

Évaluer des compétences, ou ne pas se tromper de cible

Suite à notre demande à Monsieur Xavier ROEGIERS directeur du BIEF, cet article a été envoyé de Bruxelles par Monsieur François-Marie GERARD, directeur adjoint.

Jamil est enseignant de français (langue seconde d'enseignement) en 9^e année de l'enseignement de base. Pour évaluer les acquis de ses élèves en langage oral, il organise une séance d'"exercices structuraux oraux". Il a construit une cinquantaine de phrases très proches de celles qui ont été travaillées préalablement en classe: il tire au sort une phrase, la lit devant tous les élèves et demande à l'un d'entre eux de la transformer selon certains paramètres correspondant aux objectifs visés. Chaque élève doit ainsi transformer oralement trois phrases.

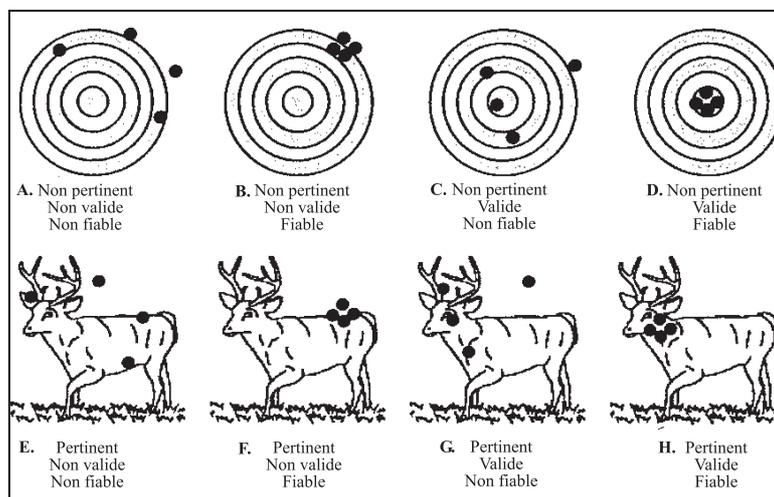
L'épreuve d'évaluation construite par Jamil est intéressante: elle porte sur l'oral alors que ce domaine fait rarement l'objet d'évaluation, elle se fonde sur une démarche systématique en fonction des objectifs poursuivis, elle offre aux élèves plusieurs possibilités de montrer ce qu'ils savent faire, elle évalue des performances sur lesquelles un apprentissage a réellement été effectué, l'aspect "tirage au sort" garantit une certaine objectivité, etc.

Néanmoins, elle présente différentes difficultés qui font qu'au bout du compte on peut émettre certaines réserves, à partir d'une triple question: cette épreuve est-elle pertinente, valide et fiable?

Il s'agit de trois caractéristiques essentielles, différentes mais complémentaires, de toute épreuve d'évaluation:

- ▶ la pertinence est le caractère plus ou moins approprié de l'épreuve, selon qu'elle s'inscrit dans la ligne des objectifs visés (DE KETELE et alii, 1989);
- ▶ la validité est le degré d'adéquation entre ce que l'on déclare faire (évaluer telle ou telle dimension) et ce que l'on fait réellement;
- ▶ la fiabilité est le degré de confiance que l'on peut accorder aux résultats observés: seront-ils les mêmes si on recueille l'information à un autre moment, avec un autre outil, par une autre personne, etc.?

Pour comprendre l'importance de ce triple questionnement, il est intéressant de se référer à l'exemple du tir à l'arc proposé par LAVEAULT et GREGOIRE (1997, 2002) et adapté par DE KETELE et GERARD (2004). Imaginons qu'on souhaite évaluer la compétence d'un trappeur à chasser du gibier grâce à son arc. Il doit pour ce faire être capable de réaliser un tir groupé sur une cible mouvante. Deux épreuves sont mises en place: d'une part, un tir de 5 flèches sur une cible classique, et d'autre part, un tir de 5 flèches sur une cible mouvante semblable à du gibier. Les situations suivantes peuvent dès lors exister:



Dans la situation A, le tir est non fiable parce que dispersé sur toute la surface de la cible fixe. Il est non valide, parce que la mesure manque de précision et ne peut attester l'atteinte de l'objectif. Il est enfin non pertinent puisqu'il ne correspond pas à l'objectif recherché et ne mesure donc de toute façon pas ce que je déclare vouloir mesurer. Dans la situation B, le tir est groupé mais rate systématiquement la cible. Il est fiable (la mesure montre que le tireur sait faire un tir groupé), mais non valide parce que

cette mesure ne me permet pas de vérifier qu'il est capable d'atteindre la cible.

Dans la situation C, le centre de la cible, c'est-à-dire l'objectif, est atteint et la mesure est donc valide, mais elle n'est pas fiable parce qu'on ne peut certifier que ce sera toujours le cas.

Dans la situation D, le tir est groupé et touche le mille, mais la cible n'est pas la bonne (par rapport à la compétence visée). La mesure est donc fiable et valide, mais non pertinente.

Les situations E à H présentent des situations pertinentes par rapport à la compétence visée. Les situations E, F et G ont cependant des problèmes de validité et/ou de fiabilité, alors que la situation H détermine une mesure à la fois fiable, valide et pertinente. Elle seule permet de certifier que le trappeur est compétent.

Ces derniers mots sont importants: l'évaluation a pour objectif de certifier que *le trappeur est compétent* pour chasser du gibier, tout comme Jamil voudrait certifier, grâce à son épreuve d'évaluation, que *les élèves sont compétents* pour s'exprimer oralement dans une conversation.

Quelle est cette compétence? Au terme de la 9^e année de l'enseignement de base dans l'apprentissage du français, la compétence suivante pourrait être visée:

Dans une situation de communication orale (exposé, conversation) et à partir d'un support, produire un message oral correspondant aux divers actes de parole de la vie courante.

Caractéristiques du support:

- écrit: documents authentiques
(courts, clairs et accessibles)
- oral: document sonore de courte durée
document audio et/ou visuel

Si Jamil évolue dans un système d'enseignement qui a adopté une telle compétence comme objectif de formation, on voit tout de suite que l'épreuve d'évaluation qu'il a conçue n'est pas pertinente par rapport à cette compétence: ce n'est pas parce que les élèves seraient à même de transformer les phrases proposées qu'ils pourraient produire des actes de parole dans une situation de communication. La maîtrise de tels exercices oraux structuraux peut bien sûr être un passage pour pouvoir produire des actes de parole: c'est une *ressource* au service de la compétence. Mais si Jamil souhaite évaluer la compétence, il ne peut se

contenter d'évaluer une ressource. Il doit - pour pouvoir avoir une évaluation pertinente - présenter une situation d'évaluation qui correspond à la compétence, c'est-à-dire ici une situation de communication complexe qui permettra aux élèves de produire le message oral attendu. Par exemple, l'élève est invité à participer à un débat à la radio sur le développement urbain. Il devra présenter oralement durant 5 minutes sa position en tenant compte de son contexte de vie. Ensuite, un débat aura lieu pendant lequel il devra interagir avec d'autres intervenants (d'autres élèves).

L'épreuve élaborée par Jamil n'est donc pas pertinente, du moins s'il s'inscrit dans une approche par les compétences, qui cherche à développer la possibilité par les apprenants de mobiliser un ensemble intégré de ressources pour résoudre une situation-problème appartenant à une famille de situations (LE BOTERF, 1994; REY, 1996; PERRENOUD, 1997; FOUREZ, 1999; DE KETELE, 2000, 2001 a et b; ROEGIERS, 2000, 2003, TILMAN, 2000; DOLZ & OLLAGNIER, 2002; JONNAERT, 2002; REY, CARETTE, DEFRANCE & KAHN, 2003, AUBRET, J. & GILBERT, P., 2003).

Cette épreuve présente également des difficultés en termes de validité, c'est-à-dire qu'elle ne permet pas d'évaluer pleinement ce qu'elle déclare vouloir évaluer, à savoir la maîtrise de la transformation de phrases. En effet, les phrases proposées sont "très proches de celles qui ont été travaillées préalablement en classe". Il est important de n'évaluer que des objectifs qui ont été préalablement appris, mais cela ne signifie pas que les matériaux doivent être identiques à ceux utilisés durant l'apprentissage. Si c'est le cas, l'évaluation porte sur des "savoir-reproduire" et non pas des "savoir-faire". En d'autres termes, l'épreuve ne permet pas de certifier que les élèves sont capables de "transformer des phrases", mais bien de "redire les transformations de phrases telles qu'elles ont été apprises". Pour toute évaluation, il importe que les situations proposées soient nouvelles, sauf bien sûr si l'objectif est d'évaluer la restitution de définitions, de propriétés, de dates... Dans une approche par les compétences, cet objectif de restitution ne sera jamais au cœur de l'évaluation, car la compétence se situe à un niveau de résolution de problèmes et non pas de reproduction.

L'épreuve proposée par Jamil présente également des difficultés de fiabilité, ce qui signifie qu'on ne peut pas avoir une confiance totale dans les résultats auxquels elle conduira. Chaque phrase

tirée au sort est lue devant toute la classe et l'élève concerné effectue les transformations. Avec de telles conditions de passation, être le premier élève qui répond ou le dernier n'est pas indifférent: il est vraisemblable que le dernier élève aura eu l'occasion de s'entraîner mentalement. Comme il n'est pas possible de certifier que sa production serait identique s'il avait été le premier élève, on peut affirmer que l'épreuve n'est pas fiable. C'est une des difficultés des épreuves orales: pour assurer leur fiabilité, il faudrait mettre en place un dispositif qui évite l'influence des autres sur la production, ce qui implique dans la plupart des cas une passation individuelle de l'épreuve orale. C'est vrai que cela alourdit le processus, mais à quoi bon évaluer si l'évaluateur ne peut pas faire confiance aux résultats de l'évaluation!

L'épreuve discutée est donc à la fois non pertinente⁽¹⁾, non valide et non fiable. Elle correspond à la situation A de l'exemple du trappeur chassant du gibier.

Il n'est certes pas facile d'élaborer des épreuves entièrement pertinentes, valides et fiables. L'approche par les compétences ne facilite certainement pas les choses: elle amène à évaluer les acquis des

élèves à travers des situations complexes exigeant une production elle aussi complexe de la part des élèves. Les techniques classiques permettant d'assurer et de vérifier la validité et la fiabilité des épreuves d'évaluation ne peuvent pas être utilisées telles quelles (DE KETELE & GERARD, 2004).

L'important nous semble être que tout évaluateur se pose des questions par rapport à l'épreuve qu'il est en train d'élaborer et/ou d'utiliser:

- ▶ l'épreuve est-elle pertinente? Permet-elle d'évaluer les objectifs visés par le système éducatif? Correspond-elle aux compétences à développer?
- ▶ l'épreuve est-elle valide? Évalue-t-elle bien ce que je déclare évaluer? Le dispositif me permet-il de le garantir? Mes critères permettent-ils de vérifier ce que je déclare vouloir vérifier?
- ▶ l'épreuve est-elle fiable? Obtiendrais-je les mêmes résultats si j'évalue à un autre moment? Si c'est quelqu'un d'autre qui fait l'évaluation? L'utilisation que je fais des critères est-elle la même pour tout le monde?

François-Marie GERARD

Directeur Adjoint du BIEF, Bureau International pour l'Enseignement et la Formation

(1) Du moins si elle s'inscrit dans un enseignement fondé sur l'approche par les compétences.

BIBLIOGRAPHIE

- AUBRET, J. & GILBERT, P. (2003). *L'évaluation des compétences*, Sprimont: Mardaga.
- DE KETELE, J.-M. (2000). En guise de synthèse: Convergences autour des compétences, in BOSMAN, C., GERARD, F.-M., ROEGIERS, X. (Éds). *Quel avenir pour les compétences?* Bruxelles: De Boeck Université, pp. 187-191.
- DE KETELE, J.-M. (2001a). Place de la notion de compétence dans l'évaluation des apprentissages, in FIGARI, G., ACHOUCHE, M. (Éds). *L'activité évaluative réinterrogée. Regards scolaires et socioprofessionnels*. Bruxelles: De Boeck Université, pp. 39-43.
- DE KETELE, J.-M. (2001b). Enseigner des compétences: repères in JADOULLE, J.-L. & BOUHON, M. (Eds). *Développer des compétences en histoire*, Louvain-la-Neuve-Bruxelles: Université catholique de Louvain et Ministère de l'Éducation, de la Recherche et de la Formation, 13-22.
- DE KETELE, J.-M., CHASTRETTE, M., CROS, D., METTELIN, P. & THOMAS J. (1989). *Guide du formateur*, Bruxelles: De Boeck Université.
- DE KETELE, J.-M. & GERARD, F.-M. (2004). La validation des épreuves d'évaluation selon l'approche par les compétences, *Mesure et Évaluation en Éducation*, soumis pour publication.
- DOLZ, J. & OLLAGNIER, E. (Éds) (2002). *L'énigme de la compétence en éducation*, Bruxelles: De Boeck Université.
- FOUREZ, G. (1999). Compétences, contenus, capacités et autres casse-têtes, *Forum des pédagogies*, mai, 26-31.
- JONNAERT, Ph. (2002). *Compétence et socioconstructivisme - Un cadre théorique*, Bruxelles: De Boeck.
- LAVEAULT, D. & GREGOIRE, J. (1997, 2^e éd. 2002). *Introduction aux théories des tests en psychologie et en sciences de l'éducation*, Bruxelles: De Boeck Université.
- LE BOTERF, G. (1994). *De la compétence. Essai sur un attracteur étrange*. Paris: Éditions de l'organisation.
- PERRENOUD, Ph. (1997). *Construire des compétences dès l'école*. Paris: ESF.
- REY, B. (1996). *Les compétences transversales en question*. Paris: ESF.
- REY, B., CARETTE, V., DEFRANCE, A. & KAHN, S. (2003). *Les compétences à l'école - Apprentissage et évaluation*, Bruxelles: De Boeck.
- ROEGIERS, X. (2000). *Une pédagogie de l'intégration*. Bruxelles: De Boeck.
- ROEGIERS, X. (2003). *Des situations pour intégrer les acquis*, Bruxelles: De Boeck.
- ROEGIERS, X. (2004). *L'école et l'évaluation*, Bruxelles: De Boeck.
- TILMAN, F. (2000). Qu'est-ce qu'une compétences?, *Exposant neuf*, n°2/2000, 28-31.