

Réflexion sur les critères de qualité d'une question

Gilles Leclerc

Rédiger des questions prend du temps. Il devient donc important que ce temps soit utilisé de manière optimale.

Dans un premier temps, et ce avant même d'écrire une question, il est nécessaire de se questionner sur la pertinence des concepts à aborder. Mais par quels critères peut-on déterminer la pertinence ? Il est de notre avis que la pertinence d'écrire une question sur un concept donné devrait reposer sur une analyse de l'authenticité du concept lui-même, sur l'authenticité dans la manière d'aborder le concept, sur l'authenticité de l'effort cognitif demandé à l'étudiant et sur l'authenticité du contexte dans lequel l'étudiant doit se plonger pour répondre à la question.

Critères de pertinence des thèmes abordés		Description
Authenticité	Cognitive	Analyse, raisonnement, synthèse, décision, prise de position <i>*L'étudiant ne devrait pas pouvoir répondre à une question par la simple mémorisation d'un élément factuel qui se retrouve dans un paragraphe isolé d'un livre de référence ou simplement par une application aveugle d'une ligne directrice;</i>
	Professionnelle	Concepts <u>importants</u> , problèmes <u>fréquents</u> ou <u>critiques</u> , étapes <u>cruciales</u> dans un raisonnement ou lors de la résolution de problème, éléments déterminants (éléments clés) Acte ou service professionnel du pharmacien Mobilisation de nombreuses connaissances et habiletés

Dans un deuxième temps, il faut s'attarder à la rédaction de la question. Pour cela, nous référons le lecteur au *Guide de rédaction de question en pharmacie*. Ce guide renferme plusieurs recommandations pour rédiger des questions visant les niveaux cognitifs supérieurs.

Finalement, la qualité d'une question s'observe par la mise à l'épreuve de celle-ci dans le cadre d'un examen. Le rendement d'une question repose principalement sur la capacité d'une question à discriminer adéquatement les étudiants sur la base de leurs habiletés.

Critère de rendement	Description
Discrimination	Discrimination élevée entre les étudiants performants et les étudiants non performants ($>0,3$)

L'analyse d'items et les indices de performance d'une question

***La discussion sur l'interprétation du rendement des questions dans le présent document est uniquement basée sur la théorie classique des tests. Ce choix est déterminé par le fait que les indices fournis présentement aux enseignants reposent exclusivement sur cette théorie.*

Indice de difficulté (p):

Pour une question à correction dichotomique, l'indice de difficulté correspond au taux de réussite de la question (valeur de 0 à 1). *Exemple : $p=0,72$ (72% des étudiants ont répondu correctement à la question)*

Pour une question à crédit partiel, la valeur moyenne du score est utilisée pour établir l'indice de difficulté

Indices de discrimination

Pour le calcul des indices de discrimination, le groupe d'étudiants soumis à l'examen est divisé en trois groupes distincts selon leur score total à l'examen.

- Groupe fort (27% supérieur)
- Groupe faible (27% inférieur)
- Autres (46% des étudiants)

L'indice de discrimination (D)

L'indice de discrimination D correspond à la différence entre l'indice de difficulté(p) calculé avec seulement les étudiants du groupe fort et l'indice de difficulté (p) calculé avec seulement les étudiants du groupe faible (valeur -1 à 1). L'indice de discrimination D peut se calculer pour une question mais également pour chacune des options de réponse. Lorsqu'appliqué à une question, cet indice peut être «positif» advenant que le groupe fort performe mieux à la question ou «négatif» advenant que le groupe faible performe mieux à la question. Lorsqu'appliqué à une option de réponse, cet indice peut être «positif» advenant que le groupe fort a choisi davantage cette option de réponse ou «négatif» advenant que le groupe faible a sélectionné davantage cette option de réponse.

Corrélation «point-bisériale» (r_{pbis})

Les indices corrélationnels de discrimination tel que la corrélation point-bisériale (r_{pbis}) cherche à établir une relation entre le score à la question et le score total à l'examen ou entre le taux de sélection d'une option de réponse et le score total à l'examen. Cet indice peut être «positif» advenant que le groupe fort performe mieux à la question ou «négatif» advenant que le groupe faible performe mieux à la question.

(Valeur théorique : -1 à 1, en pratique : -0,5 à 0,5)

Les ajustements à apporter à une question

Quelle décision prendre selon l'indice de difficulté (p) d'une question ?

Attention :

L'indice de difficulté devrait être considéré avec précaution lorsqu'un nombre important d'étudiant n'a pas répondu à la question. Dans le cas d'une question qui propose des options de réponses, la sélection au hasard ou par chance de la réponse affecte la valeur réelle de l'indice de difficulté.

La cible à atteindre pour l'indice de difficulté d'une question est entre 0,4-0,8. C'est dans cet intervalle que le maximum de variance est observé entre étudiants. C'est ces questions qui amènent le plus d'information au score total de l'examen, c'est-à-dire qui offrent le meilleur rendement. À l'opposé, tant les valeurs trop élevées ($>0,8$) que les valeurs trop faibles ($<0,4$) de l'indice de difficulté traduisent une faible variance entre les étudiants et ainsi un faible rendement envers le score total à l'examen. Une analyse des indices de difficulté permet de repérer rapidement les questions qui ne correspondent pas à l'intervalle optimal. Pour ces questions, une analyse plus approfondie devient nécessaire pour:

- comprendre les causes affectant le niveau de difficulté de la question
- guider les décisions quant au maintien ou non de la question à l'examen (Nous référons le lecteur à *l'Algorithme pour interpréter une analyse d'items*).
- guider les modifications à apporter à la question advenant une utilisation subséquente.

Quelle décision prendre selon les indices de discrimination d'une question?

Quel indice de discrimination utilisé ?

Lorsque le niveau de difficulté de l'item correspond à l'intervalle cible soit entre 0,4 et 0,8, il vous est possible d'utiliser l'un ou l'autre des indices de discrimination.

Dans toute autre circonstance, il est préférable d'utiliser la corrélation «point-bisériale» puisqu'il est possible d'établir un intervalle de confiance et de statuer sur le caractère de signification statistique avec cette statistique.

Comment interpréter les indices de discrimination

Pour les indice de discrimination, la cible visée est une valeur $>0,3$. Les questions dont l'indice de discrimination est $<0,1$ ou négatif méritent une analyse sérieuse quant à la pertinence de conserver ces questions à l'examen. Les questions dont la valeur de D est entre 0,1 et 0,3 présentent une discrimination, jugée limite. Idéalement, toute question dont l'indice de discrimination est $<0,3$ devrait faire l'objet d'une révision avant un usage subséquent.

Pourquoi analyser les options de réponse option ?

Une question à bon rendement présente habituellement une distribution «équilibrée» des mauvaises réponses sur les options de réponses inexactes communément appelées les leurres. Si certains leurres exercent un pouvoir d'attraction trop important comparativement aux autres, une analyse approfondie de la question et des options de réponse s'avère nécessaire pour exclure toutes causes d'erreurs, de biais ou d'iniquités.