



Contrôle de la validité des examens
Implantation d'un PQAP harmonisé

Août 2014

Contexte

D'ici janvier 2016, le nouveau programme de qualification pancanadien qui remplacera les programmes actuels sera implanté. Il importe de planifier la transition vers le Programme de qualification en assurance de personnes (PQAP) harmonisé afin d'assurer le bon déroulement de sa mise en œuvre.

Le PQAP tire ses caractéristiques des programmes de qualification actuels et l'ampleur des changements qu'il amènera varie d'une juridiction à l'autre. L'adoption d'une approche modulaire reposant sur la consultation de documents et un matériel d'étude normalisé exigera certains ajustements de la part tant des prestataires de cours que des candidats. Il sera par conséquent important de faire en sorte que le nouvel examen évalue adéquatement le niveau de conformité aux normes de qualification révisées.

Les mécanismes de contrôle de la validité de l'examen ci-après visent donc à s'assurer que les candidats bien préparés qui maîtrisent toutes les compétences évaluées peuvent réussir tous les modules.

D'abord et avant tout, la validité du nouvel examen sera contrôlée par un processus d'élaboration des questions d'examen rigoureux. Ce processus garantira que toutes les questions retenues qui constitueront l'examen répondent aux normes d'examen les plus élevées. Le processus d'élaboration a commencé il y a plus de deux ans, à l'occasion de la présentation des ateliers d'analyse de la profession qui ont aidé à définir les aspects que le PQAP harmonisé doit évaluer. En d'autres termes, les mesures prises avant la mise en œuvre de l'examen auront déjà contribué à assurer sa validité.

Malgré cela, le niveau de difficulté réel (par opposition à théorique) de l'examen ne sera confirmé qu'une fois que celui-ci aura été administré. L'enjeu concerne plus précisément les premiers candidats qui seront soumis aux nouveaux examens. Après quelques semaines seulement, les spécialistes de l'éducation auront recueilli suffisamment de données pour analyser les caractéristiques des examens. Il est prévu que les données confirmeront les résultats attendus quant au niveau de difficulté de l'examen. Néanmoins, les données permettront aux spécialistes d'effectuer des ajustements, au besoin.

Les mesures transitoires ci-après, qui seront mises en place à l'occasion de l'administration du premier examen, garantiront que les premiers candidats seront traités équitablement. Elles serviront à contrôler la validité de l'examen et à déterminer le niveau de difficulté des questions d'examen. Cela garantira par ailleurs que les scores du PQAP sont valides dès l'administration du premier examen.

Sans égard à ces mesures transitoires, les qualités statistiques des questions d'examen du PQAP continueront de faire l'objet d'un contrôle permanent tout au long de l'administration de l'examen.

Mécanismes de contrôle transitoires de la validité des examens

Pour que la validité des examens soit assurée, les résultats des candidats seront, dans un premier temps, analysés conformément aux principes énoncés ci-après, jusqu'à ce qu'une quantité suffisante de données statistiques ait été recueillie au sujet de toutes les questions d'examen. Il se peut qu'à la suite de cette analyse, les questions d'examen potentiellement problématiques soient supprimées (et que les résultats soient ajustés) s'il s'avère qu'elles ne permettent pas de mesurer de manière sûre la compétence des candidats.

Les critères suivants permettront de déterminer si des mécanismes de contrôle transitoires doivent être adoptés. Comme nous l'avons déjà mentionné, des mécanismes de contrôle standards de la validité des questions d'examen, qui sont décrits plus loin dans ce document, seront utilisés en permanence.

- Les résultats ne seront ajustés (par l'élimination des questions potentiellement problématiques) que si la note de passage exigée pour un module est inférieure à 70 %.
- Les résultats seront ajustés jusqu'à ce que les conditions suivantes soient satisfaites, et ils seront revus chaque semaine par les spécialistes en mesure et évaluation :
 - Au moins 300 réponses ont été recueillies pour chaque question d'examen.
 - Une note de passage moyenne de 70 % ou plus a été obtenue pendant quatre semaines consécutives pour chaque format d'examen modulaire.

Il sera important de réagir rapidement et de recueillir une quantité suffisante de données pour produire des indices fiables qui permettront de cerner et de corriger les questions problématiques. Par conséquent, les mesures suivantes seront mises en œuvre pour que la note de passage ciblée soit obtenue le plus rapidement possible :

- Le nombre de formats d'examen administrés pour chaque module sera limité jusqu'à ce qu'une quantité suffisante de données statistiques ait été recueillie pour chaque question en vue de l'établissement de formats supplémentaires.
- Des mécanismes de contrôle transitoires peuvent s'appliquer à un seul format d'un module si ce format s'avère plus difficile que les autres.

Les mécanismes de contrôle tant standards que transitoires dépendront des indices de difficulté et de discrimination. Ceux-ci seront utilisés pour déterminer les questions qui devraient faire l'objet d'un contrôle plus étroit ou être carrément supprimées ou remplacées sans tarder.

- L'indice de difficulté est une mesure du nombre de candidats qui répondent correctement à une question. Les pourcentages notés dans le diagramme ci-après indiquent la proportion des candidats qui ont bien répondu à la question.
- L'indice de discrimination renvoie à la capacité d'une question à distinguer les candidats forts et des candidats faibles.
 - Un indice positif indique que les candidats forts ont tendance à répondre correctement à la question, contrairement aux candidats faibles.
 - Un indice nul indique que la question n'établit aucune distinction.
 - Un indice négatif indique que les candidats forts ont tendance à échouer la question, alors que les candidats faibles la réussissent.

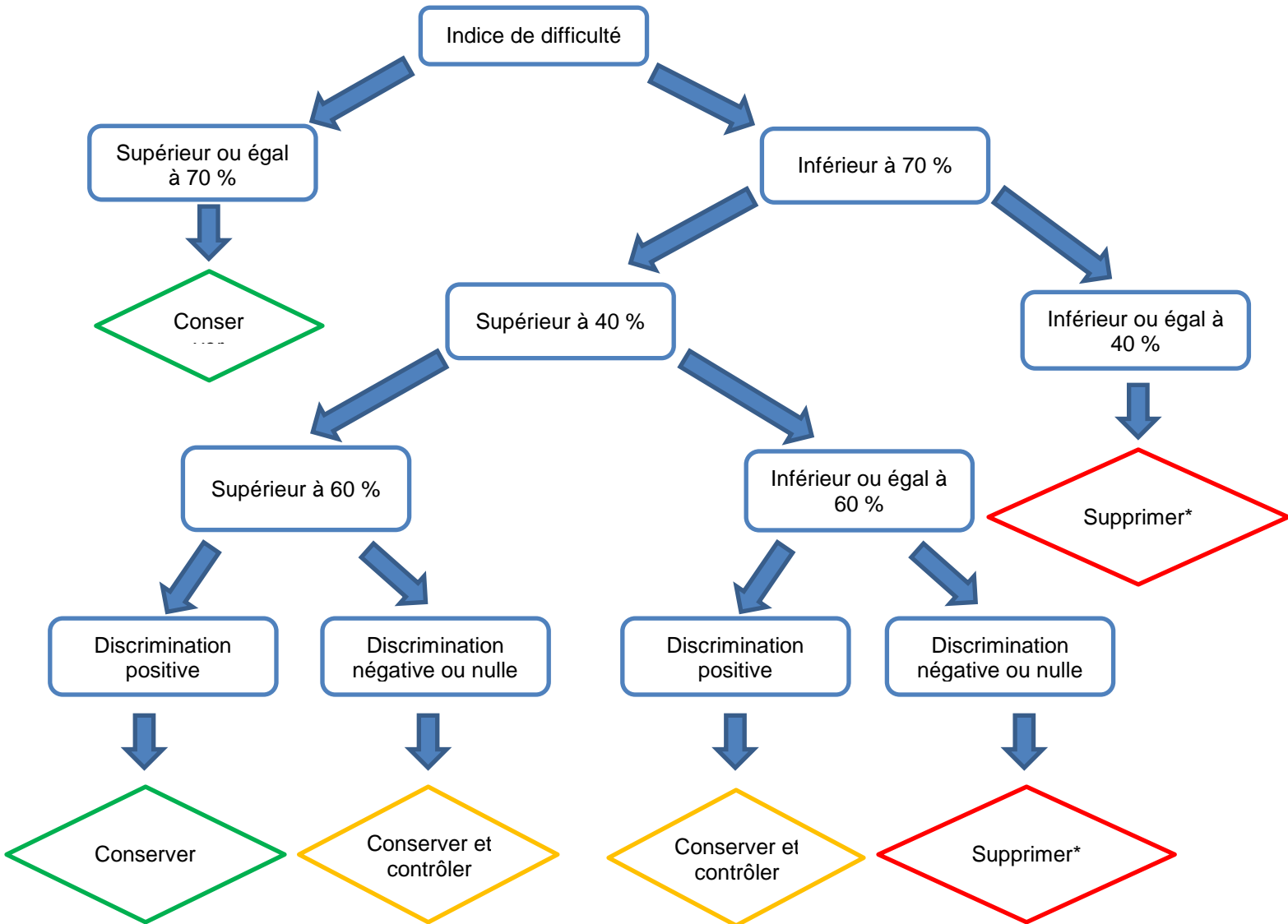
Les points de référence suivants seront utilisés pour déterminer l'état d'une question :

- Toutes les questions auxquelles 70 % des candidats ou plus ont répondu correctement devraient être conservées.
- Toutes les questions auxquelles 40 % des candidats ou moins ont répondu correctement devraient être supprimées ou remplacées.
- Les questions auxquelles entre 40 et 70 % des candidats ont répondu correctement devraient être conservées, contrôlées et potentiellement révisées, supprimées ou remplacées, selon l'indice de discrimination :
 - Les questions dont l'indice de discrimination est positif devraient être conservées.
 - Les questions dont l'indice est nul devraient être contrôlées et potentiellement révisées.
 - Les questions dont l'indice est négatif devraient être supprimées ou remplacées.

Une attention particulière sera accordée aux questions assorties de la mention « Conserver et contrôler » pour cerner les moyens d'améliorer leur qualité.

De plus, une analyse plus poussée peut être effectuée pour assurer la fiabilité des données utilisées pour déterminer s'il convient de supprimer une question. Par exemple, si un fournisseur de cours ou une juridiction est surreprésenté dans les données recueillies, des évaluations comparatives seront réalisées pour s'assurer que les résultats des candidats sont représentatifs de ceux des autres prestataires de cours ou juridictions. Cela évitera la suppression de questions valides à cause de données biaisées.

Indicateurs de la validité des questions d'examen



* Une analyse plus poussée peut être nécessaire pour confirmer qu'une question devrait être supprimée.

L'adoption des mécanismes de contrôle qui précèdent aura pour effet d'exclure des scores des premiers candidats les questions trop difficiles et celles qui ne permettent pas d'établir une distinction entre les candidats bien préparés et ceux qui ne le sont pas.

Autres mesures de contrôle de l'examen du PQAP harmonisé

L'ajustement des résultats d'examen initiaux ne sera pas la seule mesure mise en place par l'équipe chargée de l'élaboration de l'examen pour s'assurer que cet outil de mesure est approprié. Des mécanismes de contrôle standards seront également utilisés.

Comme nous l'avons mentionné précédemment, le processus d'élaboration de l'examen proprement dit est structuré de manière à soutenir la fiabilité et la validité de l'examen. Depuis leur lancement, les ateliers d'analyse de la profession ont réuni un échantillon représentatif de représentants en assurance de personnes dont l'apport a mené à l'établissement du profil de compétences et à la rédaction du curriculum. Ces documents portent respectivement sur les tâches qu'un représentant en assurance de personnes devrait être apte à remplir et sur les compétences en début de carrière qui devraient être évaluées. De plus, les questions seront soigneusement rédigées et validées grâce au concours d'experts de contenus chevronnés œuvrant dans ce secteur d'activité, issus de différentes juridictions. Les questions seront sélectionnées par des spécialistes en mesure et évaluation en fonction des caractéristiques du curriculum. L'ensemble de ces mesures garantira que l'examen est valide et pertinent par rapport à l'exercice de la profession.

Une fois mis en place, l'examen fera l'objet d'un contrôle continu et très étroit, par-delà l'analyse nécessaire pour contrôler les résultats d'examen. Les questions qui présentent des faiblesses mesurables, observées au moyen d'une analyse statistique, seront révisées ou remplacées. De plus, les commentaires et les suggestions des candidats et des parties prenantes donneront lieu à des vérifications et peut-être même à des validations additionnelles qui seront effectuées par des experts en la matière. En revanche, il se peut que ces commentaires et ces suggestions mènent à la révision, voire à la suppression, de questions en dépit de leurs caractéristiques statistiques favorables.

Conclusion

Nous recommandons l'administration directe de l'examen du PQAP harmonisé plutôt qu'une mise à l'essai. Cela ne signifie cependant pas que les qualités de l'examen seront tout à fait inconnues et incontrôlées. Des mécanismes de contrôle tant quantitatifs que qualitatifs seront mis en œuvre.

Dès le début, l'élaboration de l'examen aura été régie par un processus rigoureux. Cela aura permis de valider son construit et son contenu à maintes reprises. Quant aux diverses étapes qui mèneront à son élaboration, elles auront garanti sa validité et sa pertinence pour l'exercice de la profession en début de carrière.

De plus, la pondération des résultats d'examen des premiers candidats atténuera l'impact des éventuelles erreurs de mesure en supprimant les questions statistiquement inappropriées. Cette mesure administrative sera appliquée jusqu'à l'atteinte d'une note de passage ciblée, grâce aux corrections résultant du suivi des formats d'examen. Par contre, cette mesure contribuera à réduire les effets de l'implantation du PQAP harmonisé sur le taux de certification des nouveaux représentants en assurance de personnes. Elle assurera une transition réussie, avantageuse pour toutes les parties prenantes, tout en garantissant que la validité et la fiabilité de l'examen d'accréditation soient rapidement confirmées.

Bibliographie

American Educational Research Association, American Psychological Association and National Council on Measurement in Education (1999). Standards for educational and psychological testing. Washington: American Educational Research Association.

Bertrand, R., et Blais, J. G. (2004). Modèles de mesure : L'apport de la théorie des réponses aux items. Sainte-Foy : PUQ.

Crocker, L. et J. Algina. (1986). Introduction to Classical and Modern Test Theory. California: Wadsworth Group/Thomson Learning.

De Landsheere, G. (1992). Dictionnaire de l'évaluation et de la recherche en éducation : 2^e édition. Paris : Presses Universitaires de France, Coll. « Les grands dictionnaires ».

Kane, M., T. Crooks et A. Cohen (1999). Validating Measures of Performances. Educational Measurement: Issues and Practice, vol. 18, no 2, p. 5-17.

Laurier, M. D. (2005). Les principes de la mesure et de l'évaluation des apprentissages : 3^e édition. Montréal : Gaëtan Morin Éditeur.

Legendre, R. (2005). Dictionnaire actuel de l'éducation : 3^e édition Montréal : Guérin, Éditeur ltée.

Louis, R. (1999). L'évaluation des apprentissages en classe : 1^{re} édition. Laval : Éditions Études Vivantes.

Morissette, D. (1984). La mesure et l'évaluation en enseignement : 1^{re} édition. Sainte-Foy : Les Presses de l'Université Laval.

Moss, P.A. (1995). Themes and Variations in Validity Theory. Educational Measurement: Issues and Practice, vol. 14, no 2, p. 5-13.