

Rapport technique et statistique

Examen des architectes du Canada (ExAC 2024)
Hiver 2025

Sommaire

SOMMAIRE	5
INTRODUCTION.....	6
RAISON D'ÊTRE DE L'ExAC.....	7
HISTORIQUE.....	8
NORMES POUR LES TESTS À ENJEUX ÉLEVÉS	9
Contenu de l'ExAC	10
Caractéristiques des tests de l'ExAC	10
Format de l'ExAC.....	11
CONCEPTION ET ÉLABORATION DES TESTS.....	12
Spécifications des tests	12
Tableau de spécifications.....	13
Élaboration des items	13
Mise à l'essai des items	14
Développeurs d'items	14
Formation des développeurs d'items	15
Le comité d'évaluation du contenu de l'ExAC	15
Révisions effectuées par le comité d'évaluation du contenu	16
Processus de test pratique pour les nouveaux items	16
Assemblage des tests : Sélection des items	16
ÉLABORATION DE L'ExAC.....	18
Cycle du calendrier de l'ExAC.....	18
Le cycle de développement dans un contexte d'amélioration.....	19
ADMINISTRATION DES TESTS.....	20
Manuel d'administration de l'ExAC.....	20
Équité et uniformité des tests.....	20
Soutien aux stagiaires nécessitant des mesures d'accommodements	21
Définition d'une mesure d'accommodement	21
Exemples de mesures d'accommodements disponibles	21

Contrôle de la qualité	22
NOTATION	23
Processus de sélection des copies types	24
La phase de sélection des copies types	24
Aperçu du processus de formation	25
PRÉPARATION DU MATÉRIEL DE FORMATION POUR LA NOTATION	26
Copies types, copies de formation, test de qualification, copies de validation	26
Salles de notation pour les items à réponse construite	26
Formation pour la notation des items à réponse construite	27
Formation quotidienne continue	27
Révision matinale quotidienne des copies types	27
Activité quotidienne d'établissement de la norme	27
Rapports pour le contrôle de la qualité de la notation des items à réponse construite	28
Rapports cumulatifs portant sur le score moyen et la répartition des scores	28
Rapports cumulatifs de fidélité	28
Rapports cumulatifs de productivité	28
Sécurité	29
PROCÉDURES EXCEPTIONNELLES LORS DE LA NOTATION	30
Indications de contenu inapproprié ou d'interférence lors de l'administration de l'ExAC	30
Matériel de test endommagé ou mal imprimés	30
LES ARGUMENTS DE VALIDITÉ FONDÉS SUR LES PROPRIÉTÉS MÉTRIQUES DE L'ExAC	31
Qualité technique de l'ExAC	31
Historical Cronbach's coefficients	32
Estimation de la fidélité sur les items à réponse construite de l'ExAC 2024 spécifiquement	33
Mise en équivalence	34
Établissement de scores seuils	35
Sommaire de l'exercice d'établissement de scores seuils de 2024	36

Résultats	37
Publication des résultats	37
Taux de succès pour l'ExAC 2024	38
Demandes de revisions	38
RÉFÉRENCES.....	39
GLOSSAIRE TERMINOLOGIQUE	40

SOMMAIRE

Ce rapport présente des arguments sur la validité et la fidélité de l'Examen des architectes au Canada (ExAC). La validité et la fidélité sont des considérations importantes pour tous les tests à enjeux élevés. Les normes relatives aux tests éducationnels et psychologiques (AERA, APA, NCME, 2014) stipulent que la fidélité fait référence à la cohérence des scores d'une administration à l'autre. La validité fait référence aux interprétations et aux décisions prises à partir des scores du test, ainsi qu'aux arguments concernant plusieurs facteurs contributifs tels que l'équité, la sécurité, la faisabilité, le contenu, la pertinence, etc.

L'objectif de ce rapport est de se conformer aux normes en fournissant « aux utilisateurs de tests les informations nécessaires pour les aider à évaluer la nature et la qualité du test, les scores obtenus et les interprétations fondées sur ces scores ».

Références

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014). Standards for Educational and Psychological Testing. Washington, DC: Author.

INTRODUCTION

Ce rapport présente des arguments de validité et de fidélité relatives à l'Examen des architectes du Canada (ExAC). La validité et la fidélité sont des considérations importantes pour tous les tests à enjeux élevés. Les normes relatives aux tests éducationnels et psychologiques (AERA, APA, NCME, 2014) stipulent que la fidélité fait référence à la cohérence des scores d'une administration à l'autre. La validité fait référence aux interprétations et aux décisions prises à partir des scores du test, ainsi qu'aux arguments concernant plusieurs facteurs contributifs tels que l'équité, la sécurité, la faisabilité, le contenu, la pertinence, etc.

Le terme « validation » peut être considéré comme l'élaboration d'un argument de validité scientifiquement solide pour appuyer l'interprétation prévue des scores du test et leur pertinence. La validation est la responsabilité conjointe du concepteur du test et de l'utilisateur du test. Kane (2006) ajoute que le concepteur du test doit démontrer la validité des interprétations et des utilisations qui en découlent.

Les sources de données couvrent le contenu des tests et leurs conséquences, ainsi que les modalités de réponse et les interactions avec d'autres tests. Bien que les sources de données soient distinctes, elles ne sont pas considérées comme des types de validité différents. Le processus de validation doit tenir compte du type de test et des données disponibles pertinentes pour ses aspects techniques et pratiques.

L'objectif de ce rapport est de se conformer à la Norme 7 (AERA, APA, NCME, 2014) en fournissant « aux utilisateurs du test les informations nécessaires pour les aider à évaluer la nature et la qualité du test, les scores obtenus et les interprétations basées sur ces scores ».

RAISON D'ÊTRE DE L'ExAC

La Norme 1.1 prescrit que le but des tests soit clairement énoncé. Pour remplir son mandat, le Comité de l'ExAC (CExAC) offre l'Examen des architectes du Canada (ExAC). L'ExAC comprend quatre (4) tests (sections) qui couvrent treize (13) thèmes, trente-huit (38) objectifs généraux et cent cinquante (150) objectifs spécifiques. Les quatre tests, d'une durée maximale de trois heures chacun, sont offerts simultanément partout au Canada sur une période de deux jours. L'ExAC est offert une fois par année dans les deux langues officielles du Canada.

L'ExAC s'adresse aux stagiaires qui souhaitent accéder à la profession d'architecte. Il s'agit d'un test d'admission à la profession reconnu par les ordres professionnels qui supervisent la profession d'architecte dans les provinces et territoires du Canada. La réussite de l'ExAC est l'une des exigences visant à évaluer la compétence des architectes, et l'ExAC constitue l'une des dernières étapes à franchir pour les stagiaires qui souhaitent accéder à la profession.

Les autorités compétentes (Ordres) en matière de délivrance de permis d'exercice de la profession d'architecte dans les provinces et territoires canadiens ont pour mandat de protéger le public en contrôlant l'accès à la profession et son exercice dans leurs juridictions respectives.

HISTORIQUE

Afin d'accorder l'accès aux stagiaires à la profession d'architecte, un test parfaitement adapté au contexte de la pratique de l'architecture au Canada, l'Alberta Association of Architects (AAA), l'Ontario Association of Architects (OAA) et l'Ordre des architectes du Québec (OAQ) ont décidé d'élaborer un nouveau test d'admission dont ils exerceraient le contrôle complet, notamment en ce qui concerne sa préparation, sa mise à jour et son administration. À ce jour, outre l'AAA, l'OAA et l'OAQ, les ordres de réglementation suivants reconnaissent et adhèrent à l'ExAC :

- Architects' Association of Prince Edward Island (AAPEI);
- Association des architectes du Nouveau-Brunswick (AANB);
- Manitoba Association of Architects (MAA);
- Northwest Territories Association of Architects (NWTAA);
- Architects Licensing Board of Newfoundland and Labrador (ALBNL);
- Nova Scotia Association of Architects (NSAA);
- Saskatchewan Association of Architects (SAA);
- Architectural Institute of British Columbia (AIBC).

Voici des liens vers des documents et ressources connexes disponibles dans les deux langues officielles du Canada :

- [Document Cadre de l'ExAC - ExAC](#)
- [Site Web de l'ExAC](#)
- [Conseil Canadien de certification en architecture \(CCCA\)](#)
- [Comité de l'Examen des architectes \(CExAC\)](#)
- [Accommodements de l'ExAC](#)
- [Examen des architectes du Canada \(ExAC\)](#)
- [Programme de stage en architecture \(PSA\)](#)
- [Guide de préparation](#)
- [Regroupement des ordres d'architectes du Canada \(ROAC\)](#)

NORMES POUR LES TESTS À ENJEUX ÉLEVÉS

Validité

Les 'Normes pour le testing éducationnel et psychologique' définissent la validité comme suit :

La validité désigne le degré auquel les preuves et la théorie appuient l'interprétation des résultats des tests pour les utilisations proposées.

Les processus de validation impliquent des arguments et des preuves obtenus par des études métrologiques des processus, du contenu et de l'interprétation des résultats.

Voici quelques concepts courants :

- Sécurité des tests (tricherie, fuite, confidentialité, destruction du matériel de test) ;
- Contrôle qualité (matériel de test, processus, intégrité des données) ;
- Équité : fonctionnement différentiel entre des groupes d'individus créant un avantage pour un groupe en raison de la nature de la tâche, du sexe, de l'origine, des problèmes médicaux, de la langue ;
- Accessibilité et aménagement ;
- Contenu du test : s'assurer que le contenu est pertinent et mesurable (se référer aux éléments obsolètes ou non pertinents par rapport à l'objectif du test) ;
- Formation des évaluateurs et efficacité des algorithmes de notation et des grilles d'évaluation ;
- Accès à des outils, un environnement et des installations appropriés ;
- Efficacité des directives et des instructions ;
- Transparence ;
- Équité.

De nombreux concepts sont traités par des comités. Certains sont mesurés et suivis empiriquement. L'application rigoureuse des protocoles et des processus est essentielle pour garantir la validité des tests. Tout comme pour la fidélité, la stratégie du cycle d'amélioration continue (PDCA/PDSA : Shewhart ; 1939 et Deming ; 1950) est rigoureusement appliquée par le CExAC.

Ce rapport présente des preuves de fidélité et de validité, ainsi que des informations générales.

Contenu de l'ExAC

La Norme 11 indique que les programmes d'examen d'accès à l'exercice de la profession doivent fournir une description du contenu et des preuves le reliant aux exigences de la profession. Le contenu des objectifs généraux et spécifiques de l'ExAC s'appuie sur le *Le Programme de stage en architecture* (PSA). Le PSA a été développé par le Regroupement des ordres d'architectes du Canada (ROAC) avec pour mandat d'offrir un parcours structuré et efficace durant le stage pratique. L'ExAC évalue l'expérience des stagiaires acquise durant le PSA afin de déterminer s'ils satisfont aux normes professionnelles requises pour être architecte au Canada. L'expérience pratique des stagiaires est complétée et appuyée par plusieurs documents clés, notamment :

- Manuel Canadien de pratique de l'architecture ;
- Code national du bâtiment du Canada ;
- Code national de l'énergie pour les bâtiments – Canada.

Ces sources sont généralement reconnues pour leurs connaissances essentielles et leurs applications pratiques en architecture au Canada. Ces documents sont reconnus par le Regroupement des ordres d'architectes du Canada (ROAC).

Caractéristiques des tests de l'ExAC

Les quatre sections de l'ExAC évaluent le degré de réussite des stagiaires par rapport à l'atteinte ou au dépassement du niveau minimal de compétence pour les attentes et domaines d'expérience pratique (13 thèmes). Globalement, ces domaines reflètent les domaines de la profession d'architecte qui peuvent être mesurés : programmation, étude de l'emplacement et de l'environnement, coordination des systèmes d'ingénierie, gestion du coût, Code national du bâtiment, esquisse de projet, conception préliminaire du projet, projet définitif, appels d'offres et négociations de contrat, phase de construction (bureau), phase de construction (chantier), gestion de projet/gestion de bureau de la pratique professionnelle et littérature en matière de conception durable. Pour plus d'information, veuillez vous référer au [Guide de préparation](#).

Format de l'ExAC

Les tests comprennent des activités axées sur la performance qui nécessitent des réponses écrites à des questions à réponse courte ainsi que des questions à choix multiples, permettant aux stagiaires de démontrer leurs connaissances et leurs aptitudes par rapport aux attentes et au contenu de l'expérience pratique mentionnés ci-dessus. Une version de chaque test est préparée pour les stagiaires francophones, et une autre pour les stagiaires anglophones. Les deux versions contiennent le même nombre d'items et impliquent les mêmes types de tâches. Une partie de la banque d'items est initialement élaborée en français, et l'autre partie est d'abord produite en anglais. Chaque groupe d'items est ensuite adapté à l'autre langue. Les propriétés métriques sont calculées indépendamment pour les versions française et anglaise des items. Il ne s'agit pas simplement d'une question de traduction, mais plutôt d'une adaptation, car les tests traduits peuvent contenir des biais favorisant les membres d'un groupe linguistique particulier. L'adaptation permet de calibrer les items indépendamment par rapport aux langues et par section.

Les résultats de test sont fournis pour chaque stagiaire, pour chaque province et territoire, et pour l'ensemble du Canada. Au nom du CExAC, les coadministrateurs publient les résultats nationaux agrégés sur son site web et dans ce rapport. Des rapports détaillés sont remis à chaque ordre professionnel. Chaque stagiaire reçoit un rapport individuel de ses résultats de la part du registraire de son ordre professionnel. En cas d'échec à au moins un des quatre tests, le rapport individuel indique le ou les objectifs d'apprentissage nécessitant une remédiation.

Les membres du CExAC s'appuient sur la contribution d'architectes reconnus de toutes les juridictions participantes pour la plupart des aspects de l'ExAC, notamment l'élaboration des questions et des grilles de notation spécifiques à chaque question, la révision de leur contenu, leur pondération et la notation des réponses des stagiaires aux questions à réponse courte. Le CExAC et les architectes participants sont toujours encadrés par des consultants externes spécialisés dans la préparation et l'administration de tests à grande échelle. Ainsi, les stagiaires qui aspirent à intégrer la profession d'architecte sont évalués par des tests conçus par ces consultants, validés et corrigés par des architectes. Les architectes souhaitant participer aux tests du CExAC peuvent soumettre une demande à leur ordre professionnel. La sélection des architectes participant aux activités du CExAC se fait selon les critères suivants :

- Situation géographique (afin de maintenir un équilibre entre les juridictions et les langues) ;
- Exercice privé et public ;

- Nombre d'années d'expérience (une certaine diversité est souhaitée) ; et
- Expérience professionnelle antérieure au sein du CExAC.

CONCEPTION ET ÉLABORATION DES TESTS

Spécifications des tests

Les tests de l'ExAC reposent sur un ensemble d'objectifs généraux et spécifiques élaborés, révisés et approuvés par un groupe d'architectes expérimentés du Canada. Ces objectifs définissent les domaines et les niveaux de connaissances nécessaires à l'établissement des compétences requises pour exercer la profession d'architecte. Ils sont publiés sur le site web de l'ExAC dans le Guide de préparation. Ils sont ensuite révisés périodiquement par des groupes d'architectes de diverses régions du Canada afin d'en vérifier la pertinence par rapport à la pratique actuelle au Canada.

Le tableau de spécifications est largement basé sur la Norme de compétence pour les architectes. Il a été élaboré avec le soutien de consultants et sert à la préparation des items, afin d'assurer une uniformité dans le nombre et le type d'items d'une année à l'autre. Il définit les concepts à évaluer par chaque test, détermine le type de réponse et présente la répartition des items dans les domaines de contenu ciblés. Les domaines de contenu couverts par les tests de l'ExAC se limitent à ceux qui peuvent être évalués par des tests écrits.

Il est à noter que tous les attentes et contenus d'apprentissage ne peuvent pas être évalués lors de tests à grande échelle. Les attentes et contenus d'apprentissage mesurables sont regroupés par sujet et les items qui s'y rattachent. Tous les attentes et contenus d'apprentissage mesurables d'un groupe ne peuvent pas être évalués lors d'un test individuel, mais, avec le temps, ils le seront tous. Par conséquent, il est important que les stagiaires intègrent tous ces attentes et contenus d'apprentissage dans leur préparation.

Tableau de spécifications

Section 1	Section 3
Environ 96 questions à choix multiples et environ 8 questions à réponse courte	Environ 96 questions à choix multiples et environ 6 questions à réponse courte
Programmation (18 items sur 95)	Projet définitif (92 items sur 107)
Étude de l'emplacement et de l'environnement (7 items sur 95)	
Coordination des systèmes d'ingénierie (37 items sur 95)	
Gestion du coût (19 items sur 95)	
Esquisse de projet (6 items sur 95)	Compétences en conception durable (15 items sur 107)
Conception préliminaire du projet (7 items sur 95)	
Section 2	Section 4
Environ 75 questions à choix multiples	Environ 130 questions à choix multiples
Code national du bâtiment du Canada Code national de l'énergie pour les bâtiments (75 items sur 75)	Appels d'offres et négociations de contrats (31 items sur 130)
	Phase de construction – Bureau (41 items sur 130)
	Phase de construction – Chantier (27 items sur 130)
	Gestion de projet et gestion du bureau (31 items sur 130)

Élaboration des items

Standard 4

Les tests et les programmes de testing doivent être conçus et développés de manière à étayer la validité des interprétations des résultats pour les utilisations. Les développeurs et éditeurs de tests doivent documenter les étapes de conception et de développement afin de fournir des preuves d'équité, de fidélité et de validité pour les utilisations prévues auprès des individus de la population évaluée.

La section suivante présente ces preuves.

Mise à l'essai des items

Une mise à l'essai des quatre tests de l'ExAC a eu lieu au printemps 2008 dans trois provinces (Québec, Ontario et Alberta). Ce processus a permis de déterminer rapidement si les items fonctionnaient comme prévu. Les informations tirées des réponses ont servi à réviser et à affiner les énoncés des items à choix multiples, les items à réponse courte, la sélection des items à choix multiples, les échelles descriptives d'appréciation, et à évaluer la durée des tests. Les résultats de la mise à l'essai ont été transmis aux responsables du contenu afin de les aider à revoir, réviser et corriger les items.

Chaque année, de nouveaux items sont élaborés et validés afin de pouvoir être intégrés aux futurs tests. Des architectes de diverses juridictions assistent le CExAC dans tous les aspects de la préparation des items, notamment :

- Trouver des situations contextuelles et des exemples pertinents conformes aux exigences du tableau de spécifications, qui comprend les objectifs généraux et spécifiques ;
- Dessiner les plans et les croquis, préparer les questions à choix multiples et les questions à réponse courte, définir les tâches et préparer les échelles descriptives d'appréciation spécifiques aux questions à réponse courte ;
- Évaluer la valeur et l'intérêt des questions en les soumettant à d'autres contributeurs;
- Réviser les questions et les échelles descriptives d'appréciation, en accordant une attention particulière au contenu et aux éventuels préjugés envers des sous-groupes d'individus.

Développeurs d'items

La Norme 4.8 indique que les qualifications, les caractéristiques et la formation reçues par l'équipe de développement sont fournies. La Norme 4.7 décrit le processus utilisé pour développer, examiner et tester les items.

Au nom du CExAC, les coadministrateurs recrutent des architectes, que le consultant forme ensuite pour participer à des comités composés de 5 à 14 architectes. Ces comités se réunissent plusieurs fois par an pour élaborer et réviser les items à choix multiples et à réponses courtes utilisés dans les tests.

Tous les items doivent être conformes aux attentes et au contenu décrits dans le tableau de spécifications. Les architectes reçoivent un exemplaire du *Guide pour le*

développement des items de l'ExAC afin de les aider à développer les items à choix multiples et à réponses courtes, ainsi que les échelles descriptives d'appréciation.

Les développeurs des items sont sélectionnés selon les critères suivants :

- Connaissance approfondie et expérience de la pratique de l'architecture au Canada, à divers envergures, types et lieux ;
- Bonne connaissance des principales sources de contenu des tests;
- Expérience avérée de collaboration avec d'autres personnes, de suivre des instructions et d'accepter la rétroaction.

Formation des développeurs d'items

Une formation de groupe, d'une durée d'environ 3 heures, abordant l'importance de développer des items valides et fidèles, conformes aux objectifs spécifiques, est dispensée aux développeurs d'items avant le début des séances de développement. Une attention particulière est portée au choix des verbes, conformément aux niveaux taxonomiques décrits dans la taxonomie de Bloom (Bloom et al., 1956). Un exemplaire du *Guide pour le développement des items de l'ExAC* est remis aux responsables du développement. Ce tableau définit l'objectif et le niveau taxonomique des exigences cognitives à cibler lors du processus de développement des items.

Le comité d'évaluation du contenu de l'ExAC

Le comité d'évaluation du contenu est composé d'environ 7 architectes issus de plusieurs organismes de réglementation, dont des membres du CExAC. Sous la supervision de consultants, les membres se réunissent deux fois par an pour vérifier que les nouvelles questions élaborées évaluent correctement les objectifs et le contenu d'apprentissage décrits dans le tableau de spécifications. Ils fournissent également des conseils sur le contenu spécialisé et l'évaluation de la qualité du matériel proposé pour les tests. Ils formulent également des suggestions concernant l'inclusion, la révision ou l'exclusion de questions. À cette fin, les membres du comité d'évaluation du contenu doivent posséder les compétences susmentionnées et une expertise approfondie de la pratique de l'architecture. Le comité examine également les questions pour déceler tout contenu inapproprié (sexisme, âgisme, etc.) et tout régionalisme (expression ou référence particulière à une région du Canada).

Révisions effectuées par le comité d'évaluation du contenu

À la fin de la séance d'élaboration des items, le consultant et les concepteurs examinent les versions préliminaires des items et des échelles descriptives d'appréciation afin d'y apporter des ajustements avant leur présentation au comité d'évaluation du contenu, composé de membres du CExAC. Des révisions supplémentaires sont apportées une fois que le comité d'évaluation a terminé son évaluation, puis les items sont soumis à un processus de sélection pour la mise à l'essai.

Pour les items à choix multiples, ils évaluent la clarté et l'exhaustivité du texte, la précision de la bonne réponse, la plausibilité des leurres et les règles applicables à l'élaboration des items. Pour les items à réponse courte, ils évaluent la correspondance entre les items et les échelles descriptives d'appréciation et déterminent si les items sollicitent l'éventail de réponses attendu. Le processus permet de vérifier que les items sont appropriés pour l'ensemble des provinces et territoires et que la bonne réponse reste la même, quelle que soit l'origine du stagiaire.

Processus de test pratique pour les nouveaux items

Les items à tester sont intégrés aux tests annuels. Les items testés sont indiscernables des items réels du test, de sorte que les stagiaires ne savent pas à quelle catégorie d'items ils répondent, ce qui évite la démotivation qui peut survenir lorsque de nouveaux items sont testés hors du contexte habituel d'un test. Les scores des items testés ne sont pas nécessairement utilisés pour déterminer les résultats des stagiaires. La mise à l'essai des items fait partie du processus de sélection d'items qui sont de qualité du point de vue métrique et métrologique.

Assemblage des tests : Sélection des items

Les items du test sont composés d'environ 30 % à 40 % d'items qui n'ont pas été utilisés lors de l'ExAC précédent. Il est important de noter que les items correspondent à des groupes de contenu, mais que le contenu d'un groupe n'est pas entièrement mesuré dans un test donné ; néanmoins, tout le contenu mesurable est éventuellement inclus dans un test. Voici des facteurs importants à prendre en compte lors de la sélection des items pour un test de portée nationale :

- Les paramètres de mesure de chaque item démontrent que l'item est équitable et de difficulté comparable à ceux utilisés dans les tests précédents ;
- Les items sont conformes aux spécifications du test, leur contenu est équilibré et il n'y a pas de différence de fonctionnement entre les items ou dans le test ; et

- La préparation des items est guidée par un tableau de spécifications similaire d'une année à l'autre.

Un comité examine les tests assemblés proposés et vérifie si les instructions sont suffisantes et que les items ne fournissent pas les réponses à d'autres items. L'assemblage final de l'ExAC est ensuite envoyé pour publication. Des échantillons peuvent être trouvés dans le [Guide de préparation](#).

ÉLABORATION DE L'ExAC

Cycle du calendrier de l'ExAC

Janvier	Présentation et acceptation des résultats de l'ExAC au CExAC.
Février	Les résultats de l'ExAC sont communiqués aux juridictions, les stagiaires sont informés de leurs résultats et la période de demande de révision est ouverte.
Mars-Avril	L'ExAC de l'année en cours est élaboré à partir des questions existantes provenant de la banque de questions et des nouvelles questions approuvées lors de la validation du printemps précédent par le CExAC. La période de demande de révision est fermée ; les révisions sont terminées et les résultats remis aux stagiaires. L'ExAC à venir est validé par le CExAC (de fin avril à mi-mai).
Mai	Sur la base des résultats de la session de validation, la version finale de l'ExAC est établie et testée sur le terrain par des architectes indépendants à des fins de contrôle qualité. Les nouvelles questions à prendre en compte pour les tests sont rédigées lors de sessions distinctes de développement d'items en français et en anglais (mai-juin).
Juin	Ouverture des inscriptions pour l'administration d'automne de l'ExAC. Correction finale de l'ExAC, y compris la révision linguistique et l'approbation des tests avant la reprographie.
Juillet	Clôture des inscriptions à l'ExAC ; une liste provisoire des inscriptions est établie et mise à jour si nécessaire. Les juridictions régionales transmettent les demandes d'aménagement et aux coadministrateurs nationaux.
Août	Le CExAC examine les demandes d'aménagement et d'exemption et les résultats sont communiqués aux juridictions. Début du recrutement des superviseurs et des surveillants.
Septembre	Confirmation définitive des superviseurs et des surveillants par les juridictions.

	Validation des questions écrites au printemps ; les questions sont soit approuvées pour inclusion dans un test ultérieur, soit rejetées, retournées pour révisions.
Octobre	Envoi des livrets et du matériel de test supplémentaire aux juridictions.
Novembre	Administration de l'ExAC (généralement les premiers lundi et mardi de novembre). Les feuilles de questions à choix multiples sont corrigées par un lecteur optique de marques et les livrets de réponses courtes sont envoyés à l'un des coadministrateurs pour correction manuelle durant la session de décembre.
Décembre	Correction des livrets de réponses courtes (généralement sur quatre jours, le premier week-end de décembre ou aux alentours).

Le cycle de développement dans un contexte d'amélioration continue

Le CExAC utilise l'approche Planifier-Faire-Vérifier-Aajuster (PDCA/PDSA/PFVA) (Shewhart, 1939 ; Deming, 1950), un cadre en quatre étapes pour la mise en œuvre de changements progressifs à des fins d'amélioration continue, basé sur la méthode scientifique « hypothèse-expérimentation-évaluation-ajustement ». Le cycle de développement de l'ExAC intègre plusieurs opportunités de vérifications afin d'identifier les redondances, les vulnérabilités et les failles, et d'améliorer les processus à l'avenir.

ADMINISTRATION DES TESTS

Manuel d'administration de l'ExAC

Le *manuel d'administration de l'ExAC* est rédigé en français et en anglais et est remis à chaque autorité compétente et à chaque site d'administration de l'ExAC. Il décrit en détail les procédures d'administration de l'ExAC.

Les procédures décrites dans le manuel doivent être suivies par les administrateurs et les superviseurs de l'ExAC afin d'assurer la cohérence, la sécurité et l'équité de l'administration des tests pour les stagiaires au Canada. Le manuel comprend :

- Les procédures à suivre (p. ex., les instructions à suivre pendant la phase de testing);
- Les aménagements et dispositions spécifiques autorisés pour les stagiaires ayant soumis une demande d'aménagement ;
- Les instructions aux stagiaires (p. ex., les instructions pour la présentation des quatre tests) ;
- Les responsabilités des personnes impliquées dans l'administration des tests.

Les stagiaires disposent d'un espace calme pour travailler et doivent être surveillés en permanence. Des sanctions sévères seront imposées si un stagiaire tente de sortir des copies de l'ExAC de la salle. Il est permis de fournir aux stagiaires des éclaircissements concernant les instructions avant le début d'un test, mais il n'est pas permis de fournir des éclaircissements concernant les éléments pendant qu'un test est en cours (par exemple une reformulation ou une explication).

Équité et uniformité des tests

Le CExAC a adopté des procédures de contrôle de qualité pour garantir l'équité et l'uniformité de l'organisation des sessions d'administration de l'ExAC, ainsi que la précision de la notation.

Un système de communication instantanée entre les sites d'administration a été mis en place afin de gérer les imprévus, tels que les erreurs d'impression. Le manuel d'administration décrit en détail les obligations des superviseurs et des administrateurs des tests, notamment :

- Les procédures à suivre (par exemple, la préparation du matériel à distribuer aux stagiaires, les directives à suivre pendant la phase d'administration) ;
- Les aménagements et dispositions spécifiques autorisés pour les stagiaires ayant soumis une demande d'aménagement ;
- Les directives à donner aux stagiaires (par exemple, les instructions initiales pour les quatre tests) ;
- Les responsabilités du personnel impliqué dans la phase d'administration.

Le manuel d'administration précise la disposition physique des salles et des tables. Les superviseurs notent la position de chaque stagiaire dans la salle et fournissent au CExAC un rapport d'incident si nécessaire. Le CExAC enquête lorsqu'il le juge approprié.

Soutien aux stagiaires nécessitant des mesures d'accommodements

Le manuel d'administration fournit des informations et des instructions pour aider le CExAC à évaluer les mesures d'accommodement recommandés par les professionnels et indique la documentation à fournir.

Les modifications ou changements apportés au contenu des tests et aux critères de performance ne sont pas autorisés, car ils compromettraient la validité et la fidélité des résultats.

Définition d'une mesure d'accommodement

Une mesure d'accommodement est à la fois un dispositif de soutien et un service permettant aux candidats nécessitant des aménagements de démontrer leurs compétences durant les tests. L'aménagement modifie uniquement la manière dont un test est administré ou la manière dont le candidat répond à certaines parties du test. Il ne modifie pas le contenu des tests et ne compromet pas leur validité ou leur fidélité.

Exemples de mesures d'accommodements disponibles

Le CExAC peut fournir les mesures d'accommodements suivants pour répondre aux besoins individuels :

- Matériel de test en caractères élargis ;

- Temps supplémentaire ;
- Accès facile au site d'administration des tests ;
- Accès à des collations, pour des raisons médicales, pendant la période de testing ;
- Accès à une salle plus petite ou à un espace séparé ; et/ou
- Accès à un transcripteur.

Contrôle de la qualité

Les consultants sont titulaires de diplômes en mesure et en évaluation et possèdent une expérience avérée en matière de tests à enjeux élevés et à grande échelle. Soucieux d'offrir un test de haute qualité, le comité de l'ExAC a mis en place des mesures de contrôle de la qualité afin d'éviter les erreurs dans le matériel de test, d'assurer l'uniformité et l'équité des tests administrés partout au Canada et de garantir la validité et la fidélité des données produites. Le CExAC a adopté des procédures afin que les personnes concernées aient confiance dans la validité et la fidélité des notes attribuées.

- Analyse des données : Les consultants effectuent plusieurs types d'analyses statistiques des données tirées des réponses des stagiaires afin de détecter la présence d'irrégularités et d'en vérifier l'intégrité.
- Vérification du matériel de test : Après chaque test, le CExAC tente de déterminer si des irrégularités ont été constatées lors de l'administration sur les lieux de testing ou lors de la notation. Pour ce faire, les consultants et les administrateurs inspectent visuellement le matériel de test à différentes étapes de l'administration et de la notation de l'ExAC. Les architectes, non impliqués dans le processus de développement, examinent les tests assemblés pour détecter les erreurs dans le matériel, les directives ou les références aux graphiques, plans et tableaux.

Remarque : L'ExAC ayant une fonction de certification pour l'accès à la profession, il est nécessaire d'optimiser la qualité des résultats obtenus et de vérifier la conformité des tests aux caractéristiques d'évaluation souhaitées. À cette fin, une analyse des caractéristiques d'évaluation de chaque section de l'ExAC et de chaque item est réalisée. À la suite de cette analyse, les items non cohérents ou jugés inadéquats sont exclus de la compilation des notes globales.

Des réviseurs linguistiques et des traducteurs professionnels sont sollicités.

NOTATION

Les items de l'ExAC peuvent être divisés en deux catégories selon le type de réponse demandé : les questions à choix multiples et les questions à réponse courte, chacune nécessitant une méthode de correction et de notation différente. Les réponses à choix multiples sont lues et corrigées par un lecteur optique de marques, qui enregistre automatiquement les données dans une base de données. Au nom du CExAC, le consultant suit des procédures de notation rigoureuses pour les questions à réponse courte afin de garantir la validité et la fidélité des scores, notamment la notation par des juges spécialement formés.

Les échelles descriptives d'appréciation et les copies types sont les principaux outils utilisés pour noter les items à réponse courte. Les échelles descriptives d'appréciation présentent une description pratique des différents niveaux de performance possibles des stagiaires, et les exemples illustrent les descripteurs de chaque pointage. Les juges sont formés à se référer aux copies types afin d'assurer une notation uniforme, en faisant correspondre le travail des stagiaires aux descripteurs pour chaque pointage. Chaque réponse fournie par un stagiaire est notée en fonction de la meilleure correspondance avec l'un des pointages définis dans l'échelle descriptive d'appréciation.

Lors du processus de notation, les consultants ont les responsabilités suivantes :

- Former les juges ;
- Superviser la notation des items ;
- Superviser l'application uniforme du matériel de notation ;
- Résoudre les problèmes survenant pendant la notation ;
- Répondre aux questions des juges ;
- Vérifier que le matériel de notation ne peut être modifié par les juges une fois le processus de notation lancé ;
- Procéder à une analyse détaillée du matériel de notation pour chaque item à noter (l'item, échelle descriptive d'appréciation, copies types et leurs justifications).

Les consultants sont également responsables du matériel de notation, de la sécurité du test, de la gestion de la salle et de tout problème survenant pendant la notation. Ils examinent et analysent les rapports de données quotidiens afin de s'assurer que la notation dans leurs salles est effectuée avec un haut niveau de qualité.

De même, les juges ont les responsabilités suivantes :

- Lire l'intégralité d'une réponse avant de prendre une décision concernant la notation;
- Considérer chaque réponse dans son ensemble, sans prêter attention aux détails (comme l'orthographe) ;
- Rester objectif et impartial en examinant chaque réponse à travers le filtre de l'échelle descriptive d'appréciation ;
- Éviter de se laisser influencer par d'autres copies corrigées (effet de halo) ; et
- Ne pas ajuster la notation pour tenir compte d'une caractéristique particulière affectant un stagiaire.

Les principales étapes du processus de notation sont détaillées ci-dessous.

Processus de sélection des copies types

Les copies types sont sélectionnées pour illustrer la fourchette de niveaux de performance acceptables pour chaque pointage d'une échelle descriptive d'appréciation et favoriser un consensus entre les juges sur la notation des réponses. Le processus se déroule en deux phases : la présélection des copies types par un consultant et leur sélection lors de la formation des juges.

La phase de sélection des copies types

La sélection des copies types :

- Implique la participation de consultants à la préparation d'outils de notation de haute qualité et de supports de formation pour les juges, afin d'assurer l'exactitude, la fidélité et l'uniformité de la notation des questions à réponse courte de l'ExAC ;
- Exige des copies types représentatives des réponses des stagiaires afin de définir et d'illustrer l'éventail de leurs niveaux de performance, correspondant aux échelles descriptives d'appréciation ;
- Établit un consensus pour le codage des réponses utilisées pour former les juges aux questions à réponse courte.

Les membres de l'équipe de notation (dont le nombre peut varier d'une année à l'autre, en fonction du nombre de stagiaires) issus de plusieurs juridictions se réunissent au début de la notation pour préparer et s'entraîner à l'utilisation des copies types lors de la notation.

Les qualifications requises pour les juges sont les suivantes :

- Expertise et expérience en architecture ;
- Capacité à expliquer clairement et de manière concise les raisons pour lesquelles une réponse fournie par un stagiaire a été classée à un niveau de performance donné dans une échelle descriptive d'appréciation.

Les juges:

- Se réfèrent aux échelles descriptives d'appréciation et utilisent leur expertise en matière de notation pour déterminer la note à attribuer à un stagiaire ;
- Collaborent avec les autres juges, sous la supervision de consultants, afin de parvenir à un consensus sur les notes appropriées pour chaque réponse ;
- Préparent des recommandations pour affiner les échelles descriptives d'appréciation, le cas échéant ;
- Formulent la liste des raisons justifiant les pointages attribués.

Aperçu du processus de formation

Les consultants organisent des réunions avec les juges :

- Les juges sont divisés en deux groupes, anglophone et francophone, afin de favoriser la productivité et l'efficacité. Avant le début du processus de notation, les deux groupes se rencontrent pour échanger des informations ;
- Chaque groupe discute d'un ensemble d'items et des échelles descriptives d'appréciation correspondantes, ainsi que des réponses appropriées convenues et des pointages associés ;
- Les discussions portent notamment sur :
 - > Le contenu et les exigences de chaque item ;
 - > Le consensus du groupe sur les notes attribuées aux réponses ;
 - > Les règles de notation, le cas échéant, pour une notation uniforme.

PRÉPARATION DU MATÉRIEL DE FORMATION POUR LA NOTATION

Copies types, copies de formation, test de qualification, copies de validation

Le consultant du CExAC prépare le matériel destiné à la formation des juges :

- Copies types qui représentent rigoureusement les points bruts dans les échelles descriptives d'appréciation;
- Copies de formation qui représentent autant les réponses correspondant clairement ou moins clairement (p. ex., réponses plus courtes que la normale, démarches atypiques, mélange d'attributs très bas et très haut ou écriture difficile à déchiffrer);
- Justifications pour chaque copie type et copie de formation utilisées; et
- Documents utilisés pour contrôler la validité.

Salles de notation pour les items à réponse construite

L'ensemble des items est noté dans une grande salle comprenant habituellement deux grandes tables. Une table pour l'équipe francophone et une table pour l'équipe anglophone. L'équipe de notation comprend toujours au moins un architecte qui pratique l'architecture dans les deux langues officielles et qui a pour rôle :

- Corriger des copies dans les deux langues;
- Faciliter la communication entre les 2 équipes;
- Clarifier des contenus, le cas échéant;
- Faciliter l'alignement et la performance entre les deux équipes.

Chaque table de notation est sous la supervision d'un consultant. Les consultants gèrent la formation, la notation et la formation d'appoint des juges. Tous les juges sont formés afin d'utiliser le guide de notation (échelles descriptives d'appréciation et copies types). Un suivi de la validité et de la fidélité de la notation est effectué en temps réel au centre de notation et les consultants déterminent si une formation d'appoint est nécessaire.

Les juges effectuent la notation individuellement. Un modèle de notation anonyme est utilisé; c'est-à-dire que les juges ne savent pas quel score a été attribué par l'autre juge. Les cahiers sont notés au moins à deux reprises (à trois reprises en cas d'absence de

concordance exacte des scores attribués) par deux (ou trois le cas échéant) juges différents. Cette stratégie favorise la fidélité du processus de notation.

Certaines copies sont corrigées plus de 3 fois lorsque la disparité entre les scores est importante. La moyenne des scores est octroyée au stagiaire.

Formation pour la notation des items à réponse construite

La formation vise à développer une compréhension claire et commune du matériel de notation de façon à ce que tous les juges interprètent et appliquent le matériel de notation de la même façon et les consultants peuvent superviser efficacement. L'objectif est d'obtenir une notation fidèle, uniforme, valide et exacte des réponses des stagiaires, peu importe le juge qui score les copies.

Les consultants encouragent les juges à rejeter toutes idées préconçues et se fier aux procédures et au matériel de notation.

Formation quotidienne continue

Les consultants donnent fréquemment des éclaircissements sur la notation d'items spécifiques et d'éléments clés des échelles descriptives d'appréciation spécifiques à un item de manière à veiller à ce que tous utilisent le matériel avec exactitude et uniformité d'un jour à l'autre ainsi qu'avant et après les pauses.

Révision matinale quotidienne des copies types

Les consultants commencent chaque journée par une révision, en totalité ou en partie, des échelles descriptives d'appréciation et des copies types. L'objectif de la révision est de réorienter les juges et de souligner toute section de l'échelle descriptive d'appréciation qui demande une attention particulière.

Activité quotidienne d'établissement de la norme

Sans être prévenus, les juges ont à corriger des copies qu'ils ont déjà corrigées eux-mêmes ou des copies qui ont obtenu une concordance parfaite chez d'autres juges. Cette procédure sert essentiellement à assurer l'uniformité de la correction chez les juges dans le temps ainsi qu'entre eux.

Rapports pour le contrôle de la qualité de la notation des items à réponse construite

Les consultants examinent quotidiennement l'ensemble des données relatives à la productivité des membres de leur équipe et à la concordance interjuge (taux de concordance parfaite entre les scores obtenus par un stagiaire provenant de deux juges différents).

Durant la notation, les juges doivent maintenir un taux d'exactitude minimal de moins de 1,5 point par rapport aux autres juges. Le juge qui ne satisfait pas aux normes requises de degré d'accord précis reçoit une formation additionnelle qui comprend une révision attentive des copies types. Si cette formation d'appoint ne corrige pas la situation, il est possible que les scores attribués par ce juge ne soient pas retenus et qu'une quatrième correction des copies concernées soit requise.

Rapports cumulatifs portant sur le score moyen et la répartition des scores

Les données cumulatives du score moyen et de la répartition des scores sont utilisées pour contrôler la déviance dans la notation d'un juge. Ces rapports de données permettent de cerner des problèmes potentiels à prendre en considération lors des analyses subséquentes.

Rapports cumulatifs de fidélité

Durant la notation, les consultants effectuent un suivi de la fidélité et de l'uniformité au moyen de la réinsertion anonyme (non identifiée) de cahiers.

Tous les items à réponse construite sont acheminés pour une deuxième correction, ce qui permet de faire le suivi de la fidélité interjuge. Les rapports indiquent et résument les niveaux quotidiens et cumulatifs de concordance interjuge pour deux notations indépendantes qui incluent une concordance parfaite ou adjacente. Les rapports sont résumés par item ou groupe d'items, groupe, échelle descriptive d'appréciation et juge. Les juges sont classés selon leur niveau de fidélité.

Rapports cumulatifs de productivité

Les rapports classent les juges selon leur niveau de productivité (de bas à élevé) et mentionnent le nom de ceux qui n'atteignent pas le taux de productivité minimal. Les

consultants passent en revue les données soulignées dans ce rapport afin de déterminer si un juge a besoin d'une formation d'appoint.

Sécurité

La Norme 6 (*Standard*) souligne l'importance de la sécurité des données personnelles et du contenu des tests. Des procédures de vérification ont été mises en place pour vérifier l'identité des stagiaires. Leurs numéros d'identification sont codés conformément à la symbologie des codes zébrés Code 39. Un réseau filaire en circuit fermé complet, équipé de formulaires de saisie, permet d'enregistrer les scores et de calculer les statistiques en temps réel, tout en limitant les risques de sécurité. L'utilisation d'un réseau filaire en circuit fermé empêche le piratage et les perturbations externes. Pendant les sessions de notation, aucun ordinateur, réseau ou base de données ne contient pas d'information personnelle et aucun matériel de test. Toutes les informations d'identification personnelle sont codées et seuls les numéros d'items et les scores obtenus sont transférés. Les réseaux d'accès publics et le Wi-Fi ne sont jamais utilisés. Un lecteur de codes zébrés électronique permet aux juges de lire et de saisir le numéro d'identification du stagiaire et celui du juge sur un formulaire informatisé (l'utilisation de la technologie des codes zébrés réduit considérablement les erreurs de transcription). Les scores sont ensuite saisis sur ce formulaire par le juge. Une fois la correction terminée, les données sont automatiquement envoyées à une base de données locale connectée au réseau en circuit fermé. Le matériel de test utilisé pour la notation (copies de l'ExAC et manuels de formation) est toujours surveillé ou enfermé dans une pièce.

Les transmissions confidentielles entre l'équipe du CExAC et les associations s'effectuent via des protocoles de communication cryptés.

PROCÉDURES EXCEPTIONNELLES LORS DE LA NOTATION

Indications de contenu inapproprié ou d'interférence lors de l'administration de l'ExAC

Il arrive que les réponses des stagiaires aux questions à réponse courte contiennent du contenu inapproprié ou offensant, semblent indiquer une modification de la calligraphie, une possible interférence du personnel impliqué dans l'administration de l'ExAC, ou tout autre problème. Les juges sont formés à détecter les anomalies et à signaler les copies qu'ils soupçonnent de révéler de tels problèmes une fois notées. Les copies sont évaluées par les membres du CExAC, qui examinent les rapports d'incident, la juridiction et/ou les administrateurs du ROAC, si nécessaire.

Matériel de test endommagé ou mal imprimés

Lorsque les copies de l'ExAC sont distribuées aux stagiaires, il est possible qu'elles soient déchirées, mal agrafées ou qu'il manque des pages. Dans ce cas, les stagiaires ne doivent pas être pénalisés et le matériel doit être remplacé au moment de l'administration de l'ExAC. Si de tels dommages sont constatés lors de la séance de notation, les membres du CExAC et les consultants examinent et déterminent la procédure de notation de ces copies.

Voir la note explicative concernant la gestion des erreurs d'impression plus haut dans ce document.

LES ARGUMENTS DE VALIDITÉ FONDÉS SUR LES PROPRIÉTÉS MÉTRIQUES DE L'ExAC

Qualité technique de l'ExAC

La méthode d'analyse de la qualité technique de l'ExAC est conforme à la théorie classique des scores (TCS) ou théorie classique des tests. Cette théorie a connu ses premiers énoncés en 1907, notamment avec les travaux de Spearman (1907) et a été formalisée au cours des années 1960. Des théories des tests et des modèles de mesure plus modernes ont été mis au point depuis, par exemple la théorie de la réponse à l'item (TRI) qui présente des avantages considérables sur le plan de l'analyse technique des items et des tests dans certains programmes de testing. Cependant, ces théories sont exigeantes en ce qui a trait à la conformité aux postulats et à certaines conditions d'utilisation. Par exemple, pour calibrer les items en fonction de quelques paramètres, un échantillon considérable de cas (stagiaires) est requis. L'achalandage annuel de l'ExAC est insuffisant pour établir ces paramètres. Nous avons donc opté pour la TCS qui est une théorie valide, connue et encore largement utilisée aujourd'hui. Elle est moins contraignante par rapport à la taille de l'échantillon.

On tient compte de plusieurs paramètres dans le choix des items qui comptent pour le score total des stagiaires. La difficulté p de l'item est d'abord considérée. Le degré de difficulté p s'étend de 0 à 1, les items près de 0 étant les plus difficiles et ceux près de 1 étant les plus faciles. On tente de conserver les items qui se situent entre 0,20 et 0,95. Par exemple, un item ayant un degré de difficulté $p = 1$ est un item réussi par tous, ce qui est somme toute peu utile pour évaluer le rendement d'un stagiaire.

Le coefficient de discrimination D est aussi considéré. La discrimination d'un item réfère à la propriété de cet item de permettre de discerner les stagiaires ayant un meilleur rendement de ceux en ayant un moins bon. Par exemple, un stagiaire qui réussit très bien l'ensemble du test devrait avoir plus de chances de réussir un item en particulier que le stagiaire ayant un moins bon rendement. Le paramètre D s'étend de -1 à 1. Les items qui ont un coefficient près de 0 ou inférieur à 0 ne sont pas de bons discriminateurs. Ceux dont le coefficient est supérieur à 0,20 sont généralement retenus puisqu'ils sont de bons discriminateurs.

Afin d'assurer qu'un sous-test est cohérent sur le plan métrique, on considère la corrélation entre les résultats de l'item et le score total du sous-test. Il s'agit de vérifier

que tous les items mesurent la même chose ou le même domaine. Un item peu corrélé avec les autres pourrait mesurer la connaissance d'un fait qui n'a rien à voir avec la pratique de la profession d'architecte. Généralement, les items qui ont une corrélation item/total supérieure à 0,20 sont retenus.

Les paramètres de difficulté p , de discrimination D et de corrélation item/total contribuent, chacun à leur manière, à la précision et à la cohérence interne d'un sous-test. Des techniques d'optimisation permettent de déceler les items qui représentent des faiblesses en fonction de l'un ou de l'autre de ces paramètres. Ces items sont retirés et révisés ultérieurement afin de ne pas perpétuer ces faiblesses. Notons qu'ils ne comptent pas dans le score total des stagiaires.

L'ExAC étant un test à enjeux élevés, des mesures de précision sont effectuées pour assurer un suivi de son évolution, ainsi que pour déterminer l'ampleur de l'intervalle de confiance qui assure que tous les stagiaires qui devaient réussir ont réussi. L'alpha de Cronbach, estimation conservatrice de la fidélité, est calculé pour chaque sous-test et dans chaque langue.

L'ExAC est un test récent et que ses coefficients de fidélité augmentent de façon continue depuis sa première édition. Même si un coefficient alpha de 0,7 ou plus est considéré acceptable, nous visons des coefficients nettement supérieurs à 0,8 en raison des enjeux élevés liés à la pratique de l'architecture. Le tableau qui suit montre la progression continue de cet indicateur pour l'ExAC.

Historical Cronbach's coefficients

	Anglais	Français	Alpha moyen
Coefficients alpha (ExAC08)			0,807
Coefficients alpha (ExAC09)			0,825
Coefficients alpha (ExAC10)			0,822
Coefficients alpha (ExAC11)			0,813
Coefficients alpha (ExAC12)			0,835
Coefficients alpha (ExAC13)			0,851
Coefficients alpha (ExAC14)			0,846

Coefficients alpha (ExAC15)			0,846
Coefficients alpha (ExAC16)			0,851
Coefficients alpha (ExAC17)			0,877
Coefficients alpha (ExAC18)			0,876
Coefficients alpha (ExAC19)			0,883
Coefficients alpha (ExAC20)			0,879
Coefficients alpha (ExAC21)			0,868
Coefficients alpha (ExAC22)			0,891
Coefficients alpha (ExAC23)			
Section 1	0,880	0,854	
Section 2	0,895	0,879	
Section 3	0,900	0,885	
Section 4	0,915	0,885	0,887
Coefficients alpha (ExAC24)			
Section 1	0,883	0,860	
Section 2	0,880	0,893	
Section 3	0,901	0,886	
Section 4	0,917	0,910	0,891

Estimation de la fidélité sur les items à réponse construite de l'ExAC 2024 spécifiquement

Nous calculons la concordance interjuge de manière à estimer la fidélité de la correction des items à réponse construite. Toutes les copies sont corrigées au moins deux fois par différents architectes. Nous pouvons ainsi vérifier si les scores sont similaires.

Les pourcentages de concordance parfaite entre les deux premiers juges pour, la section 1 varient entre 63,8 % et 95,0 % pour les formes anglaise et française. Dans l'ensemble, les pourcentages de concordance parfaite pour les items de la section 1 dans les tests en anglais et en français sont supérieurs aux normes généralement acceptées.

Les pourcentages de concordance parfaite entre les deux premiers juges pour la section 3 varient entre 68,2 % et 94,6 % pour les formes anglaise et française sauf pour un item. Dans l'ensemble, les pourcentages de concordance parfaite pour les items de la section 3 dans les tests en anglais et en français sont supérieur aux normes généralement acceptées.

Notons que lorsque deux juges ne concordent pas à 100 % lors de l'attribution d'une note à une réponse d'un stagiaire, la réponse est corrigée à nouveau par un troisième juge. Les juges ne connaissent pas la note accordée par un autre juge et ne savent pas combien de fois une copie est corrigée.

D'autres analyses sont effectuées comme l'erreur type de mesure et le fonctionnement différentiel du test. Nous effectuons également un suivi des stagiaires qui ont des scores très faibles afin de vérifier si un groupe de stagiaires est désavantagé pour des raisons linguistiques ou autres.

Comme nous l'avons mentionné plus haut, plusieurs facteurs contribuent à la précision de l'ExAC : la qualité des items, la précision et l'uniformité de la notation, et la corrélation entre les items. Tous les items de l'ExAC sont liés directement aux objectifs généraux et spécifiques. Le consultant sélectionne des items de difficulté pertinente et qui permettent de discriminer les élèves ayant un fort niveau et ceux ayant un faible niveau de rendement. Plusieurs pratiques contribuent également à la précision et à l'uniformité de la notation.

Mise en équivalence

Une procédure de mise en équivalence permet de rendre comparable le niveau de difficulté de l'ExAC d'une année à l'autre ainsi entre les deux langues. Nous pouvons postuler qu'il y a une relation linéaire entre les éditions annuelles de l'ExAC dans la mesure où le tableau de spécifications, les types d'items et les modalités de réponses demeurent relativement constants d'une année à l'autre. Puisque c'est le cas, les méthodes de mise en équivalence linéaire sont indiquées. De manière à déterminer le facteur multiplicateur utilisé pour effectuer la transformation, un échantillon d'items est répété et les scores moyens obtenus de ces items sont pistés d'une année à l'autre.

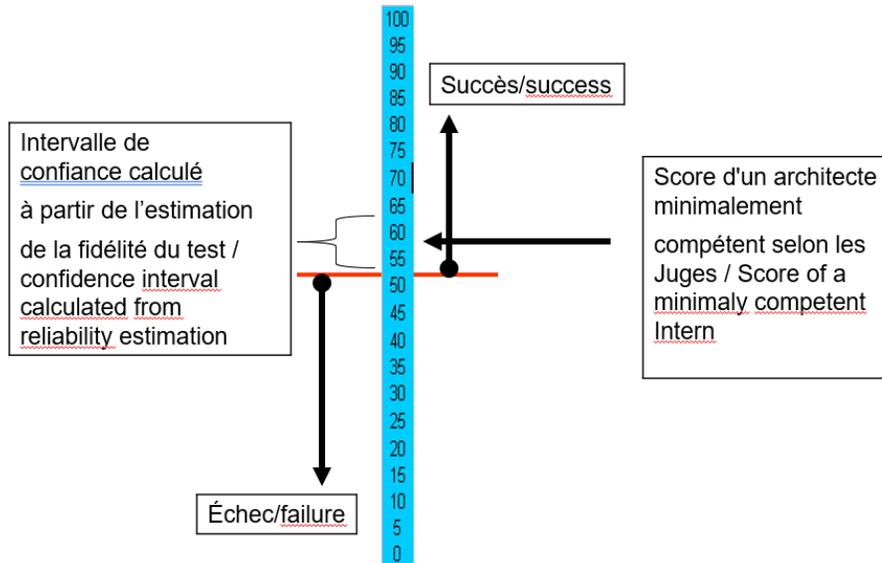
Cette méthode permet également de vérifier si les fluctuations des scores moyens des stagiaires sont dues au degré moyen de difficulté des tests ou si elles sont dues aux variations du niveau de performance moyen d'une cohorte de stagiaires d'une année par rapport à l'autre.

Établissement de scores seuils

Une série d'étapes permet d'établir le score de passage pour chaque section de l'ExAC. Le processus nécessite de prendre en compte les références contextuelles et historiques, l'évolution de la difficulté moyenne (p) des items communs, ainsi qu'une estimation de la probabilité de réussite d'un candidat ayant un niveau minimal de compétence. Ce processus est réalisé en totalité ou en partie toutes les quelques années, ou lorsqu'un événement majeur survient — comme la COVID-19 ou un changement substantiel de contenu, comme une nouvelle version du code du bâtiment. Le dernier exercice remonte à l'automne 2024. Après avoir reçu une formation, plusieurs architectes ont évalué individuellement la probabilité de réussite des items par un architecte minimalement compétent. Lorsque les jugements étaient significativement différents pour certains items, ils ont débattu.

La moyenne des estimations des juges sur le total des tests constitue en quelque sorte le score minimal que devrait obtenir à chaque test le stagiaire ayant une compétence minimale. Il s'agit d'une démarche conforme à celle d'Angoff (1971). Nous utilisons également le pistage d'un ensemble d'items répétés pour assurer que les scores seuils demeurent équivalents d'une année à l'autre. Ensuite, comme les tests sont des instruments de mesure ayant une fidélité imparfaite, nous calculons un intervalle de confiance qui se situe autour du score limite estimé par les juges (la ligne rouge dans la figure qui suit montre la limite inférieure). De cette façon, nous donnons le bénéfice du doute aux stagiaires de compétence minimale apparente. Un stagiaire qui échoue à un test obtient donc un score inférieur au score limite estimé par les juges, et inférieur au point le plus bas de l'intervalle de confiance (point de césure) pouvant être considéré à titre de marge d'erreur.

La figure qui suit illustre les facteurs qui sont pris en considération.



Sommaire de l'exercice d'établissement de scores seuils de 2024

Facteurs à considérer lorsqu'on révisé les scores seuils :

- Distribution des scores;
- Précision de l'instrument;
- Données historiques et contextuelles;
- Évolution du degré de difficulté relative des items communs dans le temps;
- Estimation de la performance des architectes minimalement compétents.

Résultats

Évolution du degré de difficulté relatif des items communs	Augmentation de 1,57 points de pourcentage
Performance des architectes minimalement compétents	Diminution du taux de passage de 2,1 points de pourcentage
L'analyse des résultats et des facteurs contributeurs suggère qu'il y a une légère baisse de performance chez les stagiaires. À la lumière de ce constat, le taux de passage a été abaissé de façon conservatrice de 1,5 points de pourcentage.	

Publication des résultats

Les résultats des tests du ExAC sont publiés en fonction de la langue des stagiaires, des provinces ou territoires et pour l'ensemble du Canada. Les résultats décrivent le rendement des stagiaires à cette étape cruciale de leur formation. Les enjeux pour les stagiaires sont très élevés puisque la réussite de l'ExAC est une condition essentielle d'accès à la profession d'architecte. Les résultats des stagiaires sont cruciaux pour les ordres professionnels sur le plan économique et pour informer la prise de décision dans une perspective de gestion axée sur les résultats et fondée sur des données probantes. Pour toutes ces raisons, un niveau de confidentialité élevé doit être maintenu de manière à protéger l'intégrité de l'ExAC et de manière à protéger la confidentialité des renseignements personnels des stagiaires.

À la demande du CExAC, les coadministrateurs publient annuellement dans les deux langues officielles ce présent rapport général ainsi qu'un rapport sommaire. Ces rapports sont publics et disponibles sur le [site Web](#) de l'ExAC après quelques mois suivant l'administration de l'ExAC. Des tableaux sont préparés et publiés dans le but de fournir des taux de réussite agrégés.

Les ordres professionnels reçoivent des fichiers de données présentant les résultats détaillés de rendement de chaque stagiaire inscrit dans leur province ou territoire. Les ordres communiquent les résultats aux stagiaires sous forme de rapport de rendement. Ce rapport fait état du statut du stagiaire concerné, des sections réussies de l'ExAC et des sections à reprendre. Les résultats sont communiqués sous la forme de réussite ou d'échec. Les scores bruts ne sont pas envoyés.

Dans le cas d'une section à reprendre, le rapport fournit des renseignements constructifs sur les portions moins bien réussies de la section en question. Veuillez consulter les règles portant sur les reprises sur le [site Web](#) de l'ExAC.

Taux de succès pour l'ExAC 2024

Dix-sept juges (juges) ont participé à la séance de notation de l'ExAC 2024. Le tableau qui suit présente le taux de succès pour chaque section. Sur les 931 stagiaires qui ont participé à l'ExAC 2024, 672 d'entre eux ont réussi tous les tests pour lesquels ils étaient admissibles. Le taux de réussite complète est de **72,2 %**. Le taux de succès moyen par section est **83,5 %**.

Taux de succès combinés pour les formes anglaises et françaises de l'ExAC 2024

Section 1	Section 2	Section 3	Section 4
705 / 839	704 / 847	699 / 835	704 / 845
84,0 %	83,1 %	83,7 %	83,3 %

Demandes de revisions

En conformité avec la Norme 8, il est possible pour un stagiaire d'effectuer une demande de révision. Les règles précises qui encadrent ce processus sont incluses sur le [site Web](#). Une fois la demande complétée, des consultants effectuent une vérification en plusieurs points des feuilles réponses à bulles. Les items à réponses courtes sont corrigés à nouveau par deux juges expérimentés. Un pistage à l'intérieur des bases de données pour les données relatives aux stagiaires est également effectué de manière à vérifier l'intégrité des données.

RÉFÉRENCES

- American Educational Research Association, American Psychological Association and National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, DC, American Educational Research Association.
- Angoff, W.H. (1971). *Scales, norms and equivalent scores*. In R.L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 508-600). Washington, DC: American Council on Education.
- Bloom, B. S.; Engelhart, M. D.; Furst, E. J.; Hill, W. H.; Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals*. Handbook I: Cognitive domain. New York: David McKay Company.
- Kane, M.T. (2006). *Validation* in R. L. Brennan (Ed.), *Educational Measurement*, Fourth Edition. Westport, CT. American Council on Education and Praeger Publishers, pp. 17-64.
- Kolesar, Peter J. (1994). *What Deming told the Japanese in 1950*. *Quality Management Journal* 2(1) (1994): 9–24.
- Shewhart, W. A. (1939). *Statistical method from the viewpoint of quality control*. Washington, DC: Graduate School of the Department of Agriculture.
- Spearman, C. (1907). *Demonstration of formulae for true measurement of correlation*. *American Journal of Psychology*, 18, 161-169.
- Swamidass, P. M. (2000). *Deming Cycle (PDCA)*. In: Swamidass, P. M. (eds) *Encyclopedia of Production and Manufacturing Management*. Springer, New York.

GLOSSAIRE TERMINOLOGIQUE

Terme	Explication
Accommodement	Soutien ou service offert sur demande à des individus qui ont une condition pouvant nuire à leur participation à un test.
Administrateurs/ Coordonnateurs nationaux de l'ExAC	Deux individus qui supervisent la préparation, l'administration, les ressources et les activités liées aux différentes phases de l'ExAC.
Agrégats	Résultats dépersonnalisés regroupés en fonction de groupes d'appartenance ou de régions.
Amorce	La question ou l'énoncé reliés à un item (sans les choix de réponses).
Banque d'items	Réserve d'items archivés ou récemment développés qui peuvent être choisis pour développer les quatre sous-tests.
Biais d'un item ou d'un test	Particularité d'un item qui favorise ou défavorise des individus en fonction d'un groupe d'appartenance particulier.
Calibrer les items	Attribution de statistiques propres à chaque item. Chaque item est connu pour son degré de difficulté, son coefficient de discrimination et son coefficient de corrélation avec les autres items du test. On parle de propriétés métriques. Le choix d'inclure un item dans le test est tributaire en bonne partie de ces statistiques.
Banque d'items	Réserve d'items archivés ou récemment développés qui peuvent être choisis pour développer les quatre sous-tests.
Biais d'un item ou d'un test	Particularité d'un item qui favorise ou défavorise des individus en fonction d'un groupe d'appartenance particulier.
Calibrer les items	Attribution de statistiques propres à chaque item. Chaque item est connu pour son degré de difficulté, son coefficient de discrimination et son coefficient de corrélation avec les autres items du test. On parle de propriétés métriques. Le choix

	d'inclure un item dans le test est tributaire en bonne partie de ces statistiques.
Centre de notation	Durant la séance de notation, une salle séparée est équipée d'ordinateurs et d'équipement permettant de faire les calculs et les suivis nécessaires au bon déroulement de la séance.
CExAC	Comité de l'Examen des architectes du Canada (ExAC)
Coefficient de corrélation	La corrélation réfère à la qualité du lien entre des entités (items, tests). Plus le lien est fort entre les items, plus le test forme un ensemble cohérent. Par exemple, il y a corrélation positive entre deux items si les individus qui réussissent l'item 22 tendent à réussir l'item 34.
Coefficient de discrimination	Un item qui discrimine bien est un item qui est généralement réussi par les individus qui réussissent bien et qui est échoué généralement par les individus qui ont des scores totaux faibles.
Concordance interjuge	Il s'agit d'un indice de comparaison entre des juges lors de la correction d'un même item. Il y a concordance lorsqu'un juge attribue le même score qu'un autre juge pour la réponse d'un individu particulier à un item en particulier.
Concordance intrajuge	Il s'agit d'un indice de comparaison entre les scores attribués aux items par le même juge. Il y a concordance lorsqu'un juge attribue le même score à un moment différent pour la réponse d'un individu particulier à un item en particulier.
Consultants.es	Spécialistes de la construction des tests et de leurs analyses. Les consultants ne sont pas des architectes.
Copies de formation	Copies types utilisées aux fins de formation. Ces copies sont des copies des réponses de stagiaires.
Copies de validation	Copies types utilisées pour vérifier la performance et la qualité du travail des juges.
Copies types	Exemple de réponse à un item qui est typique d'un niveau de rendement en particulier.

Degré de difficulté	Rapport entre le nombre d'individus qui ont réussi un item et le nombre de stagiaires qui ont participé au test. Ce rapport est utilisé pour évaluer la difficulté d'un item ou le niveau d'atteinte d'un objectif.
Docimologie	Science qui porte sur l'élaboration, l'administration et l'interprétation des tests.
Échelle descriptive d'appréciation spécifique à un item	Grille d'évaluation utilisée durant la séance de notation des items à réponse construite. Cette grille décrit à l'aide de descripteurs les différents niveaux de performance à un item.
Édumétrie	Science qui porte sur l'élaboration, l'administration et l'interprétation des tests en éducation.
Établissement de scores seuils	Processus permettant de déterminer et de confirmer le seuil minimal acceptable de performance.
ExAC	Examen des architectes du Canada
Examen	Dans ce rapport, le terme « examen » est synonyme de test. On tend à utiliser le mot test lorsqu'on parle de l'instrument de mesure en soi.
Feuille réponse à bulles	Feuille réponse qui permet aux individus de consigner leurs réponses aux items à choix multiples. Les feuilles sont rigides et peuvent être lues par un lecteur optique de marques. Les scores sont consignés électroniquement dans un système de gestion de bases de données.
Fidélité	La fidélité est un indice. Il s'applique à un groupe d'items et indique dans quelle mesure il y a cohérence interne entre les items. En d'autres mots, les items mesurent-ils la « même chose »? La fidélité se distingue de la validité qui vise à déterminer si le groupe d'items mesure ce qu'il doit mesurer.
Fonctionnement différentiel	Biais qui favorise ou défavorise un groupe d'individus en fonction de leur groupe d'appartenance.
Forme du test	L'ExAC a une forme française et une forme anglaise.
Groupe d'appartenance	Groupe identifiable selon le sexe, la religion, l'ethnie, la culture ou un handicap.

Item	Question dans un test. On parle aussi de « tâche » lorsqu'on réfère à un item à réponse courte.
Item à choix multiple	Item qui comporte une amorce, des leurres et une bonne réponse.
Item à réponse courte ou construite	Item qui exige une réponse élaborée sans option de réponse.
Juge	Synonyme de correcteur ou de correctrice dans ce document.
Lecteur optique de marques	Appareil qui peut lire des feuilles réponses à bulles. Il s'agit d'un balayeur optique spécialisé.
Leurre	Options de réponse fausse dans le cas d'un item à choix multiple.
Mise à l'essai des items	Processus par lequel les items sont testés sur le terrain de manière à vérifier leurs propriétés métriques et leur fonctionnement.
Mise en équivalence	Méthode permettant de s'assurer que les résultats aux tests sont comparables dans le temps et entre les formes du test.
NCARB	Test provenant du National Council of Architectural Registration Boards. L'ExAC est une alternative au NCARB depuis 2008.
Niveaux taxonomiques	Réfère à la hiérarchie des exigences cognitives évoquées par les items. Les items cherchent à mesurer une connaissance, la compréhension, la capacité à appliquer des concepts, la capacité d'analyse, de synthèse et d'évaluation.
Ordre	Association professionnelle règlementée par la loi. L'architecture est une profession règlementée par des lois provinciales.
Pondération d'un item	Nombre de points attribuable à un item. Réfère au poids d'un item par rapport au score total.
Propriétés métriques	Plusieurs statistiques qui portent sur le test et ses items (fidélité, validité, etc.).

Psychométrie	Science qui porte sur l'élaboration, l'administration et l'interprétation des tests en psychologie et en éducation. Le mot Édumétrie peut aussi être utilisé.
Remédiation	Stratégie visant à compenser ou corriger en cas d'échec ou de faiblesse dans un domaine.
Salle de notation	Lors de la séance de notation, une salle séparée est réservée pour les juges qui corrigent les copies.
Scores	Réfère aux points obtenus et au nombre de bonnes réponses.
Seuil de compétence minimal	Score total à un test qui représente le minimum qu'un individu doit obtenir pour être considéré compétent. Réfère au point de césure qui distingue entre les individus qui passent ou échouent.
Superviseur de test	Individu qui supervise l'administration d'un test.
Tableau de spécifications	Grille utilisée durant l'élaboration des items et l'assemblage du test. Le tableau de spécifications stipule explicitement tous les contenus à mesurer, leur lien avec les objectifs spécifiques et quels items correspondent à quel contenu.
Test à enjeux élevés	Expression utilisée lorsque les résultats au test affectent sérieusement les humains. Les tests certificatifs comme l'ExAC sont forcément des tests à enjeux élevés.
Test à grande échelle	Se distingue du test de salle de classe de par l'envergure de son déploiement. Les tests certificatifs comme l'ExAC sont forcément des tests à grande échelle.
Test certificatif ou d'admission à une profession	Test qui a pour intention de certifier un individu en fonction de critères bien définis. Puisque les enjeux sont très élevés, la précision, la pertinence et les processus entourant la conception, l'administration, la notation, la sécurité et l'interprétation des scores sont prioritaires. L'ExAC est un test certificatif.
Test de qualification	Des copies types sont utilisées pour vérifier si les juges ont bien assimilé les échelles descriptives d'appréciation.
Test diagnostique, test d'aiguillage, de placement	Test qui a pour intention de déceler les forces et les difficultés des étudiants, d'aiguiller des étudiants dans un programme en particulier ou d'évaluer le rendement d'un programme ou d'une stratégie.

Validité	La validité est un concept large et englobant qui réfère à la notion de pertinence du test et de ses processus. Par exemple, un test est valide lorsqu'il mesure ce qu'il doit mesurer, et rien d'autre. Le test est valide quand l'interprétation des résultats et son utilisation sont équitables et pertinentes.
Version	Édition d'un test. « Nous avons une version française et une version anglaise de l'ExAC ».

