The purpose of this document is to provide a detailed description of the College’s certification examination development process, through a question and answer format, for the benefit of interested stakeholders (e.g. schools, the public, exam candidates, and massage therapists). The College is committed to excellence in the creation and execution of our certification examinations and strict adherence to international testing standards ensures the development of standardized, fair, valid, reliable and defensible examinations.

**Why does the College of Massage Therapists of Ontario conduct certification examinations?**

Massage Therapy is one of the regulated health professions in Ontario. All regulated health professions in Ontario are required to self-regulate through a regulatory College (e.g. the College of Massage Therapists of Ontario) according to the Regulated Health Professions Act (RHPA), 1991, profession specific legislation (e.g. the Massage Therapy Act (MTA), 1991), other related legislation (e.g. the Health Care Consent Act (HCCA), 1996), the Personal Health Information Protection Act (PHIPA), 2004), regulations, policy and by-laws. Conducting certification examinations is one of the required elements of self-regulation.

The overarching objective of conducting certification examinations is to protect the public through the maintenance of standards for the profession. The specific goal of the certification examinations is to identify individuals who exhibit a predetermined level of competency by evaluating a candidate’s proficiency level against a standard that differentiates competence from incompetence. Notably, we are not concerned with how candidates perform relative to other candidates, but rather with how a candidate performs relative to a defined passing standard.

Specific legislative direction with regards to the certification examinations can be found in both the RHPA and the MTA. According to the RHPA, each regulated health profession in Ontario is
required to “develop, establish and maintain standards of qualification for persons to be issued certificates of registration.” According to the MTA, “in setting the examinations to be taken by applicants to the College for registration, the College shall specify the general areas of competency to be examined and shall ensure that the examinations provide a reliable and valid measure of a candidate’s competency in knowledge, skills and ability for the practice of massage therapy in Ontario.”

**Who is on the certification services examination content development team?**

The certification examination content development team is composed of the College’s Content Specialist, Subject Matter Experts (SMEs), and Schroeder Measurement Technologies, Inc. (SMT).

The Content Specialist and the OSCE and MCQ SMEs are Registered Massage Therapists (RMTs) in good standing with the College with demonstrated content mastery in relation to the certification examination content outlines. Typically, these individuals have several of the following qualifications: a) extensive clinical experience, b) relevant additional post-secondary education, c) abundant continuing massage therapy education (i.e. CEUs), d) prior related teaching experience, e) test development experience, and f) understanding of the function of the College and the examinations demonstrated through prior roles with the College (e.g. peer assessment, examiner, etc).

Schroeder Measurement Technologies (SMT) is a professional testing company that works with the College on all of the steps of certification examination development (e.g. job analysis, item writing, item review, test development, standard setting, psychometric research, and consulting) to ensure that RMT candidates are provided with a fair, valid, reliable, and effective means of registration in Ontario. Lee L. Schroeder Ed.D. founded SMT in 1995 and has over 40 years of experience in the field of professional assessment and psychometrics. SMT administers certification examinations in over 70 professional categories and administers/scans/scores over 150,000 examinations per year through its global testing center network. Key SMT personnel maintain active membership in research and development organizations, including the American Educational Research Association (AERA), The National Council on Measurement in Education (NCME), The National Council on Licensure Enforcement and Regulation (CLEAR), the National Organization for Competency Assurance (NOCA), and the Association of Test Publishers (ATP). SMT personnel perform active and contributory roles within these organizations; serving on the executive boards and subcommittees, regularly appearing as speakers, and contributing to their publications. SMT works with several Ontario healthcare colleges on examination, delivery, quality assurance, and research projects and has a comprehensive understanding of the Ontario statutory and regulatory environment. SMT staff members include those with experience as
psychometricians, test developers, program managers, editors, computer programmers, test administrators, and support staff for processing, finance, shipping, printing, and security.

**How does the College protect the stakeholders in the certification examination process?**

The College’s commitment to providing the most defensible certification examination program is built upon the concept of transparency. Whenever feasible, within the confines of examination security, information about the development process is openly shared with massage therapy schools, instructors, RMT candidates and the public.

Examination content security is paramount, involving careful screening of examination staff, and standard setting meeting participants. Examination administration security policies, including examination candidate requirements, item bank protection and access are strictly monitored and enforced. Statistical analysis of item performance is routinely conducted to identify possible examination security breaches.

**How is the approved references list generated?**

The College’s approved references list is generated by taking a survey of all of the Ontario massage therapy schools every 3-5 years. This survey asks each school to identify the references that they are using to teach their massage therapy curriculum. This information is then used to generate an updated approved references list. In this fashion, it is the schools that are responsible for the content of the CMTO approved references list, as they are best positioned to evaluate and select their own teaching instruments.

It is important to note that schools are not required to teach from materials on the approved references list. Schools are free to select their own suitable references to support the delivery of their curriculum. Essentially, the development and publication of the approved references list is designed to create transparency with respect to the references that are being used to support the test development process.

References that multiple schools (i.e. more than two) are using are automatically considered for inclusion on the approved references list. The Content Specialist and the Subject Matter Experts (SMEs) consider the following when reviewing a reference:

1. The quality of the referencing;
2. Its evidence base;
3. The expertise of the author(s) and;
4. Whether the reference covers content that is not well covered by another reference.

If more than one edition of a book is being used by the schools then only the most recent edition is included on the approved references list. If only one school is using a given reference then the Content Specialist and SMEs will review the reference utilizing the criteria listed above to make a decision regarding its suitability. An exception to this rule may occur if a specific exam content area is identified that does not have a source on the approved references list to develop new content. In this case, the content specialist and SMEs may initiate a search for an appropriate reference, find one that best suits our needs, and immediately place the reference on the approved references list. The schools and registered examination candidates are notified of any such addition.

What are the steps that the College follows to develop defensible certification examinations?

The six main steps involved in certification examination development are:

1. Job Analysis Study/Content Outline Development
2. Item Writing
3. Item Review
4. Test Development
5. Standard Setting/Score Scaling/Scoring and Reporting

Step 1 – Job Analysis Study/Content Outline Development

The foundation for defensible certification examinations is the Job Analysis Study (JAS). The JAS provides research-based support for the relevance, validity, and legal defensibility of the certification examinations by establishing the link between job performance of critical tasks, and successful performance on the certification examinations. Adhering to international testing standards, the College has adopted a JAS model to establish a means of identifying the shifts and changes that are natural to our profession and to provide a means of translating those shifts to the OSCE and MCQ examination content outlines. The College has adopted an ongoing-cyclical model, which allows the review of JAS data and update of the content outlines every 12-24 months. In this way the College is able to identify very subtle shifts and changes to the role which translate into very minor changes to the examination content outlines. Keeping with its policy of transparency, this information is shared with schools, candidates and the public as part of this ongoing cycle.

The first phase of the JAS includes conducting a comprehensive review of the professional role of Massage therapists. Periodicals, provincial and federal rules and regulations, approved texts,
College policy and bylaws, job descriptions, interviews of professionals and the College Core Competency Standards are used to develop an exhaustive list of the tasks required of competent practice. Next, a group of RMTs in good standing that characterize the profession is assembled (e.g. representing the diversity of practice, education, experience, training, age, gender, ethnicity, and geographic practice settings) to evaluate which tasks are appropriate for inclusion on the task list and identify any that may have been left off.

After the task list is generated, it is placed into an on-line survey format. Each month, a group of randomly selected RMTs are invited to participate in this job analysis survey. Respondents are asked to rate each task according to its importance to the role of massage therapy, and the frequency of the task performance. In this fashion, a mechanism is established for massage therapists to identify the knowledge, skills and abilities that they require to safely and effectively perform their job in the province of Ontario. Demographic and practice related information is collected to provide a means of evaluating the representativeness of the respondent sample, and enable a further sub-group analysis of the data to ensure that no demographic group has undue influence on the results.

Once a sufficient sample size of completed surveys is accumulated the data is tabulated and analyzed. Typically, this occurs every 12-24 months. A representative group of RMTS in good standing is assembled to review the data and they utilize decision criteria to determine which tasks will be included in the certification examination content outlines (an example of which is found at the back of the candidate handbook). At this time, adjustments are made to both the task listing and survey to reflect shifts in practice and respondent comments, as required. Decisions are also made regarding whether tasks should be examined in the written (MCQ) examination, the practical (OSCE) examination, or both. Weights are then assigned to the three main sub-sections of the content outline (i.e. health science, professional development, and clinical science) to reflect the percentage of items or tasks that will be devoted to these areas on each examination.

In short, candidates are only examined on material that their fellow massage therapist peers have identified as essential to everyday practice. In a broad sense, the body of practicing massage therapists in Ontario is responsible for the content of the certification examinations through their participation in the JAS.

**Step 2 – Item Writing**

Once the content outline is delineated from the job analysis study, test items are assigned and written to fully cover this outline. Each test item (i.e. each MCQ question and each individual task in an OSCE case) is “linked” to a specific content area on the content outline. Item banks are constantly assessed to ensure that they contain adequate pools of questions to support content,
item age and exposure requirements, item statistical fit, and content relevance. Items for the MCQ and scenarios on the OSCE are reviewed, updated and retired as needed to keep the banks “fresh” and content-relevant. New OSCE and MCQ items are continuously being written in order to provide greater test development flexibility and avoid over-exposure of items.

Item development is a team effort, utilizing the expertise of the College’s Content Specialist, OSCE and MCQ SMEs, and the College’s psychometric consultants at SMT. Initially, a psychometrician trains item writers on how to properly write new items (e.g. no tricky or confusing formats, no humor or slang, no idioms, elimination of bias, appropriate language level, etc). The goal of the training is to define high functioning item formats, train the item writers on what makes a strong test item, and review various item types and their related cognitive complexities. The Content Specialist and the OSCE and MCQ SMEs are then responsible for writing new examination content that is accurate, clear and unambiguous, properly referenced, psychometrically sound, free from bias, clinically relevant, and linked to a specific content area on the content outline.

Key components of OSCE case writing include creation of candidate instructions (i.e. the stem), standardized client instructions, and examiner marking criteria. Specific and detailed instructions are written to ensure that client performance is standardized and that examiner marking is based on specific and objective “yes” or “no” criteria for each task.

Each test item must be supported by two approved references at minimum, when available. Certain questions (e.g. those related to a specific College Standard of Practice or legislation) may have only one reference. If a topic is covered by more than two references then all are reviewed to ensure that the correct answer is unequivocally supported.

**Step 3 – Item Review**

For the MCQ, each new draft item is first reviewed by a psychometric consultant from SMT. When required, input is provided and items are either edited or deleted. For the OSCE, each new draft item is created from a station template that has been psychometrically reviewed and formalized. These processes are designed to ensure the psychometric conformity of each new item.

The Content Specialist, MCQ and OSCE SMEs, and SMT psychometric staff review the content of each newly written item to ensure that it fits all of the criteria discussed in the section above.

Approved MCQ and OSCE items undergo a thorough translation process. The items are first translated into French by the College’s experienced French translator. Then, the College’s French SME, a French-trained bilingual massage therapist, reviews and ensures that the translation is appropriate within a massage therapy context. Bilingual SMT psychometric staff ensures that the mechanics (e.g. examination instructions and tutorials) of the testing process that are presented in French are standardized and appropriate. The College has developed a French lexicon that is
utilized to ensure consistency of translation across time. This lexicon has been reviewed and approved by all Ontario French massage therapy schools.

**Step 4 – Test Development**

**MCQ Examination**

In 2012 a breach of the MCQ Examination was discovered. To continue the MCQ, it was necessary to change a number of aspects of the MCQ program. The following describes the development and administration of the MCQ program following this breach.

The MCQ examination presently consists of 150 questions. Of these, candidate scores are based upon 125 questions and 25 questions are included on the examination for research purposes. The content outline for the examination was revised for 2013 and all information for candidates published about the examination reflects this revised outline.

The examination is to be administered four times per year. Beginning in March 2013, the administration will be computer-based at a single location in Toronto. For each administration, a new form of the examination will be developed. A statistical method referred to as equating will be used to assure that candidates taking a new form of the examination will be evaluated using the same standard as was in place for prior forms of the examination.

Examination forms are assembled of questions meeting the above criteria that have been approved by a panel of RMTs. Questions that are used to determine a candidate score on each new form will also have been pre-tested. The pre-testing process assures that all questions appearing on each new examination form will not exhibit extremely high or low difficulty or have other aberrant properties.

The minimum passing score for the examination has also been established by, and approved by, a panel of RMTs based upon such considerations as the relevance of each question to competent practice, the difficulty of each question, and other considerations.

**OSCE**

Each year, the Content Specialist and SMT generate a new set of OSCE examinations. For each of the seven stations, six cases are selected for inclusion in the OSCE (i.e. a total of 42 cases). Both new and repeat cases are utilized each year. The examination is reviewed to ensure that it matches content outline requirements. Cases are randomly assigned to scheduled exam days in order to avoid over-exposure. Given the matrix of seven stations, each with six scenarios, there are a total of 279,936 different possible examination combinations. Obtaining forecasts for anticipated graduate numbers and timing from each school enables the creation of an OSCE exam
day schedule (i.e. number of days and timing) that is sensitive to the needs of massage therapy diploma graduates. Computerized score sheets are generated for each case to enable computerized scoring on exam days.

Annual OSCE training is conducted over several days prior to the first OSCE administration. The Content Specialist and OSCE SMEs train examiners and clients on their respective standardized roles utilizing the objective written case material. Examiners and clients are trained together to enable examiner familiarity with the client role and provide the means to conduct and thoroughly review mock demonstrations and examinations for each OSCE case. A psychometric consultant from SMT is in attendance during these training days in order to a) support and direct appropriate training efforts, b) collect required psychometric data on new OSCE cases, and c) complete a formal OSCE training session report.

**Step 5 – Standard Setting/Score Scaling/Scoring and Reporting**

Standard setting refers to methods used to establish legally defensible passing points for examinations. As with other steps in the examination development process, standard setting is a fundamental component in establishing a claim of validity in terms of pass/fail decisions. International testing standards dictate that standard setting must be performed in a standardized, defensible manner that is neither arbitrary nor capricious.

As mentioned earlier, the goal of the certification examinations is to identify candidates possessing an established level of minimum competency for entry-level practice. Minimum competency represents a level of proficiency that enables protection of the public and maintenance of professional standards of practice. It is not appropriate to expect candidates to demonstrate competence in areas to which they might be exposed only after significant experience on the job, or through voluntary continuing education or research. By establishing the criteria required of the minimally competent entry-level practitioner and judging candidate performance against those criteria, we have confidence that all candidates who demonstrate the required minimum competencies will pass and those candidates that fail to demonstrate minimum competency will fail.

For the MCQ, the College utilizes a standard setting approach referred to as the Angoff Method, which is the most widely used and recognized method within the international professional testing industry. For the OSCE, a different approach is used that is referred to as the Bookmark Method.

Both the Angoff and Bookmark Methods rely upon the judgments of RMTs to define minimum competence. The RMTs recruited to participate in the process are in good standing with the College and as a group they characterize the profession in terms of diversity of practice, education, experience, training, age, gender, ethnicity, and geographic practice settings. This
diversity and expertise provides the foundation for the establishment of an appropriate passing standard. Each member of this group brings a unique perspective, be it as a highly experienced practitioner nearing retirement, to one who has recently achieved certification.

The Angoff Method requires that each of the RMTs on the panel to individually define minimum competence in terms of examination performance. That is, if an RMT on the panel believes a given examination question reflects an essential element of minimal competence, that question is given more weight in the determination of the minimum passing score. If another question does not appear to reflect an essential element of minimal competence, that question will receive less, or no, weight in the establishment of the minimum passing score. A psychometric consultant from SMT trains the RMTs in the procedure, leads the meeting and facilitates the process.

A standard setting meeting begins by identifying the attributes of a minimally competent entry-level practitioner. Discussion focuses on the knowledge, skills and abilities that are required to practice safely while serving and protecting the public. The opinions of the RMTs (the weights mentioned above) for each examination question are discussed by the RMT panel and participants are given the opportunity to change their mind in the interest of building consensus. The final opinions of the RMTs are analyzed and a minimum passing score is determined.

The Bookmark Method used for the OSCE follows a different approach. The OSCE provides each candidate with the opportunity to provide a behavior-based performance in each of seven standardized stations. The candidates are observed by two examiners and each examiner determines if the candidate has, or has not, performed each of 10 to 16 behavioral criteria associated with that station.

For the purposes of standard setting, the behavioral criteria for the specific performance under consideration are presented to the RMT panel in the order of actual difficulty (if prior candidate data exists) or as rank ordered according to difficulty by the RMT panel (if no prior candidate data exists). The RMTs on the panel are asked to define minimum competence by finding the place in the list that divides the list into two parts; what a minimally competent candidate must perform and what a minimally competent candidate may perform.

It is reasonable to define minimum competence, in part, by expecting a minimally competent candidate to perform the easiest behavioral criteria. It may also be reasonable to define minimum competence, in part, by not expecting a minimally competent candidate to perform the most difficult behavioral criteria. Of the behavioral criteria in the list arranged from easiest to most difficult, the goal is for the RMT panel to reach consensus on the point in the list, working from easiest to most difficult, where the expectation changes from “must be performed” to “may be performed.”
Inherent in the concept of standard setting is the recognition that alternative forms on an examination may differ in difficulty. When this happens, in the interest of fairness, the minimum passing score is adjusted to reflect the difference in difficulty. It would be confusing if two examinations of slightly differing difficulty had different minimum passing scores and candidates were told that they passed or failed in relationship to these different standards. Such a circumstance may lead to one candidate passing one form of an examination with a score of, for example, 68, while another candidate may fail an alternative form of the examination with the higher score of 69.

In response to this potential, a score scale is developed for the purpose of reporting scores. The scale developed by the College goes from 1 to 99 for both the MCQ and the OSCE. For both exams, a fixed passing scaled score of 70 was adopted. A conversion equation converts raw scores into reported or scaled. This conversion equation is developed so as to eliminate the effect of any differences in test difficulty on scores reported to candidates.

**Step 6 - Statistical Analysis, Review and Feedback**

The final and ongoing process involves maintenance of the standards established for the examination program to ensure that examinations are performing appropriately and continuing to serve as a fair and valid measure of what massage therapists do on the job. Notably, various formal periodic reports are completed by SMT in order to officially document the steps and processes that are followed and include the Job Analysis Study report, OSCE and MCQ standard setting reports, technical test reports, and the OSCE training assessment reports.

When a Job Analysis Study reveals shifts or changes in the criteria required of the minimally competent practitioner, there is a corresponding alteration of the examination content outlines and content. Based upon these changes, passing standards are re-established as part of program support and maintenance. Since the College conducts Job Analysis relatively frequently (i.e. every 12 to 24 months versus an industry standard of every 4-5 years), shifts or changes tend to be minor and are easily incorporated into the examination process. While some drift in pass/fail rates is expected, ongoing analysis enables the identification of any sudden changes so that they may be properly evaluated and addressed, if necessary.

For the MCQ, live and pretest item performances are statistically analyzed on an ongoing basis. Items that are too easy or difficult, or are not discriminating among candidates of varying abilities are highlighted for review by the MCQ SME team. Pre-test items that have been sufficiently tested are statistically reviewed to determine if they are appropriate for inclusion into the live examination. Candidate comments that are collected on individual items during test taking are also reviewed to identify potential problem items. Items may be left unchanged, permanently retired, or edited depending on the specific issue(s) identified. Edited items are considered new
items and moved back into the pre-testing stage. Periodically, reviews of entire test banks ensure that the banks represent current practice, referencing and general suitability in a consistent manner. Candidate test site exit surveys are utilized to ensure overall quality control of the examination process.

For the OSCE, custom software has been developed to enable real-time statistical monitoring of essential examination components including examiner inter-rater reliability; examiner harshness; pass rates by individual case, station, and overall; ease or difficulty of individual tasks; and discrimination of individual tasks among candidates of varying abilities. Examiner re-training is provided in cases where inter-rater reliability or rater harshness is outside of acceptable ranges. Daily OSCE staff evaluation forms are also utilized to direct feedback and re-training, when required. Pass rates for individual OSCE cases that differ from other cases at a statistically significant level (i.e. are much easier or harder than the average case) are reviewed to ensure there is no confusion or fault. If deemed appropriate, cases are modified, and/or new cut score studies undertaken. Individual tasks that are very easy, very difficult or fail to discriminate are also reviewed. Decisions may be made to edit, delete or leave a task unchanged, depending on the specific issue(s) identified. Typically, alterations in the examination are made at the end of an OSCE season.

As outlined earlier in this monograph, transparency is a cornerstone of the College’s certification examinations. Candidates who fail the examinations receive detailed information concerning their performance outlining sub-content areas of strength and weakness, designed to help in future study and success. Also, schools are encouraged to direct questions about the examinations to the Content Specialist at any time. Both questions and answers are captured in an ongoing correspondence log that is shared with all of the schools. Each year, schools are provided with individualized reports that outline how their students performed at the MCQ and OSCE certification examinations. Examination content outline areas of strength and weakness are highlighted by comparing individual school performance with the performance of all schools combined. Finally, OSCE and MCQ pass rates and psychometric statistics are published each year in the College newsletter, the TouchPoint.