

Q.C.M.

Guide à l'usage des E.E. pour la création de questionnaires à choix multiples

Auteurs : Baillifard A. et Thurre-Millius A.
Date : Avril 2020
Temps de lecture : 30 minutes

Sommaire

Table des matières

1	Avant-propos	2
2	Les quatre visées des QCM	3
2.1	Diagnostiquer : un QCM au service de la curiosité	4
2.2	Détecter : le QCM « formatif » au service de l'effort et de la motivation	4
2.3	Former : le QCM « formateur » au service de l'apprentissage	5
2.4	Évaluer : le QCM « sommatif » (validité, étalonnage et notation)	6
3	Conception	10
3.1	La consigne	10
3.2	L'énoncé	10
3.3	La ou les bonne(s) réponse(s) (ou items corrects)	10
3.4	Les distracteurs (leurres, items faux)	11
3.5	Rétroaction formative (feedback)	11
4	Idées reçues	13
4.1	Idée reçue n°1 : la réussite peut être due au hasard	13
4.2	Idée reçue n°2 : La réponse correcte peut se déduire par élimination des distracteurs	14
4.3	Idée reçue n°3 : la conception d'un bon QCM demande du temps à fabriquer	15
4.4	Idée reçue n°4 : Le QCM ne permet pas de mesurer si l'étudiant est capable d'analyser ou de prendre du recul sur la matière.	15
5	Récapitulatif : pièges à éviter et pistes	18
6	Avantages et limites	19
7	Références bibliographiques	20
7.1	Références scientifiques ou pédagogiques	20
7.2	Bibliographie complémentaire	21
7.3	Références non scientifiques	21

1 Avant-propos

En quoi les tests (mini-quiz, QCM, ...) appartiennent au club très sélectif des plus efficaces outils d'apprentissages ?

- a) Ce n'est pas vrai.
- b) Ils obligent l'étudiant à chercher la réponse en mémoire.
- c) Ils consolident les connaissances en les restructurant.
- d) Ils obligent l'étudiant à tâtonner.
- e) Ils augmentent le temps de travail des étudiants.

Réponses correctes : b) et c)

L'enseignant pose une question, l'étudiant répond ; l'enseignant approuve ou désapprouve, prend note des progrès ou des lacunes. L'évaluation des connaissances acquises est un souci central pour tout enseignant consciencieux.

- Comment évaluer ? Comment être sûr que l'on évalue bien ce que l'on veut évaluer ?
- Quelle forme d'évaluation choisir ?
- Comment être certain que l'évaluation créée mesure bien ce que l'enseignant souhaite mesurer ?
- Quel seuil de réussite adopter ?
- Comment vérifier si l'étudiant est capable d'appliquer les connaissances ?
- Etc.

Ici, nous centrerons la discussion autour d'une seule forme d'évaluation : le questionnaire. Nous n'allons pas penser le questionnaire uniquement en tant qu'outil d'évaluation, mais en termes plus larges, car les QCM servent à tout le processus d'apprentissage. Pour Brown et al. (2014), les mini-quiz sont bien plus efficaces que le fait de réviser en relisant. « En nous amenant à rappeler les choses, ils renforcent la mémoire et interrompent le processus d'oubli. Un seul et simple quiz après lecture d'un texte ou à la sortie d'un cours fournit un meilleur outil d'apprentissage et de mémorisation que la relecture du texte ou de ses notes » (p.23).

Vous l'aurez compris, le QCM, outil pédagogique autant répandu que décrié (Bertrand, 2014), ne participe pas qu'à l'évaluation, il permet d'enseigner, de former, de susciter la curiosité, de faciliter les apprentissages, de stimuler, de créer de la motivation... Nous allons voir comment. Ce document vise à servir de vade-mecum à l'usage des enseignants : il n'impose aucune pratique miracle, mais suggère des pistes. En particulier, il invite à recourir au questionnaire comme outil d'apprentissage. Trois idées principales sont véhiculées dans ce document :

Idées clés	Quel chapitre en parle le plus ?
Les hauts niveaux cognitifs (analyse, synthèse, jugement) sont testables via des QCM, bien que ces derniers soient généralement utilisés pour mesurer l'acquisition de connaissances	4.4 Idée reçue n°4
Les feedbacks consolident l'apprentissage, mais aussi la motivation et l'enthousiasme. Les feedbacks, déchargent l'étudiant tantôt de son illusion de savoir tantôt de l'inconfort de faire faux sans comprendre : ils lui permettent d'être un observateur avisé de lui-même.	3.5 Rétroaction formative (feedback)
Un QCM peut servir d'outil d'évaluation sommative à condition d'être bien construit et de respecter des règles de <i>scoring</i> mathématiquement cohérentes.	2.4 Evaluer : le QCM « sommatif » et ses enjeux de validité et d'étalonnage

2 Les quatre visées des QCM

Quel est le but d'un QCM ? En fonction de sa visée, il ne sera pas construit de la même manière. On peut différencier quatre visées :

1. **Diagnostiquer** : susciter la curiosité, introduire un thème, faire le point sur les connaissances de départ, motiver à s'intéresser à une matière en montrant que le bon sens ne suffit pas toujours ;
2. **Détecter** : Faire prendre conscience à l'étudiant qu'il a des notions à apprendre. Il arrive que certains étudiants ne se rendent compte que lors de l'examen qu'ils ne comprennent pas la matière ;
3. **Former** : optimiser l'apprentissage grâce à des feedbacks rétroactifs. Quand il dispose d'un retour sur ses erreurs, l'étudiant apprend plus efficacement ;
4. **Evaluer** : si la question porte sur un objectif d'apprentissage et qu'elle est posée lors d'un examen.

Une même question peut remplir des visées différentes selon le contexte : elle peut revenir à différents moments de l'apprentissage. Par exemple, les deux questions ci-dessous peuvent servir à susciter la curiosité, à situer l'étudiant dans ses connaissances, à le former ou à évaluer ses connaissances.

Combien pèse un nuage ?

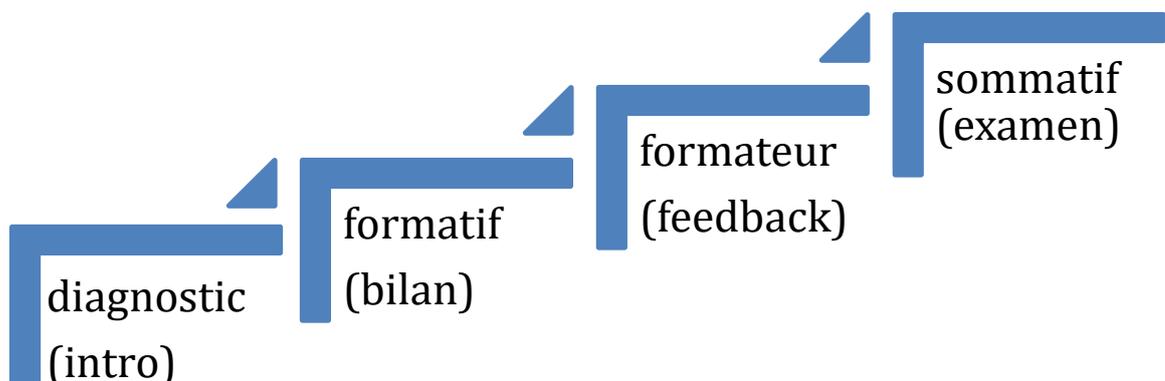
- a) Environ une tonne
- b) Environ 25 tonnes pour un petit
- c) Environ 1000 tonnes pour un petit
- d) Les nuages, même gros, ne pèsent rien
- e) On ne sait pas, on n'a jamais pesé un nuage

Feedback en page 5

Il fait plus froid au pôle Sud qu'au pôle Nord :

- a) À cause des glaciers
- b) À cause des icebergs
- c) À cause de l'altitude
- d) À cause de la latitude
- e) À cause de l'inclinaison terrestre

Feedback en page 5



2.1 Diagnostiquer : un QCM au service de la curiosité

Les tests sont en réalité très utiles avant que l'on soit exposé au contenu à apprendre ou avant de l'avoir assimilé, pourvu qu'un feedback correctif soit fourni par la suite. (Brown, Roediger et McDaniel, 2014)

Il est utilisé en début de formation pour faire le point sur les connaissances de départ ou avant d'introduire un nouveau thème pour voir où se situent les étudiants. Peu utilisé, il peut être utile :

- Pour montrer une évolution : après un présentiel, avant de commencer une période d'apprentissage. On peut rendre disponible un QCM diagnostique à passage unique, puis demander aux étudiants de le compléter à nouveau 3 ou 4 semaines plus tard. En rendant visible les progrès, un QCM diagnostique peut devenir un vecteur de motivation. Les paramétrages de Moodle permettent de rendre visible un QCM durant un laps de temps choisi au préalable ou de fixer un délai entre deux passations.
- Pour adapter son cours : si des prérequis importants ne sont pas acquis ou qu'au contraire des éléments du cours sont déjà acquis.
- De prédisposer les étudiants à s'intéresser / se focaliser sur les bons éléments quand il se pencheront sur le cours.
- Pour l'enseignant qui veut s'adapter à ses étudiants lors de présentiel ou de classe virtuelle (un sondage par exemple).

Etonnamment délaissés, les QCM diagnostiques méritent qu'on s'y intéresse davantage.

Cas particulier : les questions rhétoriques

Les auteurs de « l'encyclopédie qui raconte » (1996) proposent des questions insolubles par leurs jeunes lecteurs, dont voici quelques exemples :

- Pourquoi les ouaouarons ont-ils une troisième paupière ?
- Comment les pigeons voyageurs font-ils pour retrouver leur chemin ?
- Que mange un ornithorynque ?
- Pourquoi les mouches se frottent-elles les pattes ?
- Si les chenilles mangent des feuilles, les papillons en mangent-ils aussi ?
- ...

Les auteurs n'attendent pas de leurs lecteurs qu'ils sachent répondre à ces questions, elles visent plutôt à susciter la curiosité, l'apprentissage et la recherche d'informations.

Avec de telles questions, le lecteur, bien certain d'avoir quelque chose à apprendre, portera toute son attention sur les paragraphes qui viennent ensuite. Ce type de questions peut être utile dans les polycopiés pour introduire chaque chapitre.

2.2 Détecter : le QCM « formatif » au service de l'effort et de la motivation

Le paradoxe des bons cours consiste à créer une illusion de savoir chez les étudiants. Ils tendent à ne pas se rendre compte qu'ils ne maîtrisent pas la matière à cause de la clarté des polycopiés et des enseignants. La facilité à suivre un cours fait naître un sentiment erroné de maîtriser la matière. Vous avez sûrement vécu cette impression désagréable d'être incapable de raconter ce que vous avez appris en cours alors que vous étiez sûr d'avoir tout compris. Le QCM formatif vise à faire le bilan en cours de module ou en cours de semestre. Il vérifie si les étudiants acquièrent bel et bien des connaissances. Il fournit une notation indicative qui ne compte pas dans l'évaluation du cours. C'est un indicateur d'avancement des étudiants. Il peut être effectué en autonomie ou en tant qu'autoévaluation.

Le QCM formatif permet donc de faire le point. En forçant l'étudiant à s'élever au-dessus de son impression de savoir, le QCM formatif lui offre une vision sur l'étendue de son ignorance. Il est donc à double tranchant : d'une part, il peut motiver à se replonger dans la matière, mais d'autre part, il peut être vecteur de découragement. A ce sujet, il peut être utile d'adapter la difficulté de l'ensemble du QCM, afin qu'il ne soit ni trop facile, ni trop décourageant. Ayons à l'esprit qu'une « excessive sévérité [...] démoralise les apprenants ; car ils se désespèrent et ils s'affligent et, ce qui est le plus nuisible, inquiets de tout, ils ne tentent plus rien. [...] Rien ne vaut l'espérance pour faire de l'étude un plaisir » (Quintilien, II.10-13).

Toutefois, le propos change en changeant de perspective. Quand on vise la profondeur d'apprentissage plutôt que la motivation, rien ne vaut la difficulté et la fréquence à laquelle les questions sont posées. En effet, les bénéfices d'apprentissage sont fonctions de la régularité et des efforts cognitifs dans l'acte de remémoration. Des questions obligeant à manipuler plusieurs concepts simultanément et à plonger au plus profond de sa mémoire conduisent le cerveau à consolider ses souvenirs. Ceci rendra plus aisé les actes ultérieurs de rappel à l'esprit. De fait, les parties d'un cours les mieux maîtrisées sont celles qui furent l'objet de petites interrogations en cours de semestre, particulièrement quand elles sont corrigées et commentées (Brown P.C. et al., 2014).

Un enseignant pourrait rendre obligatoire de réussir un QCM avec un certain seuil de réussite pour que l'étudiant puisse se rendre à l'examen ou rendre une dissertation. Par cela, il transmettrait le message selon lequel certains contenus doivent être maîtrisés avant de pouvoir être manipulés.

2.3 Former : le QCM « formateur » au service de l'apprentissage

Des ajouts pertinents de textes (piste, indice, feedback, commentaire, aide) autour du questionnaire facilitent la révision et l'entraînement, aident à progresser et réduisent l'opacité de l'évaluation. Dans ce cas, le QCM formatif devient QCM formateur. Un QCM formateur intègre une forme particulière de paratexte : des **feedbacks formatifs**. Nous séparons QCM formatif et QCM formateur pour insister sur l'importance de ces feedbacks (aussi appelés « **rétroactions formatives** »). Il s'agit pour l'enseignant d'apporter quelques indications pour faciliter les apprentissages. Ces indications peuvent être des aides de résolution, des éclaircissements, des réponses détaillées, des perspectives, des prises de recul sur l'objectif traité, des prolongements... Ces rétroactions sont souvent offertes en cas de réponses fausses, mais il est aussi possible d'offrir un feedback positif ouvrant vers une perspective nouvelle en cas de réponse positive.

Feedbacks facilitatifs pour les questions de la page 3 :

Comment échapper à la nécessité de placer après certaines questions, un commentaire qui permette d'en juger la subtilité ? Les feedbacks ci-dessous importent plus que les questions. Les questions suscitent la curiosité du lecteur et les feedbacks le motivent à aller plus loin. Si avant même d'y répondre, l'étudiant en connaissait les perspectives ou les enjeux, les questions seraient superflues.

a) Poids d'un nuage

Environ 25 tonnes pour un petit. Sans avoir placé un nuage sur une balance, on peut en estimer la masse. On évalue son volume et donc sa quantité d'eau et de glace (un nuage contient le millionième de son poids en eau). Un petit nuage d'un volume de 20 à 30 millions de m³ ne contiendra qu'un volume de 20 à 30 m³ d'eau.

b) Pôle Sud glacial

Il fait plus froid au pôle Sud qui est plus élevé que le pôle Nord qui est à plat sur la banquise, où il fait tout de même moins 30 degrés. Au pôle Sud (2800m), la température descend à moins 60 degrés.

Questions et réponses adaptées de Quéva R., (2008), *La science est un jeu*, pp. 120-128.

Les QCM formateurs réduisent l'opacité des objectifs et des attentes de l'enseignant, augmentent l'intérêt pour la tâche et pour le domaine (en élargissant, en donnant des perspectives, en proposant un exemple pour illustrer la règle), diminuent les taux d'abandon ou d'échec. C'est principalement vers ce type de QCM

que nous aimerions sensibiliser les équipes enseignantes. La notion de feedback « facilitatif » (décrite par Shute, 2008) vise à guider l'étudiant avancé dans sa compréhension plutôt qu'à lui dire ce qu'il doit faire. On évitera des feedbacks du type : « La réponse est fausse » ou « Voir tel livre à tel page ». Bransford et al. (2000) proposent de :

- Donner des pistes (aide dans l'énoncé ou aide facultative, indice) pour accomplir la tâche ;
- Indiquer l'écart entre la réponse fournie par l'étudiant et la réponse attendue ;
- Définir clairement ce qui est attendu dans l'activité d'apprentissage (notation, barème, niveau d'exigence, nombre de réponses correctes).

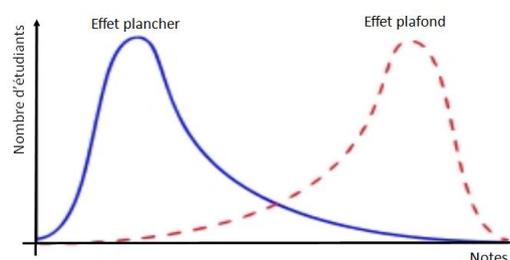
Bien sûr, le feedback facilitatif ne devrait pas permettre d'être copié pour répondre à la question et apparait après que l'étudiant ait répondu. Voici quelques pistes :

- Donner un moyen mnémotechnique pour se souvenir d'une subtilité ;
- Donner la raison pour laquelle telle réponse est correcte ;
- Donner des informations sur le contenu du cours étudié auquel fait référence la question ;
- Donner la démarche à suivre sans donner la réponse ;
- Expliquer ce qui est faux et pourquoi : donner des informations sur les erreurs ;
- Indications stratégiques sur la manière de procéder : « Pour trouver le maximum de cette fonction « f », il convient de passer par sa dérivée. La dérivée d'une fonction étant sa pente, quand la dérivée de « f » vaut zéro, alors la pente de la fonction « f » est *horizontale*. Ensuite, à vous de détecter si vous êtes face à un maximum ou à un minimum » ;
- Offrir un éclairage sur un cas réel qui s'articule avec la question ;
- Offrir un éclairage sur un débat scientifique en lien avec le thème abordé.

2.4 Evaluer : le QCM « sommatif » (validité, étalonnage et notation)

En fin de semestre, le QCM sommatif est un outil d'évaluation des connaissances acquises. Sa visée n'est plus de transmettre une connaissance mais de la mesurer le plus objectivement possible en créant un classement des résultats. Discriminer et mesurer l'acquisition des compétences sont les deux dimensions d'un QCM sommatif. On peut parler d'étalonnage et de validité.

Un test est bien **étalonné** s'il est capable de différencier les résultats de chaque étudiant. Idéalement, il distribue les étudiants en une gaussienne. La forme de la distribution des résultats d'un QCM trop facile sera asymétrique à droite (effet plafond) et inversement pour un QCM trop difficile. La forme de la distribution indique si un test discrimine bien les résultats, mais ne dit rien sur l'adéquation entre la difficulté du test et les objectifs atteints (validité).



Un effet plafond peut tout à fait signifier que la volée d'étudiants est plus homogène et compétente qu'une autre volée. La notion d'étalonnage pose un problème moral. En classant les résultats selon une loi normale, en attribuant à chacun une lettre ou un nombre, on projette chaque étudiant sur une seule dimension, on perd de l'information. A partir des aptitudes (en logique, en expression, ...) et des savoirs (disciplinaires ou non) enchevêtrés de l'étudiant, l'enseignant en donne une synthèse : la note. Mais à partir de la note, il est impossible de revenir aux aptitudes de départ. Gardons cela à l'esprit autant lors de l'élaboration de nos tests que de l'interprétation des résultats.

Le test est **valide** s'il mesure bien ce qu'il est censé mesurer. Le QCM sommatif ne saurait valider la somme des acquis cognitifs, méthodologiques et techniques de l'étudiant qu'en posant des questions centrées sur les objectifs travaillés durant le semestre. Il ne comporte ni indice ni ambiguïté. Chaque test non valide donne raison à Kalz (dans Dessus et Tarbouriech, 2019) : « tout étudiant qui connaît sur le bout du doigt son système éducatif sait que, s'il est facile d'éviter un mauvais cours, il est en revanche plus difficile d'éviter une mauvaise évaluation » (p.91). Un test peut être valide et non étalonné si les étudiants maîtrisent parfaitement la matière et obtiennent tous un excellent résultat. Il est ainsi plus important de proposer des tests valides que des tests bien étalonnés.

Il existe différents **types de questions** possibles : les QCM, les vrai/faux, les questions à appariement (où on associe deux informations entre elles), les questions à réponse numérique et les questions à réponse courte. Il est possible d'automatiser la correction de réponses courtes en prévoyant à l'avance toutes les réponses possibles. Pour les QCM diagnostique, formatif ou formateur, nous encourageons prioritairement l'utilisation des questions à correction automatisée, pour encourager l'autonomie de l'apprenant. Pour les QCM sommatif, nous déconseillons les vrai/faux (la probabilité de répondre juste au hasard est plus grande que s'il y a plusieurs propositions). A l'inverse, nous encourageons l'intégration de questions à réponses courtes (avec correction de l'enseignant) : elles permettent à l'étudiant de rédiger une réponse personnelle, diminuent les chances de répondre au hasard et limitent les possibilités de tricher.

Une des craintes des QCM sommatif est que le test ne soit pas représentatif du niveau de l'étudiant et que le résultat soit, en partie, dû au **hasard**. Il y a plusieurs solutions pour limiter l'influence du hasard dans un QCM sommatif :

- Statistiquement, plus il y a de propositions par question, moins il y a de chance de répondre juste au hasard. Nous conseillons entre 4 et 5 items par question.
- Il est possible de demander à l'étudiant de préciser son **degré de certitude** pour chacune de ses réponses et d'en tenir compte dans la notation. Le tableau ci-dessous (Prosperi, 2015) donne un exemple de la manière de compter les points en fonction du degré de certitude de l'étudiant et de la justesse de sa réponse.

Score attribué à un étudiant en fonction de l'exactitude de la réponse et du degré de certitude choisi

		Degrés de certitude					
		0	1	2	3	4	5
Réponse	Correcte	+13	+16	+17	+18	+19	+20
	Incorrecte	+4	+3	+2	0	-6	-20

- La manière d'attribuer des points (notation), permet également de limiter le hasard en pénalisant les mauvaises réponses. Si l'étudiant perd des points en répondant faux, il prendra le temps de la réflexion et ne prendra pas le risque de répondre au hasard.

Une deuxième crainte liée au QCM sommatif est la possibilité de **tricherie** par les étudiants. Différentes fonctionnalités (sur Moodle ou ailleurs) permettent de limiter cette problématique :

- Il est possible de choisir de manière aléatoire un certain nombre de questions dans une base de données. Ainsi, tous les étudiants ne répondront pas aux mêmes questions. Attention à créer des questions d'un niveau comparable pour que les résultats puissent être comparables.
- Il est possible de rendre l'ordre d'apparition des questions aléatoire, pour que chaque étudiant découvre les questions dans un ordre chaque fois renouvelé.
- Les réponses à chaque question peuvent apparaître dans un ordre aléatoire.
- Nous vous conseillons, à l'inverse d'un QCM formatif ou formateur, de ne pas afficher la bonne réponse à l'étudiant, ni de proposer un feedback pendant l'évaluation certificative. Ce feedback peut venir plus tard.
- Moodle offre la possibilité de ne rendre le QCM disponible que sur une période courte (une heure ou une demi-journée, par exemple) pour limiter la possibilité de communiquer entre les étudiants ou pour empêcher des stratégies de recherche sur le Web.
- Proposer des questions à réponses courtes pour avoir une réponse personnelle de l'étudiant.

La question centrale pour un QCM sommatif est celle de la **notation**. Pour les QCM diagnostique, formatif ou formateur, la notation importait peu, car ces outils visaient l'apprentissage. En revanche, la notation devient indispensable quand l'objectif est de vérifier les acquis de l'étudiant.

Il y a plusieurs questions à se poser pour créer son barème de notation :

- Y a-t-il une ou plusieurs réponses possibles ?
- Toutes les réponses sont-elles valorisées de la même manière ou certaines sont plus importantes que d'autres, voire éliminatoires ? La pondération des questions dépend de la pondération des objectifs d'apprentissage : quelle notion est centrale ? Laquelle est plus difficile ou moins importante ? Il est possible soit d'avoir le même barème pour toute l'évaluation ou alors de pondérer par question (les questions plus longues, plus compliquées ou plus importantes valent plus de points) ou par proposition (pénaliser une erreur grave mais pas une autre).
- Une mauvaise réponse est-elle pénalisée ? Si oui, de quelle manière ?
- Est-ce que l'absence de réponse est pénalisée ou non ?

Le tableau ci-dessous (inspiré de Bouvy et Vanhove, 2005) représente les différentes manières de noter, selon qu'il y a une ou plusieurs réponses correctes.

Barèmes classiques pour des QCM ($k = \text{nombre de solutions proposées}$, $c = \text{nombre de solutions correctes}$)							
Barèmes	Nombre de points				Avantages	Inconvénients	
	Une réponse correcte	Réponse correcte		Absence de réponse			Réponse incorrecte
		Plusieurs réponses correctes	Toutes les bonnes réponses				Une partie des bonnes réponses
Simple	Une réponse correcte	+1		0	0	Simple d'utilisation. Encouragé pour QCM diagnostique, formatif, formateur. Possibilité de mettre des points à chaque bonne réponse ou seulement quand elles y sont toutes.	Manque de subtilité pour départager les candidats. Favorise le hasard. Déconseillé pour QCM sommatif.
	Plusieurs réponses correctes	+1	+1/c	0	0		
			0				
Symétrique	Une réponse correcte	+1		0	-1	Pénalise le hasard.	La note peut facilement être nulle voire négative.
	Plusieurs réponses correctes	+1	+1/c	0	-1/c		
Double	Une réponse correcte	+1		0	-0,5	Pénalise le hasard de manière moins brutale. Idéal pour les QCM sommatifs	Pas toujours populaire
	Plusieurs réponses correctes	+1	+1/c	0	-0,5/c		
Avec valorisation de l'absence de réponse	Une réponse correcte	+1		+1/k	0	Evite le hasard, car pénalise les réponses fausses mises au hasard. Idéal pour les QCM sommatifs.	Pas toujours populaire.
	Plusieurs réponses correctes	+1	+1/c	+1/k	0		
Barème selon concordance des réponses	Le nombre de point dépend du nombre d'erreurs global, indépendamment des questions.				Le candidat qui fait plus d'erreurs est plus pénalisé que celui qui en fait moins.	Les erreurs sont quantifiées, on ne s'intéresse pas ici au contenu des erreurs.	
Conformité relative	Résultat = $\text{nbr. réponses justes du candidat} / \text{nbr. réponses justes totales} * \text{point à la question} + \text{nbr erreurs du candidat} / \text{nbr erreurs totales} * \text{point de pénalité}$				Si les points par question sont égaux aux points de pénalité, le barème est équilibré.	Il est impossible d'avoir aucune réponse vraie ou aucune réponse fausse pour ce barème.	
Conformité relative avec coefficient sur réponse	Résultats = $\sum \text{des coefficients de réponses justes du candidat} / \sum \text{des coefficients des réponses justes totales} * \text{point à la question} + \sum \text{des coefficients des erreurs du candidat} / \sum \text{des coefficients des erreurs totales} * \text{points de pénalité}$				Permet de donner plus ou moins d'importance à une proposition de réponse : une erreur grave ou une réponse fondamentale auront un fort coefficient et inversement.		

3 Conception

3.1 La consigne

La consigne décrit le questionnaire, le mode d'évaluation, les modes de réponses voire les attentes de l'équipe enseignante. A tort, les consignes sont souvent délaissées, plongeant l'étudiant dans l'opacité et l'inconfort. Ignorant les attentes et la sauce à laquelle il sera mangé, l'étudiant ne peut consacrer l'entièreté de ses ressources cognitives à la résolution du QCM. L'apport principal des consignes est de diminuer les malentendus en augmentant la transparence :

- Expliciter le nombre de réponses correctes (une seule, plusieurs ou encore une seule en principe bien qu'une indication supplémentaire puisse déroger à la règle pour certaines questions)
- Expliciter les principes de notation : donner le barème du QCM (y a-t-il des points négatifs en cas de réponse incorrecte ou d'absence de réponse, certaines questions sont-elles pondérées différemment, certaines sont-elles éliminatoires, faut-il que tous items soient justes si plusieurs réponses sont correctes, ... ?)
- Visée du QCM : il est bon de préciser si le QCM vise à faire le bilan des connaissances (diagnostic), à consolider les apprentissages (instructif/formateur), à faire un bilan des apprentissages (formatif) ou à noter l'étudiant (sommatif).

La consigne devrait être écrite, transparente et valable pour l'ensemble du QCM.

3.2 L'énoncé

Avant de rédiger l'énoncé, il convient d'avoir à l'esprit l'objectif qu'il vise. On se concentre sur un objectif par énoncé. On préférera :

- Un énoncé concis et complet, en principe sans information superflue ;
- Que les éléments nécessaires pour répondre se trouvent dans l'énoncé plutôt que dans les items ;
- Un énoncé qui évite une réponse en vrai / faux ;
- On évite généralement la forme négative. On peut utiliser la forme interrogative : cela stimule directement l'étudiant et facilite la construction des réponses (moins de soucis grammaticaux ou d'indices dans la bonne réponse). La forme affirmative permet notamment de ne pas répéter une structure commune à toutes les propositions.

Exemple d'un énoncé complet :

En 1898, Emile Zola adresse une lettre ouverte au président de la République française. Son « J'accuse... ! » dénonce la condamnation du capitaine Alfred Dreyfus. Quel quotidien en fit sa une ?

a. Le Matin

b. Aujourd'hui

c. L'Aurore

3.3 La ou les bonne(s) réponse(s) (ou items corrects)

- Les bonnes réponses doivent être incontestablement **exactes**.
- Si possible, elles **ressemblent aux distracteurs** (longueur, construction, précision, ...), sans quoi la dissemblance pourrait fournir un indice à l'apprenant.
- Sa position parmi les propositions doit varier aléatoirement au long du test. On peut favoriser une logique interne dans certaines questions (ordre alphabétique, numérique, chronologique, ...)
- Il est possible que la réponse correcte soit « aucune de ces réponses ». L'option « toutes ces réponses » n'est pas conseillée si l'étudiant sait qu'il n'y a qu'une réponse correcte. En effet, s'il connaît deux des items corrects, alors il choisira « toutes ces réponses ».

- Mieux vaut éviter les indications focalisant l'attention sur 2-3 items (deux réponses qui se ressemblent beaucoup risquent d'éliminer les autres). Evitez également les convergences entre les réponses (la réponse avec le plus grand nombre d'élément en commun avec les autres réponses a de grandes chances d'être la bonne réponse) ou entre l'énoncé et la bonne réponse (la réponse qui reprend le plus de mots-clés compris dans l'énoncé peut donner l'impression d'être la bonne réponse).

Exemple d'une réponse exacte ressemblant aux distracteurs

Qu'est-ce que le montant auquel un vendeur est payé moins son coût ?

- a) Le surplus du consommateur
- b) Le surplus du producteur
- c) Le bénéfice net
- d) L'excédent net d'exploitation.

3.4 Les distracteurs (leurres, items faux)

Les distracteurs sont les mauvaises réponses. La difficulté du QCM repose en grande partie sur les distracteurs qui doivent :

- Être incontestablement faux... mais subtils. Idéalement, le distracteur ne répond qu'à une partie du problème ou semble juste avant analyse.
- Eviter les distinctions minimales entre eux.
- Être nombreux (au moins 3) pour diminuer les chances de répondre au hasard.
- Ressembler à la bonne réponse (longueur, complexité, style, formulation, ...).
- Eviter les complexités qui ne sont pas en lien avec la question.
- Eviter les termes vagues (habituellement, parfois, ...) ou absolus (toujours, jamais, ...)
- Un item peut être « Je ne sais pas » et être moins pénalisé si l'étudiant le choisit, sauf peut-être dans un QCM sommatif. Le premier pas vers le savoir est souvent un aveu d'ignorance.

Par quel capteur percevons-nous le chaud ?

- a. Bulbes de Krause
- b. Corpuscules de Meissner
- c. Terminaison nerveuse libre
- d. Corpuscule de Ruffini
- e. Corpuscules de Golgi-Mazzoni
- f. Disques de Merkel
- g. Corpuscule de Pacini

Dans l'exemple proposé, l'étudiant averti comprend qu'il doit aussi étendre son apprentissage aux autres récepteurs cutanés (pression, douleur, ...). Pour éprouver la qualité d'un distracteur, Johnstone (2003) suggère qu'environ 5% des utilisateurs devraient le choisir. Si personne ne choisit un distracteur, il peut être enlevé. On peut s'inspirer des mauvaises réponses données habituellement par des étudiants, des confusions et erreurs fréquentes ou encore jouer sur les fausses idées reçues.

3.5 Rétroaction formative (feedback)

La plateforme Moodle permet de fournir un feedback automatisé. Cette option est fondamentale : elle transforme un QCM formatif (QCM bilan) en QCM formateur. Les rétroactions donnent à l'étudiant des indications sur ses échecs ou sur ses mécompréhensions. Elles lui font gagner du temps sur ses apprentissages en rendant explicites ses difficultés. Quand l'étudiant n'a pas conscience de la profondeur de son ignorance, il ne peut se mobiliser pour la pallier. Il y a deux composantes de rétroaction : celle qui valide ou non la réponse (« réponse correcte » ou « réponse fausse ») et celle qui explique. Cette seconde nous intéresse davantage : elle vise à faire comprendre à l'étudiant ses erreurs.

Les feedbacks formatifs explicatifs se veulent utiles, courts et complets, c'est-à-dire : spécifiques au niveau attendu du cours, fréquents, **immédiats**¹. L'enseignant peut :

- Expliquer, comme il l'a déjà fait dans son cours. On peut copier le paragraphe pertinent du cours.
- Préciser un raisonnement ou une étape piège d'un raisonnement.
- Exclure les distracteurs pernecieux en expliquant pourquoi. Dans le cadre d'un QCM formateur, il peut être efficace de choisir exceptionnellement des distracteurs portant à confusion accompagné ensuite d'un feedback clair pour dissiper la confusion. Dans ce cas, le distracteur alambiqué vise à susciter la curiosité, l'intérêt ou la pensée critique chez l'apprenant.

Deux exemples de rétroaction

Pour l'audition, le mécanisme de transduction est réalisé par les

- a. photo-récepteurs
- b. papilles
- c. terminaison nerveuse libre
- d. cellules ciliées
- e. bulbes de Krause

La réponse correcte est « cellules ciliées »

ou

La réponse correcte est d.

A long terme, tous les coûts sont variables et aucun coût n'est fixe.

a. Vrai

b. Faux

La distinction entre le court terme et le long terme (en économie) est la suivante : à court terme certains facteurs de production sont fixes tandis qu'à long terme la quantité de tous les facteurs de production peut être modifiée. En d'autres termes, à long terme tous les facteurs de production sont variables. La notion de long terme dépend du problème que nous analysons. Le long terme sera la période de temps nécessaire à l'entreprise pour modifier le facteur de production qui à court terme est fixe. Les coûts fixes de l'entreprise sont les coûts correspondant à des rémunérations de facteurs que l'entreprise ne peut pas ajuster à court terme par conséquent ces coûts ne varient pas avec le niveau de production.

Dans le premier exemple proposé, le feedback n'apporte aucune plus-value à l'étudiant, tandis que le second est instructif. Ceci ne signifie pas que le premier exemple est mauvais : l'enseignant ne fournit pas systématiquement des précisions, il choisit les questions pour lesquelles les précisions rétroactives lui semblent précieuses et profitables.

Mauvaise pratique : une rétroaction inutile proposée par Fernand Raynaud, 1976 :

Toto : Je voudrais te demander : les sous-marins atomiques, comment ils peuvent fonctionner dans l'eau ? Parce que, il paraît que d'un côté, il n'y a pas de moteur Diesel, il n'y a pas de moteur électrique... On en parlait justement encore à la récréation... Comment ils peuvent fonctionner dans l'eau, dis papa, les sous-marins atomiques ?

Le père : Hum... hum... Mon vieux... C'est étudié pour !

Combien de fausses explications, de « *Hum... hum... Mon vieux... c'est étudié pour !* », l'étudiant a-t-il eu l'impression d'entendre au cours de son parcours académique ?

Conclusion :

Une rétroaction utile permet à l'étudiant de mesurer l'écart entre son niveau et le niveau de référence attendu par l'enseignant. L'information donnée dans la rétroaction permettra à l'étudiant de réduire cet écart.

¹ Moodle propose l'option feedback immédiat ou après avoir complété le quiz entier. Le feedback immédiat est plus efficace, car il permet de corriger en direct les fausses conceptions de l'apprenant.

4 Idées reçues

Nous allons présenter ci-dessous les idées reçues les plus fréquemment rencontrées à propos des QCM, qui sont notamment des craintes que vous partagez peut-être, et nous les commenterons en disant en quoi elles sont vraies, fausses ou les deux.

4.1 Idee reçue n°1 : la réussite peut être due au hasard

VRAI : Quel correcteur saurait différencier une réponse « juste par hasard » d'une réponse « juste par connaissance » ? Ce problème diminue (sans disparaître) si le nombre de distracteurs augmente. Ce problème disparaît quand on quitte la forme QCM et que l'étudiant s'exprime à travers une réponse écrite élaborée.

FAUX : Cette idée reçue s'effondre dès le premier exemple (René Krebs, 2004) :

Un patient de 47 ans se présente aux urgences après une chute à vélo. Il est agité et présente un intense foetor éthylique. Il se plaint de douleurs thoraciques dans la région de la base gauche et dans l'hypochondre gauche. Le score de Glasgow est à 15. Il n'y a aucune plaie apparente. La fréquence respiratoire est de 38, les pulsations sont à 140, la tension artérielle à 90/60 mmHg. Quelle est l'explication la plus vraisemblable du tableau hémodynamique ?

- (A) un choc cardiogène
- (B) une hypovolémie
- (C) une embolie pulmonaire
- (D) une conséquence de l'intoxication éthylique
- (E) un réflex vagal post-traumatique

On voit ici qu'un énoncé et des distracteurs bien choisis rendront une question inaccessible pour des étudiants n'ayant pas suffisamment intégré la matière. Remarquons également que cette question repose sur un traitement cognitif de haut-niveau : l'étudiant doit synthétiser les informations et appliquer ses connaissances pour y répondre.

VRAI : Si l'enseignant propose deux items par question, qu'il offre un point par question juste et zéro point par réponse fausse, alors les résultats obtenus par les étudiants ne reflèteront pas leurs connaissances. En effet, si l'étudiant choisit systématiquement au hasard, il obtiendra la moitié des points (c'est-à-dire l'équivalent de la note 3.5). De plus la distribution des résultats sera asymétrique à droite (effet plafond). On peut limiter les effets du hasard en augmentant le nombre de propositions à 3 puis à 5, mais il restera malgré tout un résidu de la note dû au hasard.

Ce mode de pondération à conformité stricte souvent utilisé pour sa facilité d'utilisation est un barème de notation sommative déraisonnable.

FAUX : Car le point développé ci-dessus est contournable via différentes **modalités de notation** qui visent à déjouer le hasard ([voir page 8](#)). Ces méthodes prennent en compte le nombre de distracteurs et nécessitent des calculs plus ou moins compliqués, nous proposerons ici une méthode alternative simple.

Avant de la présenter, rappelons les deux fonctions d'un test sommatif : mesurer ce qu'il est censé mesurer (validité) et discriminer les résultats (étalonner). Une méthode consiste à définir un seuil minimal puis à **étalonner** les scores bruts en scores prédéfinis. A partir du seuil minimal (0 point) et le maximum possible, on distribue les scores intermédiaires de façon linéaire.

Supposons qu'un questionnaire comporte 20 questions. On peut définir un seuil minimal de réussite à 10 questions correctes. Ensuite, par une relation linéaire, on fixe les paliers suivants :

- | | |
|---------------------------------------------------|-----------|
| - 20 questions correctes : maximum | 10 points |
| - 18 questions correctes : | 8 points |
| - 15 questions correctes : à mi-chemin du maximum | 5 points |
| - 10 questions correctes : seuil minimal | 0 point |
| - 5 questions correctes : insuffisant | 0 point |

Cette procédure évite, en partie, de noter le hasard, remplit la fonction de discriminer les résultats entre les étudiants, a l'avantage d'être simple à mettre en place et laisse à l'enseignant fixer le seuil du niveau attendu. Pourtant, elle peut ne pas être populaire, car l'étudiant sera déçu d'obtenir seulement 1 point sur 10 pour 11 réponses correctes. Il oublie que son résultat découle d'une maîtrise partielle de la matière. Il a probablement répondu avec certitude à certaines questions, par élimination à d'autres et par hasard aux dernières.

A chaque enseignant la responsabilité de choisir sa méthode de notation, en gardant à l'esprit que le risque de donner des résultats biaisés n'est pas négligeable.

4.2 Idee reçue n°2 : La réponse correcte peut se déduire par élimination des distracteurs

VRAI : Combien d'étudiants ont obtenus des notes non conformes à leurs connaissances en devinant certaines réponses ? Probablement tous ! Combien d'apprenants se sont fait piéger par des distracteurs ou une formulation alambiquée alors qu'ils connaissaient la bonne réponse ? Probablement tous !

FAUX : Si les distracteurs sont bien choisis, la réponse ne peut pas être déduite.

<u>Mauvais exemple</u>	<u>Meilleur exemple</u>
<p>La baleine bleue est parmi les plus longs animaux vivants sur Terre. Quelle peut-être sa taille ?</p> <ul style="list-style-type: none"> - Plus de 3 kilomètres - Plus de 30 mètres de longueur - Plus de 30 centimètres de longueur - Plus de 300 mètres de longueur <p><i>A moins d'ignorer les unités de longueur, on peut évidemment deviner la réponse</i></p>	<p>La baleine bleue bat des records de taille et de poids. Parmi ces propositions, laquelle est correcte ?</p> <ul style="list-style-type: none"> - Plus de 30 mètres, mais moins de 170 tonnes - Plus de 40 mètres et plus de 170 tonnes - Plus de 40 mètres, mais moins de 170 tonnes - Plus de 30 mètres et plus de 170 tonnes. <p>Les distracteurs empêchent de facilement déduire la bonne réponse.</p>

FAUX : Si la question nécessite un traitement (calcul, réflexion), l'étudiant ne peut plus deviner par élimination. Exemple : « Quel âge a un homme s'il est 4 fois plus âgé que son fils et que dans 20 ans, il aura 2 fois son âge ? »

Malheureusement FAUX : Chaque fois que l'enseignant met un piège que l'étudiant ne décèle pas. Ce dernier fera faux par automatisme, par un traitement cognitif manquant de profondeur. Le déclenchement de mécanismes automatiques est une cause fréquente de bévues. L'étudiant traite rapidement la question et répond à ce qu'il pense avoir compris, plutôt qu'à ce que la question signifie. Borst G. (2016) remarquait :

Faire prendre conscience à un enfant [étudiant] qu'il est devant un piège est fondamental pour qu'il puisse le déjouer. En imagerie fonctionnelle, nous avons montré que le cerveau passe de l'erreur à la réussite en se reconfigurant. Il y a un basculement de l'activation cérébrale de la partie

postérieure du cortex - impliquée dans les automatismes - au cortex préfrontal, zone du blocage des heuristiques.

L'étudiant devrait pouvoir bloquer ses mécanismes de réponse automatiques pour activer sa pensée profonde. Dire à un étudiant : « attention, piège » suffit à le faire basculer d'une pensée automatique à une réflexion plus profonde, à passer d'une attitude passive à un apprentissage actif.

4.3 Idee reçue n°3 : la conception d'un bon QCM demande du temps à fabriquer

VRAI : La conception et la fabrication d'un QCM demande du temps. En particulier quand il s'agit d'un QCM sommatif. Toutefois, quand il s'agit d'un QCM formatif ou formateur, on peut être plus nuancé pour quatre raisons :

- Le niveau de perfection dans la conception d'un QCM formatif n'est pas le même que pour un QCM sommatif sur lequel repose la note finale de l'étudiant.
- Moodle propose de remplir des feedbacks et des aides qui apportent autant de béquilles aux imperfections de la question. Les maladroites de formulation ou les ambiguïtés dans les distracteurs peuvent être dissipées par un feedback éclairant ou une aide mettant en garde.
- Le QCM formatif ou formateur est réutilisable chaque année donc améliorable et enrichissable au fil des semestres.
- Si dans les consignes, l'enseignant invite les étudiants à être co-acteurs de ses QCM (en signalant des erreurs, des coquilles, des remarques), il disposera à la fois d'un retour utile et d'un contact avec ses étudiants qu'ils ne rencontrent que peu.

FAUX : Si on élargit l'échelle temporelle jusqu'à la correction, alors on peut prendre en compte le gain de temps en termes de feedback automatisé et d'apprentissages autonomes par les étudiants. Tout ce temps est libéré pour l'enseignant.

FAUX : Nombre de questions, une fois éditées, peuvent être déclinées en plusieurs versions.

4.4 Idee reçue n°4 : Le QCM ne permet pas de mesurer si l'étudiant est capable d'analyser ou de prendre du recul sur la matière.

TROMPEUR : Cette idée reçue est autant tenace qu'insidieuse. L'interrogation est légitime, mais mal posée. La question n'est pas de savoir si les étudiants doivent connaître, comprendre, appliquer des concepts ou porter des jugements. La question n'est pas non plus de savoir s'il faut favoriser l'acquisition de connaissances ou la capacité à former des jugements. La question est de savoir comment développer des connaissances ET former des jugements, « car sans connaissances on ne dispose pas des fondements nécessaires au développement de capacités plus avancées telles qu'analyse, synthèse et raisonnement créatif » (Brown P.C. et al., p.39).

A VOUS DE JUGER : Pour répondre à cette idée reçue, on peut choisir deux approches qui nous conduiront à des conclusions voisines : une approche cognitive (1) et une approche pédagogique (2) :

1. Pour développer ses aptitudes, l'étudiant mobilise nombre de processus mentaux et de facultés cognitives (mémoire, raisonnement, résolution de problèmes, prise de décision, attention, ...).
2. Les pédagogues aiment se référer à une classification par niveau d'acquisition des connaissances : les domaines de la connaissance, de la compréhension, de l'application (analyse), de la synthèse et de l'évaluation. (Taxonomie de Bloom).

Dans le contexte de la création de QCM, on peut donc se poser deux questions :

1. Les QCM ne mobilisent-ils que partiellement les facultés cognitives et déboucheraient ainsi sur un apprentissage parcellaire ou superficiel ?
2. Les QCM permettent-ils d'atteindre les niveaux supérieurs de connaissance (analyse, évaluation et synthèse en particulier) ?

Dans les deux cas, le souci de l'enseignant est d'éviter un apprentissage superficiel, manquant de profondeur. L'enseignant aimerait former des citoyens capables de nuances plutôt que des hommes savants. L'enseignant soucieux marche sur les traces de Montaigne :

Que le maître ne demande pas seulement à son élève de lui répéter les mots de sa leçon, mais de lui en donner le sens et la substance. Et qu'il juge du profit qu'il en aura tiré, non par le témoignage de sa mémoire, mais par celui de son comportement. [...] Régurgiter la nourriture telle qu'on l'a avalée prouve qu'elle est restée crue sans avoir été transformée : l'estomac n'a pas fait son travail, s'il n'a pas changé l'état et la forme de ce qu'on lui a donné à digérer. (Essais I,26).

Le QCM peut-il apporter du sens et de la substance à la leçon ? Le QCM fait-il changer l'état et la forme de ce que l'étudiant apprend ? Les QCM peuvent-ils mesurer les capacités d'analyse ou d'application des étudiants ? On aimerait répondre trois fois non et pourtant il convient de se méfier des évidences ! Tous les QCM ne mesurent pas imparfaitement les connaissances en se focalisant sur la mémorisation, car ils peuvent atteindre d'autres niveaux cognitifs. Regardons de plus près trois exemples :

	Enoncé	Items proposés à relier avec les...	...possibilités
Question 1	<p>Pour chaque définition (Fromkin, 1973) retrouvez le type de lapsus concernés.</p>	<p>a) Le locuteur remplace un mot par un autre dont la sonorité est similaire mais le sens différent.</p> <p>b) Le locuteur substitue un élément de la langue à la place d'un autre.</p> <p>c) Le locuteur intervertit les positions de deux éléments de la langue.</p> <p>d) Le locuteur utilise un élément de la langue avant son emploi approprié dans la phrase parce qu'il correspond à un élément qui doit apparaître plus tard dans l'énoncé.</p> <p>e) Le locuteur utilise un élément de la langue qui convenait plus tôt dans la phrase mais qui est impropre par la suite.</p>	<ul style="list-style-type: none"> • perversion, • inversion • substitution • anticipation • incongruité.
		<p>Cette question mobilise des facultés de raisonnement, d'attention, de mémoire, ... Elle se situe au niveau cognitif de l'application, car l'étudiant doit inférer à partir d'une règle, il quitte les niveaux connaître et comprendre pour entrer dans l'application. Ainsi en est-il à chaque fois qu'il faut relier l'exemple à la règle.</p>	

	Enoncé	Items
Question 2	<p>D'après les linguistes, la différence entre <i>lexique</i> et <i>vocabulaire</i> se situe dans le fait que :</p>	<p>a) Les deux peuvent être considérés comme synonymes.</p> <p>b) Le lexique représente la totalité des morphèmes d'une langue donnée alors que le vocabulaire représente le répertoire de mots.</p> <p>c) Le lexique représente l'ensemble des mots utilisés par une personne alors que le vocabulaire représente l'ensemble des mots d'une langue.</p>
		<p>Ici, la question se réfère à un contexte (les linguistes), l'étudiant est ainsi invité à contextualiser un savoir à une situation donnée</p> <p>Quand une question commence par « Selon tel scientifique (année), ... », l'enseignant hisse ses étudiants au-delà de l'application : il leur permet d'analyser voire d'évaluer.</p>

	Enoncé	Items à relier avec les...	...possibilités
Question 3	<p>Pour chaque faiblesse nous vous invitons à trouver la méthode de recherche correspondante parmi cette liste :</p>	<p>a) Impossibilité de décrire des processus non-conscients. L'auto-observation de l'activité cognitive peut elle-même représenter un biais d'observation.</p> <p>b) Absence de contrôle expérimental, biais relatif à l'observation.</p> <p>c) Pas toujours possible de généraliser les résultats au-delà d'un lieu, d'une durée et d'une tâche spécifique ; écarts entre des comportements de la vie courante et ceux manifestés au laboratoire.</p> <p>d) Limitations imposées par les limites du hardware, problème de généralisation des résultats à la cognition naturelle.</p> <p>e) Résultat non généralisable.</p> <p>f) Accessibilité limitée pour le chercheur, équipement rare et très onéreux, problème de généralisation si observation de cerveaux anormaux.</p>	<ul style="list-style-type: none"> • comptes-rendus introspectifs • observations naturalistes • études de cas • simulation par ordinateur, • recherche psychobiologique • expériences contrôlées en laboratoires.
	<p>Est-ce plus difficile de nommer les faiblesses de chaque méthode de recherche séparément ou de choisir parmi les items de ces deux listes dans le désordre ? Cela dépend des étudiants.</p> <p>Nous sommes à un niveau d'application / d'analyse, voire de synthèse ou d'évaluation (tout dépend de l'attitude de l'étudiant qui répond).</p> <p>Des facultés d'attention, de raisonnement et de prise de décision sont mobilisées.</p>		

Un QCM peut engager des activités cognitives de haut niveau si les questions conduisent à considérer les connaissances de façon plus profonde. On peut par exemple :

- Interroger des relations (temporelles, logiques, ...) entre des concepts.
- Tester la capacité à appliquer les connaissances à des situations non présentes dans le cours.
- Proposer des analogies pour forcer l'étudiant à comparer et différencier des concepts.
- Proposer des questions qui nécessitent plusieurs étapes de raisonnement pour répondre.
- Impliquer plusieurs compétences cognitives en « demandant une évaluation des connaissances acquises pour parvenir à une décision logique » (Dessus et Tarbouriech 2019)

En conclusion, l'enseignant est invité à diversifier ses façons de solliciter l'apprenant, de varier les niveaux de complexité. Par exemple, il peut faire appliquer des représentations complexes dans des cas particuliers ou éclaircir la hiérarchie relative des idées.

5 Récapitulatif : pièges à éviter et pistes

On ne peut appliquer ce qu'on connaît de manière judicieuse si on n'a rien à appliquer.
Sternberg R. et al. (2008).

- **Manque de clarté** : formulation claire sans ambiguïté (faire passer le QCM à une tierce personne).
- **Granularité incohérente** : choisir des questions générales si l'on vise une compréhension générale et des questions portant sur des détails si l'on veut s'assurer d'une compréhension fine.
- **Pauvreté** : viser la variété avec des questions aux caractéristiques différentes (longueur, thème, niveau cognitif, ...) entre elles pour éviter des stratégies occluses d'apprentissage (mémoriser plutôt qu'apprendre, par exemple). Il ne faut pas hésiter à entremêler des sujets distincts, mais liés. En effet, des matières entremêlées lors de l'étude s'acquièrent mieux que quand elles sont abordées l'une après l'autre (Brown P.C. et al., 2014).
- **Négliger les consignes** : expliciter les barèmes, les visées, les attentes, les objectifs.
- **Rester à un niveau « connaissance »** : pour pousser l'étudiant à l'analyse, la synthèse ou le jugement, on peut intégrer des questions évoquant les controverses scientifiques, touchant des aspects nuancés, évoquant les divergences de point de vue entre tel et tel auteur, utiliser des exemples qu'il faut associer à des théories, proposer des phrases à associer à des auteurs, ...
- **Renforcer les fausses idées** : l'étudiant ne fournit pas toute son attention s'il n'a pas conscience d'ignorer. Alors, il répond rapidement en parcourant la question rapidement. En utilisant des rétroactions immédiates, on l'aide à prendre conscience de ses lacunes. De même, il vaut la peine d'avertir si une question comporte un piège : l'étudiant mobilisera toutes ses ressources cognitives pour répondre.
- **Se perdre dans les paramétrages de Moodle** : il est préférable de concevoir les questions séparément, puis de les placer dans Moodle. Un cas particulier n'a pas été évoqué : il est possible de faire définir un concept dans un quiz Moodle. L'étudiant définit la notion clé proposée par l'enseignant puis découvre la définition de l'enseignant. Faire reformuler les idées centrales d'un cours est propice à l'apprentissage !
- **Proposer des questions proches** : si l'intitulé d'une question offre la réponse d'une autre, le test n'est plus valide. On évitera donc les indices mutuels de solutions et, sur Moodle, on éloignera l'une de l'autre deux questions trop proches.
- **Peur de l'ennui** : les propositions fantaisistes ne sont pas nécessaires, à moins qu'elles aient une valeur humoristique, l'ennui se combat plus efficacement par le choix des sujets, par des questions d'un niveau élevé de traitement cognitif ou encore par des paratextes qui élargissent le champ de la question.
- **Temps de fabrication** : un temps considérable est nécessaire pour créer de bonnes questions formatrices. On peut, chaque semestre, enrichir sa base de données de questions. Peu importe si la première année, il n'y a qu'une vingtaine de questions, les apprentissages seront déjà facilités.

6 Avantages et limites

Le tableau des apports et faiblesses des QCM devrait être lu en pensant aux bonnes et mauvaises pratiques décrites durant les pages précédentes. Certains avantages s'estompent dans un mauvais QCM et certaines limites disparaissent dans un bon QCM. Aucune méthode, aucun outil, aucune approche ne pourrait devenir le graal pédagogique. Le QCM est un **outil d'apprentissage et d'évaluation parmi d'autres**. Rappelons que la pédagogie est un art et que chaque enseignant reste libre de le pratiquer selon ses sensibilités, sa matière, ses affinités. L'implication de l'enseignant reste le meilleur prédicateur de la réussite de ses étudiants.

Avantages	Limites
Correction rapide Correction d'un grand nombre de personnes en peu de temps.	Conception d'un bon QCM demande du temps.
Correction objective et fidèle (Sabatier P., 2004).	Réponse exacte peut être due hasard.
Evaluer une large variété de compétences : connaissance, compréhension, analyse, synthèse, évaluation.	Pas de test de la capacité à rédiger, créer ou produire des raisonnements complexes.
Constitution progressive d'une banque de questions facilite l'élaboration ultérieure de nouveaux questionnaires.	Pas très populaire en phase certificative.
Se rendre compte des attentes de l'enseignant.	Ne suffit pas en phase certificative pour évaluer les capacités à former des jugements ou à avoir un regard critique sur la matière. Parcellisation des connaissances.
Couvre un plus large éventail de matière qu'une question.	Le candidat n'a pas besoin de connaître la réponse s'il arrive à écarter les distracteurs.
Le feedback concernant l'erreur est efficace pour la corriger et aide à apprendre la bonne solution.	L'étudiant risque de mémoriser les mauvaises réponses. S'exposer à l'erreur augmente les chances d'apprendre l'erreur.

Terminons par une comparaison entre les QCM et les exercices de rédaction² (dissertation, analyse de cas, réponses sur un casus, question ouverte, question de développement, ...) Quand un étudiant doit construire sa réponse (CR), il mobilise des compétences cognitives moins nécessaires à la réalisation des QCM (organiser sa pensée, structurer sa pensée, faire preuve de pensée divergente, induire, déduire, abduire, ...) Les QCM, au contraire, permettent de ratisser plus large, mais moins en profondeur. Les QCM, en peu de temps peuvent interroger sur des dizaines de thèmes, mobiliser des points de vue très variés, tandis que les CR obligent à préciser la pensée de l'étudiant et lui donnent l'occasion de s'exprimer.

Ce tutoriel aimerait que les exercices nécessitant une construction de réponse restent au centre de vos préoccupations, mais que vous n'hésitez pas à mettre en place des QCM formateurs, formatifs ou diagnostics en cours de semestre. Leurs effets sont extrêmement positifs sur les apprentissages dans un contexte universitaire hybride (Spanjers et al., 2015).

² Ou exercices oraux qui sont moins fréquents.

7 Références bibliographiques

7.1 Références scientifiques ou pédagogiques

- Bertrand, M. (2014). *L'utilisation du QCM est-elle un problème ?* PEPS : pour une pédagogie participative et sociale.
<https://pepsagogie.wordpress.com/2014/07/24/lutilisation-du-qcm-est-elle-un-probleme/>
- Bransford, J., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience and school*. Washington: NAP.
- Bravard, S. (2004). *Usages pédagogiques des QCM. Un guide pour la mise en place d'un questionnaire à choix multiple*. Mémoire. Université de Poitiers.
<https://docplayer.fr/5833326-Usages-pedagogiques-des-qcm.html>
- Brown, P. C., Roediger, H.L., & McDaniel, M.A. (2014). *Mets-toi ça dans la tête ! Les stratégies d'apprentissage à la lumière des sciences cognitives*. Genève : Markus Haller.
- Dessus, P., Tarbouriech, N. (2019). *Evaluer les apprentissages*. MEEF-PIF : Univ. Grenoble Alpes.
- Johnstone, A. (2003). *Effective practice in objective assessment. The skills of fixed response testing*. LTSN Physical Sciences.
- Krebs, R. (2004). *Mode d'emploi pour la création de questions et d'examens à choix multiples pour l'enseignement de la médecine*. Institut für Medizinische Lehre.
https://www.iml.unibe.ch/attachment/56/download/Manuel_QCM_fr.pdf
- Prospero, O. (2015). Le réalisme avec degrés de certitude. *Mesure et évaluation en éducation*, 38 (1), 121–140. <https://doi.org/10.7202/1036553ar>
- Sabatier, P. (2004). Service d'Appui à la Pédagogie, *Guide d'élaboration de questionnaires d'évaluation des connaissances sous forme de QCM*. Université Toulouse III.
http://sup.ups-tlse.fr/documentation/docs/fich_272.pdf et http://sup.ups-tlse.fr/documentation/docs/fich_271.pdf
- Shute, V. J. (2008). *Focus formative feedback*. *Review of Educational Research*, 78(1), pp. 153–189.
- Spanjers, I. A.E., Könings, K. D., Leppink, J., Verstegen, D. M.L., Jong, N. d., Czabanowska, K., & Merriënboer, J. J.G. v. (2015). *The promised land of blended learning: quizzes as a moderator*. *Educational Research Review*, 15, pp. 59–74.
- Sternberg, R.J., Grigorenko, E.L., & Zhang, L. (2008). *Styles of learning and thinking matter in instruction and assesment*, *perspectives on psychological Science* 3, pp. 486-506.
- Unisciel (L'Université des sciences en ligne). *Banque de tests de positionnement. Guide de rédaction des questions d'évaluation*.
<http://unisciel.fr/wp-content/uploads/2017/11/GuideRedactionQCM2018.pdf>

7.2 Bibliographie complémentaire

- Appleby, D. C. (2008). *A cognitive taxonomy of multiple-choice questions*. In L. T. Benjamin (Ed.), *Favorite activities for the teaching of psychology* (pp. 119-123). Washington, DC: American Psychological Association. (Traduction Dessus P., 2019).
- Bouillet, D., & Monget, M.-C. (2016). *De l'usage des quiz dans Moodle : retour d'expériences et questions posées*. HAL, archives-ouvertes.
<https://hal.archives-ouvertes.fr/hal-01314850/document>

7.3 Références non scientifiques

- Borst, G. cité par Sandrine, C. (23.05.2016). *Les neurosciences peuvent-elles sauver l'école*. Le Monde.
- Montaigne, M. (1580/1595), *Essais* I,26, texte établi et traduit par Naya, E., Reguig, D., & Tarrête, A. Paris : Gallimard (2009).
- Quéva, R. (2008). *La science est un jeu*. Paris : Librio.
- Quintilien, M.F. *De l'Institution Oratoire*, texte établi et traduit par Cousin, J. (1976/1980). Paris : les belles lettres, (livres II et X).
- Reynaud, F. (1976). *Heureux !* Paris : Gallimard.

5 BIENFAITS DES QCM

Rendre son enseignement meilleur



1

CONSOLIDER

QCM formatif : En obligeant l'étudiant à se remémorer ce qu'il a appris, on lui permet de consolider ses souvenirs. On lui évite de former ses apprentissages dans du sable. Grâce à ses efforts, il se souviendra demain de ce qu'il a vu aujourd'hui.



2

INDIQUER

QCM formateur : quand l'étudiant fournit un effort pour répondre et qu'il obtient une rétroaction de l'enseignant, il fait face à la distance entre ce qu'il sait et ce que l'on attend de lui. La direction lui est donc donnée pour progresser et donner du sens à ses efforts.



3

ENTRAINER

QCM formateur : par des aller-retours entre les concepts et des cas réels, entre la théorie et la pratique, l'étudiant contextualise son savoir. C'est à dire qu'il élargit l'horizon de sa maîtrise. qu'il donne un sens de plus en plus profond à la matière.



4

METTRE EN GARDE

QCM diagnostic : quel plaisir d'être surpris, de découvrir que la réalité va à l'encontre du sens commun, de se trouver face à un piège ou confronté à une difficulté. L'une des fonctions de la science consiste, tel un diagnostic, à révéler ce qui est caché.



5

STIMULER LA MOTIVATION

QCM diagnostic : après avoir pris conscience de son ignorance vient la phase de questionnement, de recherche et de curiosité. La question posée par l'enseignant soulève celles qu'adresseront les étudiants.