

QCM ET POINTS NÉGATIFS

Eric Uyttebrouck – version 2.0 – 4/1/2023

« Formula testing [correction for guessing] is generally inappropriate for classroom testing. If instruction has been reasonably effective and the test represents a reasonable sample of what examinees should know, it is unlikely that very many examinees will be totally ignorant on very many items. »

(Frary, 1988, p. 37)

➤ INTRODUCTION

Dans le cas d'épreuves recourant aux Questions à Choix Multiples (QCM), le barème de notation le plus fréquent est sans doute celui qui consiste à attribuer 1 point à une réponse correcte et 0 point à une réponse incorrecte ou à une omission. D'autres barèmes plus complexes existent cependant, dont celui, encore relativement populaire, de la « correction pour devinette », traduction de l'anglais *correction for guessing* (Leclercq, 1987).

L'objectif de la correction pour devinette, comme son nom l'indique, est de compenser les avantages qui résulteraient de réponses effectuées au hasard. On recourt pour cela à la formule classique suivante :

$$\text{Score final} = \text{Nombre_réponses_correctes} - \text{Nombre_réponses_incorrectes} / (\text{nombre_choix} - 1)$$

Pour prendre un exemple simple, imaginons un questionnaire de 20 QCM à 4 propositions de réponses. Un étudiant répondant totalement au hasard aurait 1 chance sur 4 de répondre correctement à chaque question, et obtiendrait statistiquement 5 réponses correctes sur 20. Avec le barème standard, l'étudiant obtiendrait 5/20. Avec la formule de correction pour devinette, l'étudiant obtiendrait $5 - 15 / (4-1)$, soit 0.

➤ LES CRITIQUES ADRESSÉES À LA FORMULE

En 2013, l'Université de Gand a banni l'usage de la correction pour devinette, sur base du fait que le système, d'après la recherche, présente plus de désavantages que d'avantages¹. La littérature sur le sujet est à la fois abondante et relativement technique, mais nous tenterons ici de résumer les critiques essentielles adressées à la procédure, en cinq points :

¹ <http://deredactie.be/cm/vrtnieuws.english/News/1.1798440>

1. la procédure s'appuie sur des présupposés notoirement faux ;
2. elle avantage généralement les étudiants plus enclins à prendre des risques ;
3. elle introduit un biais de genre ;
4. elle induit les étudiants en erreur sur la meilleure stratégie à adopter ;
5. elle n'offre pas d'avantages réels en termes de fidélité ou de validité.

1. La procédure s'appuie sur des présupposés notoirement faux

Comme le rappelle par exemple Roberts (s.d.), la correction pour devinette repose sur les présupposés implicites suivants :

- a. les étudiants connaissent la réponse à une question ou ne la connaissent pas, il n'y a pas de position médiane ;
- b. l'étudiant qui connaît la réponse donne la réponse juste ; s'il ne la connaît pas, il tente sa chance au hasard ;
- c. toutes les réponses incorrectes sont le résultat d'un choix effectué au hasard.

Or ces présupposés sont faux : les étudiants ont généralement une **connaissance partielle** qui leur permet souvent d'**éliminer un ou plusieurs distracteurs**. Les choix effectués totalement au hasard représentent un cas particulier, dont nous ne pouvons présager de la fréquence.

2. La procédure avantage généralement les étudiants plus enclins à prendre des risques

De ce qui précède, on pourrait en conclure un peu vite que la correction pour devinette pénalise les étudiants qui tentent de répondre sans être sûrs de leur fait. En réalité, c'est typiquement l'inverse qui se produit. En effet, puisque les étudiants sont souvent en mesure d'éliminer un ou plusieurs distracteurs, ils effectuent un choix non entre 4 solutions (dans le cas le plus classique), mais entre 3, voire 2. Dès lors, la formule sous-corrige les scores. L'exemple ci-dessous l'illustrera : imaginons une épreuve de 20 QCM à 4 choix possibles, où deux étudiants, Fred et Karl, possèdent les mêmes connaissances : ils connaissent les réponses à 10 questions (en vert), croient erronément connaître la réponse à deux questions (en rouge) et ont une connaissance partielle des 8 autres questions où ils sont en mesure d'éliminer 2 distracteurs sur 4 (en orange). Fred, plus enclin à la prise de risques, répond au hasard aux questions « orange », avec une chance sur deux de répondre correctement ; plus prudent, Karl s'abstient lorsqu'il n'est pas sûr de lui. Lors de la correction, Fred, qui a commis 6 erreurs, subit une pénalité de $6 / (4-1)$ soit 2 points, pour un score final de 12. Karl, qui a joué la prudence, ne perd « que » 0,66 points pour un score final de... 9,33. Fred a donc eu « raison » de tenter sa chance, malgré la correction pour devinette.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Total	Correction	Résultat
Fred	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	0	0	14	2,00	12,00
Karl	1	1	1	1	1	1	1	1	1	1	/	/	/	/	/	/	/	/	0	0	10	0,67	9,33

3. La procédure introduit un biais de genre

Corollaire de ce qui précède, plusieurs recherches ont montré que lorsque les mauvaises réponses sont pénalisées, les femmes seront plus enclines à s'abstenir que les hommes, ce qui peut notamment s'expliquer par des différences relatives à la prise de risque (Brogniaux, 2021 ; Baldiga, 2014).

4. La procédure induit les étudiants en erreur sur la meilleure stratégie à adopter

Autre corollaire de ce qui précède, le recours à la correction pour devinette **induit les étudiants en erreur**, puisqu'elle prétend décourager la prise de risques alors qu'au final, c'est la prise de risques qui s'avère la stratégie la plus payante. L'exemple supra l'illustre parfaitement, au point de friser le problème déontologique : Fred, qui a eu dans une certaine mesure recours à la chance, obtient un meilleur résultat que Karl, qui s'est abstenu de « deviner » suivant les consignes explicites de l'enseignant ! Il y a même pire : Karl échoue, et échoue à cause même de la pénalité pour devinette... **alors qu'il n'a à aucun moment fourni de réponse au hasard !**

5. La procédure n'offre pas d'avantages réels en termes de fidélité ou de validité

Enfin, la recherche n'a pu établir d'avantages réels de la méthode en termes de fidélité ou de validité (Frary, 1988 ; Leclercq, 1987).

› QUELLES AUTRES SOLUTIONS POUR CORRIGER LE HASARD ?

D'autres possibilités pour réduire l'effet du hasard peuvent être trouvées dans le *Guide pratique pour l'élaboration d'une évaluation par QCM* conçu par le Centre d'appui pédagogique (De Clerck, 2021).

› ET SI L'ON SOUHAITE, MALGRÉ TOUT, UTILISER LES POINTS NÉGATIFS ?

Dans le cas où l'on souhaiterait absolument, malgré les mises en garde de la littérature sur le sujet, appliquer la formule de correction pour devinette, les deux précautions suivantes paraissent des précautions minimales :

- le principe de triple concordance entre intentions pédagogiques, méthodes et évaluation (voir Uyttebrouck, 2009) veut que les étudiants soient au préalable entraînés à répondre à des épreuves similaires ;
- l'utilisation de la procédure de « correction pour devinette » doit être clairement annoncée dans les consignes, avec le barème précis en cas de réponse fausse ou d'omission.



➤ ➤ RÉFÉRENCES

- Baldiga, K. (2014). Gender differences in willingness to guess. *Management Science*, 60(2), 434–448.
- Brogniaux, A. (2021). *Gender bias in learning evaluation: Negatively Marked Multiple Choice Questions*. Mémoire présenté en vue de l'obtention du Master en Ingénieur de Gestion, finalité spécialisée. Université libre de Bruxelles.
- De Clerck, O. (2021). *Guide pratique pour l'élaboration d'une évaluation par questions à choix multiples (QCM)*. Université libre de Bruxelles, Centre d'appui pédagogique.
<https://uv.ulb.ac.be/mod/resource/view.php?id=826926&forceview=1>
- Frary, B. (1988). Formula scoring of multiple-choice tests (correction for guessing). *Educational Measurement: Issues and Practices*, 7(2), p. 33-38.
- Leclercq, D. (1986). *La conception des Questions à Choix Multiple*. Bruxelles : Labor.
- Leclercq, D. (1987). *Qualité des questions et signification des scores (avec application aux QCM)*. Bruxelles : Labor.
- Roberts, D. (s.d.). Let's talk about the 'correction for guessing' formula.
<http://www.personal.psu.edu/users/d/m/dmr/papers/CORR4GUS.pdf>.
- Uyttebrouck, E. (2009). La triple concordance dans une action de formation. ULB : PRAC-TICE (Fiche-outil PRAC-TICE). https://portail.ulb.be/medias/fichier/2-triple-concordance_1570453113859-pdf
-