



PQAP - Rencontre avec les intervenants

Toronto, 29 juin 2016

Ordre du jour

1. Accueil et présentation de l'ordre du jour
2. Statistiques sur les examens
3. Analyse indépendante
4. Période de rétroaction et fin de la période de transition
5. Révision du montant des redevances
6. Prochaines rencontres

2. Statistiques sur les examens

- Les formats révisés et mis à jour sont administrés depuis le 16 mai :
 - Contrôles en place, mais pas de questions annulées ou de résultats révisés depuis l'implantation;
 - Le module de fonds distincts est très fiable, la nature du contenu explique les taux de réussite actuels.
- Changements additionnels prévus pour la mi-juillet pour corriger des coquilles et mieux ancrer les questions les plus performantes.
- Exemples de questions supplémentaires disponibles pendant l'été.

2.1 Données pour trois formats

- Résultats nationaux modulaires (tous essais confondus – incluant les essais compensatoires)

Modules	Formats	Nb. Candidats	Nb. Résultats ajustés	Taux de réussite
Déontologie – Common Law	Original	1,187	S.O.	82%
	Corrigé-1	2,103	S.O.	83%
	Corrigé-2	1,866	S.O.	83%
Déontologie – Code civil	Original	495	97	68%
	Corrigé-1	615	S.O.	78%
	Corrigé-2	532	S.O.	80%
Assurance-vie	Original	1,417	222	66%
	Corrigé-1	2,505	8	71%
	Corrigé-2	2,208	S.O.	74%
Maladie-Accident	Original	1,659	112	76%
	Corrigé-1	2,643	S.O.	83%
	Corrigé-2	2,313	S.O.	81%
Fonds distincts	Original	1,385	266	64%
	Corrigé-1	2,500	193	67%
	Corrigé-2	2,308	S.O.	65%

2.2 Données globales

- Résultats modulaires nationaux pour les formats révisés seulement (incluant les essais compensatoires)

Modules	Essai No.	Nb. Candidats	Taux de réussite
Déontologie – Common Law	1	1,676	84%
	2	170	73%
	3	20	70%
Assurance-vie	1	1,825	75%
	2	311	63%
	3	69	77%
	4	3	100%
Maladie et accident	1	2,072	82%
	2	205	70%
	3	36	67%
Fonds distincts	1	1,863	67%
	2	367	58%
	3	77	61%
	4	1	100%

2.2 Données globales (suite)

- Taux de réussite au premier essai pour l'ensemble des modules depuis l'implantation (essais modulaires seulement)

Formats	Discipline complète		Maladie et accidents	
	Nb. Candidats	Taux de réussite	Nb. Candidats	Taux de réussite
Original	899	46%	259	63%
Corrigé-1	1,395	58%	242	82%
Corrigé-2	1,221	60%	212	75%

2.2 Données compensatoires

- Analyse de l'option compensatoire – Janvier à Juin
 - 1171 candidats pour la discipline complète
 - Taux de réussite de 70%
 - 25% de toutes les réussites comprennent en fait au moins un module échoué (2.5% ont échoué deux modules)
 - Sur une base modulaire, le taux de réussite serait de 53%
 - Le module le plus souvent échoué est Fonds distincts (46% des échecs modulaires), suivi d'Assurance vie (32%).

2.2 Données compensatoires (suite)

- Impact de l'option compensatoire
 - Modules échoués, mais comptabilisés dans la réussite (236)
 - Modules réussis, mais non comptabilisés en raison d'un échec (456)

2.3 Rapports statistiques

- Deuxième rapport semestriel prêt d'ici la fin juillet
- Objectif pédagogique, pour aider les prestataires à améliorer la formation PQAP
- Portera sur tous les postulants (options de correction modulaires et compensatoires)
- Les données ciblées demeureront les mêmes :
 - Taux de réussite global au premier essai;
 - Taux de réussite par module par essai;
 - Moyenne et distribution des scores;
 - Taux de réussite par élément de compétence.

3. Analyse indépendante

- Processus de sélection
 - Soumissions par les parties prenantes
 - Analyse des offres de service
- Expert retenu : Yardstick – Testing & Training Experts
- Portée de l'analyse psychométrique :
 - Comparaison des processus du PQAP par rapport aux meilleures pratiques (structure des questions, développement des examens, contrôle de la validité, etc.)
 - Révision des questions d'examen et des résultats
 - Publications des constats et des recommandations

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.

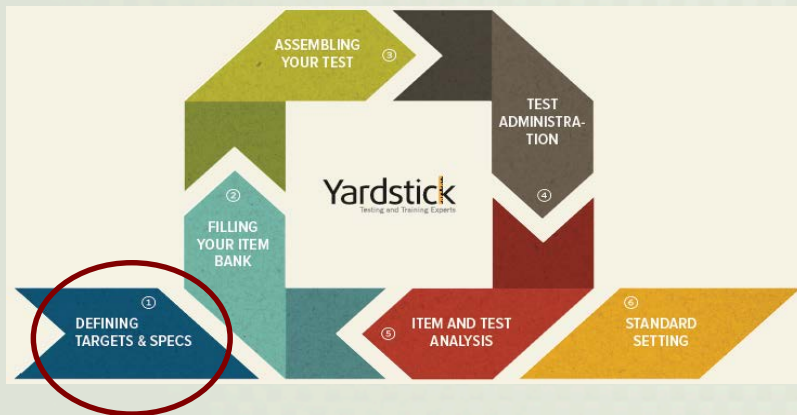


Overview of Results

- 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. Période de rétroaction

- 45 jours à compter de la publication – jusqu’au 8 août 2016
- Soumissions écrites exigées afin d’assurer une documentation et un suivi appropriés
- Rétroaction demandée en particulier pour :
 - Établissement des normes (seuil de réussite)
 - Évaluation préalable des questions (questions “pilotes”)
 - Intérêt pour des rencontres individuelles

4.1 Transition – modèle choisi

- Correction modulaire au Québec (pas de changement au modèle actuel)
- Projet pilote en Alberta
- Correction compensatoire optionnelle dans la plupart des autres juridictions pendant la période de transition

4.1 Durée de la transition

- Temps suffisant pour recueillir une quantité significative de données statistiques
 - 1,000 répondants pour 3 formats par module
 - Ajustements en continu par l'Autorité
 - Rapport d'analyse psychométrique indépendant
 - 45 jours de préavis pour annoncer la fin de la période de transition après la publication du rapport de l'expert
- Fin anticipée de la transition : 8 août 2016

5. Révision des redevances

- L'écart favorable entre les revenus et les dépenses justifie une baisse des frais au cours de 2016 :
 - Les montants et les dates seront confirmés en septembre.

6. Prochaines rencontres

- Calendrier suggéré
 - Rencontres individuelles en août pour revoir les soumissions écrites
 - Décembre : bilan de la première année d'administration des nouveaux examens et planification de l'implication continue des intervenants



Questions?



Merci!