

**Recommending a Passing Score on the National Assessment Collaboration (NAC)  
Examination: Procedures and Results**

**Ronald K. Hambleton, PhD**

*Distinguished Professor of Education*

**University of Massachusetts Amherst, USA**

**André De Champlain, PhD**

*Acting Director of Research & Development*

**Medical Council of Canada**

**April 9, 2013**

## **Acknowledgements**

### **MCC Staff**

Ms. Ingrid de Vries, *Manager Clinical Assessment Group, Evaluation Bureau*

Ms. Alexa Fotheringham, *NAC Associate Manager, Evaluation Bureau*

Dr. Andrea Gotzmann, *Research Psychometrician, Research and Development*

Mr. Yves Lafortune, *Director, Evaluation Bureau*

Ms. Caroline Marshall, *NAC Production Assistant, Evaluation Bureau*

Ms. Judy McCormick, *NAC Test Development Officer, Evaluation Bureau*

Ms. Lisa Pistilli, *NAC Administration Assistant, Evaluation Bureau*

Mr. Sirius Qin, *Statistical Analyst, Research and Development*

Dr. Marguerite Roy, *Research Analyst, Research and Development*

Ms. Sarah Waddell, *NAC Production Assistant, Evaluation Bureau*

Ms. Natalie Woodstock, *NAC Senior Information Analyst, Evaluation Bureau*

## Table of Contents

PURPOSE OF THE MEETING.....	4
PROCEDURES .....	5
Selecting Panelists .....	5
Description of the “Just Qualified” Candidate.....	6
Two Sub-Panels.....	8
Selecting a Method to Set a Passing Score .....	8
Hofstee Method .....	11
Review of the Two-Day Agenda .....	12
RESULTS.....	14
OSCE and MCQ Findings.....	14
Hofstee Results .....	15
Summary of Evaluation Survey Findings.....	16
Conclusion.....	18
Appendix A.....	19
A Search for Panel Members .....	19
Appendix B.....	24
Description of the “Just Qualified” or “Borderline” Candidate.....	24
Appendix C.....	28
The Two-Day Meeting Agenda and the List of Participants.....	28
Appendix D.....	31
Summary of the Sub-Panel 1 and Sub-Panel 2 Responses to the Post-Meeting Survey .....	31

## PURPOSE OF THE MEETING

### *Background*

On March 27 and 28, 2013, 18 physicians from across Canada met at the Medical Council of Canada (MCC) offices in Ottawa to participate in a standard setting exercise that ultimately led to the recommendation of a passing score for the National Assessment Collaboration (NAC) Examination. The National Assessment Collaboration (NAC) is an alliance of Canadian organizations whose aim is to streamline the evaluation process for international medical graduates (IMGs) seeking a licence to practise medicine in Canada. The NAC examination assesses the readiness of an IMG for entrance into a Canadian residency program. It is a national, standardized examination that tests the knowledge, skills and attitudes at the level of a graduate from a Canadian medical school essential for entrance into postgraduate training in Canada. The NAC examination is composed a series of Objective Structured Clinical Examination (OSCE) stations as well as a MCQ therapeutic (TPx) component. The examination may include, but is not limited to, problems in medicine, pediatrics, obstetrics, gynecology, psychiatry and surgery.

Our standard setting exercise resulted in recommended passing scores for the NAC Examination Committee to consider before approval of results for the March 2013 NAC examination. Panelists were informed that their role was one of *recommending* a passing score rather than *setting* a passing score. Final approval of the recommended passing score is the responsibility of the NAC Examination Committee who is responsible for the creation and maintenance of NAC Examination content and performance. The NAC Examination Committee holds the ultimate authority to approve examination results.

## **PROCEDURES**

The present standard setting exercise was preceded by a review of potential methods and related issues to consider for setting passing scores on exams such as the NAC. Planning of the standard setting exercise as well as the review of materials and documents was conducted by the two exercise facilitators. MCC staff supported the preparation and delivery of the standard setting exercise.

### *Selecting Panelists*

Many features of a standard setting exercise can influence the validity of the recommended passing score as well as its associated process and results. One of these features is the selection of well-qualified panelists. In December 2012, the MCC sent a message via email to individuals and groups around the country, soliciting participation in our standard setting exercise, which resulted in more than 60 physicians being nominated. Each nominee completed a demographic form. The original invitation and demographic form are shown in Appendix A.

On the basis of this demographic information, staff selected 18 participants and assigned them to two panels denoted 1 and 2 in Table 1. While a multitude of background information was collected, we focused the assembly of the two panels using those variables listed in Table 1 based on their importance. Every effort was made to match both panels as closely as possible on the following key descriptors: gender, region of the country, ethnic background, specialty, and number of years in practice.

### *Advanced Mailing*

Prior to the standard setting exercise, panelists received the following documents: (1) a letter describing meeting logistics; (2) an agenda for the meeting; (3) a description of the *unacceptable/poor*, *just qualified/borderline*, and *acceptable/good* candidates; and (4) two papers which provided an overview on standard setting written by one of the exercise facilitators.

### *Description of the Just Qualified Candidate*

Working with the NAC Examination Committee, MCC staff members generated a description of *unacceptable/poor*, *just qualified/ borderline*, and *acceptable/good* candidates. These descriptors are defined in Appendix B. Panelists were asked to adopt these descriptions and identify candidates who they thought reflected a *just qualified* or *borderline* performance. The actual performances for these *just qualified* candidates on the exam was then used for arriving at a passing score on the OSCE and the TPx components of the NAC exam.

Table 1. Demographic variables and associated descriptive statistics by standard setting panel

Variable of interest	Group	Sub-panel 1	Sub-panel 2	Total
<b>Gender</b>	Male	5 (55.5%)	5 (55.5%)	10 (55.5%)
	Female	4 (44.4%)	4 (44.4%)	8 (44.4%)
<b>Geographic Region</b>	West	1 (11.1%)	2 (22.2%)	3 (16.6%)
	Prairies	2 (22.2%)	1 (11.1%)	3 (16.6%)
	Ontario	4 (44.4%)	4 (44.4%)	8 (44.4%)
	Quebec	1 (11.1%)	1 (11.1%)	2 (11.1%)
	Atlantic Prov.	1 (11.1%)	1 (11.1%)	2 (11.1%)
<b>Ethnic Background</b>	Caucasian	6 (66.6%)	6 (66.6%)	12 (66.6%)
	Asian	3 (33.3%)	3 (33.3%)	6 (33.3%)
<b>Specialty</b>	Primary care	5 (55.5%)	4 (44.4%)	9 (50%)
	Other care	4 (44.4%)	5 (55.5%)	9 (50%)
<b>Number of Years in practice post-residency</b>	0 to 10	5 (55.5%)	5 (55.5%)	10 (55.5%)
	11 to 20	4 (44.4%)	4 (44.4%)	8 (44.4%)

### *Two Panels*

Panelists were assigned to one of two panels to assess generalizability of the cut-score across groups, i.e. can we replicate the recommended passing scores across two more or less equivalent panels? Demonstrating this equivalence across two independent panels lends considerable credibility to the ensuing recommendations. When panel recommendations are highly related, they are usually averaged to produce a final recommended cut-score. Conversely, in instances where they diverge, a discussion among panelists can follow to clarify any reasons that might account for this discrepancy (e.g., one facilitator was pushing the sub-panel to adopt higher or perhaps lower results, the two panels simply had very different ideas, or one or two panelists were exerting substantial influence on the ratings of others, etc.). The use of parallel panels in recommending passing scores has become commonplace in recent years due to the added value of these two sets of recommendations.

Replicating a standard setting exercise across panels carries additional appeal for another reason. If a panel is too big, it becomes more difficult for individual panelists to share their views due to competing availability of time and other group dynamic factors. Splitting panelists into groups tempers this concern.

### *Selecting a Method to Set a Passing Score: The NAC OSCE Component*

Several methods have been proposed for setting passing scores on exams such as the NAC exam. Methods that work well with one examination format may be inappropriate with another. The NAC exam is a mixed format assessment that includes both a performance assessment component (OSCE) as well as a traditional multiple-choice question (MCQ) component (TPx). Based on this consideration, we decided to select a tailored method for setting a passing score on each component of the NAC examination.



For the OSCE stations, we chose an examinee-centered method (the Borderline Group method) which is particularly well-suited to the complex, multidimensional nature of performance assessments. The Borderline Group method requires that panelists provide a holistic judgment of each candidate score sheet, assigning each to one of three levels. These three levels (1-3) corresponded to *unacceptable/poor*, *just qualified/borderline*, or *acceptable/good* performance on an OSCE station.

Prior to commencing the standard setting exercise, a thorough training session was conducted, utilizing video as well as live performances on a pilot station from one NAC examination form administered in March 2013. In this training session, two videos of candidate performances were shown to the entire group of panelists; one performance had previously been rated as *good* by the Chair and Vice-Chair of the NAC Examination Committee while the remaining performance was judged to be *poor*. Ample time for discussion of this station and the performance was allotted. Next, two live performances of the same pilot station were conducted *in vivo*, reflecting *very good* and *poor* performances, respectively. Time for discussion was also provided prior to the start of the standard setting exercise.

The two panels were assigned to different rooms and facilitated by different experts. Panelists were familiarized with each station through discussion with MCC test development staff. Test development staff then reviewed the competencies targeted by the case and panelists observed a *very good* candidate's performance on the station presented in a video. Subsequently, the panelists independently reviewed a set of 50 candidate score sheets, ordered from highest to lowest score by station, and assigned a rating from 1 to 3 (again, either *unacceptable/poor*, *just qualified/borderline*, or *acceptable/good*). There was no limit specified on the number of borderline judgments that they could provide. A panelist's passing score on an OSCE station

corresponded to the median candidate score for those they identified as *just qualified/ borderline*.

For each panelist, the recommended passing score on the OSCE component was the median of the pass scores they set across the 10 stations. The median of the panelists' passing scores was then calculated and used as the recommended passing score for the OSCE component. Since panelists were organized into two groups, a recommended passing score was obtained for each panel. These two estimates were then averaged to obtain a recommended OSCE passing score from the total panel.

To illustrate, assume that panelists classified the following score sheets as *just qualified/ borderline*: 3.5, 3.6, 3.7, and 4.0. Note that these values are the actual 1-5 scores assigned by physician examiners to the performances during the actual examination. Computing the median of these score sheets, yields the estimate of the cut-score for the *just qualified or borderline* candidate on that station. This task was replicated across the 10 stations by all 18 panelists to arrive at a recommended passing score. More precisely, the median cut-score value across judges was calculated for each of the 10 stations. The median of these station values (medians) was then computed and treated as the OSCE recommended passing score for this overall component of the NAC examination. It is important to reiterate that throughout the two days, panelists were routinely prompted to keep in mind the *just qualified/ borderline* candidate definition as they were carrying out the task of recommending a passing score.

#### *Selecting a Method to Set a Passing Score: The NAC Therapeutics Component*

For the 22 multiple-choice items on the TPx component, panelists used both the classical Angoff method as well as two established variants of the procedure. In the classical Angoff method, panelists are asked to estimate, on an item-by-item basis, the number of candidates out of 100

*just qualified/borderline* examinees who *would* correctly answer the item. For example, they might estimate that 80 *just qualified* candidates would answer what they perceived to be an easy question, while perhaps 30 *just qualified* candidates would answer a much more difficult item. Within each panel, and for each test item, the mean estimate was computed for each panelist. Effectively, panelists were predicting the expected scores of *just qualified* candidates. An individual panel estimate of the passing score was obtained by averaging the values of the individual panelists. Finally, the two panel estimates of the passing score were averaged to produce an overall therapeutics recommended cut-score.

In the revised Angoff method (round 2), we provided the panelists with the actual difficulty levels of the test items (proportion of correct responses on each item or  $p$ -values) from candidate results on the March 2013 exam administration. This allowed panelists to integrate candidate performances into their final judgments, *if desired*. Additionally, within each panel, panelists discussed their ratings on an item-by-item basis, prior to offering their final estimates. Note that these final ratings were used in determining the recommended passing score on the TPx component of the NAC examination.

#### *Incorporating Political and Other Considerations: The Hofstee Method*

Prior to the conclusion of the standard setting exercise, we asked panelists to answer four specific questions which define the *Hofstee method*. The latter is generally viewed as a procedure which allows judges to gauge the appropriateness of standards in light of a *reality check* which includes both criterion-referenced (acceptable cut-score value) and norm-referenced (acceptable failure rate) considerations. Specifically, panelists were asked to specify the lowest and highest passing scores that they believed were reasonable for both the OSCE and the TPx components of the examination. Additionally, panelists were asked to provide the

lowest and highest failure rates that they felt were tolerable in light of past historical trends and other policy considerations. On the OSCE component, panelists provided acceptable low and high passing score values on the actual score scale, i.e., between 1.0 and 5.0. For the multiple-choice therapeutics items, panelists set a passing score on a percent-correct scale.

Hofstee ratings permit the integration of both criterion-referenced and norm-referenced boundaries to gauge the appropriateness of passing scores derived using both the Borderline Group method (OSCE) and the Angoff family of procedures (TPx). The hope is that the final recommended passing scores would fall within the range of acceptable values considered by members of the panel (i.e. their "gut" estimates).

#### *Review of the Two-Day Agenda*

The two day agenda appears in Appendix C. The meeting began with an introduction of panelists as well as an overview of the primary purpose of the meeting. This was followed by an outline of the NAC examination and its content. The next section of the exercise was devoted to a thorough discussion of the *just qualified/borderline* candidate, framed around the nine competencies measured by the NAC OSCE, as detailed in Appendix B. The training session that provided was extensive and entailed a demonstration of a pilot station, both via videos and *in vivo*. Presentations of both *excellent* and *very poor* performances were provided to panelists.

The meeting then proceeded with a detailed presentation of the following four-step standard setting process: (1) description of the content of each case; (2) a presentation of a sample video; (3) discussion of the case; and finally; (4) the actual provision of the 1-3 standard setting judgments (*unacceptable/poor*, *just qualified/borderline*, or *acceptable/good*) for each of 50 candidate score sheets (more or less uniformly distributed across the score scale from highest to lowest scores by station). Initially, panelists were given about 28 minutes to

complete the rating task for the first few cases. Over the course of reviewing the 10 NAC OSCE stations, these allotments were reduced to about 10 minutes, based on observed pacing. Panelists were always allowed more time if required, and all ratings were provided independently of other panelists. No discussions of ratings took place during this part of the exercise. Ratings were electronically entered into an MCC-designed standard setting database. All OSCE ratings were completed by mid-morning on the second day of the exercise. Thus, we proceeded to training panelists on the Angoff method at this point in the schedule. Following this full-group training session and a short practice exercise, panelists returned to their respective panels and completed the following three-step process: (1) in a first round, they provided their first set of Angoff ratings (prior to the provision of performance data); (2) in a second round, they examined item difficulty statistics, and the range of item level ratings of their panel members with ensuing discussion and; (3) following this discussion, they provided their second and final item judgments.

Once the Angoff ratings were finished, we moved on to training panelists on the Hofstee method. We then collected panelists' Hofstee ratings, as previously described. Both panels reconvened for a brief presentation of their recommended OSCE and TPx passing scores. No feedback was given on the Hofstee ratings because there was no time to perform the necessary computations.

The standard setting exercise ended by asking all panelists to complete an evaluation survey which gauged their impressions of various aspects of the exercise as well as their confidence in the ultimate recommended passing score value for each component of the examination and the total score.

## RESULTS

### *OSCE and TPx (MCQ) Findings*

Cut-scores that were computed for each panel as a function of the NAC examination component are shown in Table 2 along with associated descriptive statistics. The two panels were nearly identical in their recommended passing scores, regardless of the component, as highlighted in Table 2.

Table 2. *Summary of recommended passing scores by panel and NAC examination component*

	OSCE	TPx		NAC Examination				
	Final	Initial	Final	Final				
Panel groups	Median cut-score	Mean cut-score	Mean cut-score	Overall cut-score	Reported cut-score	SD	Min	Max
Panel 1	3.39	50.17	51.06	58.95	66.00	4.29	51.02	63.52
Panel 2	3.31	55.26	53.20	57.28	64.00	3.92	52.81	63.15
<b>Mean of both panels</b>	<b>3.35</b>	<b>52.72</b>	<b>52.13</b>	<b>58.12</b>	<b>65.00</b>	<b>4.00</b>	<b>51.02</b>	<b>63.52</b>

For the OSCE component, the final recommended passing score from panel 1 was equal to 3.39, whereas it was equal to 3.31 for panel 2. The overall cut-score for the OSCE component was equal to 3.35, based on the 1-5 scale.

For the TPx component, the final recommended passing score for panel 1 was 51.06 while it was equal to 53.20 for panel 2. The average or overall cut-score was equal to 52.13 for the TPx component of the NAC examination.

More importantly, on the overall percent-correct scale (obtained by converting the OSCE score to a percent-correct score and then combining the OSCE and TPx scores weighting them, respectively, 90% and 10% as per previous policy decisions), panel 1 recommended a total passing score of 58.95% whereas panel 2 recommended a total passing score of 57.28%. Averaging these two scores, the total group recommended a passing score of 58.12%. When the percent-correct passing score is converted to the final reporting metric of 0 to 100, with a mean of 70.00 and standard deviation of 8.00 on the March 2013 NAC exam test form, the passing score for panel 1 is equal to 66 while it corresponds to 64 for panel 2.

The final recommended cut-score on the NAC reporting scale was equal to 65.

### *Hofstee Results*

With respect to Hofstee OSCE results, the highest acceptable passing score, on average, was judged to be 3.76, and the lowest, 2.87. The recommended passing score of 3.35, obtained with the Borderline Group method, is within this acceptable range.

For the TPx portion of the exam, the highest acceptable passing score on average was judged to be 61.94% whereas the lowest was 41.85%. The recommended passing score value of 52.13%, obtained via Angoff standard setting methods, is again within this acceptable range.

For the OSCE portion of the exam, the highest acceptable failure rate was on average judged to be 41.41% while the lowest was 13.89%. For the multiple-choice portion of the exam, the

highest acceptable failure rate on average judged to be acceptable was 44.63%, whereas the lowest was 12.51%.

### *Summary of Evaluation Survey Findings*

The evaluation survey was divided into sections that largely reflect major activities that occurred over the two-day meeting. Overall findings, of the report, indicate that:

1. ALL panelists were *very clear, clear, or somewhat clear* on the definition of the *just qualified/borderline* candidate. About 78% indicated they were *very clear* or *clear*.
2. Eighty-three percent of the panelists indicated that they benefitted from the discussion of the *just qualified/borderline* candidate early in the meeting. Eighty-three percent of the panelists thought the time spent on the definition was about right; 11% would have been happier with less time, while 6% (n=1) would have liked more time devoted to this activity.
3. Eighty-nine percent of panelists felt that the training session devoted to the OSCE portion of the NAC exam was appropriate. Similarly, 89% indicated that the clarity of scoring procedures for the OSCEs was *excellent* or *very good*. All panelists rated the training session on the scoring component of the NAC exam as *good* or better. Eighty-nine percent of the panelists rated the training of the process for setting the passing score on the OSCEs as *excellent* or *very good*.
4. Panelists were asked what factors they considered in providing their OSCE judgments. All of the factors we considered important were used: the definition of the *just qualified*



candidate as well as their experience and knowledge of the field. Least frequently cited by the panelists was the discussion phase (11%). This was probably due to our decision to limit discussion to the content of the stations and not the ratings themselves.

5. With regard to allotted time, 61% of the panelists judged the time as sufficient for rating the OSCE score sheets; the remaining 39% felt too much time was allowed. This is an important point to integrate for future NAC standard setting exercises. No panelist noted feeling “rushed” to complete their ratings.
6. In regard to the training of the Angoff method for rating the multiple-choice items, 83% of the panelists felt the time devoted was *appropriate* whereas 17% felt it was *somewhat appropriate*. With respect to training on the Angoff method, 94% of the ratings ranged from *good* to *excellent*. As for time allocated to training, 94% of the panelists felt it was *about right*.
7. Again, as was previously the case with the OSCE ratings, we were interested in the factors panelists took into consideration when offering their judgments on the therapeutics component of the NAC examination. Panelists indicated that they made use of everything we had provided: the definition of the *just qualified* candidate (100%), perception of item difficulty (100%), item statistics (61%), other panelists (81%), experience (83%), knowledge and skills measured by the questions (83%), and quality of the distractors (83%). The number of distractors was not reported as a major factor in the judgments of most panelists. Only 22% of panelists indicated they did consider the number of distractors in their judgments.

8. Eighty-three percent of panelists considered the time allocated to judgments was *sufficient* while another 11% felt that they had *too much* time.
9. Similarly, 83% of the panelists were *very comfortable* with the individual panel discussions while the remaining panelists reported being *comfortable* participating in the discussions.
10. On the important question of comfort with the resulting passing score on the multiple-choice questions, 95% were either *confident* (56%), or *very confident* (39%).

Finally, with respect to the most important question, i.e., "What level of confidence do you have in the final recommended passing score?", 88% of the panelists indicated they were *very confident* (44%), or *confident* (44%). One panelist indicated *somewhat confident* while only a single panelist indicated that he/she was *not at all confident*. A breakdown of results for each individual panel is provided in Appendix D.

### *Conclusion*

Overall the recommended passing scores from each panel were very similar, for both the OSCE and therapeutics components of the NAC examination. The two panels, who worked independently of each other in providing their ratings, provided nearly identical recommended standards. At the NAC reported score level, the two panels differed by 1 score point, a difference that is well within the range of sampling error expected in any study aimed at recommending passing scores. These results, taking in the aggregate, provide strong evidence of reliability as well as validity for the resulting recommended standard.

**Appendix A**

**A Search for Panel Members**

The letter was on MCC Letterhead

December 6, 2012

Dear Prospective Panelist:

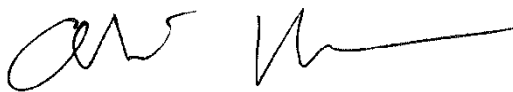
In an effort to set the performance standard for the National Assessment Collaboration (NAC) examination, the governing bodies of the NAC have decided to launch a standard setting exercise. To begin this process, the Research and Development directorate at the Medical Council of Canada (MCC) is soliciting participation for a panel to recommend passing scores. It is expected that the final passing score will be used for the examination starting with the spring 2013 administration.

This invitation and the attached Demographic Information Sheet are being distributed to solicit a group of physician volunteers from whom we will select the final panel. We are approaching you based on your knowledge and expertise and hope you will consider volunteering to participate on the panel. Panelists will be selected to ensure that the diversity of medical experts and practice contexts across Canada are well represented.

The panel will carry out this task on March 27 & 28, 2013 at the Medical Council of Canada (MCC) office in Ottawa. Panelists will be trained to evaluate examination materials and will be guided through a set of procedures to set the passing score. An honorarium of \$500 per day (2 day meeting) plus reasonable travel expenses to Ottawa will be provided.

Should you be interested in participating we ask that you fill out the enclosed Demographic Information Sheet and return it to MCC by mail or electronically and that you reserve these March 2013 dates in your calendar. Responses are requested by January 4, 2013 and your participation will be confirmed by January 14, 2013.

Thank you very much for your interest and support in this important endeavor.



**André De Champlain, PhD**  
Consulting Chief Research Psychometrician  
Acting Director Research & Development  
Medical Council of Canada

## Demographic Information Sheet

The information requested below is being collected to help the MCC obtain a nationally representative panel to recommend a passing score on the National Assessment Collaboration OSCE. This information will only be used to select the panel members so that we can represent the diversity of physicians across the country. The information will not be linked in any way to the collection of data for setting the passing score. Let us remind you the meeting will take place on March 27-28, 2013, so only complete the form if you are available on those dates and you are interested in attending the meeting.

Please provide you name and contact information, and check a box to each of the questions.

---

Name (please print): \_\_\_\_\_

Preferred contact information (mailing address and email address): \_\_\_\_\_

\_\_\_\_\_  
\_\_\_\_\_

---

**1. Number of years in practice post residency:**

- 1-5 years
- 6-10 years
- 11-20 years
- 21-30 years
- More than 30 years

**2. Number of years' experience supervising residents:**

- 1-5 years
- 6-10 years
- 11-20 years
- 21-30 years
- More than 30 years

**3. Do you have experience supervising International Medical Graduates:**

- Yes   
No

**4. Have you ever been an examiner for an Objective Structured Clinical Examination:**

- Yes   
No

**5. Region of the country in which you live:**

- Alberta   
British Columbia   
Manitoba   
Maritimes   
Ontario   
Quebec   
Saskatchewan   
Territories

**6. First Language:**

- English   
French   
Other (\_\_\_\_\_ )

**7. Sex:**

- Male   
Female

**8. Ethnicity:**

- Asian   
Black   
Causation   
First Nations   
Hispanic

**9. Medical Specialty:**

- Pediatrics
- Internal Medicine
- Psychiatry
- Obstetrics and Gynecology
- Surgery
- Family Medicine
- Other \_\_\_\_\_

**10. Type of community in which you work:**

- Urban
- Rural

**11. Type of care setting:**

- Hospital-based
- Community-based

**Appendix B**

**Description of the *Just Qualified* or *Borderline* Candidate**



**STANDARD SETTING 2013  
GUIDELINE TO RATING DESCRIPTORS**

The criteria below are meant to help support and guide discrimination between “Acceptable” and “Unacceptable” candidate performance.

<b>Competency Measured</b>	<b>Unacceptable / Poor</b>	<b>Just Qualified/ Borderline</b>	<b>Acceptable / Good</b>
<b>History Taking</b>	Incomplete data gathering; superficial questioning; disordered; fails to elicit key points; questions appear irrelevant; uses “scattershot” approach; operates by rote; unable to consolidate information; unable to answer patient’s questions; approach is illogical or inappropriate.	Elicits some (but not most) of the essential elements; gains <i>just</i> enough information to provide adequate care; may be barely able to “get the job done”; some key elements are recognized, others are overlooked; may be unable to consolidate or process information completely or thoroughly; there may be some misunderstanding of the patient problem or presenting complaint; some organizational approach but may be “formulaic” with no flexibility (e.g. lists of questions for patient).	Includes most or all of the essential elements; acquires from the patient, family or other source a progressive, chronologic, medically logical description of pertinent events; questions clearly relevant; obtains additional information to fully explore patient’s problem; acquires information in sufficient breadth and depth to permit clear definition of patient’s problem(s).
<b>Physical Examination</b>	Physical examination technique is poor and/or physical examination is incomplete; procedures demonstrated poor understanding of patient problem; actions may cause patient unnecessary discomfort; insensitive to patient comfort and/or modesty; approach seems illogical or inappropriate; abnormalities not detected or considered normal.	Physical examination technique is marginal or inconsistent; includes some (but not most) of the essential elements; barely able to “get the job done”; may be disorganized; is not harmful to the patient but may not recognize patient’s discomfort or need for modesty; lacks focus and may “cast a net” in hopes of getting the marks.	Physical examination technique is consistently good; includes most or all of the essential elements; elicits physical findings in an efficient logical sequence that documented the presence or absence of abnormalities and supported their theory of the patient’s problem; sensitive to the patient’s comfort and modesty; explains actions to the patient.
<b>Organization &amp; Time Management Skills</b>	Disorganized, disjointed or disordered, jumps back and forth; runs out of time or does not use enough time.	Minimally adequate; some disjointed elements; may use too much or too little time for some elements; somewhat disorganized approach.	Logical flow, naturally progressive, completes tasks in a timely fashion; no wasted time; approach coherent and succinct.

<b>Competency Measured</b>	<b>Unacceptable / Poor</b>	<b>Just Qualified/Borderline</b>	<b>Acceptable / Good</b>
<b>Communication Skills &amp; Language Fluency</b>	Inadequate interpersonal skills; language skills inadequate for good communication; unable to establish rapport or trust with patient; inappropriate or overly medical vocabulary; poor non-verbal skills (e.g., eye contact, gesture, posture, use of silence); provides unclear or confusing information or instructions to patient; unable to answer patient's questions.	Marginal interpersonal skills; language is <i>just</i> good enough to meet the patient's needs; may demonstrate effective verbal and non-verbal communication skills inconsistently or with lack of skill; at times may make inappropriate comments.	Appears comfortable and confident; uses a patient focused approach; shows respect, establishes trust; attentive to patient's needs of comfort, confidentiality, information; provides appropriate, clear information; confirms patient's understanding throughout clinical encounter; encourages questions; demonstrates appropriate non-verbal communications (e.g., eye contact, gesture, posture, use of silence); speaks clearly; language skills do not hinder interaction; provides easily understood instructions, comments and questions; uses appropriate choice of words and expressions for the context; avoids jargon/slang.
<b>Diagnosis</b>	Diagnosis and/or differential diagnosis is inappropriate or incorrect; misses the substance of the case; demonstrates lack of knowledge to address complaint; provides misinformation.	Diagnosis and/or differential diagnosis <i>just</i> adequate; possibly only partially correct or incomplete; may not demonstrate total understanding of the clinical presentation; may be non-specific and vague or general.	Diagnosis/differential diagnosis logical and demonstrates a good understanding of the presentation; discriminates important from unimportant information and reaches a reasonable differential diagnosis or diagnosis.
<b>Data Interpretation</b>	Is incorrect or inadequate; candidate misunderstands or misinterprets results or findings.	Data interpretation only partially correct or incomplete; may have overlooked an essential element on data gathering so are unable to process all data.	Organizes pertinent data in a logical manner and synthesizes the data into an integrated concept that defines the problem; succinct and correct.
<b>Investigation Plan</b>	Is inappropriate; may be detrimental to the patient's care or safety.	Selects mainly appropriate laboratory or diagnostic studies to elucidate or confirm the diagnosis, but some may be unnecessary or inappropriate in this patient's context; may over- or under-use resources; may "cast a net" to encompass many possibilities.	Selects appropriate laboratory or diagnostic studies to elucidate or confirm the diagnosis; considers risks and benefits; demonstrates judicious use of resources.
<b>Management</b>	Is not appropriate to the patient and the clinical case; risk/benefit inappropriate or poorly considered; inefficient.	Selects treatments or approaches which may not be wrong, but may be unnecessary or inappropriate in this patient's context; may over- or under-use resources.	Selects appropriate treatments (monitoring, counselling, medications, follow-up); considers risks of therapy and instructed the patient appropriately; negotiates a mutually acceptable plan of management and treatment.

## General Performance Descriptors

<b>Unacceptable / Poor</b>	<b>Just Qualified/Borderline</b>	<b>Acceptable / Good</b>
<p><i>The candidate performs below a level compatible with entry into postgraduate training. The candidate demonstrates clear deficiencies. Their patient care is inadequate. Their questions and/or procedures demonstrate that they do not understand the problem. The candidate seems uncertain and lacks confidence.</i></p>	<p><i>The candidate performs at a level marginally compatible with entry into postgraduate training. The candidate may have just enough skills and knowledge to meet the patient's needs. The candidate may have some difficulty acquiring, consolidating and processing information, but still may be able to "get the job done." Candidate may be able to commence station but is uncertain and may struggle to proceed to completion. No evidence of reasoning/discrimination when answering questions in the station (e.g., unstructured "lists").</i></p>	<p><i>The candidate performs at a level compatible with entry into postgraduate training. They demonstrate a clear understanding of the problem and the action necessary to meet the patient's needs. They are knowledgeable, caring, interact well with the patient and are organized.</i></p>

**Appendix C**

**The Two-Day Meeting Agenda**



### AGENDA – Wednesday March 27, 2013

AGENDA – Wednesday March 27, 2013		
8:00am	Continental breakfast at MCC/Welcome	
8:20am	Purpose of the meeting/agenda	
8:30am	Overview of NAC examination	
9:10am	Overview of setting a passing score on the NAC Examination	
9:20am	Just qualified candidate discussion	
<b>9:50am</b>	<b>BREAK</b>	
10:00am	Training station 10	
11:00am	Four-step process for setting a passing score	
11:20am	Split into sub-panels (OSCE—setting a passing score)	
<b>12:10pm</b>	<b>LUNCH</b>	
1:10pm	Split into sub-panels (OSCE-setting a passing score) continued	
<b>3:10pm</b>	<b>BREAK</b>	
3:30pm	Split into sub-panels (OSCE-setting a passing score) continued	
4:45-5:00pm	Wrap-up and Overview of the next day	
<b>5:00pm</b>	<b>FINISHED</b>	
<b>6:30pm</b>	<b>Dinner at Courtyard, 21 George Street</b>	

<b>AGENDA – Thursday March 28, 2013</b>		
<b>8:00am</b>	<b>Continental breakfast at MCC (meeting begins at the same time)</b>	
8:15am	Split into sub-panels (OSCE-setting a passing score) continued	
<b>10:30am</b>	<b>BREAK</b>	
10:45am	Administration of the 24 items from the MCQ portion of the exam.	
11:30am	Review scoring key and discussion of the items	
<b>12am</b>	<b>LUNCH</b>	
1:00pm	Housekeeping – Expense forms, taxis	
1:15pm	Training on the Angoff method	
1:45pm	Angoff ratings of the 24 multiple-choice items-both rounds of ratings (in subpanels)	
<b>3:15pm</b>	<b>BREAK</b>	
3:30pm	Hofstee judgments	
4:00pm	Panel feedback on the results and discussion	
4:30pm	Evaluation of the meeting	
<b>5:30pm</b>	<b>FINISHED – Panelists leave for airport or home</b>	

**Appendix D**

**Summary of Total Group, Panel 1 and Panel 2 Responses to the Post-Meeting Survey**

**Summary Report—Total Panel (N=18)**

**1. How clear were you about the description of the *Just Qualified* (or sometimes called *Borderline*) candidate on the EXAM as you began the task of setting a passing score following the training on Wednesday morning? (Select ONE)**

Response	Chart	Percentage	Count
Very clear		61%	11
Clear		17%	3
Somewhat clear		22%	4
Not clear		0%	0
<b>Total Responses</b>			<b>18</b>

**2. Did you feel the discussion of the *Just Qualified* (or sometimes called *Borderline*) candidate on the EXAM was helpful during the training on Wednesday morning? (Select ONE)**

Response	Chart	Percentage	Count
Yes, very helpful		33%	6
Yes, helpful		50%	9
Yes, somewhat helpful		17%	3
Not helpful at all		0%	0
<b>Total Responses</b>			<b>18</b>



**3. How would you judge the length of time spent (about 60 minutes on the agenda) introducing, discussing and editing the definition of the *Just Qualified* or *Borderline* candidate? (Select ONE)**

Response	Chart	Percentage	Count
About right		83%	15
Too little time		6%	1
Too much time		11%	2
<b>Total Responses</b>			<b>18</b>




**4. What is your impression of the training you received for setting a passing score on the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Appropriate		89%	16
Somewhat appropriate		11%	2
Not appropriate		0%	0
<b>Total Responses</b>			<b>18</b>








**5. What is your impression of the clarity of the information that was provided regarding the scoring procedures for the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Excellent		22%	4
Very good		67%	12
Good		11%	2
Fair		0%	0
Poor		0%	0
<b>Total Responses</b>			<b>18</b>

**6. What is your overall evaluation of the training that was provided for setting a passing score on the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Excellent		22%	4
Very good		67%	12
Good		11%	2
Fair		0%	0
Poor		0%	0
<b>Total Responses</b>			<b>18</b>




**7. What factors influenced the ratings you made of just qualified (or borderline) candidate responses on the OSCE portion of the exam? (Select ALL choices that apply)**

Response	Chart	Percentage	Count
The description of the just qualified or borderline candidate		89%	16
My perception of the difficulty of the questions		56%	10
The scoring of the individual questions		56%	10
Other panelists during any discussion		28%	5
My experience in the field		67%	12
Knowledge and skills measured by the case		67%	12
Other (please specify)		11%	2
<b>Total Responses</b>			<b>18</b>





**8. What factors influenced the ratings you made of just qualified (or borderline) candidate responses on the OSCE portion of the exam? (Select ALL choices that apply) (Other (please specify))**

#	Response
1.	overall gestalt of the balance of ratings over all the competencies
2.	experience writing and scoring past exams!



**9. How would you judge the length of time for completing the ratings for each of the stations? (Select ONE)**

Response	Chart	Percentage	Count
About right		61%	11
Too little time		0%	0
Too much time		39%	7
<b>Total Responses</b>			<b>18</b>





**10. What level of confidence do you have that the method produced an appropriate passing score for the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Very confident		33%	6
Confident		56%	10
Somewhat confident		6%	1
Not at all confident		6%	1
<b>Total Responses</b>			<b>18</b>



**11. What is your impression of the training you received for setting a passing score on the multiple-choice questions? (Select ONE)**

Response	Chart	Percentage	Count
Appropriate		83%	15
Somewhat appropriate		17%	3
Not appropriate		0%	0
<b>Total Responses</b>			<b>18</b>









**12. What is your evaluation of the training that was provided for the ratings of the multiple-choice questions? (Select ONE)**

Response	Chart	Percentage	Count
Excellent		33%	6
Very good		39%	7
Good		22%	4
Fair		6%	1
Poor		0%	0
<b>Total Responses</b>			<b>18</b>

**13. How would you judge the length of time provided for training for the ratings of the multiple-choice questions? (Select ONE)**

Response	Chart	Percentage	Count
About right		89%	16
Too little time		0%	0
Too much time		11%	2
<b>Total Responses</b>			<b>18</b>




**14. What factors influenced the ratings you made on each of the multiple-choice questions? (Select ALL choices that apply.)**

Response	Chart	Percentage	Count
Description of the just qualified or borderline candidate		100%	18
My perception of the difficulty of the multiple-choice questions		100%	18
The question statistics		61%	11
Other panelists during the discussion		78%	14
My experience in the field		83%	15
Knowledge and skills measured by the multiple-choice questions		83%	15
The quality of the distractors to the multiple-choice questions		83%	15
The number of answer choices to the test questions		22%	4
Other (please specify)		0%	0
<b>Total Responses</b>			<b>18</b>



**15. What factors influenced the ratings you made on each of the multiple-choice questions? (Select ALL choices that apply.) (Other (please specify))**

# Response

**16. How would you judge the length of time for completing the multiple-choice question ratings? (Select ONE)**

Response	Chart	Percentage	Count
About right		83%	15
Too little time		6%	1
Too much time		11%	2
<b>Total Responses</b>			<b>18</b>

**17. How did you feel about participating in the group discussions conducted with the multiple-choice question ratings? (Select ONE)**

Response	Chart	Percentage	Count
Very comfortable		83%	15
Somewhat comfortable		17%	3
Unsure		0%	0
Somewhat uncomfortable		0%	0
Very uncomfortable		0%	0
<b>Total Responses</b>			<b>18</b>

**18. What level of confidence do you have that the ratings process was appropriate for setting the passing score on the multiple choice questions? (Select one)**

Response	Chart	Percentage	Count
Very confident		39%	7
Confident		56%	10
Somewhat confident		0%	0
Not at all confident		6%	1
<b>Total Responses</b>			<b>18</b>

**19. What level of confidence do you have that the consequential data and final discussion this afternoon helped the panel arrive at a defensible passing score? (Select one)**

Response	Chart	Percentage	Count
Very confident		50%	9
Confident		39%	7
Somewhat confident		6%	1
Not at all confident		6%	1
<b>Total Responses</b>			<b>18</b>





**20. What level of confidence do you have in the final recommended passing score? (Select one)**

Response	Chart	Percentage	Count
Very confident		44%	8
Confident		44%	8
Somewhat confident		6%	1
Not at all confident		6%	1
<b>Total Responses</b>			<b>18</b>




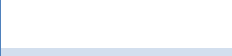
## Summary Report—Panel 1

---

**1. How clear were you about the description of the *Just Qualified* (or sometimes called *Borderline*) candidate on the EXAM as you began the task of setting a passing score following the training on Wednesday morning? (Select ONE)**


Response	Chart	Percentage	Count
Very clear		67%	6
Clear		0%	0
Somewhat clear		33%	3
Not clear		0%	0
<b>Total Responses</b>			<b>9</b>

**2. Did you feel the discussion of the *Just Qualified* (or sometimes called *Borderline*) candidate on the EXAM was helpful during the training on Wednesday morning? (Select ONE)**



Response	Chart	Percentage	Count
Yes, very helpful		33%	3
Yes, helpful		56%	5
Yes, somewhat helpful		11%	1
Not helpful at all		0%	0
<b>Total Responses</b>			<b>9</b>





**3. How would you judge the length of time spent (about 60 minutes on the agenda) introducing, discussing and editing the definition of the *Just Qualified* or *Borderline* candidate? (Select ONE)**

Response	Chart	Percentage	Count
About right		100%	9
Too little time		0%	0
Too much time		0%	0
<b>Total Responses</b>			<b>9</b>

**4. What is your impression of the training you received for setting a passing score on the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Appropriate		89%	8
Somewhat appropriate		11%	1
Not appropriate		0%	0
<b>Total Responses</b>			<b>9</b>

**5. What is your impression of the clarity of the information that was provided regarding the scoring procedures for the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Excellent		33%	3
Very good		67%	6
Good		0%	0
Fair		0%	0
Poor		0%	0
<b>Total Responses</b>			<b>9</b>

**6. What is your overall evaluation of the training that was provided for setting a passing score on the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Excellent		33%	3
Very good		67%	6
Good		0%	0
Fair		0%	0
Poor		0%	0
<b>Total Responses</b>			<b>9</b>

**7. What factors influenced the ratings you made of the *just qualified* (or *borderline*) candidate responses on the OSCE portion of the exam? (Select ALL choices that apply)**

Response	Chart	Percentage	Count
The description of the just qualified or borderline candidate		100%	9
My perception of the difficulty of the questions		44%	4
The scoring of the individual questions		67%	6
Other panelists during any discussion		44%	4
My experience in the field		44%	4
Knowledge and skills measured by the case		67%	6
Other (please specify)		11%	1
<b>Total Responses</b>			<b>9</b>

**8. What factors influenced the ratings you made of the *just qualified* (or *borderline*) candidate responses on the OSCE portion of the exam? (Select ALL choices that apply) (Other (please specify))**

#	Response
1.	experience writing and scoring past exams!



**9. How would you judge the length of time for completing the ratings for each of the stations? (Select ONE)**

Response	Chart	Percentage	Count
About right		78%	7
Too little time		0%	0
Too much time		22%	2
<b>Total Responses</b>			<b>9</b>




**10. What level of confidence do you have that the method produced an appropriate passing score for the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Very confident		22%	2
Confident		67%	6
Somewhat confident		11%	1
Not at all confident		0%	0
<b>Total Responses</b>			<b>9</b>



**11. What is your impression of the training you received for setting a passing score on the multiple-choice questions? (Select ONE)**

Response	Chart	Percentage	Count
Appropriate		89%	8
Somewhat appropriate		11%	1
Not appropriate		0%	0
<b>Total Responses</b>			<b>9</b>

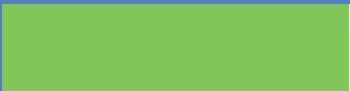







**12. What is your evaluation of the training that was provided for the ratings of the multiple-choice questions? (Select ONE)**

Response	Chart	Percentage	Count
Excellent		56%	5
Very good		22%	2
Good		22%	2
Fair		0%	0
Poor		0%	0
<b>Total Responses</b>			<b>9</b>

**13. How would you judge the length of time provided for training for the ratings of the multiple-choice questions? (Select ONE)**

Response	Chart	Percentage	Count
About right		89%	8
Too little time		0%	0
Too much time		11%	1
<b>Total Responses</b>			<b>9</b>


**14. What factors influenced the ratings you made on each of the multiple-choice questions? (Select ALL choices that apply.)**

Response	Chart	Percentage	Count
Description of the just qualified or borderline candidate		100%	9
My perception of the difficulty of the multiple-choice questions		100%	9
The question statistics		89%	8
Other panelists during the discussion		67%	6
My experience in the field		78%	7
Knowledge and skills measured by the multiple-choice questions		89%	8
The quality of the distractors to the multiple-choice questions		78%	7
The number of answer choices to the test questions		11%	1
Other (please specify)		0%	0
<b>Total Responses</b>			<b>9</b>


**15. What factors influenced the ratings you made on each of the multiple-choice questions? (Select ALL choices that apply.) (Other (please specify))**

#	Response
---	----------


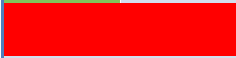
**16. How would you judge the length of time for completing the multiple-choice question ratings? (Select ONE)**

Response	Chart	Percentage	Count
About right		100%	9
Too little time		0%	0
Too much time		0%	0
<b>Total Responses</b>			<b>9</b>

**17. How did you feel about participating in the group discussions conducted with the multiple-choice question ratings? (Select ONE)**

Response	Chart	Percentage	Count
Very comfortable		100%	9
Somewhat comfortable		0%	0
Unsure		0%	0
Somewhat uncomfortable		0%	0
Very uncomfortable		0%	0
<b>Total Responses</b>			<b>9</b>

**18. What level of confidence do you have that the ratings process was appropriate for setting the passing score on the multiple choice questions? (Select one)**

Response	Chart	Percentage	Count
Very confident		33%	3
Confident		67%	6
Somewhat confident		0%	0
Not at all confident		0%	0
<b>Total Responses</b>			<b>9</b>

**19. What level of confidence do you have that the consequential data and final discussion this afternoon helped the panel arrive at a defensible passing score? (Select one)**

Response	Chart	Percentage	Count
Very confident		56%	5
Confident		44%	4
Somewhat confident		0%	0
Not at all confident		0%	0
<b>Total Responses</b>			<b>9</b>




**20. What level of confidence do you have in the final recommended passing score? (Select one)**

Response	Chart	Percentage	Count
Very confident		44%	4
Confident		56%	5
Somewhat confident		0%	0
Not at all confident		0%	0
<b>Total Responses</b>			<b>9</b>




## Summary Report-Panel 2

---




**1. How clear were you about the description of the *Just Qualified* (or sometimes called *Borderline*) candidate on the EXAM as you began the task of setting a passing score following the training on Wednesday morning? (Select ONE)**

Response	Chart	Percentage	Count
Very clear		56%	5
Clear		33%	3
Somewhat clear		11%	1
Not clear		0%	0
<b>Total Responses</b>			<b>9</b>

**2. Did you feel the discussion of the *Just Qualified* (or sometimes called *Borderline*) candidate on the EXAM was helpful during the training on Wednesday morning? (Select ONE)**



Response	Chart	Percentage	Count
Yes, very helpful		33%	3
Yes, helpful		44%	4
Yes, somewhat helpful		22%	2
Not helpful at all		0%	0
<b>Total Responses</b>			<b>9</b>

**3. How would you judge the length of time spent (about 60 minutes on the agenda) introducing, discussing and editing the definition of the *Just Qualified* or *Borderline* candidate? (Select ONE)**




Response	Chart	Percentage	Count
About right		67%	6
Too little time		11%	1
Too much time		22%	2
<b>Total Responses</b>			<b>9</b>






**4. What is your impression of the training you received for setting a passing score on the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Appropriate		89%	8
Somewhat appropriate		11%	1
Not appropriate		0%	0
<b>Total Responses</b>			<b>9</b>








**5. What is your impression of the clarity of the information that was provided regarding the scoring procedures for the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Excellent		11%	1
Very good		67%	6
Good		22%	2
Fair		0%	0
Poor		0%	0
<b>Total Responses</b>			<b>9</b>

**6. What is your overall evaluation of the training that was provided for setting a passing score on the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Excellent		11%	1
Very good		67%	6
Good		22%	2
Fair		0%	0
Poor		0%	0
<b>Total Responses</b>			<b>9</b>



**7. What factors influenced the ratings you made of just qualified (or borderline) candidate responses on the OSCE portion of the exam? (Select ALL choices that apply)**

Response	Chart	Percentage	Count
The description of the just qualified or borderline candidate		78%	7
My perception of the difficulty of the questions		67%	6
The scoring of the individual questions		44%	4
Other panelists during any discussion		11%	1
My experience in the field		89%	8
Knowledge and skills measured by the case		67%	6
Other (please specify)		11%	1
<b>Total Responses</b>			<b>9</b>

**8. What factors influenced the ratings you made of just qualified (or borderline) candidate responses on the OSCE portion of the exam? (Select ALL choices that apply) (Other (please specify))**

#	Response
1.	overall gestalt of the balance of ratings over all the competencies

**9. How would you judge the length of time for completing the ratings for each of the stations? (Select ONE)**

Response	Chart	Percentage	Count
About right		44%	4
Too little time		0%	0
Too much time		56%	5
<b>Total Responses</b>			<b>9</b>

**10. What level of confidence do you have that the method produced an appropriate passing score for the OSCE portion of the exam? (Select ONE)**

Response	Chart	Percentage	Count
Very confident		44%	4
Confident		44%	4
Somewhat confident		0%	0
Not at all confident		11%	1
<b>Total Responses</b>			<b>9</b>



**11. What is your impression of the training you received for setting a passing score on the multiple-choice questions? (Select ONE)**

Response	Chart	Percentage	Count
Appropriate		78%	7
Somewhat appropriate		22%	2
Not appropriate		0%	0
<b>Total Responses</b>			<b>9</b>









**12. What is your evaluation of the training that was provided for the ratings of the multiple-choice questions? (Select ONE)**

Response	Chart	Percentage	Count
Excellent		11%	1
Very good		56%	5
Good		22%	2
Fair		11%	1
Poor		0%	0
<b>Total Responses</b>			<b>9</b>

**13. How would you judge the length of time provided for training for the ratings of the multiple-choice questions? (Select ONE)**

Response	Chart	Percentage	Count
About right		89%	8
Too little time		0%	0
Too much time		11%	1
<b>Total Responses</b>			<b>9</b>




**14. What factors influenced the ratings you made on each of the multiple-choice questions? (Select ALL choices that apply.)**

Response	Chart	Percentage	Count
Description of the just qualified or borderline candidate		100%	9
My perception of the difficulty of the multiple-choice questions		100%	9
The question statistics		33%	3
Other panelists during the discussion		89%	8
My experience in the field		89%	8
Knowledge and skills measured by the multiple-choice questions		78%	7
The quality of the distractors to the multiple-choice questions		89%	8
The number of answer choices to the test questions		33%	3
Other (please specify)		0%	0
<b>Total Responses</b>			<b>9</b>



**15. What factors influenced the ratings you made on each of the multiple-choice questions? (Select ALL choices that apply.) (Other (please specify))**

# Response

**16. How would you judge the length of time for completing the multiple-choice question ratings? (Select ONE)**

Response	Chart	Percentage	Count
About right		67%	6
Too little time		11%	1
Too much time		22%	2
<b>Total Responses</b>			<b>9</b>

**17. How did you feel about participating in the group discussions conducted with the multiple-choice question ratings? (Select ONE)**

Response	Chart	Percentage	Count
Very comfortable		67%	6
Somewhat comfortable		33%	3
Unsure		0%	0
Somewhat uncomfortable		0%	0
Very uncomfortable		0%	0
<b>Total Responses</b>			<b>9</b>

**18. What level of confidence do you have that the ratings process was appropriate for setting the passing score on the multiple choice questions? (Select one)**

Response	Chart	Percentage	Count
Very confident		44%	4
Confident		44%	4
Somewhat confident		0%	0
Not at all confident		11%	1
<b>Total Responses</b>			<b>9</b>

**19. What level of confidence do you have that the consequential data and final discussion this afternoon helped the panel arrive at a defensible passing score? (Select one)**

Response	Chart	Percentage	Count
Very confident		44%	4
Confident		33%	3
Somewhat confident		11%	1
Not at all confident		11%	1
<b>Total Responses</b>			<b>9</b>

**20. What level of confidence do you have in the final recommended passing score? (Select one)**

Response	Chart	Percentage	Count
Very confident		44%	4
Confident		33%	3
Somewhat confident		11%	1
Not at all confident		11%	1
<b>Total Responses</b>			<b>9</b>