoscopy practice patterns (Hilsden & Tepper, 2007), the results of the Heneghan et al (2005) study showed that rural American surgeons performed significantly more endoscopic procedures than their urban counterparts (220 vs 77 procedures per year respectively, $p < 0.0001$). Rural surgeons were also more likely than urban surgeons to say they wished they had had more training in flexible endoscopy during residency (62% vs 46%, $p = 0.0009$). However, it is interesting that 46% of urban surgeons still would have liked more endoscopy training, even though it is less likely to be part of their practice.

These studies highlighted several key factors in confirming the need for endoscopy training within general surgery residency programs. First, in both Canadian and American settings, rural surgeons are key providers of endoscopy services, most likely because smaller communities cannot attract and support a full-time gastroenterologist. Due to the breadth of services provided by general surgeons (e.g., endoscopy, breast cancer, hernias and gallbladders, major abdominal surgery, and in some cases even basic orthopedics and obstetrics cases), smaller communities often can support at least one general surgeon. Because diagnostic and therapeutic flexible endoscopy is a key component of health care delivery for many diseases, continued training of endoscopy for general surgeons practicing in rural areas is essential throughout North America (American Board of Surgery, 2011).

The above-described studies also showed that endoscopy is a component, albeit less prominent, in the practice of urban general surgeons as well. There are many reasons for this given the variety of niche practice patterns for urban general surgeons, but certainly some

disciplines within general surgery value endoscopy highly as part of their practice. These indisputably include colorectal surgery and any general surgeon who treats colon cancer, as well as those involved in colon cancer screening programs of which colonoscopy is a key diagnostic tool. Furthermore, those surgeons who treat benign and malignant upper gastrointestinal disorders often also perform endoscopy (Hazey et al., 2013).

Given these findings, it is clear that flexible endoscopy remains a key tool for many general surgeons in all practice settings, leading the American Board of Surgery to emphasize the necessity of continued training in flexible endoscopy during surgical residency (American Board of Surgery, 2011). However, given the recent changes in work-hour restrictions throughout the United States and Canada, as well as increasing patient-safety concerns impacting all aspects of hands-on clinical medicine training, concerns that graduating residents do not have sufficient endoscopy experience have been raised (Cass, Freeman, Peine, Zera, & Onstad, 1993; Vassiliou et al., 2010b). This has led to a recent focus among medical educators regarding how best to ensure and verify competence in flexible endoscopy for both general surgery and gastroenterology trainees. These efforts are summarized below and include work to define what actually comprises flexible endoscopy skills, training guidelines available, and the development of assessment tools that can accurately measure proficiency.

**Defining Basic Endoscopic Skills**

Skills needed to perform flexible endoscopy encompass both cognitive and manual elements. Around the world, several prominent specialist societies have defined goals of flexible endoscopy training programs in order to encapsulate the broad range of knowledge and technical skills needed to perform endoscopy.

The American Society of Gastroenterologists (ASGE) is a specialist organization representing gastroenterologists in North America that has set out to define the principles of training for flexible endoscopy. Similarly, the American Board of Surgery, an organization whose members comprise general surgeons from all around the world, and the sole licensing body for general surgeons practicing in the United States, has also defined goals of flexible endoscopy training. The European Society of Gastroenterologists (ESGE) is the European equivalent to the ASGE. While the ESGE has not published comprehensive endoscopy training guidelines, they provide a wealth of training material to their members in the form of guidelines and e-learning DVDs aimed at furthering member education in flexible endoscopy. Another major society in the field of gastrointestinal endoscopy is the Japanese Society of Gastroenterology, which also provides education and training materials to its members.

Although each society individually provides educational and training materials and guidance to its members, the nature of these does not vary much between societies. All agree as to what constitutes basic endoscopic skills, both cognitive and technical. In general, the cognitive aspects of both upper (gastroscopy) and lower (colonoscopy) endoscopy involve knowledge of indications and contraindications, knowledge of risks and complications, understanding of patient preparation, positioning, and monitoring, lesion recognition and indications for intervention, and knowledge of appropriate follow up (American Board of Surgery, 2014; Eisen et al., 2002; Rembacken et al., 2012).

Technical skills necessary for upper endoscopy include safe and comfortable esophageal intubation. scope navigation to the second stage of the duodenum;. instrumentation (such as polypectomy or targeted biopsy), and retroflexion (turning the scope tip 180 degrees to perform examinations and instrumentation while looking backwards). For lower endoscopy, the technical

skills are similar but also include loop reduction (undoing a coil in the scope without removing it completely), techniques of patient positioning and scope manipulation to navigate through the long and tortuous colon, and intubation of the ileocecal valve (American Board of Surgery, 2014; Walsh, Sherlock, Ling, & Carnahan, 2012).

**Current Endoscopy Training Guidelines**

Numerous organizations throughout North America have published guidelines with respect to endoscopy training. These vary in their depth and specificity, with some being quite broad and imprecise while others are more explicit and well-defined.

The Royal College of Physicians and Surgeons of Canada, the sole credentialing body for medical trainees in Canada, specifies that general surgery trainees must "demonstrate proficient and appropriate use of procedural skills, both diagnostic and therapeutic, specifically: demonstrate effective, appropriate, and timely performance of diagnostic procedures relevant to General Surgery including, but not limited to upper gastrointestinal endoscopy and colonoscopy (Objectives of Training in the Specialty of General Surgery, 2010). The College does not specify how these training requirements are to be met, but does stipulate that up to six months of the five-year training period may be spent on rotations related to general surgery, including gastroenterology (Specialty Training Requirements in General Surgery, 2010). Some general surgery training programs, such as at the University of Toronto and McGill University, require their residents to spend at least two months on a dedicated gastroenterology rotation (Ahmed, 2014; Fata, 2014).

The American Board of Surgery, the main credentialing organization for general surgeons in the United States, provides considerably more specific guidelines with respect to endoscopy training than its Canadian counterpart. The ABS has developed a full endoscopy

curriculum through which surgical trainees are expected to progress during their residency training (American Board of Surgery, 2014). The curriculum describes time commitments, necessary skills of educators and trainers, as well as equipment needed to administer the entire curriculum. Furthermore, recommendations regarding assessment of trainee progress are made to enable programs to ensure their residents are acquiring the recommended knowledge and skills. The program details cognitive and technical skills commensurate with each level of residency training and makes recommendations regarding ways of acquiring those skills. Finally, a summary of resources (textbooks, didactic materials, and training platforms for technical skills) can also be found in curriculum outline.

Finally, the American Association of Gastroenterologists (ASGE) published a summary of the principles of training in flexible endoscopy 15 years ago (Adler et al., 2012), and it along with the updates made to it over time, have served as a guideline for endoscopy training for both gastroenterology and general surgery programs. In the original article, the authors outlined in broad terms the key elements of endoscopy training, both cognitive and technical. However, because at the time of publication very few inanimate endoscopy training platforms were available, the main recommended method of acquiring procedural skills remained graduated responsibility within an apprenticeship-type program. This differs substantially from the curriculum outlined by the American Board of Surgery, that emphasizes the training of basic skills on virtual reality or physical endoscopy simulator models, before allowing novices to practice procedural skills on real patients (American Board of Surgery, 2014).

Another key difference between the ASGE recommendations and the ABS curriculum is that the ASGE document is merely a summary of guiding principles in endoscopy training, rather than a true curriculum. The ASGE has since published a true curriculum, however,

encompassing not only skills in flexible endoscopy, but all cognitive and procedural skills needed to practice gastroenterology as a specialty (ASGE Task Forces, 2003). This guideline is geared towards the training of gastroenterologists, however, the section regarding gastrointestinal endoscopy may also serve as a guideline for the training of surgeons who perform flexible endoscopy. Similar to the ABS document, the ASGE curriculum highlights the importance of skilled trainers and adequate training facilities, however does not provide details regarding the sequential acquisition of skills and knowledge, or provide recommended resources for study and practice the way the ABS curriculum does. Furthermore, the ASGE curriculum does not explicitly indicate which assessment tools are to be used to determine competence.

In summary, although there are several key stakeholders in the field of endoscopy training within North America, clear resources detailing endoscopy training are few and vary considerably between regulatory organizations. The recently-published American Board of Surgery endoscopy curriculum is presently the most comprehensive and specific training manual for flexible endoscopy currently available in North America, however it should be noted that this curriculum focuses on procedural endoscopy and does not cover the entire breadth of the speciality of gastroenterology.

**Measures of Proficiency in Flexible Endoscopy**

In parallel to the recent increase in interest in endoscopy curricula, the last 10 to 15 years have seen a surge of interest among the medical community with respect to documenting proficiency in flexible endoscopy and thereby having benchmarks against which to determine who is eligible for hospital privileges to perform endoscopy. This has stemmed in no small measure from an increasing focus on patient safety across all medical disciplines, heightening the importance of credentialing criteria for specialized procedures.

**Case numbers as a measure of proficiency.** Historically, the most commonly-used methods of determining proficiency in endoscopy were by documenting number of cases performed in training, total time spent in training, or both. The most recent ASGE gastroenterology curriculum published in 2003 (ASGE Task Forces, 2003) stipulated that a minimum of 130 upper and 140 lower endoscopies should be performed by trainees before competence could be assessed, and stated that all trainees required a minimum of at least these many procedures ("never less") to achieve competence. However, the curriculum does not go on to describe how competence should be assessed after completion of the requisite number of procedures, except to say that the training faculty are to determine competence of their trainees. In contrast to this, the Accreditation Council for Graduate Medical Education (ACGME), the organization responsible for accreditation of all residency-training programs in the United States, requires general surgery residents to complete only 50 colonoscopies and 35 upper endoscopies to achieve competence in flexible endoscopy (ACGME Case Log System; Surgery Policy Information, 2005). The reason for the discrepancy between these credentialing organizations regarding the minimum number of cases needed to achieve competence is due conflicting and sparse data in this field.

The minimum numbers for upper and lower endoscopy required for credentialing by the ASGE were derived from a study by Cass et al (1993), in which seven gastroenterology fellows and five fourth-year general surgery trainees were evaluated using a simple scoring system for upper and lower endoscopy proficiency (trainees were scored either "yes" or "no" based on their ability to complete the each procedure and to note abnormalities along the way). This study found that by 50 upper endoscopies, trainees could successfully initiate and complete the procedure over 95% of the time, however these numbers did not improve and even decreased

when procedure volumes increased to 140 cases. Similar results were found for lower endoscopy (which is more technically challenging than upper endoscopy). After 50 procedures, trainees were able to complete the procedure 86% of the time, but this rate did not improve with increased experience even up to 110 cases. The study went on to document that more experienced trainees who had completed >270 cases, and faculty who have performed over 1000 cases in total and perform approximately 100+ cases each year, could complete lower endoscopies 97% of the time. This study also found no difference in the speed of acquisition of skills between general surgery and gastroenterology trainees.

It is interesting that this study formed some of the basis by which case volumes for determining proficiency in endoscopy were determined by the ASGE. First, it involved a relatively small number of trainees (12) and was performed at a single training center. Second, the difficulty level of the cases the trainees were involved in was not standardized or documented. Moreover, the results suggested that competence can be achieved with as few as 50 of each upper and lower endoscopies, however the authors concluded that at least 100 procedures for each are necessary before competence can be achieved.

There are unfortunately few other studies in the literature addressing the question of case numbers needed to achieve procedural competence in flexible endoscopy. One of the most-commonly cited is a prospective study of quality and outcomes in colonoscopy performed by surgeons that revealed a high rate of successful completion and a low complication rate (Wexner et al,. 2001). In this study, 13,580 colonoscopies were self-evaluated by the surgeon performing the procedure according to a standardized questionnaire. Experience ranged from under 10 to over 1000 prior procedures. In this study, a minimum of 50 prior cases was associated with significantly lower procedure times and complication rates, however the completion rate was not

significantly associated with any experience level (staff, fellow, or resident) or the number of prior cases performed. The authors concluded that, based on their findings, case numbers alone cannot be used to determine competence in flexible endoscopy.

**Objective measures of proficiency in flexible endoscopy.** The lack of strong evidence supporting case volume alone as a measure of competence has caused the ASGE to explicitly state that "performance of an arbitrary number of procedures does not guarantee competence" (Eisen et al., 2002), even though the organization continues to support the use of minimum case numbers for credentialing. Similarly, case numbers are still used in guiding who is eligible for credentialing in Canada as well (Armstrong et al., 2007). Perhaps the reason for this is the lack of objective assessment tools to determine competence in flexible endoscopy, and so for the time being case numbers offer at least a rough approximation of who is likely to possess adequate endoscopy skills (Dunkin & Vargo, 2008). The American Board of Surgery has called for more objective measures of competency than case numbers or time in training (American Board of Surgery, 2011), and there is a general consensus among medical educators that such assessment tools of endoscopic skills have been needed greatly.

In the last four years, educators in endoscopy in North America and Europe have answered this need with the development of several different assessment tools for flexible endoscopy. Following is a summary of these instruments and the evidence behind each.

***Mayo Colonoscopy Skills Assessment Tool (MCSAT).*** This assessment tool was developed by a group of gastroenterologists at the Mayo Clinic in Rochester, Minnesota (Sedlack, 2010). It was developed as a computer-based assessment, and breaks colonoscopy skills down into component competencies that were derived from published expectations of skills by prominent endoscopy organizations and expert consensus of the nine staff endoscopists at the

Mayo Clinic. The group defined 14 specific subskills (six cognitive and eight motor) that they felt encapsulated the core skills needed for competence in routine colonoscopy. The wording of each assessment component was then debated by the committee and by surveying other endoscopy teachers for their feedback and opinions. The MCSAT breaks each component down into four skill categories: novice, intermediate trainee, advanced trainee, and competent, with anchors (or benchmark performance) describing the achievement level for each to avoid ambiguity and standardize scoring between assessors. The tool was designed to be administered only by trained endoscopists who are familiar with teaching endoscopic skills to trainees.

The initial study validating the use of the MCSAT for assessment of colonoscopy skills involved scoring 41 gastroenterology fellows over three years in the performance of 6390 colonoscopies. All subskill scores correlated highly ($r > 0.59$) with overall cognitive and motor scores and total scores, demonstrating strong internal structure. Furthermore, average scores in each subdomain significantly improved with procedure number from 20 to 150 to 300 colonoscopies, demonstrating that each subdomain reliably measured a component of competence that increases as expected with increasing experience.

The advantages of the MCSAT assessment tool is that it was validated for use in an educational setting as part of a continuous assessment program. It can therefore be used to track learning trajectories as well as to determine competency level at a particular point in time. Furthermore, the MCSAT allows for feedback and procedure take-over by the supervising physician as would normally happen in any training environment, with these events being accounted for in the scores.

Limitations of this assessment tool are that it is quite lengthy to administer if it is to be used for each procedure a trainee performs. Also, the Mayo Clinic group who developed it

incorporated it into their computer-based endoscopy charting system, allowing for scores to be instantaneously and privately recorded through password protection. However, this technique limits the use of the MCSAT in other centers who use different endoscopy charting software or still rely solely on paper charts. Finally, the MCSAT was designed as a measure of colonoscopy proficiency only, rather than endoscopy as a whole. It therefore incorporated measures of competency only relevant to colonoscopy (such as "*cecal intubation time*") into the total score, limiting its use in measuring competency in other endoscopic procedures.

***Assessment of Competence in Endsocopy (ACE).*** To compensate for some of these limitations in the MCSAT, the ASGE has recently developed a modified version of the assessment tool that they have termed the Assessment of Competency in Endoscopy (ACE) score (Sedlack et al., 2014). This instrument now comprises two separate scales, one for upper and one for lower endoscopy, addressing the issue of the limited scope of the previous MCSAT tool. Furthermore, some of the wording of the Likert-scale anchors was changed to be clearer, and subscales previously found to be too vague were divided into narrower categories, clarified, or both. The ASGE also published a paper version of both tests to facilitate dissemination of the metric to all training centers. The metric currently is only available in English.

***Direct Observation of Procedural Skills (DOPS).*** The Direct Observation of Procedural Skills endoscopy proficiency metric was developed in the United Kingdom by the Joint Advisory Group for Gastrointestinal Endoscopy (JAG). The purpose of the assessment tool was to determine competence in colonoscopy for practicing endoscopists who wanted to participate in the national Bowel Cancer Screening Program (BCSP). The DOPS was developed through consultation with expert endoscopists, endoscopy nurses and endoscopy educators in an iterative process. The final assessment tool breaks endoscopy skills into four sections: pre-procedure

preparation, sedation and monitoring, diagnostic procedural skills, and therapeutic ability. In total, 20 individual items were developed across all four categories, and measured on a four-point Likert scale with detailed anchors describing the level of skill attributable to each level.

Study participants were expert endoscopists who met certain skills criteria for inclusion (such as high polyp detection rate, more than 500 lifetime colonoscopies performed, and use of appropriate average medication dosages). In total, 147 colonoscopists were evaluated using the DOPS instrument. Participants also received global evaluations by trained study personnel. Each participant was scored by two assessors on two colonoscopies. The expert global assessments mirrored the DOPS scores in 97% of assessments. Furthermore, there was good reliability between scores on the two colonoscopies done by each participant. Interestingly, the DOPS scores did not correlate with participant self-reported skill levels (such as polyp detection rate or procedure completion rate).

The strengths of the DOPS tool include the use of experienced colonoscopists as raters, and that the test assesses both cognitive and motor skills necessary to perform endoscopy. The limitations of the study to assess the validity of the DOPS is that only expert endoscopists were enrolled, so the true ability of the test to discriminate between endoscopists of varying skill levels is limited. Furthermore, the test is designed to only evaluate proficiency in colonoscopy, limiting its use to evaluate endoscopic skills in general.

***Global Assessment of Gastrointestinal Endoscopic Skills (GAGES).*** The Global Assessment of Gastrointestinal Endoscopic Skills (GAGES) assessment tool was developed in cooperation between centers in North America and Sweden for assessment of the core motor skills necessary to perform both upper and lower endoscopy (Vassiliou et al., 2010a).

Although the other assessment tools described above assess both cognitive and motor skills involved in the performance of colonoscopy, this final assessment instrument assesses only motor skills. There is one version for upper endoscopy (GAGES-UE) and another for colonoscopy (GAGES-C). Both assessment tools were developed by consensus with input from expert endoscopists and educators. Each tool breaks procedural skills down into five component parts, and rates performance from 1 to 5 on a Likert scale with clear anchors at 1, 3 and 5. Expert endoscopists performed the assessments. During the procedure, proctor prompting and procedure take-over are allowed with these events taken into account by the scoring system.

The validation of both assessment tools involved 139 evaluations collected at 11 institutions in the United States, Canada, and Sweden. Correlation between GAGES-UE and GAGES-C was high (0. 75), and experienced endoscopists performed significantly better than novices on both tests ($p < 0.0001$). Furthermore, comparison of scores obtained from observers and attending endoscopists showed high reliability, with interclass correlation coefficients of 0. 96 (GAGES-UE) and 0. 97 (GAGES-C).

The strengths of the GAGES assessment tools are that they allow evaluation of both upper and lower endoscopy, rather than just one type of procedure, and use a specific metric for each. Furthermore, the test takes into consideration procedure difficulty level and the need for prompting or take-over by the proctor, making it a safe evaluation tool for all skill levels. The tool is also able to discriminate between experts and novices, demonstrating the validity of the skills being assessed in contributing to endoscopy proficiency. Finally, the Likert-scale anchors allow for formative feedback to be provided to trainees, potentially improving learning.

The limitations of this tool are that it does not evaluate the cognitive aspects of endoscopy (such as indications for the procedure, or patient preparation). Also, reliability was

tested using two observers scoring the same participant, but participants were only evaluated on one procedure. Finally, participants were evaluated on "*uncomplicated*" procedures, but the determination of this level of complexity was not standardized and was left to the discretion of the attending endoscopists.

***The Fundamentals of Endoscopic Surgery (FES).*** The Fundamentals of Endoscopic Surgery is a comprehensive assessment tool of basic endoscopic proficiency (Hazey et al., 2013) developed by the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES). The program comprises a didactic cognitive curriculum (regarding indications for endoscopic procedures, identification of abnormalities, patient preparation and monitoring, etc), as well as a multiple-choice knowledge test (Poulose et al., 2014) and a hands-on skills assessment (Vassiliou et al., 2014). With the inclusion of the didactic training modules, the FES program is intended to be used as both a training tool and evaluation instrument.

The hands-on component includes five tasks that test core skills needed to be proficient in flexible endoscopy: scope navigation, loop reduction, retroflexion, mucosal evaluation, and instrument targeting (Hazey et al., 2013; Poulose et al., 2014; Vassiliou et al., 2014). Unlike the other assessment tools described previously, the FES manual-skills tasks are assessed not on live patients but on a virtual-reality endoscopy-simulation platform (GI Mentor$^{TM}$ II, Simbionix, Israel). The original validation studies of the manual skills tasks enrolled 111 participants from gastroenterology and general surgery programs at 11 different academic training centers in the United States and Canada at all levels of training. In this study, the FES hands-on task scores were found to have good internal consistency (internal consistency coefficient = 0. 82) and correlated highly with experience level ($r$ = 0. 73). The passing score of the FES hands-on skills

test was thus calculated using receiver-operator curves to maximize sensitivity and specificity. The passing score is proprietary and not revealed by the study authors (Vassiliou et al., 2014).

# Colonoscopy Performance Correlates With Scores On The FES Manual Skills Test

Carmen L Mueller, Pepa Kaneva, Gerald M Fried, Liane S Feldman, Melina C Vassiliou

*Steinberg-Bernstein Centre for Minimally Invasive Surgery, McGill University Health Centre, and Department of Surgery, McGill University, Montreal, Canada*

## Rational for the Present Study

Although the reviewed data support the use of the FES manual-skills tasks as a test of endoscopic technical proficiency, correlation of performance on the FES hands-on skills-test tasks with a measure of real-life clinical performance has not yet been examined. Such evaluation would provide additional validity evidence to support FES as a measure of endoscopic skills. Thus, the objective of this study was to correlate FES hands-on test scores with clinical colonoscopy performance.

## Methodology

### Participants

This was a single-institution prospective study. Participants were trainees (general surgery residents and gastroenterology fellows) and attending physicians who regularly (at least once per month) performed colonoscopy in their practice. Participation was entirely voluntary. The study protocol was approved by the Research Ethics Board of McGill University and the FES Task Force of the Society of Gastrointestinal and Endoscopic Surgeons (SAGES).

Participants were required to be naïve to the five FES hands-on modules. Participants need not have performed endoscopy in the clinical setting before, but were required to be physicians in a specialty that would include flexible endoscopy as part of their training or future practice.

Once recruited, participants provided basic demographic data using a standardized questionnaire. Subjects were then scheduled to perform either the FES hands-on test first or the live colonoscopy, depending on availability. Live colonoscopy was performed on elective outpatients only (i. e. , emergency colonoscopies were excluded). To decrease the potential learning effects between tests, and from one test to another, trainees were required to perform

both tests a minimum of 48 hours and a maximum of two weeks apart. Because the learning effect of frequent clinical exposure was felt to be more important than the learning effect from one test to another, trainees on a clinical rotation in which they would perform endoscopy on a regular basis were required to complete both tests within 48 hours to minimize a major change in skill level between tests.

**Assessments**

**Clinical performance.** Clinical colonoscopy performance was measured using the Global Assessment of Gastrointestinal Endoscopic Skills (GAGES). As described previously, GAGES is a global rating scale with validity evidence supporting its use as a measure of upper and lower endoscopy skills in the clinical setting (Vassiliou et al., 2010a). The scoring system was designed for use in routine, uncomplicated upper and lower endoscopic procedures with a specific instrument for each: GAGES-UE (upper endoscopy) and GAGES-C (colonoscopy). The GAGES-C instrument is comprised of five skills, scored on a five-point Likert scale with a description to provide anchors at 1, 3, and 5. The minimum composite score is 5 and the maximum score is 25. The five skills are: scope navigation, use of strategies, keeping a clear endoscopic field, instrumentation, and quality of the examination. Because not all patients require instrumentation (such as targeted biopsy or polypectomy), this skill was omitted from this analysis and a maximum total score of 20 was used.

Evaluations using GAGES-C were completed by experienced endoscopists. These individuals were either trained observers who perform endoscopy, second-year gastroenterology fellows, or the attending physician who was supervising the study participant during the endoscopic procedure. In total, scoring was done by nine different proctors. To help determine the reliability of the GAGES-C scoring system among novice users, study participants were also

asked to score themselves on the GAGES-C with minimal instructions regarding the use of the scoring metric. Study participants were blinded as to proctor scores.

**FES testing.** The FES hands-on test was performed on the GI Mentor™ II virtual reality endoscopy simulator (Simbionix™, Airport City, Israel). Exam proctoring was done by study personnel trained in administrating the FES examination. The FES test is comprised of five individual tasks designed to test core endoscopic skills, and scoring was done according to a computerized scoring system. The FES tasks are assessed for both efficiency and precision according to a proprietary algorithm, and the passing score was previously determined from the individual task scores to maximize sensitivity and specificity (Vassiliou et al., 2014). Clinical assessors of the GAGES-C scores were blinded to the participants' performance on the FES tests.

All statistical analyses were performed using SPSS version 21 (IBM). GAGES-C and FES scores, as well as proctor and participant self-scores on the GAGES-C instrument, were compared using Pearson's correlation coefficient. Trainee likelihood of passing the FES test in relation to them having completed a dedicated endoscopy rotation was determined using Fischer's exact test. Data are expressed as means (95% confidence interval) unless otherwise indicated.

## Results

Participant characteristics are presented in Table 1. A total of 24 participants were enrolled, 18 (75%) of whom had completed at least two months of dedicated endoscopy training. The mean age was 30 years (range: 23 to 52), 54% were male and 96% were right hand dominant. The participants included both general surgery (19; 79%) and gastroenterology (5; 21%) trainees and practitioners.

Participants had a broad range of endoscopy experience (Table 2). Seventeen percent reported no experience, 21% reported fewer than 25 upper endoscopies and 54% had done fewer than 25 colonoscopies. Twenty-five percent reported having done more than 100 upper endoscopies while 21% reported having done more than 100 colonoscopies. Study subjects had greater overall experience with upper endoscopy than lower endoscopy. Participants from General Surgery were mostly residents in training (16; 67%) and from Gastroenterology participants were mostly first year fellows (4; 17%). Prior exposure to virtual reality endoscopy simulators was low in the study population, with 20 (83%) reporting no prior experience of any kind, and only 1 (4%) reporting greater than one hour of prior exposure.

The range of FES and GAGES proctor scores reflected the range of endoscopy experience in the study group (FES score range: 32-105 ; GAGES score range: 5-20). The correlation between GAGES-C proctor scores and FES manual skills scores was 0. 78 (0. 54-0. 90, p<0. 0001). (Figure 1). All eight participants with GAGES-C scores greater than 15 out of 20 achieved a passing FES hands-on test score. Furthermore, the median GAGES-C score of subjects who obtained a passing FES score was 18 out of 20 (range: 11 to 20). Subjects who had completed at least a two-month endoscopy rotation were significantly more likely to pass FES ($p$ = 0. 035, Fisher's exact test).

The range of proctor GAGES-C scores and GAGES-C participant self-scores was 5 to 20 in both cases. The correlation between proctor GAGES-C scores and participant self-scores was $r$ = 0. 91 (0. 80-0. 96, $p$ < 0. 0001) (see Figure 2).

## Discussion

Fundamentals of Endoscopic Surgery (FES) is the only objective test of proficiency in endoscopy currently available with validity evidence to support its use as a high-stakes

examination. The test comprises both cognitive and manual skills assessments. Although considerable validity evidence has been collected for the hands-on assessment (Vassiliou et al., 2014), the relationship to performance in the clinical environment had not previously been reported. The present study showed that FES manual skills scores are strongly correlated with clinical colonoscopy performance. Also, satisfactory clinical performance (GAGES-C score > 15) was highly predictive of passing the FES hands-on test. Furthermore, the median clinical score of those who obtained a passing score on the FES test was very high (18 out of 20; range: 11 to 20). Together, these findings provide good evidence of predictive validity of the FES manual skills test with respect to actual clinical performance.

Historically, procedural training in medicine has been a true apprenticeship, with the focus being on achieving high case volumes as a way to ensure competence. Patient safety concerns and work-hour modifications have changed the way physicians are trained and credentialed. The focus is turning towards more objective measures of proficiency. As such, the current impetus in medical education is to design valid measures of competence rather than to rely on exposure time or volume as surrogates for proficiency (Leach, 2004). This is certainly true for endoscopy training because case numbers alone have been shown to correlate poorly with skill, and the number of cases needed to achieve proficiency is a matter of ongoing debate among professional societies ("ACGME Memorandum. Changes in minimum requirements for laparoscopy and endoscopy. ," 2006; Adler et al., 2012; Cass et al., 1993; Wexner et al., 2001).

In the current reality of surgical education, there is a need for objective measures of endoscopy knowledge and skill and the FES test was designed to help fulfill this need. The FES hands-on test has undergone rigorous testing in order to obtain evidence to support validity of the test scores in accordance with what would be required for a high stakes examination. In these

studies, “expert” endoscopists were selected to represent the cohort with greatest proficiency, essentially by the opinion of their peers, time in practice, and case numbers (Vassiliou et al., 2014). The present study provided additional evidence to support the validity of the FES hands-on test. The high correlation between clinical performance scores and FES test scores, as well as the finding that subjects who had completed an endoscopy rotation were significantly more likely, but not guaranteed, to pass FES, supports the use of the FES hands-on test as a measure of endoscopy proficiency.

One of the advantages of the FES test is that the scores are computer-generated according to a predetermined algorithm (Vassiliou et al., 2014) and are therefore not subject to rater error or bias. In contrast, the GAGES scoring system relies on the use of human observers to perform the scoring (Vassiliou et al., 2010a), which may be influenced by various sources of error such as the difficulty of an individual clinical case and interactions between individuals. In this study, participants were given no specific training with respect to the use of the GAGES tool, but were blinded to proctor scores and asked to use the Likert scale to score themselves. Their scores correlated very highly with proctor scores ($r$ = 0. 91). This suggests that, although the GAGES metric requires a human observer to determine the score, the subscale descriptions and Likert scale anchors are sufficiently self-explanatory to limit inter-observer variability and provide a reliable score. Although this study was carried out at a single institution, the finding that novice scorers’ results correlated strongly with experienced endoscopists’ scores suggests the GAGES-C instrument could be used to self-assess clinical endoscopy performance in other settings.

One of the limitations inherent in the use of virtual-reality simulators to recreate clinical environments is the extent to which they can truly mimic reality. Indeed, the simulator chosen to create the FES test does not provide haptic feedback, or suction and irrigation functionality as is

normally present in live endoscopy. This study strengthens the validity of the FES hands-on skills tests by demonstrating high correlation between measures of real-life endoscopy skill with scores achieved in the simulated environment, indicating that, although some elements of real colonoscopy are not simulated by the FES modules, the test is nevertheless able to measure endoscopic skills. This also suggests that the simulation experience provides salient experiences in the learning of endoscopic techniques.

Another limitation of virtual reality is the reliance on computer software and complex simulator platforms. These can go through revisions or require recalibration over time. Certainly any major changes to the FES modules, scoring system, or platform of administration would require the acquisition of additional validity evidence. This study was performed on the most current version of the FES hands-on test modules and no major changes have been instituted between data collection and publication of these results.

One of the strengths of this study is that the study group was representative of the populations most likely to take the FES test. These included both general surgeons and gastroenterologists, in training and in practice, with varying levels of experience. Furthermore, multiple proctors performed the GAGES-C scoring, which is akin to real-life training and credentialing environments in which evaluations are often performed by multiple observers.

The limitations of this study included the small sample size, the use of one-time assessments for both FES performance and clinical performance scores, and that the study was conducted at a single center. Further investigations with greater subject enrolment and assessments carried out at other centers would be beneficial to corroborate the findings reported here.

## Conclusion

There is a strong correlation between clinical colonoscopy performance and scores achieved on the FES hands-on test. These data support the validity of the FES hands-on test as a measure of proficiency in flexible endoscopy.

References

*ACGME Case Log System; Surgery Policy Information*. (2005). Retrieved April 27, 2014, from http://www. dconnect. acgme. org/acWebsite/RRC_440/440_policyArchive. asp

*ACGME Memorandum. Changes in minimum requirements for laparoscopy and endoscopy*. (2006). Retrieved April 29, 2014, from http://www. dconnect. acgme. org/acWebsite/RRC_440/440_policyArchive. asp

Adler, D.G., Bakis, G., Coyle, W.J., DeGregorio, B., Dua, K.S. , Lee, L.S., … Faulx, A.L. (2012). Principles of training in GI endoscopy. *Gastrointestinal Endoscopy*, *75,* 231–235. doi:10. 1016/j. gie. 2011. 09. 008

Ahmed, N. (2014). *University of Toronto General Surgery.* Retrieved May 01, 2014, from http://generalsurgery. utoronto. ca/Assets/GenSurgery+Digital+Assets/Education/Residency/Structure+and+Organization+of+the+Program/Mandatory+Content+and+Rotations. pdf

American Board of Surgery. (2011). *Position Statement on GI Endoscopy.* Retrieved April 22, 2014, from http://www. absurgery. org/default. jsp?newsgiresponse

American Board of Surgery. (2014). *The American Board Of Surgery Flexible Endoscopy Curriculum For General Surgery Residents* (pp. 1–26). The American Board of Surgery Inc. Retrieved from http://www. absurgery. org/default. jsp?certgsqe_fec

Armstrong, D., Enns, R., Ponich, T., Romagnuolo, J., Springer, J., & Barkun, A.N. (2007). Canadian credentialing guidelines for endoscopic privileges: an overview. *Canadian Journal of Gastroenterology = Journal Canadien de Gastroenterologie*, *21,* 797–801.

ASGE Task Forces. (2003). Training the gastroenterologist of the future: the Gastroenterology Core Curriculum. *Gastroenterology*, *124,* 1055–1104. doi:10. 1053/gast. 2003. 50160

Cass, O., Freeman, M., Peine, C., Zera, R., & Onstad, G. (1993). Objective Evaluation of Endoscopy Skills During Training. *Annals of Internal Medicine*, *118*, 40–44.

Dunkin, B. J. , & Vargo, J. J. (2008). Measuring procedural competence in endoscopy: what do the numbers really tell us? *Gastrointestinal Endoscopy*, *68,* 1063–1065. doi:10. 1016/j. gie. 2008. 06. 053

Eisen, G.M., Baron, T.H., Dominitz, J.A., Faigel, D.O., Goldstein, J.L., Johanson, J.F., … Wheeler-Harbough, J. (2002). Methods of granting hospital privileges to perform gastrointestinal endoscopy. *Gastrointestinal Endoscopy*, *55,* 780–783.

Fata, P. (2014). *McGill General Surgery.* Retrieved May 01, 2014, from http://www. mcgillgeneralsurgery.com/index.php?option=com_content&view=article&id=78&Itemid=95

Hazey, J.W., Marks, J.M., Mellinger, J.D., Trus, T.L., Chand, B., Delaney, C.P., … Melvin, W.S. (2013). Why fundamentals of endoscopic surgery (FES)? *Surgical Endoscopy*, 13–15. doi:10.1007/s00464-013-3299-3

Heneghan, S.J., Bordley, J., Dietz, P.A., Gold, M.S., Jenkins, P.L., & Zuckerman, R.J. (2005). Comparison of urban and rural general surgeons: motivations for practice location, practice patterns, and education requirements. *Journal of the American College of Surgeons*, *201,* 732–736.

Hilsden, R., & Tepper, J. (2007). Who provides gastrointestinal endoscopy in Canada? *Can J Gastroenterology*, *21,* 843–846.

Leach, D.C. (2004). A model for GME: shifting from process to outcomes. A progress report from the Accreditation Council for Graduate Medical Education. *Medical Education*, *38*, 6–8. doi:10. 1046/j.1365-2923.2004.01733.x

*Objectives of Training in the Specialty of General Surgery.* (2010). The Royal College of Physicians and Surgeons of Canada. Retrieved April 29, 2014, from http://generalsurgery.utoronto. ca/Assets/GenSurgery+Digital+Assets/Education/otr.pdf

Poulose, B.K., Vassiliou, M.C., Dunkin, B.J., Mellinger, J.D., Fanelli, R.D., Martinez, J.M., … Marks, J.M. (2014). Fundamentals of Endoscopic Surgery cognitive examination: development and validity evidence. *Surgical Endoscopy*. *28,* 631-638. doi:10.1007/s00464-013-3220-0

Rembacken, B., Hassan, C., Riemann, J.F., Chilton, A., Rutter, M., Dumonceau, J.M., … Ponchon, T. (2012). Quality in screening colonoscopy: position statement of the European Society of Gastrointestinal Endoscopy (ESGE). *Endoscopy*, *44,* 957–68. doi:10.1055/s-0032-1325686

Sedlack, R.E. (2010). The Mayo Colonoscopy Skills Assessment Tool: validation of a unique instrument to assess colonoscopy skills in trainees. *Gastrointestinal Endoscopy*, *72,* 1125–33. doi:10.1016/j.gie.2010.09.001

Sedlack, R.E., Coyle, W.J., Obstein, K.L., Al-Haddad, M.A., Bakis, G., Christie, J. A., … Rajan, L. (2014). ASGE's assessment of competency in endoscopy evaluation tools for colonoscopy and EGD. *Gastrointestinal Endoscopy*, *79,* 1–7. doi:10.1016/j.gie.2013.10.003

*Specialty Training Requirements in General Surgery.* (2010). Royal College of Physicians and Surgeons of Canada. Retrieved April 25, 2014, from http://generalsurgery.utoronto.ca/Assets/GenSurgery+Digital+Assets/Education/str. pdf

Vassiliou, M.C., Dunkin, B.J., Fried, G.M., Mellinger, J.D., Trus, T., Kaneva, P., … Marks, J.M. (2014). Fundamentals of endoscopic surgery : creation and validation of the hands-on test. *Surg Endosc*. *28,* 704-711. doi:10.1007/s00464-013-3298-4

Vassiliou, M.C., Kaneva, P., Poulose, B.K., Dunkin, B.J., Marks, J. M., Sadik, R. , … Fried, G.M. (2010a). Global Assessment of Gastrointestinal Endoscopic Skills (GAGES): a valid measurement tool for technical skills in flexible endoscopy. *Surgical Endoscopy*, *24,* 1834–41. doi:10.1007/s00464-010-0882-8

Vassiliou, M.C., Kaneva, P., Poulose, B.K., Dunkin, B.J., Marks, J.M., Sadik, R., … Fried, G.M. (2010b). How should we establish the clinical case numbers required to achieve proficiency in flexible endoscopy? *American Journal of Surgery*, *199,* 121–5. doi:10.1016/j.amjsurg.2009. 10. 004

Walsh, C., Sherlock, M., Ling, S., & Carnahan, H. (2012). Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy (Cochrane Review ). *Cochrane Collaboration*, *6,* 1–92.

Wexner, S., Garbus, J., Singh, J., & Group, T.S. (2001). A prospective analysis of 13,580 colonoscopies; Reevaluation of credentialing guidelines. *Surgical Endoscopy*, *15,* 251–261.

## Tables

***Table 1.*** Characteristics of study participants

| | |
|---|---|
| **Demographic Information** | |
| Number of subjects | 24 |
| Completed Endoscopy Rotation | 18 (75 %) |
| Mean Age (range) | 30 (23 to 52) |
| Female/Male | 13/11 |
| Right-Handed | 23 (96%) |
| | |
| **Specialty** | |
| *General Surgery* | |
| PGY 1-2 (no endoscopy rotation) | 6 (25%) |
| PGY 2-6 (endoscopy rotation) | 11 (50 %) |
| Attending | 2 (8%) |
| *Gastroenterology* | |
| Fellow (year 1) | 4 (17 %) |
| Fellow (year 2) | 1 (4 %) |
| | |
| **Previous Endoscopy Simulator Exposure** | |
| No Exposure | 20 (83%) |
| < 1 hour | 3 (12%) |
| 1 to 5 hours | 1 (4 %) |

***Table 2***. *Upper and lower endoscopy experience of study participants.* Data represent the number of study participants in each experience category. (Upper = esophagogastroduodenoscopy; Lower = colonoscopy).

| Procedure Experience Range | Upper | Lower |
|---|---|---|
| 0 | 4 | 4 |
| 1-25 | 1 | 7 |
| 26-50 | 5 | 6 |
| 51-100 | 8 | 2 |
| 101-200 | 5 | 3 |
| >200 | 1 | 2 |

## Figures

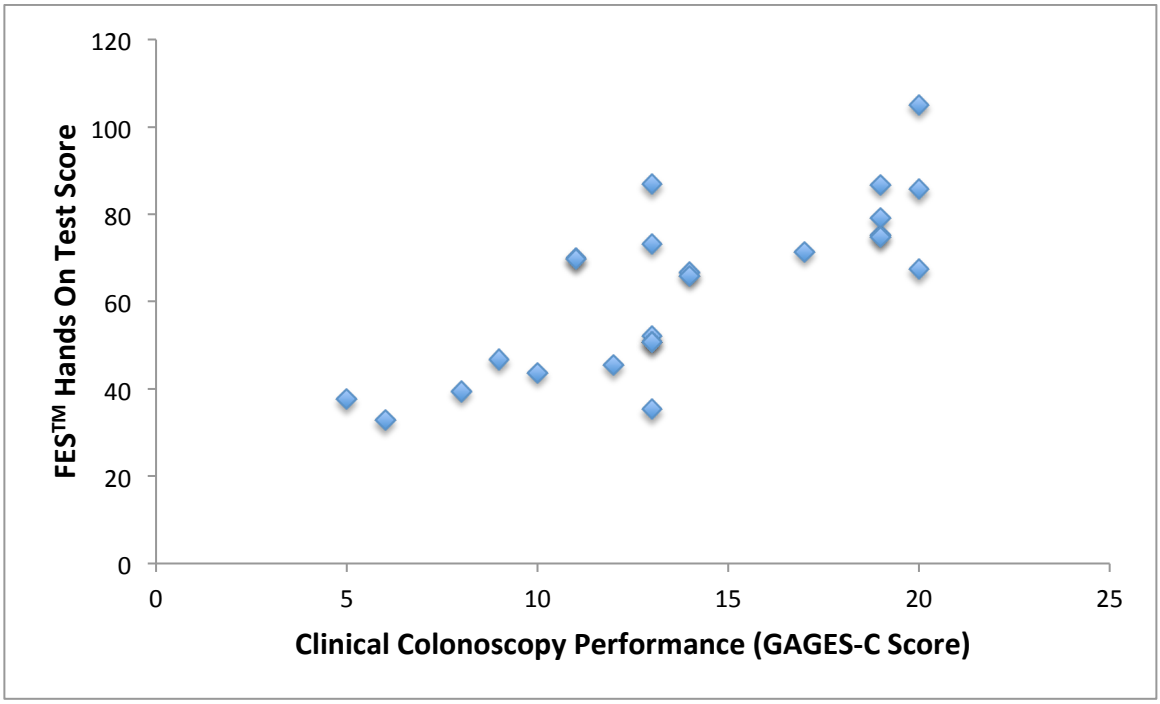


***Figure 1.*** Correlation between clinical colonoscopy performance (GAGES-C score) and FES™ manual skills task scores: r=0. 78 (0. 54-0. 90, p <0. 0001)

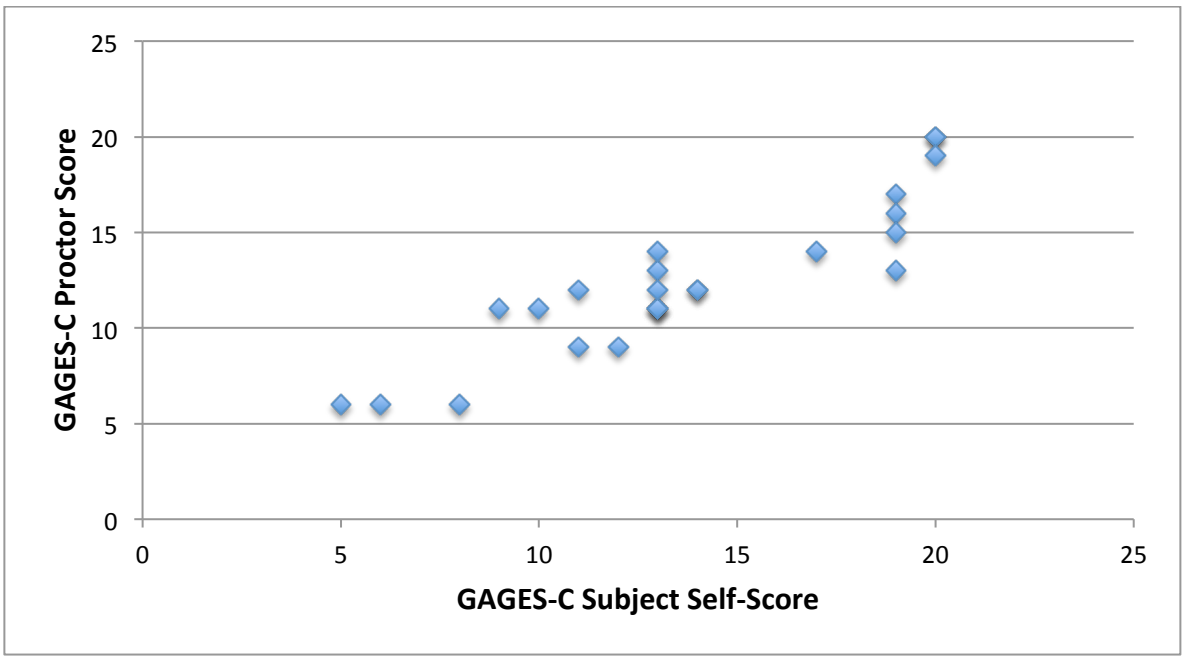


***Figure 2 –*** Correlation between clinical colonoscopy scores recorded by proctors and participant self-scores: $r = 0.91$ (0. 80 to 0. 96, $p < 0.0001$)