

Lignes directrices destinées à la révision des questions à choix multiples des  
examens des EMPC

Dr Timothy J. Wood

2 mars 2020

Je remercie tout particulièrement les Drs Michelle Anawati, Isabelle Burnier, Heather MacLean, Laurie McLean et Chris Ramnanan des commentaires qu'ils ont formulés concernant les premières versions du présent document.

Vos questions et vos suggestions en vue d'une amélioration peuvent être adressées à Tim Wood ([twood@uottawa.ca](mailto:twood@uottawa.ca)).

Il arrive que des questions d'examen affichent un mauvais rendement pour diverses raisons : erreurs typographiques, problèmes linguistiques, mauvaise rédaction ou évaluation de sujets qui n'ont pas été abordés dans le cadre du cours. Après un examen, il est habituel de passer en revue le rendement des questions et de cerner tous les items dont l'indice de performance est faible. Le fait de dégager ces questions et de tenter de comprendre leur piètre rendement veille à ce que les résultats des étudiants soient le miroir exact de leurs compétences et de leurs connaissances.

Le présent document a pour but de fournir des lignes directrices aux responsables d'unité et aux experts de contenu lorsqu'ils examinent le rendement des questions à choix multiples dont l'indice de performance est faible dans le cadre d'un examen. Notre attention portera sur des indices statistiques servant à dégager les questions à faible rendement et sur les mesures à prendre en fonction de ces informations.

### *Indices statistiques utilisés pour le signalement de questions à faible rendement*

Le tableau 1 présente un exemple de rapport concernant des questions à choix multiples, qui est fourni aux responsables d'unité et aux experts de contenu après un examen. Le présent tableau comporte deux indices statistiques qui sont utilisés pour repérer les questions à faible rendement. Le premier est l'indice de difficulté (surligné en jaune), qui indique le nombre d'étudiants ayant répondu correctement à la question. Cet indice est rapporté sous forme de taux de réussite à la question. Par exemple, 0,75 ou 75 % des étudiants ont répondu correctement à la question 1, alors que 0,96 ou 96 % des étudiants ont répondu correctement à la question 2. Si l'indice de difficulté est inférieur à 0,30, la question est dite très difficile. Si l'indice de difficulté est supérieur à 0,95, la question est dite très facile.

Tableau 1. Exemple de statistiques générées au sujet des questions à choix multiples

Question	Réponse	Indice de difficulté	Corrélation bisériale	Corrélation bisériale de point	Indice de discrimination
1	E	0,75	0,45	0,38	0,47
2	D	0,96	0,11	0,05	0,08
3	C	0,25	0,22	0,15	0,12
4	E	0,25	0,12	0,07	0,07
5	D	0,25	-0,10	-0,12	-0,11

Le deuxième indice présenté dans le tableau 1 est l'indice de discrimination, qui, pour une question donnée, permet de différencier les étudiants ayant obtenu le meilleur résultat à l'examen (ci-après, les étudiants forts) des étudiants ayant obtenu les résultats les plus faibles à l'examen (ci-après, les étudiants faibles). En réalité, deux valeurs présentées dans le tableau servent à cette fin. La première est le coefficient de corrélation bisériale de point (surligné en vert). Cette corrélation mesure la relation entre les résultats obtenus à une question donnée et les résultats obtenus à l'examen dans son ensemble. La deuxième est appelée indice de discrimination (surlignée en gris). Afin de générer cette

valeur, on identifie les étudiants dont les résultats globaux à l'examen se trouvent dans le quart (25 %) supérieur et ceux dont les résultats se trouvent dans le quart (25 %) inférieur. Pour toute question donnée, le nombre d'étudiants faibles ayant répondu à la question correctement est soustrait au nombre d'étudiants forts ayant répondu correctement à la question, puis il est divisé par le nombre total d'étudiants. Bien que la corrélation bisériale figure dans le tableau, elle n'est pas utilisée parce que les hypothèses de cette corrélation ne s'appliquent pas aux questions à choix multiples.

Le coefficient de corrélation bisériale de point et l'indice de discrimination peuvent être utilisés indifféremment pour mesurer la capacité d'une question de distinguer les étudiants forts des étudiants faibles, et le choix de l'indice relève simplement de la préférence de chacun. Les deux produisent des résultats similaires et les deux valeurs sont interprétées de la même façon. Un indice de discrimination supérieur à 0,10 indique que les étudiants forts ont généralement répondu correctement à la question, tandis que les étudiants faibles ont généralement échoué à cette question. Ce cas de figure serait considéré comme étant une question de niveau optimal. Un indice de discrimination entre 0 et 0,09 indique qu'il y a peu de différence entre le résultat à la question des étudiants forts et celui des étudiants faibles, et donne à penser que la question est problématique. Un indice de discrimination négatif indique que les étudiants faibles ont répondu correctement à la question, tandis que les étudiants forts y ont répondu incorrectement. Une question dont l'indice de discrimination est négatif est toujours jugée à faible rendement. Dans l'exemple présenté dans le tableau 1, la question 5 a une corrélation bisériale de point négative et serait automatiquement signalée.

L'indice de difficulté et l'indice de discrimination interagissent également l'un avec l'autre. Si une question est très difficile, alors un grand nombre d'étudiants forts y répondront incorrectement et l'indice de discrimination aura tendance à être bas. Si une question est très facile, alors même les étudiants faibles y répondront correctement et l'indice de discrimination aura aussi tendance à être bas. La question 2 du tableau 1 illustre cette interaction. L'indice de difficulté est élevé (0,96) et, de ce fait, la corrélation bisériale de point est inférieure à 0,10, ce qui indique que même les étudiants faibles ont répondu correctement à la question parce qu'elle était relativement facile.

### *Quels aspects doit-on considérer lorsque l'on révise une question signalée?*

La première étape permettant de déterminer si le rendement d'une question est faible consiste à relever les questions qui affichent des statistiques médiocres. Pour comprendre les résultats statistiques, toutes les questions signalées doivent être révisées. À ce stade, la prise en considération des commentaires des étudiants pourrait s'avérer utile, car cette rétroaction pourrait fournir une explication des statistiques qui n'était pas évidente au moment de la révision.

Les lignes directrices ci-après dressent une liste des informations dont il faut tenir compte au moment d'examiner une question.

1. Si une question est difficile (taux de réussite inférieur à 0,30) :
  - vérifiez le corrigé et assurez-vous que la bonne réponse a été correctement indiquée aux fins de notation;
  - lisez la question et confirmez que la bonne réponse figure parmi les choix et, le cas échéant, vérifiez son exactitude et si d'autres bonnes réponses sont possibles;

- lisez la question et assurez-vous qu'il n'y a pas d'indice (p. ex., erreurs typographiques ou grammaticales) ou d'autre ambiguïté qui inciterait les étudiants à choisir la mauvaise réponse;
- confirmez que la question évalue bel et bien un objectif abordé dans le cadre du cours;
- assurez-vous que les versions française et anglaise de la question sont équivalentes;
- vérifiez si les étudiants ont signalé des problèmes relatifs à cette question.

Si toutes les vérifications ci-dessus sont négatives, alors tournez-vous vers l'indice de discrimination. S'il est supérieur à 0,10, il indique que la question est difficile, mais que les étudiants forts y ont généralement répondu correctement et, par conséquent, que la question est probablement adéquate. Si l'indice de discrimination se trouve entre 0 et 0,09, alors il est possible que la question soit inadéquate et que l'on doive l'éliminer (l'exclure du calcul de la note finale) ou l'envoyer à un expert de contenu aux fins de révision avant qu'elle ne soit utilisée à nouveau.

2. Si une question est extrêmement facile (indice de difficulté supérieur à 0,95) :

- lisez la question pour vous assurer qu'il n'y a aucun indice (p. ex., erreurs typographiques ou grammaticales) qui révélerait la bonne réponse.

Si la vérification ci-dessus n'explique pas clairement la facilité de la question, alors il serait possible de la conserver en tant qu'item noté, mais elle devrait être envoyée à un expert de contenu aux fins de révision avant qu'elle ne soit utilisée à nouveau.

3. Si l'indice de discrimination est négatif :

- vérifiez le corrigé et assurez-vous que la bonne réponse a été correctement indiquée aux fins de notation;
- lisez la question et confirmez que la bonne réponse se trouve parmi les choix et, le cas échéant, vérifiez son exactitude et si d'autres bonnes réponses sont possibles;
- lisez la question et assurez-vous qu'il n'y a pas d'indice (p. ex., erreurs typographiques ou grammaticales) ou d'autre ambiguïté qui inciterait les étudiants à choisir la mauvaise réponse;
- confirmez que la question évalue bel et bien un objectif abordé dans le cadre du cours;
- assurez-vous que les versions française et anglaise de la question sont équivalentes;
- vérifiez si les étudiants ont signalé des problèmes relatifs à cette question.

Si toutes les vérifications ci-dessus sont négatives, il se peut qu'un autre aspect de la question soit inadéquat. Il arrive, bien que rarement, qu'une question dont l'indice de difficulté est négatif soit conservée à l'examen (p. ex., la valeur était proche de 0 et un petit nombre d'étudiants ont été évalués), mais, dans la plupart des cas, la question devrait être exclue du calcul de la note finale et envoyée à un expert de contenu aux fins de révision avant qu'elle ne soit utilisée à nouveau.

*Quelles décisions faudra-t-il prendre concernant une question?*

Quatre possibilités s'offrent à la personne chargée de la révision d'une question signalée comme étant à faible rendement.

1. Supprimer la question

La suppression de la question l'exclut de la note finale, de telle sorte que, même si un étudiant y répond correctement, il n'obtient pas de point pour celle-ci. Cette mesure est prise : a) lorsqu'il n'y a pas de bonne réponse à la question (c.-à-d., toutes les réponses sont incorrectes) ou b) lorsque la question évalue un objectif qui n'a pas été abordé dans le cadre du cours. La question supprimée devrait être renvoyée à l'expert de contenu à des fins d'amélioration ou de suppression de la banque de questions.

2. Attribuer les points pour la question

L'attribution des points pour la question signifie que les points de la question sont déduits du total des points, mais que les étudiants ayant répondu correctement à la question obtiennent tout de même des points pour celle-ci. Cette mesure est prise lorsque la révision indique qu'il y a un problème avec la question (p. ex., erreur typographique, ambiguïté, problème de traduction). Ces questions devraient être renvoyées à l'expert de contenu aux fins de modification avant qu'elles ne soient utilisées à nouveau.

3. Modifier le corrigé

Il arrive que la réponse qui est indiquée comme étant correcte ne le soit pas, et qu'une autre le soit. Dans ce cas, la réponse doit être modifiée et l'examen doit être noté à nouveau au moyen du bon corrigé. Remarquez que la personne chargée de la révision peut décider qu'il y a plus d'une bonne réponse, auquel cas l'examen devrait être noté à nouveau en utilisant toutes les bonnes réponses indiquées. Dans ce dernier cas, la question devrait être envoyée à l'expert de contenu afin qu'il règle le problème des bonnes réponses multiples avant qu'elle ne soit utilisée à nouveau.

4. Conserver la question à l'examen, mais la renvoyer à l'expert de contenu aux fins de modification

Cette mesure est prise lorsqu'un problème a été soulevé au sujet d'une question, sans qu'il s'agisse d'un défaut majeur.

5. Conserver la question à l'examen

Cette option est utilisée lorsque la question ne contient pas de problème de contenu ni de structure et qu'elle est pertinente pour le cours.

### *Exemples de modèles statistiques courants et de marches à suivre*

Les exemples ci-après présentent des situations courantes qui peuvent survenir en lien avec les questions et explicitent l'interprétation des statistiques. Tous les exemples se rapportent aux questions et statistiques présentées dans le tableau 1.

Question 1 : indice de difficulté = 0,75, indice de discrimination (corrélation bisériale de point) = 0,38  
Comme l'indice de difficulté se trouve entre 0,30 et 0,95, et comme l'indice de discrimination est supérieur à 0,10, cette question serait considérée comme une question à bon rendement et ne serait pas signalée.

Question 2 : indice de difficulté = 0,96, indice de discrimination (corrélation bisériale de point) = 0,05  
L'indice de difficulté pour cette question est supérieur à 0,95, par conséquent, cette question serait considérée comme étant très facile. L'indice de discrimination est inférieur à 0,10, ce qui indique qu'il y a peu de différence entre les étudiants forts et les étudiants faibles. Passez en revue la question pour vous assurer qu'il n'y a pas d'erreur typographique ou d'autre indice concernant la bonne réponse. S'il n'y en a pas, conservez la question à l'examen en tant qu'item noté, mais envoyez-la à l'expert de contenu aux fins de révision avant qu'elle ne soit utilisée à nouveau.

Question 3 : indice de difficulté = 0,25, indice de discrimination (corrélation bisériale de point) = 0,15  
Cette question serait considérée comme étant très difficile, mais l'indice de discrimination supérieur à 0,10 indique que les étudiants forts ont mieux réussi cette question que les étudiants faibles. Examinez le corrigé et vérifiez si les étudiants ont tendance à choisir une autre réponse. Assurez-vous que la question est bien rédigée et vérifiez-en la traduction pour exclure un problème linguistique. Confirmez que la question évalue bel et bien un objectif abordé dans le cadre du cours. S'il n'y a pas de problème évident qui pourrait expliquer l'indice de difficulté bas, cette question pourrait être conservée à l'examen.

Question 4 : indice de difficulté = 0,25, indice de discrimination (corrélation bisériale de point) = 0,07  
Cette question serait considérée comme étant très difficile. Comme l'indice de discrimination se trouve entre 0 et 0,10, il indique que les étudiants forts répondent similairement aux étudiants faibles et ceci pourrait indiquer un problème. Examinez le corrigé et vérifiez si les étudiants ont tendance à choisir une autre réponse. Assurez-vous que la question est bien rédigée et vérifiez-en la traduction pour exclure un problème linguistique. Confirmez que la question évalue bel et bien un objectif abordé dans le cadre du cours. Vérifiez si les étudiants ont fourni de la rétroaction au sujet de cette question. Si aucun problème n'est décelé, la question pourrait être conservée dans l'examen si elle est jugée importante. Dans le cas contraire, envisager de la supprimer tout en attribuant aux étudiants les points correspondants.

Question 5 : indice de difficulté = 0,25, indice de discrimination (corrélation bisériale de point) = -0,12  
Cette question a été notée avec le mauvais corrigé, n'a pas de bonne réponse, évalue un objectif qui n'a pas été abordé dans le cadre du cours, est mal rédigée, ou contient un problème de traduction. Si elle a été notée avec le mauvais corrigé, elle devrait être notée à nouveau en fonction de la bonne réponse. Dans d'autres cas, cette question devrait soit être supprimée ou exclue tout en attribuant aux étudiants les points correspondants.