National Assessment Collaboration

A GUIDE FOR CONTENT AUTHORS OF THE NAC EXAMINATION

MEDICAL COUNCIL OF CANADA   LE CONSEIL MÉDICAL DU CANADA

19 12

JANUARY 2023

Welcome, NAC content authors

Thank you for agreeing to be a content author of the National Assessment Collaboration (NAC) Examination. The NAC Examination is an Objective Structured Clinical Examination (OSCE). This document will help guide you through the process and is organized as follows:

Part 1: Content authors

Part 2: Policies and other information

Part 3: NAC Examination information

Part 4: NAC Examination Blueprint

Part 5: Psychometrics

PART 1: CONTENT AUTHORS

YOUR ROLES AND RESPONSIBILITIES AS A CONTENT AUTHOR

You will be expected to do the following:
- Attend workshops for which you have signed up
- Develop content as required
- Work with standardized participant (SP) trainers on scripts
- Participate in content reviews
- Play the role of the candidate during run-throughs with the SP trainer as the patient

You may also be asked to do the following:
- Attend ad hoc meetings based on the needs of the examination
- Review new content that was piloted
- Participate in the approval of exam test forms
- Mentor new authors during a workshop

WHAT TO EXPECT DURING A CONTENT WORKSHOP

The workshops are held virtually or in-person roughly once a month over a two- or three-day period. The group consists of eight to 10 authors, two SP trainers and two or three Medical Council of Canada (MCC) staff. Authors focus on writing new content and usually work in teams of two or three. The teams will write a draft and staff, peers and trainers will review it. There will also often be run-throughs with the SP trainers and one of the authors playing the role of the candidate. After the run-throughs, the team completes a thorough review of the SP script and the checklist.

CRITERIA FOR DEVELOPING CONTENT

Consider the following before writing a case:
- Topics must be relevant and must map to the NAC Blueprint needs and examinable objectives
- Content must be plausible and realistic
- Content and the task difficulty level must be appropriate to assess the ability and clinical performance level of the candidates
- Tasks must be accomplishable in the allotted time (11 minutes)

CANDIDATE INSTRUCTIONS

Authors must write candidate instructions to provide information to candidates before they enter a station. The instructions include the clinical stem and task.

SCORING INSTRUMENTS

Once the candidate instructions have been developed, the instruments that will be used to score performance should be developed. Scoring instruments consists of a key feature checklist and rating scales (e.g., history taking, physical examination, communication skills).

KEY FEATURES (KFS)

KFs are essential elements of a clinical problem crucial to its management or to patient safety. KFs can also be understood as clinical decisions that lead to the resolution of a problem and avoid common errors or avoid going down the wrong path in resolving a clinical problem.

When creating the KF checklist, consider the following questions:
- What does an examiner need to observe to know if a candidate has demonstrated that they have met each key feature?
- Given the clinical problem, what must candidates do?
  - Behaviours?
  - Tasks and actions?
  - Decisions?

ORAL QUESTIONS

In any station, a case author may include two or more oral questions that examiners must ask all candidates to answer during the last allotted minutes of the station. Alternatively, the SP may ask oral questions during the encounter.

RATING SCALES

Rating scales evaluate complex skills like communication, history taking, and management. They capture behaviours and the quality of the performance on a wider spectrum.

THE SP STORY

The SP information should be complete yet concise. It should generally be written in plain language and be based on the patient's profile (e.g., educational level). This makes it easier for SPs to use language tailored to the patient and helps ensure that SP trainers and SPs understand the problem from the patient's perspective. Here are some guidelines to follow:
- Give relevant information only
- Include critical and pertinent negatives
- Suggest noncritical information
- Articulate the SP's beliefs and/or agenda
- Ensure findings and affects can be simulated
- Link the SP information provided with the scoring instruments (i.e., provide responses for all checklist items)
- Provide responses to likely candidate actions even if not listed on the scoring instruments

PROPS AND EQUIPMENT

Consider providing reference material for the trainers, SPs and/or examiners. This could include photographs, diagrams, figures or journal articles (referenced as appropriate).

PART 2: POLICIES AND OTHER INFORMATION

TRAVEL POLICY

The MCC encourages travel arrangements be made through our preferred travel agency, which provides full travel services including air travel, hotel accommodations and car rentals.

Centrum Travel
Attn: John Ostris
12-300 EARL GREY DRIVE
OTTAWA ON K2T 1B8

Email: *john@centrumtravel.ca*

Tel.: 613-592-4144 or 1-866-366-8834

If you do travel to MCC for a workshop, we will send you our latest travel policy and preferred hotels for your review.

CODE OF BUSINESS CONDUCT

Participants are required to sign the MCC Code of Business Conduct once a year. You will receive an email invitation to complete the MCC Code of Business Conduct and Security Video course from *notification@LearnUpon.com* a few weeks before your first workshop of the year.

CONTINUING PROFESSIONAL DEVELOPMENT (CPD) CERTIFICATE

As a content author, you will also receive a letter of participation to claim continuing professional development credits. The letters are sent twice a year, in June and December.

ASSESSMENT DEPARTMENT

The Assessment department oversees the development and administration of MCC examinations and provides support to the test committee. The team includes the following:

- Chief Assessment Officer
- Manager of Content Development and Psychometrics
- Medical Education Advisor
- Team Leader, Content Development
- Research Psychometricians
  - Their role is to assemble the examinations according to the Blueprint and test constraints. They also calculate exam-related statistics.
- Analysts
- Assessment Content Developers (ACDs)
  - Their role is to work closely with the content authors to support the creation of examinations content. ACDs attend all content-creation-related activities to ensure that everything runs smoothly for participants.
- Coordinators of Assessment Content
- Their role is to work closely with the ACDs and to offer administrative support to the content authors during a workshop

EXAMINATION OVERSIGHT COMMITTEE (EOC)

The EOC ensures that MCC examinations are appropriate for their intended purposes and administered through validated, standardized, and reliable processes and formats. They approve Blueprint adjustments and examination formats. They are also responsible for approving exam results and reviewing special cases.

PART 3: NAC EXAMINATION

The National Assessment Collaboration (NAC) Examination assesses the readiness of an international medical graduate (IMG) for entrance into a Canadian residency program. It is a pan-Canadian standardized examination that provides residency program directors with objective information on the skills, attitudes and knowledge level of IMGs applying for postgraduate training in Canada.

PURPOSE

The NAC Examination was developed to reduce duplication among provincial IMG assessment programs and offer standardized results to residency program directors across the country. The exam results provide a comprehensive assessment of an IMG's skills, attitudes and knowledge level as compared to the level of a recent Canadian medical graduate and their readiness for entry into postgraduate training. This information could be used with other information, such as the Canadian Resident Matching Service (CaRMS) online application or results from the Medical Council of Canada Qualifying Examination (MCCQE) Part I, to obtain a comprehensive view of an applicant's skill set.

CONTENT DEVELOPMENT

Content for the NAC Examination is developed by physicians from across the country with expertise in multiple medical disciplines. ACDs oversee content development to ensure that it is appropriate for the expected performance of a graduate from a Canadian medical school. The content is created to evaluate candidates' ability to assess, diagnose, manage, and communicate in a range of clinical situations commonly encountered by physicians.

EXAMINATION SCOPE AND FORMAT

The NAC Examination is a one-day clinical skills examination that consists of 12 different 11-minute clinical stations, each with a standardized participant (SP) and a physician examiner (PE) or non−physician examiner (NPE). In total, the duration of the exam is approximately three hours. All candidates rotate through the same series of stations.

At each station, a brief written statement introduces a clinical problem and outlines the candidate's tasks (e.g., take a history, conduct a physical examination).

An OSCE includes a series of stations where candidates are presented with typical clinical scenarios. It includes problems in the following areas:
- Medicine
- Surgery
- Pediatrics
- Obstetrics and Gynecology
- Psychiatry
- Preventive Medicine and Public Health

Approximately 1,700 candidates challenge the NAC Examination annually. The exam is offered in English across the country or in French at designated centres.

The NAC Examination is assembled based on the following blueprint.

DISCIPLINE	Recommended stations, No.	SYSTEM	Recommended stations, No.
Medicine	2–4	Respiratory	≥ 1
Surgery	2–4	Cardiovascular	≥ 1
Psychiatry	1–2	Gastrointestinal	≥ 1
OB/GYN [a]	1–2	Musculoskeletal	
Pediatrics	1–2	Genitourinary	
Geriatric medicine	1–2	Endocrine	2-3
Urgent care	1	Neurologic	
		Mental health	
		Reproductive Health	2-3
		Multisystem	

[a] OB/GYN: Obstetrics and Gynecology

CLINICAL COMPETENCY	Recommended stations, No.	AGE [b]	Recommended stations, No.
History taking	6–7	0–2 mo (newborn)	
Physical examination	1	2–23 mo (infant)	
Combined history and physical examination	2–3	2–5 yr (preschool child)	1-2
Communication skills	≥ 6	6–12 yr (child)	
Diagnosis	≥ 3	13–17 yr (adolescent)	1-2
Data interpretation	≥ 3	18–44 yr (young adult)	
Investigations	≥ 3	45–64 yr (adult)	4-5
Management [c]	≥ 3	≥ 65 yr (older adult)	2-3
		GENDER [d]	
		Of 10 stations, no more than 60% should be male or female	

[b] AGE of actual participant, not necessarily the SP's age
[c] Up to 20% must be therapeutics-specific
[d] GENDER of actual participant, not necessarily the SP's gender

The following defines the clinical competencies:

1. **ASSESSMENT AND DIAGNOSIS**
   covers the following physician activities:

   - **History Taking:** Acquires from the patient, family, or other source a chronologic, medically logical description of pertinent events; gathers information in sufficient breadth and depth to permit a clear definition of the patient's problems

   - **Physical Examination:** Elicits physical findings in an efficient logical sequence that documents the presence or absence of abnormalities and supports a definition of the

patient's problems; sensitive to the patient's comfort and modesty; explains actions to the patient

- **Diagnosis:** Discriminates important from unimportant information and reaches a reasonable differential diagnosis and/or diagnosis

- **Data Interpretation:** Interprets investigative data appropriately in the context of the patient's problems

- **Investigation:** Selects suitable laboratory or diagnostic studies to elucidate or confirm the diagnosis; takes into consideration associated risks and benefits

2. **MANAGEMENT:**

Discusses therapeutic management, including but not limited to pharmacotherapy, adverse effects and patient safety, disease prevention and health promotion when appropriate; selects appropriate treatments (including monitoring, counselling, follow-up); considers risks and benefits of therapy and instructs the patient accordingly.

3. **COMMUNICATION SKILLS:**

Uses a patient-centred approach; establishes trust and respect, and shows sensitivity to the patient's needs; provides clear information; confirms patient's understanding (encourages questions and uses repetition and summarizing to confirm and/or reinforce understanding); respects confidentiality when appropriate; speaks clearly (volume and rate); avoids use of jargon and slang and uses vocabulary appropriate to the patient; demonstrates appropriate nonverbal communication (e.g., eye contact, gesture, posture, and use of silence).

EVALUATION AND SCORING

The result (i.e., pass, fail) is based on whether a candidate's examination score is equal to or higher than the established cut score for the NAC Examination. The standard is based on the level of performance compared to a graduate from a Canadian medical school.

Candidates are assessed by PEs on up to seven different competencies per station. These competencies are as follows:

- History taking
- Diagnosis
- Management
- Communication skills
- Physical examination (assessed in a modified format)
- Investigations
- Data interpretation

Standardized guidelines are used for exam administration, the training of PEs and SPs, and the use of predetermined scoring instruments for the NAC Examination.

Each station is worth the same as every other station. A candidate's score for the entire exam is the average of the station scores.

PART 5: PSYCHOMETRICS

Information in this section may be helpful when participating in post-pilot reviews, flagged items review meetings, or test form approvals.

The MCC Senior Research Psychometrician provides support and education to content development staff and subject matter experts for the NAC Examination, including guidance on psychometrics, scoring, and standard setting. The focus of research is in several areas: test content analysis and blueprinting, test development, scoring, equating, standard setting, test fairness, and longitudinal studies.

The Senior Research Psychometrician gathers various sources of validity evidence to support the interpretation and uses of test scores. One source of evidence relates to statistical measures, including the reliability (i.e., reproducibility) of an examination and measures of the station and item functioning, including difficulty and discrimination.

Stations and items that perform well make it possible to discriminate between higher-ability candidates and lower-ability candidates based on their demonstrated clinical skills for the NAC Examination. A procedure known as a Station and Item Analysis is performed to evaluate the quality of a station and item before it is used in scoring. Stations and items with poor statistics may indicate (but do not dictate) potential content problems (e.g., script or issues with the standardized portrayal, scoring instructions, key feature item, oral question text). Accordingly, these stations and items need to be reviewed by subject matter experts, who decide whether specific items on a station are included in scoring or if future modifications to the stations or items are needed before using them operationally in the future. The test theory that is currently being used for the NAC Examination is classical test theory. The Senior Research Psychometrician uses statistics produced from this theory to analyze the performance of stations and items and scoring related outcomes for the candidates.

STATION AND ITEM ANALYSIS

Several statistical properties are reviewed during the station and item analysis. Two important measures that provide information about how an item is functioning are station/item difficulty (measured using p-value) and the discrimination index (measured using Station Total Correlation (STC) / Item Total Correlations (ITC) and Item Station Correlations (ISC)). In addition, histograms of the station show the spread of station scores that can indicate bimodal or skewed distributions of station scores.

P-VALUE

There are two sets of p-values that are calculated for each OSCE station, station-level p-value, and item-level p-value. For the station-level p-value, it is the average station scores that the candidates achieved on each of the stations. P-values indicate station difficulty and range between 0.00 to 1.00. Station p-values that are low (< 0.30) indicate a difficult station and those that are high (> 0.80) indicate an easy station.

P-value is the average percentage score of an item that the candidates achieved on the item on that station. If the p-value is high (> 0.90), it indicates, on average, candidates did well on this item. If the p-value is low (< 0.10), it indicates, on average, candidates did poorly on this item. Hard items with p-values smaller than 0.10 and easy items with values greater than 0.90 need to be reviewed to ensure the items are correct and fair.

P-values are sample dependent. That is, comparisons of p-values across different samples of candidates do not take into account potential differences in overall candidate ability. Rather, p-values provide a general sense of the range of difficulty of stations or items on a particular test form. Therefore, evaluating both the p-value and discrimination indexes is helpful in determining if there are potential content issues. For example, an item with a low p-value may also be flagged with lower ITCs or ISCs due to low variability of scores. In some cases, there may not be an obvious content-related issue and may be cohort dependent.

STATION AND ITEM DISCRIMINATION

There are two sets of discrimination indexes that are calculated for each OSCE station, station-level STCs, and item-level ITCs and ISCs. STCs, ITCs, and ISCs measure the relationship between candidate station or item scores and the total score. The total score is based on the station or items minus the station or item. That is, the total score is corrected to reduce double counting of the station or item in the calculation of the discrimination index. The relationship can be positive (0.00 to 1.00), negative (between −1.00 and 0.00), or neutral (0.00).

STCs are indicators of discrimination between low- and high-ability candidates for a given station. A low positive or negative STC (< 0.30) indicates that there is a weak or negative relationship between the station score and the overall exam score. This information is useful in flagging stations that should be reviewed by content experts and possibly removed from scoring. A negative STC would indicate that lower-ability candidates are doing well on that particular station as opposed to higher-ability candidates obtaining higher station scores as would be expected. Lower STCs could indicate a content, script, or other scoring-related issue with a particular station or that station is measuring a fairly unique uncommon construct. A moderate to high STC (≥ 0.30) indicates that high-ability candidates are performing well on a given station. Flagged and reviewed stations may still be included on an exam when the content is deemed relevant, important, and has been verified to be correct.

ITCs are indicators of discrimination between low- and high-ability candidates' total score and a given item on a particular station. This discrimination index is calculated between the item scores and the total scores. ISCs are indicators of discrimination between lower-ability and higher-ability station scores and a given item on that station. This discrimination index is calculated between the item scores and the station scores. ITCs and ISCs are flagged in the values between 0.00 and 0.05 which indicate low discrimination, and less than 0.00 as being a poorly discriminating item(s). Stations and items that are flagged with low discrimination should be reviewed before finalizing the scoring of a particular operational exam, after operational or pilot use in preparation for future test forms, and to apply potential adjustments as needed.

ITEM SCORING AND EQUATING

The NAC Examination total score is the average of 10 station scores (a percentage). A candidate's total scores are adjusted to account for test form difficulty differences over time using a Tucker observed-score method (Kolen & Brennan, 2014), also called test score linking. After test score linking, the candidate's linked scores are converted to a score scale between 1300 to 1500, with a pass score of 1374[1] . The pass score was established using a panel-based standard-setting

---

[1] Score scales for the NAC Examination have and will continue to change with exam delivery, content and scoring changes.

exercise, October 2020 for the current exam. Information on the various score scales used by the NAC Examination can be found in *Score Interpretation Guide* on the MCC's website. In addition, various technical reports on standard-setting activities and annual validity, reliability, and psychometric information can be found on the *Technical reports* webpage.

CONTACT US

If you have any questions, please do not hesitate to reach out to us by email.

- **Hélène Martineau**
  Coordinator, Assessment Content
  hmartineau@mcc.ca

- **Marieke Kalkhove**
  Assessment Content Developer
  mkalkhove@mcc.ca

We appreciate your contribution!