



LLQP Stakeholder Meeting

Toronto, June 29, 2016

Agenda

1. Welcome & Agenda Review
2. Exam data
3. Independent review
4. Consultation period and end of transition
5. Licence fee update
6. Next meetings

2. Exam data

- Updated revised formats administered since May 16
 - Ongoing monitoring but no questions cancelled/scores reviewed since implementation
 - Segregated funds is highly reliable, nature of contents support current pass rates
- Additional changes set for mid-July to correct typos and anchor better performing questions
- Additional examples available during the Summer

2.1 Data for three sets of formats

- National modular results (all attempts combined – including modular writes)

Modules	Formats	No. Candidates	No. Revised Scores	Pass rate
Ethics – Common Law	Original	1,187	N.A.	82%
	Revised-1	2,103	N.A.	83%
	Revised-2	1,866	N.A.	83%
Ethics – Civil Code	Original	495	97	68%
	Revised-1	615	N.A.	78%
	Revised-2	532	N.A.	80%
Life Insurance	Original	1,417	222	66%
	Revised-1	2,505	8	71%
	Revised-2	2,208	N.A.	74%
A&S Insurance	Original	1,659	112	76%
	Revised-1	2,643	N.A.	83%
	Revised-2	2,313	N.A.	81%
Seg. funds	Original	1,385	266	64%
	Revised-1	2,500	193	67%
	Revised-2	2,308	N.A.	65%

2.2 Overall data

- National modular results for revised formats only (including compensatory writes)

Modules	Attempt No.	No. Candidates	Pass rate
Ethics – Common Law	1	1,676	84%
	2	170	73%
	3	20	70%
Life Insurance	1	1,825	75%
	2	311	63%
	3	69	77%
	4	3	100%
A&S Insurance	1	2,072	82%
	2	205	70%
	3	36	67%
Seg. funds	1	1,863	67%
	2	367	58%
	3	77	61%
	4	1	100%

2.2 Overall data (cont.)

- First attempt pass rate on all modules since implementation (modular writes only)

Formats	Full Life		A&S	
	No. Candidates	Pass rate	No. Candidates	Pass rate
Original	899	46%	259	63%
Revised-1	1,395	58%	242	82%
Revised-2	1,221	60%	212	75%

2.2 Compensatory data

- Compensatory analysis – January to June
 - 1171 candidates for Full Life
 - 70% pass rate
 - 25% of the passes actually include at least one failed module (2.5% failed two modules)
 - Pass rate would be 53% on a modular basis
 - Module failed most often is Segregated Funds (46% of modular fails), followed by Life Insurance (32%)

2.2 Compensatory data (cont'd)

- Impact of Compensatory Option
 - Modules failed but credit given (236)
 - Modules passed but credit not given (456)

2.3 Results reports

- Second quarterly reports due by end of July
- Pedagogical objective, to help improve LLQP training
- Will cover all candidates (modular and compensatory scoring options)
- Targeted data will remain the same:
 - First attempt pass rates overall
 - Modular pass rates split by attempt
 - Average scores and distribution of scores
 - Pass rates per competency component

3. Independent review

- Selection process
 - Submissions by Stakeholders
 - Analysis of service offerings
- Selected expert: Yardstick – Testing & Training Experts
- Scope of psychometric review:
 - Measurement of LLQP processes against best practices (question structure, exam development, validity controls, etc.)
 - Review of exam questions and exam results
 - Publication of findings and recommendations

Yardstick

Testing and Training Experts

Psychometric Audit of the LLQP Exam

By Natasha Parfyonova, Ph.D.



Psychometric Audit

- **Purpose:**

- To examine evidence of validity, reliability, and fairness of the LLQP Exam to ensure that appropriate conclusions can be made on the basis of exam scores

- **Evaluation Criteria**

- Standards for Educational and Psychological Testing
- The NCCA Standards for the Accreditation of Certification Programs

Psychometric Audit

- **Sources of Information**

- Written documentation on exam design and administration
- Interviews with the exam development team
- Psychometric review of exam questions
- Statistical analyses of exam data

Assessment Life Cycle

- The results of the psychometric audit were organized into six sections, which correspond to the stages of Yardstick Assessment Life Cycle.



Overview of Results

- CISRO uses sound psychometric procedures to develop items for the LLQP Exam.
- Most of the recommendations for improvement focus on exam administration, scoring and reporting procedures.

<input checked="" type="checkbox"/>	<i>STAGE 1: Exam Purpose, Content, and Specifications</i>
<input checked="" type="checkbox"/>	<i>STAGE 2: Item Development</i>
<input checked="" type="checkbox"/>	<i>STAGE 3: Exam Assembly</i>
<input type="checkbox"/>	<i>STAGE 4: Exam Administration, Scoring, and Reporting</i>
<input checked="" type="checkbox"/>	<i>STAGE 5: Exam and Item Analysis</i>
<input type="checkbox"/>	<i>STAGE 6: Standard Setting</i>



Stage 1: Exam Purpose, Content, and Specifications



- Development and validation of the *Life Insurance Agent Competency Profile*
- Development and validation of the Exam Specifications known as the *Curriculum*

Exam Purpose and Intended Uses of Scores

- ✓ The information on exam purpose and the construct being assessed is provided in different technical documents.

Recommendation

- To comply with the *Standards*, it is important that CISRO provide the information on exam purpose, the intended population of examinees, the construct being assessed, and the intended use(s) of exam scores in one document.

Competency Profile

- ✓ A psychometrically sound procedure was used to develop and validate *the Life Insurance Agent Competency Profile*.
- ✓ The process underlying competency profile development was documented in detail.

□ Recommendation

- CISRO should continue collecting relevant data on the qualifications and demographic characteristics of participants in exam validity studies.

Exam Specifications

- ✓ A solid psychometric procedure was used to design and validate detailed exam specifications.
- ✓ The coverage of exam specifications is incomplete.

□ Recommendation

- CISRO should add to the exam specifications the information on:
 - item type;
 - criteria for exam assembly;
 - exam administration requirements; and
 - general description of the plan for scoring and equating the exam.

Stage 2: Item Development



- Processes involved in the selection and training of item writers and reviewers
- The psychometric review of items against the principles of writing multiple-choice items
- Practices for ensuring exam fairness

Selection and Training of SMEs

- ✓ In general, basic item development processes are followed.
- ✓ Item writers and reviewers are provided with the appropriate training and support to assist them in developing multiple-choice items.
- ✓ Item writers and reviewers provide the information on their qualifications and demographics in the application process.

□ Recommendation

- It is recommended that CISRO continue collecting relevant data on the qualifications and demographic characteristics of item writers and reviewers.

Selection and Training of SMEs

- ✓ In general, basic item development processes are followed.
- ✓ Item writers and reviewers are provided with the appropriate training and support to assist them in developing multiple-choice items.
- ✓ Item writers and reviewers provide the information on their qualifications and demographics in the application process.

Recommendation

- It is recommended that CISRO continue collecting relevant data on the qualifications and demographic characteristics of item writers and reviewers.

Psychometric Review of Items

- ✓ The LLQP Exam items in English and French were written in accordance with psychometric standards.
- ✓ The stems are clear, focused, and concise, and answer options are parallel to one another.

□ Recommendation

- A few minor issues that should be addressed in the items include 1) the use of the personal pronoun “you”; 2) the use of items that require a “yes” or “no” answer; and 3) the inconsistent use of professional titles

Exam Fairness

- ✓ It is the responsibility of an exam developer to remove any content in exam items that may be considered stereotypical, offensive, or unfamiliar to some groups of examinees.
- ✓ The LLQP Exam items are currently reviewed for clarity and simplicity of language.

□ Recommendation

- It is worth considering the inclusion of “other” types of insensitive content in the item drafting and review guidelines (e.g., emotionally provocative content, unequal referrals to men and women).

Stage 3: Exam Assembly



- Processes involved in putting together defensible exam forms
- Equivalence of alternate exam forms
- Statistical analyses of exam and item performance (English items)
- Exam translation and adaptation
- Exam form equating

Exam Assembly

- ✓ It is important to establish that an exam adheres to its specifications.
- ✓ The LLQP Exam is assembled using both competency information from the exam specifications and theoretical indices of item difficulty.
- ✓ The alignment of exam content with the exam specifications is well documented.
- ✓ However, theoretical indices of item difficulty are not reliable indicators of item quality.

□ Recommendation

- It is recommended that CISRO field test new items to obtain statistical indices of item quality for exam assembly.

Exam Form Equivalence

- ✓ Alternate exam forms are considered equivalent when they are built to the same content and statistical specifications.
- ✓ The alternate forms of the LLQP modular exams are similar in content and have similar exam- and item-level statistics.
- ✓ The statistical analyses revealed acceptable levels of exam reliability. Note that short exam tend to have low reliability.

□ Recommendation

- CISRO is advised to monitor statistical performance of alternate exam forms over time to gather evidence of exam form equivalence.

Exam Translation and Adaptation

- ✓ It is important to establish that scores on English and French exams have the same meaning.
- ✓ CISRO uses the professional judgment of several subject matter experts and translators to ensure that English and French exams are equivalent in content.

□ Recommendation

- It is recommended that CISRO conduct an empirical evaluation of English and French exams and document relevant process and outcomes.

Exam Form Equating

- ✓ Anchor items are used to equate alternate forms of the LLQP Exam.
- ✓ An anchor set should reflect the total exam in terms of content and psychometric properties.
- ✓ Anchor sets consisting of four items are not suitable for meaningful evaluation of exam quality.

□ Recommendation

- CISRO is advised to increase the size of the anchor set.
- Alternatively, CISRO may consider setting standards on items to calibrate them in terms of their difficulty level. Item standards (or passmarks) can then be used to put together alternate exam forms with the same passing standard.

Stage 4: Exam Administration, Scoring and Reporting



- Exam administration conditions and policies across the jurisdictions
- Processes involved in scoring exams and reporting results to examinees

Exam Administration

- ✓ The accuracy and usefulness of exam scores depends on the extent to which exam administration, scoring, and reporting procedures are standardized.
- ✓ At present, exam administration conditions and policies vary across the participating jurisdictions.

□ Recommendation

- The recommendation is to standardize exam administration policies and procedures across the jurisdictions as much as possible to ensure consistent (fair) treatment of examinees, e.g.,
 - Criteria for exam eligibility
 - Candidate identification procedures
 - Policy on exam accommodations

Exam Scoring

- ✓ Exam scoring is done by the jurisdictions using the answer key provided by AMF.
- ✓ It is not clear if there are any quality control procedures in place to ensure the integrity of exam scores.

□ Recommendation

- CISRO should synchronize exam scoring and quality assurance processes across the jurisdictions to ensure consistent and accurate scoring.

Exam Reporting

- ✓ There is no consistency in the type of exam information provided to examinees and in the way in which it is provided.
- ✓ Exam reporting varies across the jurisdictions.

□ Recommendation

- To ensure the validity of exam score interpretation, exam reporting should be standardized across the jurisdictions. Exam score reports should include at minimum the information on exam content coverage, exam result (e.g., pass or fail), and the total exam score.

Stage 5: Exam and Item Analysis



- Procedures involved in evaluating the statistical properties of exams and items

Exam and Item Analysis

- ✓ The statistical performance of the LLQP exams is monitored on an on-going basis to ensure exam quality.
- ✓ Exam and item analyses are conducted on large samples using reasonable criteria for evaluation of statistical indices.
- ✓ At present, only exams with pass rates below 70% are scrutinized.

□ Recommendation

- Exam and item performance should be evaluated at regular intervals that are not contingent on the exam pass rate.
- Should the criteria for psychometric evaluation of exams change in the future, they should be documented.

Stage 6: Standard Setting



- Processes used to set the passmark on the LLQP Exam

Standard Setting

- ✓ Standard setting produces an exam passmark that serves as important evidence of validity of exam score interpretation.
- ✓ The LLQP Exam uses a “historical” passmark that may be outdated.
- ✓ There is no evidence to suggest that the passmark was tied to any kind of performance standard.

□ Recommendation

- It is recommended that CISRO conduct a new standard setting study to link the LLQP Exam passmark to the absolute performance standard for entry-level life insurance agents in the industry.
- The process followed and the outcomes of the standard setting must be documented.

4. Feedback period

- 45-day as of publication – until August 8, 2016
- Written submissions required to ensure proper documentation and follow-up
- Feedback requested specifically on:
 - Standard setting opportunity
 - Pilot testing questions
 - Desire for individual meeting

4.1 Transition – selected model

- Modular marking in Quebec (no change to current model)
- Pilot in Alberta
- Compensatory marking option in most other jurisdictions for transition period.

4.1 Length of transition

- Sufficient time to gather statistically significant data
 - 1,000 writes on 3 sets of formats per module
 - Ongoing adjustments by AMF
 - Independent psychometric report
 - 45-day notice period of end of transition following publication of expert report offerings
- Anticipated end of transition: August 8, 2016

5. Licence fee update

- Balance between earnings and expenses support decrease in fee during the course of 2016
 - Amounts and dates to be confirmed in September

6. Next meetings

- Suggested timing
 - Individual meetings in August to review written submissions
 - December: for recap of 1st year of new exams and planning for ongoing stakeholder engagement



Questions?



Thank you!