

La complexité d'une évaluation des compétences à travers des situations complexes : nécessités théoriques et exigences du terrain¹

François-Marie GERARD, BIEF

Quel que soit leur cadre de référence, tous ceux qui inscrivent leur réflexion, ou mieux leur pratique pédagogique, dans l'approche par compétences s'accordent aujourd'hui pour dire que celles-ci s'exercent lorsqu'il s'agit de résoudre des situations-problèmes nécessitant la mobilisation de plusieurs ressources. Que ce soit pour l'apprentissage des compétences ou pour leur évaluation, il convient donc de situer l'élève dans un environnement complexe et de l'amener à mobiliser ses différents acquis pour résoudre une situation complexe.

L'évaluation des compétences à travers des situations complexes est elle-même un processus complexe, ce qui peut décourager les enseignants lors de sa mise en œuvre. Malgré cette complexité, le recours à des situations complexes est une modalité indispensable et économique pour l'évaluation des compétences, en complément d'autres pratiques comme le portfolio par exemple. Sa faisabilité est liée à la connaissance et à la maîtrise des éléments qui constituent sa complexité.

1. Compétences et complexité

La complexité est au cœur de l'approche par les compétences. Être compétent, c'est pouvoir gérer le complexe, parce que la vie, les groupes d'individus ou l'environnement sont d'une complexité étonnante. Qu'on nous comprenne bien : être compétent, ce n'est pas gérer le complexe en permanence ni ne gérer que lui, comme CRAHAY (2006) a dénoncé une possible dérive. Être compétent, c'est bien sûr aussi gérer des situations coutumières ne nécessitant que des micro-adaptations et l'application, de manière relativement routinière, d'un certain nombre de procédures simples et répétitives. Néanmoins, être compétent exige aussi de pouvoir faire face à l'imprévu, au complexe, même en dehors de toute situation de crise. Pour reprendre l'exemple proposé par CRAHAY (p. 99), il nous semble évident qu'un chirurgien qui réussit pour la quarantième fois une transplantation cardiaque fait preuve de compétence. Et qu'il fait encore preuve de la même compétence lorsqu'à la quarante-et-unième fois, il doit modifier quelque peu sa manière de procéder parce que le cas qui se présente à lui est un petit peu plus complexe que d'habitude. C'est justement parce qu'il a cette faculté de s'adapter en cas de nécessité et de gérer la complexité que ses patients font appel à lui, lui reconnaissant sa véritable compétence. En réalité, on peut même être compétent au sens de pouvoir gérer une situation imprévue et complexe, et n'avoir jamais de sa vie l'occasion d'exercer cette compétence. Combien de pilotes d'avion parfaitement compétents n'ont ainsi jamais dû – heureusement – montrer qu'ils pouvaient réagir en situation de crise ? Personne ne songerait cependant à remettre en cause le fait qu'ils doivent apprendre à le faire !

¹ Référence : GERARD, F.-M. (2007), La complexité d'une évaluation des compétences à travers des situations complexes : nécessités théoriques et exigences du terrain, *Actes du Colloque international « Logique de compétences et développement curriculaire : débats, perspectives et alternative pour les systèmes éducatifs »*, Montréal : ORÉ, 26 et 27 avril 2007.

Mais qu'est-ce que la complexité ? Par quoi se caractérise le complexe ? Qu'est-ce qui le différencie du compliqué ?

La complexité est une notion assez... complexe, dont la théorie reste à construire en grande partie ! Il semble que celle-ci se structure autour de deux paradigmes qui s'opposent en partie :

- d'une part, un paradigme « objectiviste » définit la complexité comme « l'irréductibilité à un modèle fini » (LE MOIGNE, 1990). La complexité est donc comprise comme quelque chose qui existe en soi, qui s'impose en tant que telle. Cette irréductibilité implique qu'il n'y a pas de solution idéale pour l'action en milieu complexe (CIES, 2006), mais uniquement quelques repères, au sens de la « stratégie tâtonnante » proposée par d'AVERNIER (1997) ;
- d'autre part, un paradigme « constructiviste » pour lequel la complexité est relative. Comme le montre Robert DELORME (1999), elle dépend de la relation sujet/objet : en 1973 déjà, ASHBY relevait qu'un boucher et un neurophysiologiste ne voient pas une cervelle d'agneau de la même manière. Pour le boucher, elle n'a rien de complexe, ce n'est qu'un objet à découper et à vendre ; alors que pour le neurophysiologiste, le fonctionnement et les caractéristiques de cette même cervelle en font un objet éminemment complexe.

La complexité se caractérise notamment par la multiplicité : plusieurs interactions s'enchevêtrent en augmentant la quantité d'informations à traiter et donc la difficulté de la prise de décision. La complexité appelle dès lors l'adaptabilité, car on ne sait jamais à quoi on s'attend exactement face à une situation complexe. La question n'est pas tant celle de l'inédit de la situation que de son incertitude ou de son imprévisibilité (MORIN, 1990).

Ce détour par la théorie de la complexité apporte quelques éléments intéressants pour la réflexion sur la complexité présente dans l'approche par les compétences :

- tout d'abord, selon le paradigme constructiviste, une situation n'est complexe que pour celui qui la déclare telle (DELORME, R., 1998, 1999, 2000). Une des difficultés en matière d'évaluation est dès lors de savoir si la situation est réellement complexe pour celui qui doit la résoudre, ou du niveau de complexité souhaité par le concepteur ;
- selon le paradigme objectiviste, la complexité n'est pas liée à la difficulté de la situation, mais au fait qu'elle est « irréductible à un modèle fini ». Si la situation est réellement complexe, l'élève ne peut lui appliquer un algorithme défini une fois pour toutes pour ce type de situations. Il doit chaque fois s'adapter et développer une nouvelle stratégie, dans l'incertitude ;
- la complexité est liée au grand nombre de parties apparemment indépendantes l'une de l'autre (en ce sens que la connaissance de l'état d'une partie ne fournit pas ou peu d'informations sur l'état des autres parties), mais qui interagissent entre elles, de telle sorte que « le tout est supérieur à la somme des parties » ;
- la réponse qui peut être apportée face à la situation complexe est une réponse en forme de réseau qui se fonde sur les liens qui existent entre les éléments. C'est par la mobilisation et l'intégration interactive de différentes ressources qu'une solution peut être apportée à la situation complexe.

2. Complexité de l'évaluation des compétences

2.1 Élaborer des situations complexes

De manière plus concrète, ROEGIERS (2003) fait des distinctions très utiles entre situations compliquées et situations complexes, dans l'univers pédagogique.

Pour cet auteur, une situation compliquée est « une situation dont la résolution met en jeu des savoirs et des savoir-faire nouveaux, peu connus de celui qui la résout, ou moins bien maîtrisés par lui » (p. 121). Par exemple, une situation dans laquelle l'élève va devoir résoudre une équation du second degré et une intégrale est – a priori en tous cas – plus compliquée qu'une situation dans laquelle l'élève doit résoudre une équation du premier degré et effectuer un calcul de pourcentage.

Toujours pour ROEGIERS, la complexité « ne dépend pas tellement du type d'activités à exercer, du type de savoirs et de savoir-faire à mobiliser, mais surtout de la quantité de savoirs et de savoir-faire à mobiliser. La difficulté vient non pas de chaque opération à exécuter, mais de l'articulation de ces opérations entre elles. » Une situation complexe combine des éléments que l'élève connaît, qu'il maîtrise, qu'il a déjà utilisés plusieurs fois, mais de façon séparée, dans un autre ordre ou dans un autre contexte. Il s'agit donc non seulement d'éléments connus de l'élève, mais qui ont été travaillés par lui, séparément ou conjointement.

Cette conception de la situation complexe est particulièrement importante relativement à l'évaluation des compétences par situations complexes : lors d'une telle évaluation, il ne s'agit pas de confronter l'élève à une situation qui serait à la fois compliquée – composée de savoirs et de savoir-faire nouveaux, encore peu maîtrisés – et complexe – composée de plusieurs savoirs et savoir-faire dont l'articulation nouvelle permettra la résolution de la situation. En cas d'échec à cette situation compliquée et complexe, il sera impossible de savoir si l'élève n'a pas été à même de la résoudre parce qu'elle était compliquée ou parce qu'elle était complexe. En d'autres termes, lorsqu'un enseignant souhaite évaluer des compétences à partir d'une situation complexe, il doit s'être assuré préalablement que toutes les ressources (savoirs et savoir-faire) à mobiliser pour résoudre la situation ont été non seulement « apprises » par les élèves, mais sont également « maîtrisées ». L'approche par les compétences n'élimine pas le travail sur les ressources, bien au contraire. Leur acquisition et leur évaluation sont des préalables indispensables tant à l'acquisition qu'à l'évaluation des compétences.

Dans ce contexte, une des complexités fondamentales de l'évaluation des compétences par situations complexes est d'élaborer des situations pertinentes et valides. Pour pouvoir y parvenir, il convient de se référer au concept de « famille de situations », alors même que ce concept pose certaines difficultés théoriques. Une **famille de situations** est un ensemble de situations d'un niveau de complexité équivalent qui se rapportent à une même compétence. Mais, si la complexité se définit comme « l'irréductibilité à un modèle fini », il peut être difficile de cerner de manière précise des familles de situations. En effet, en vertu même de l'incertitude,

il n'est pas possible de définir les familles sur la base d'algorithmes communs de résolution des situations concernées².

Dans l'univers professionnel, le problème se pose en termes différents. En effet, lorsqu'une organisation professionnelle souhaite s'inscrire dans une gestion des compétences, elle commence la plupart du temps par définir un « référentiel métier » qui décrit et organise toutes les situations auxquelles les travailleurs seront confrontés dans leur pratique professionnelle. Sur la base de ce référentiel métier, on élabore ensuite un « référentiel de compétences » qui permet de dégager toutes les compétences nécessaires pour résoudre les situations décrites dans le référentiel métier. Ensuite, on définit d'une part un « référentiel de formation » qui permettra de savoir comment on va agir pour que les travailleurs maîtrisent effectivement les compétences nécessaires et d'autre part un « référentiel d'évaluation » qui permettra l'évaluation des compétences réelles des travailleurs afin d'élaborer des plans individuels de formation.

Dans l'univers de l'enseignement, la plupart du temps, le « référentiel métier » n'est pas défini, ou alors insuffisamment, et rarement avant d'écrire le référentiel de compétences. Plus souvent, les réformes curriculaires commencent par l'écriture d'un référentiel de compétences – quand ce n'est pas directement par celle du référentiel de formation – confiée à des enseignants, des inspecteurs ou des conseillers pédagogiques (c'est-à-dire des acteurs internes au système). Il manque ainsi souvent le « référentiel des situations » (ou « profil de sortie ») qui devrait être constitué de deux dimensions :

- d'une part et de manière relativement classique, la description de ce qui est attendu de l'élève pour pouvoir gérer la suite de sa scolarité ;
- d'autre part, la description précise des situations auxquelles les élèves peuvent être confrontés en dehors du système éducatif, c'est-à-dire dans la vie sociale, familiale, professionnelle... Cette description devrait partir de la vie réelle et être réalisée en collaboration avec des acteurs externes au système.

Ce n'est que sur la base de ce « référentiel de situations » qu'il devrait être possible d'élaborer de manière pertinente les référentiels de compétences, de formation et d'évaluation (ROEGERS, 2000 ; JONNAERT, BARRETTE, MASCIOTRA & YAYA, 2006).

Lorsqu'un tel référentiel de situations existe, il est alors possible de définir des « familles de situations » auxquelles correspondront les compétences visées et évaluées. Ces familles de situations pourront se concrétiser dans la définition des paramètres qui sont les caractéristiques communes à l'ensemble des situations relatives à une compétence. Ils regroupent essentiellement deux catégories (GERARD & ROEGERS, 2006) :

- les paramètres liés à l'univers de référence en termes de ressources, relatifs à la quantité et à la nature des ressources à mobiliser dans chaque situation, ainsi qu'à leur combinaison ;
- les paramètres liés à la tâche demandée à l'élève, qu'ils soient liés au contexte de la situation, à la démarche de résolution (nombre d'étapes et nature), aux supports, au type de tâche attendue, aux conditions de résolution...

² Remarquons d'ailleurs que s'il était possible de définir des familles de situations sur la base d'algorithmes constants, on ne se situerait plus au niveau de la résolution de problèmes, mais au niveau de l'application.

Cette difficulté de définir les situations et les paramètres des familles de situations a amené certains à critiquer, voire à rejeter, le concept de « famille de situations » (REY, CARETTE, DEFRANCE & KAHN, 2003 ; CRAHAY, 2006). Cette position n'est pas sans poser de difficulté en matière d'évaluation. En effet, s'il n'est pas possible de relier une situation d'évaluation à une famille de situations reliée elle-même à une compétence, alors il n'est pas possible de tirer la moindre conclusion en termes d'évaluation de la compétence. La seule conclusion possible est qu'un élève a pu résoudre telle ou telle situation, telle ou telle tâche. Mais il est impossible d'inférer à partir de cette situation d'évaluation (ou d'un ensemble restreint de situations) la maîtrise ou non de la moindre compétence, puisque chaque situation serait totalement isolée l'une de l'autre. C'est d'ailleurs une des difficultés que rencontrent ceux qui entrent dans l'approche par les compétences par les compétences dites *transversales*. Celles-ci ne permettent pas de circonscrire des familles de situations de manière suffisamment concrète, ce qui rend quasi impossible leur évaluation.

Pour en terminer avec la complexité de l'élaboration des épreuves d'évaluation par situations complexes, ajoutons qu'il est très difficile d'élaborer du premier coup une situation qui corresponde à toutes les exigences nécessaires d'une telle épreuve :

- difficile de garantir – même en présence de paramètres clairement définis – que la situation est équivalente à une autre situation de la même famille de situations, évaluant la même compétence ;
- difficile d'être sûr que le contexte et le support de la situation ainsi que la tâche demandée seront significatifs pour les élèves ;
- difficile d'estimer avec précision le temps nécessaire pour réaliser la tâche ;
- difficile d'assurer que toutes les informations nécessaires sont contenues dans la situation, parce qu'elles sont peut-être simplement contenues dans la tête du concepteur, mais pas dans celle des élèves ;
- difficile de formuler une consigne qui soit intégralement comprise en vue de mobiliser l'action des élèves, suffisamment contraignante tout en leur laissant des marges de créativité, en ne disant pas tout ce qui doit être fait et comment mais en guidant suffisamment ;
- difficile d'identifier a priori – et nous y reviendrons – tous les bons indicateurs ;
- etc.

Il nous semble dès lors indispensable de tester les situations afin de s'assurer que tout fonctionne bien, de confronter la situation à la réalité des élèves, et de recueillir l'information nécessaire pour réguler l'épreuve.

Cette « expérimentation » peut parfois être systématique, avec un dispositif important, dépassant largement le cadre de l'enseignant confronté à ses élèves. DE KETELE et GERARD (2005) ont montré que la validation des épreuves d'évaluation par situations complexes était une opération difficilement réalisable sur la base des techniques classiques de validation des tests pédagogiques, essentiellement en raison du caractère multidimensionnel de ce type d'épreuves, alors que la théorie classique des scores ou les théories de réponse à l'item sont fondées sur un postulat d'unidimensionnalité. Ces auteurs proposent cependant quelques pistes de validation :

- une validation *a priori*, par recours à des « juges » qui devront vérifier que la ou les situations appartiennent bien à la famille de situations en fonction des paramètres définis ;
- une validation empirique interne qui vise à vérifier l'unité conceptuelle de l'épreuve en elle-même ;

- une validation empirique externe qui tend à montrer la convergence de l'épreuve avec d'autres épreuves similaires elles-mêmes validées ou encore à confronter les résultats avec d'autres épreuves qu'on sait avoir une relation inverse ou non avec la compétence évaluée.

Les techniques pour assurer de telles validations ne sont bien sûr pas accessibles, ni même utiles, à l'enseignant qui souhaite simplement valider l'épreuve d'évaluation qu'il élabore pour ses élèves. Cela ne signifie pas que cet enseignant ne peut pas procéder à un travail de validation. Il suffit parfois de tester la situation auprès de 4 ou 5 élèves. Un enseignant peut ainsi s'organiser pour disposer progressivement d'un ensemble de situations validées :

- utiliser une situation dans une classe en guise d'apprentissage de l'intégration, puis l'utiliser dans une autre classe en guise d'évaluation ;
- tester une situation dans sa classe en lui donnant une valeur formative, puis l'utiliser l'année suivante – après l'avoir améliorée – pour une évaluation certificative ;
- s'organiser entre deux enseignants de deux classes ou écoles différentes pour tester des situations qui seront ensuite utilisées à titre d'évaluation dans l'autre classe ou école ;
- capitaliser au fil du temps différentes situations élaborées par l'enseignant lui-même ou empruntées à des collègues ou à toute autre source ;
- ...

Enfin, lorsqu'une expérimentation, même légère, n'est pas possible, il est toujours possible et intéressant de montrer l'épreuve à un collègue en lui demandant d'en faire une lecture critique. Cela suffit souvent pour apporter à l'épreuve les améliorations indispensables.

Pour conclure notre réflexion sur l'élaboration des épreuves par situations complexes, il nous semble important d'insister sur le fait qu'un enseignant – ou un système d'évaluation – ne devraient jamais confronter des élèves à des épreuves d'évaluation par situations complexes qu'à trois conditions :

- d'abord, et nous l'avons déjà dit, il faut s'être assuré que l'ensemble des ressources qui seront à mobiliser pour résoudre les situations appartenant à la famille de situations ont effectivement été apprises par les élèves et sont maîtrisées et/ou disponibles en tant que telles. Comment un maçon pourrait-il monter un mur s'il ne possède pas une truelle et s'il n'a pas appris à la manier ?
- ensuite, ce n'est pas au moment de l'évaluation que l'on apprend à être compétent. Avant la situation d'évaluation, les élèves devraient donc avoir été confrontés à des fins d'apprentissage à au moins deux situations de même niveau de complexité et appartenant à la même famille de situations. La mobilisation et l'intégration des ressources pour résoudre une situation complexe doit faire l'objet d'un apprentissage spécifique, au même titre que celui des ressources ;
- enfin, si la (ou les) situation(s) d'évaluation doi(ven)t appartenir à la famille de situations relative à la compétence évaluée, et donc correspondre aux paramètres de celle-ci, si elle doit être du même niveau de complexité, il faut veiller à ce qu'elle ne soit pas plus difficile que les situations qui ont été abordées lors de l'apprentissage. Au contraire, elles pourraient même être plus faciles – ce qui ne veut pas dire moins complexes –, car il est assez logique d'aller plus loin en situation d'apprentissage qu'en situation d'évaluation.

2.2 Corriger des productions complexes

La complexité de l'évaluation de compétences par des situations complexes n'apparaît pas seulement au moment d'élaborer les épreuves, mais également lorsqu'il faut corriger les productions que les élèves ont réalisées à partir de la situation complexe. Si la situation, ou la tâche demandée à l'élève, est réellement complexe, il est vraisemblable que la production sera elle-même complexe. Il ne suffit pas dès lors de se demander si la réponse de l'élève est *bonne* ou *mauvaise*, car la plupart du temps, il n'existe pas une et une seule bonne réponse, clairement identifiable. Quand bien même une telle bonne réponse unique existerait, comme ce peut être le cas dans les disciplines scientifiques ou en mathématiques, il est peu intéressant de ne s'intéresser qu'à cette seule réponse, sans se préoccuper de la manière dont elle a pu être obtenue.

Le regard à porter sur la production complexe de l'élève devra donc être multiple, organisé à travers des critères d'évaluation. Chaque critère est un regard différent que l'on porte sur l'objet évalué, un *point de vue* auquel on se place pour évaluer l'objet. Le critère correspond à une **qualité** de cet objet. C'est à travers ces critères que la production sera à évaluer, c'est-à-dire à « en faire sortir la valeur » (*evaluere*). C'est sur leur base qu'on décidera si la compétence est maîtrisée ou non, qu'on déterminera les difficultés rencontrées par l'élève et qu'on lui proposera éventuellement une remédiation.

Le choix des critères est donc un élément essentiel de l'évaluation des compétences par situations complexes. Les critères d'évaluation devraient être (GERARD & VAN LINT-MUGUERZA, 2000) :

- **pertinents**, c'est-à-dire qu'ils doivent permettre d'évaluer vraiment que la compétence est maîtrisée ou non et de prendre la bonne décision ;
- **indépendants**, ce qui signifie que l'échec, ou la réussite, d'un critère ne doit pas entraîner automatiquement l'échec, ou la réussite, d'un autre critère ;
- **pondérés**, en définissant par exemple des critères minimaux, qui doivent absolument être maîtrisés pour certifier la maîtrise de la compétence, à côté de critères de perfectionnement qui concernent des qualités dont la présence est préférable, mais non indispensable ;
- **peu nombreux**, non seulement pour éviter la multiplication des points de vue étudiés et l'« infaisabilité » de l'évaluation, mais aussi parce qu'on risque de chercher en vain la perfection, l'oiseau rare satisfaisant à une quantité invraisemblable de critères dont l'indépendance serait d'ailleurs douteuse. En matière d'évaluation des compétences, l'expérience que nous avons acquise nous pousse à dire que **l'idéal consiste à utiliser 3 critères minimaux et 1 critère de perfectionnement**.

Selon nous, il est impossible et inutile de proposer une liste finie de critères qui seraient pertinents et exhaustifs pour toutes les disciplines et tous les niveaux.

Néanmoins, certains critères semblent ressortir (GERARD, 2005 ; ROEGIERS, 2004) :

- la pertinence, ou l'adéquation de la production à la situation (et notamment à la consigne), qui consiste à répondre aux questions « est-ce que l'élève fait bien ce qu'il doit faire ? N'est-il pas *hors sujet* ? » ;
- l'utilisation correcte des concepts et des outils de la discipline, qui concerne la question « est-ce que l'élève fait correctement ce qu'il fait ? », même si ce n'est pas cela qu'il doit faire ;

- la cohérence, c'est-à-dire l'utilisation d'une démarche logique, qui ne présente pas de contradictions internes (même si elle n'est pas pertinente), le choix cohérent des outils, l'enchaînement logique de ceux-ci, l'unité de sens de la production, etc. ;
- la complétude, c'est-à-dire le caractère complet de la réponse, pour autant bien entendu qu'il soit possible de déterminer ce qu'est une réponse ou une production complète.

On peut aussi identifier un certain nombre de critères dont la pertinence variera en fonction des objectifs que l'on poursuit, des valeurs qu'on souhaite privilégier, d'exigences spécifiques, etc. La plupart du temps, ces critères seront considérés comme critères de perfectionnement :

- la qualité de la langue, pour les disciplines non littéraires ;
- la qualité de la présentation du travail, de la copie de l'élève ;
- l'utilité sociale, l'intérêt ou la profondeur des propositions émises ;
- l'originalité de la production ou son caractère personnel, spécialement pour les disciplines littéraires ou sociales ;
- la précision, spécialement dans les disciplines scientifiques ;
- ...

La liste des critères découle toujours d'un choix des concepteurs-évaluateurs. Ce choix sera effectué en fonction de leurs objectifs (et de ceux de leur système éducatif), des accents qu'ils souhaitent mettre, en tenant compte bien sûr de la spécificité de chaque discipline et de chaque niveau. Même si des listes limitées de critères ont été proposées (BÉLAIR, 1999), il n'y a, selon nous, **pas de liste idéale de critères, pas de modèle** qui s'imposerait de manière absolue. La réalité est toujours complexe. Une modélisation vise à introduire un peu de simplicité, mais elle ne peut jamais être au détriment de la complexité de la réalité. D'autres critères que ceux qui sont proposés ici peuvent être utilisés, avec d'autres formulations, mais il faut alors toujours s'assurer que le critère est pertinent par rapport à la compétence et vérifier l'indépendance entre les critères.

Les critères doivent donc permettre de porter différents regards sur la production de l'élève afin de l'évaluer. Cependant, il ne suffit pas de constater – une fois et une seule – qu'une production est, par exemple, pertinente par rapport à la tâche demandée pour pouvoir certifier que l'élève maîtrise le critère de pertinence. Plus grave même, il ne suffit pas de constater – une fois et une seule – qu'une production n'est pas pertinente pour certifier que l'élève est incapable de réaliser une production pertinente. Il y aurait là une généralisation abusive excessivement dangereuse, d'autant plus que l'erreur de l'élève peut être simplement liée à une mauvaise formulation de la consigne, à une distraction passagère, une fatigue, etc.

Face à cette difficulté, DE KETELE (1996) et ROEGIERS (2000) ont proposé d'appliquer la « règle des 2/3 ». Il s'agit de donner à l'élève au moins trois occasions indépendantes l'une de l'autre de démontrer sa compétence et donc de vérifier chaque critère. On considèrera qu'un élève maîtrise un critère lorsqu'il réussit au moins 2 des 3 occasions qui lui sont offertes. Autrement dit, le professeur doit se donner au moins trois occasions de vérifier si un critère est rencontré et il peut considérer que c'est le cas si l'élève a réussi au moins deux fois sur les trois. Cette règle concilie à la fois le droit à l'erreur (indispensable en situation d'apprentissage), le fait que pour des raisons indépendantes de sa compétence une

tâche peut ne pas être accessible à un élève et le besoin d'établir un seuil de réussite relativement élevé (la proportion 2/3 est équivalente à 66,6%).

Concrètement, ces occasions peuvent être :

- une production d'au moins trois phrases pour vérifier la correction orthographique ;
- trois textes de même difficulté à résumer ;
- trois occasions d'exécuter un parcours en éducation physique et sportive ;
- trois morceaux de musique à interpréter ;
- trois consignes différentes et indépendantes à partir d'une même situation ;
- ...

Ces occasions doivent être réellement distinctes et indépendantes les unes des autres. En d'autres termes, la non-réussite d'une occasion ne doit pas entraîner automatiquement l'échec à une autre occasion.

Notre expérience nous a montré que l'idéal est qu'à l'occasion d'une **même situation** et d'un **même ensemble de supports**, on propose **trois tâches distinctes, trois consignes indépendantes l'une de l'autre**, tout en veillant à ce que celles-ci soient chacune **complexes** et permettent de vérifier chacune les différents critères. Cette proposition – qui est pourtant quasiment indispensable à prendre en compte – n'est évidemment pas faite pour simplifier le travail d'élaboration d'une épreuve d'évaluation par situations complexes. Elle oblige en effet les concepteurs à imaginer trois tâches différentes, indépendantes l'une de l'autre et de complexité équivalente. L'exemple suivant (GERARD, 2006b) illustre cette exigence :

1. La compétence à évaluer : produire dans une situation significative pour l'élève (dans un court message écrit) un énoncé écrit court et simple, dans l'intention de faire agir quelqu'un.

2. La situation et les consignes

Ton école envisage d'organiser une excursion, dans l'une des localités de la région, riche en paysages attirants. Afin de préparer cette excursion, le maître vous demande 3 tâches, dans le respect du règlement de l'école.

Consignes :

- A. Décris en quelques lignes l'endroit où tu voudrais aller. Écris pourquoi et ce qu'on pourrait y faire.
- B. Établis la liste du matériel qu'il faudra prendre avec soi pour l'excursion.
- C. Écris une petite lettre au directeur de l'école pour lui demander l'autorisation de partir en excursion.

Règlement de l'École (extrait)

5. Excursions

- Une fois par trimestre, les élèves d'une classe peuvent aller en excursion pour découvrir la région.
- Une semaine avant l'excursion, une demande écrite doit être adressée au directeur de l'école, en précisant les objectifs de l'excursion, la date proposée et le moyen de transport envisagé.

6. Activités sociales

3. Les critères et les indicateurs

Cette situation peut par exemple être évaluée à travers 3 critères minimaux :

- la pertinence de la production (l'élève fait-il ce qu'on lui demande ?),
- l'utilisation correcte des outils de la discipline, ici surtout la correction orthographique (l'élève écrit-il correctement ?), et
- la cohérence de la production (ce que l'élève écrit comporte des informations qui se répètent (répétition), comporte dans son développement des informations nouvelles (progression), n'a aucune contradiction entre les informations (non-contradiction), comporte des informations qui peuvent être reliées (relation)³ ;

et un critère de perfectionnement :

- l'originalité de la production (l'élève apporte-t-il des idées originales ?).

Dans l'exemple ci-dessus, il existe une différence importante entre le critère de « cohérence » et les trois autres. Ceux-ci sont simplement cités et assortis d'une question relativement générale, alors que le critère de cohérence est précisé par un certain nombre d'éléments qui vont permettre au correcteur de mieux cerner ce qu'il doit observer lors de la correction. Ces éléments sont des indicateurs du critère d'évaluation.

Si un critère est général, abstrait et de l'ordre de l'idéal, l'indicateur est contextualisé, concret et de l'ordre de la réalité (GERARD, 2002) : il dit ce qu'il faut regarder concrètement dans la production de l'élève pour l'évaluer. L'indicateur apporte de l'information sur la maîtrise de la compétence par l'élève, mais ce n'est que de l'information, une indication. Un danger, fréquemment observé, est d'accorder à l'indicateur plus de valeur qu'il n'en a.

Dans l'évaluation d'une situation complexe en vue d'inférer la maîtrise d'une compétence, on évalue la **maîtrise du critère**, et non pas la maîtrise de l'indicateur. L'indicateur apporte une information à mettre en relation avec d'autres informations, d'autres indicateurs. En soi, un indicateur n'est qu'un élément parmi d'autres qui permet à l'évaluateur d'estimer, d'apprécier (par un jugement de valeur) le degré de maîtrise du critère (GERARD, 2005). Pour pouvoir estimer ce degré, il faut un faisceau d'indices qui, s'ils convergent, permettront à l'évaluateur de décider si le critère est maîtrisé ou non.

Dans ce cadre, deux remarques importantes doivent être faites :

- d'une part, un indicateur ne devrait normalement pas être un « indicateur minimal », c'est-à-dire quelque chose qui devrait absolument être présent pour décider de la maîtrise du critère : à lui seul, un indicateur n'est jamais une preuve. C'est le critère qui peut être minimal. Pas l'indicateur. Celui-ci ne fait que donner une indication. Certains indicateurs sont bien sûr plus importants que d'autres (notamment en matière de sécurité pour les disciplines expérimentales), mais cela ne devrait théoriquement jamais rendre un indicateur indispensable pour attester de la réussite d'un critère (et donc de la réussite de la compétence, et donc de la réussite de la discipline, et donc de la réussite du cycle...) ;
- d'autre part, cela implique le rejet d'une pratique pourtant courante : la notation à partir de la présence ou non des indicateurs. Il arrive souvent que les grilles de correction soient constituées d'une liste d'indicateurs reliés à des critères, voire même à des questions (sans passer par des critères). L'enseignant se contente ensuite de constater si l'indicateur est présent ou non, de cocher l'indicateur présent, et puis de compter le nombre d'indicateurs

³ CHAROLLES, M. (1988).

présents ou encore d'additionner le nombre de points affectés à chaque indicateur. Cette pratique, certes commode même si elle se révèle parfois fastidieuse, est à éviter, car elle déplace l'évaluation : ce n'est pas le critère qui est évalué, mais l'évaluateur se contente de constater que des éléments prédéfinis sont présents ou non, sans réelle possibilité de se laisser interpeller par la réponse apportée par l'élève.

Il faut aussi se méfier de considérer la liste des indicateurs définie a priori comme exhaustive. C'est parfois le cas, mais en réalité il est plus souvent impossible d'imaginer toutes les réponses possibles des élèves⁴. Il arrive dès lors que la production d'un élève, face à une situation complexe, contienne des éléments auxquels on n'avait pas pensé, mais qui montre sa maîtrise (ou non) du critère. L'évaluateur doit prendre en compte ces « indicateurs inédits » pour apprécier le degré de maîtrise du critère. Il peut donc être dangereux d'adopter des formulations telles que « **tous** les indicateurs doivent être présents pour avoir la maîtrise maximale du critère ».

Dans le même esprit, il est préférable de ne pas déplacer la « règle des 2/3 » à la gestion des indicateurs. La règle des 2/3 propose d'offrir à l'élève au moins 3 **occasions** différentes et indépendantes de montrer sa maîtrise d'un critère et de considérer que celui-ci est maîtrisé si l'élève réussit au moins 2 de ces 3 occasions. Il est préférable d'éviter d'avoir un certain nombre d'indicateurs pour un critère donné et de considérer que celui-ci est maîtrisé si au moins 2/3 des indicateurs sont présents. D'une part, on ne dispose pas dans ce cas de 3 occasions différentes et indépendantes, mais d'indicateurs différents à propos d'une même occasion. D'autre part, à nouveau, ce ne sont pas les critères qui sont évalués dans ce cas, mais bien les indicateurs.

La gestion des indicateurs est un des éléments les plus complexes de l'évaluation des compétences par situations complexes. Il n'y a sans doute pas une seule formule à privilégier et toutes les autres à rejeter. Les remarques formulées ici doivent être plus perçues comme des pistes de réflexion pour enrichir et questionner les pratiques en vigueur que comme des critiques péremptoires.

Ce qui est important est d'aboutir à l'évaluation la plus pertinente, la plus valide et la plus fiable de la maîtrise des critères. Les indicateurs sont indispensables pour réaliser ce travail, mais ils n'automatisent pas la démarche d'évaluation. Celle-ci ne pourra jamais être réalisée par une machine, mais nécessitera toujours le travail d'un expert de l'évaluation (l'enseignant). Les indicateurs informent l'évaluateur de ce qu'il doit regarder pour réaliser son inférence à propos de la maîtrise du critère. Grâce à ce faisceau d'indices, il prendra une décision sur la base d'un jugement de valeur (c'est-à-dire d'une évaluation). Il est bien sûr essentiel que cette décision ne soit pas arbitraire, laissée à la libre volonté du correcteur, sans le respect de la moindre règle. Le « jugement de valeur » doit être conduit et maîtrisé, grâce aux indicateurs et à des règles d'utilisation de ceux-ci ; il doit être lucide et pertinent ; l'évaluateur doit pouvoir à tout moment expliquer et justifier son choix... mais au bout du compte, l'évaluation reste – et doit rester – un processus éminemment subjectif, en ce sens qu'elle repose sur une série de choix de l'évaluateur

⁴ Il faut sans doute distinguer les situations « ouvertes » où les indicateurs sont donnés à titre d'exemples des situations « fermées » où les indicateurs peuvent être plus précis et plus définis.

permettant de mettre en valeur les compétences des élèves. Ces choix doivent être contrôlables et contrôlés, pertinents, systématiques. Ce sont eux qui donnent toute sa valeur au processus d'évaluation. Ils sont *l'indispensable subjectivité de l'évaluation* (GERARD, 2002).

3. La faisabilité d'une évaluation par situations complexes

Le moins que l'on puisse dire est que l'évaluation par situations complexes est un processus complexe qui présente beaucoup d'embûches pour l'enseignant désireux de s'y lancer !

Il n'est effectivement pas simple :

- d'identifier des situations suffisamment complexes et significatives pour les élèves ;
- de s'assurer que ces situations appartiennent à la famille de situations liée à la compétence qu'on souhaite évaluer ;
- de proposer dans la situation d'évaluation trois occasions indépendantes et de même niveau de complexité qui permettront d'évaluer chaque critère retenu ;
- d'identifier les meilleurs indicateurs pour chaque occasion et chaque critère ;
- de gérer ces différents indicateurs pour établir le niveau de maîtrise du critère ;
- ...

De plus, certaines problématiques pourtant importantes n'ont pas été abordées ici :

- comment identifier les ressources à mobiliser dans une situation d'évaluation pour être sûr qu'on évalue bien la compétence voulue ?
- comment prendre en compte toutes les dimensions de situations réellement significatives, dimensions cognitives, socioaffectives, conatives, sensori-psychomotrices... ?
- comment préparer les élèves à résoudre des situations complexes pour développer leurs compétences ?
- faut-il évaluer autrement selon que l'on veuille réguler le processus d'apprentissage ou certifier la maîtrise de compétences ?
- faut-il évaluer uniquement des compétences ou aussi les ressources ?

L'évaluation par situations complexes est un processus complexe, mais elle est pourtant indispensable pour évaluer des compétences. C'est toute la question de la pertinence de l'évaluation, et des outils qu'elle utilise, par rapport aux objectifs qui sont poursuivis par un système éducatif (DE KETELE & GERARD, 2005). Si celui-ci souhaite que les élèves deviennent compétents, c'est-à-dire qu'ils soient à même de résoudre des situations-problèmes en mobilisant différentes ressources, il ne suffit pas d'évaluer la maîtrise de ces ressources... Encore faut-il que les élèves soient réellement à même, face à une situation-problème complexe, de mobiliser les ressources nécessaires, c'est-à-dire de les identifier et de les utiliser de manière conjointe pour résoudre la situation. Le seul moyen de le savoir est que l'élève soit confronté à des situations complexes.

Pour autant, l'évaluation par « situations complexes construites » n'est pas la seule modalité possible pour évaluer des compétences. Si la mise en « situations complexes naturelles » est peu pertinente et difficilement faisable pour l'éducation

de base, au contraire des formations professionnalisantes, le recours au **portfolio** en tant qu'assemblage finalisé des travaux de l'élève qui démontrent ses efforts, ses progrès et ses acquisitions dans un ou plusieurs domaines (ALLAL & al., 1998 ; SCALLON, 2004) est, par exemple, un autre moyen très utilisé. Le portfolio offre plusieurs avantages en tant qu'outil d'évaluation. D'une part, il permet de partir des compétences réelles des élèves plutôt que de partir des compétences visées pour évaluer les compétences réelles (GERARD, 2006a). En ce sens, il est un véritable outil d'évaluation *authentique*. D'autre part, il implique la participation de l'élève à la sélection des contenus, à la définition des critères de sélection et d'appréciation des travaux, ainsi que le développement de pratiques métacognitives (NUNZIATI, 1990). Le portfolio devient ainsi lui-même outil d'apprentissage : en le créant, l'élève est supposé apprendre mieux et plus en profondeur en le rendant conscient de ses apprentissages. Cette idée généreuse, renforcée par le taux élevé de satisfaction des utilisateurs, n'a pas encore fait l'objet, à notre connaissance, de validations empiriques.

Même si l'évaluation des compétences par situations complexes se révèle un processus complexe, elle reste une méthode relativement économique. En effet, une fois le processus maîtrisé, les enseignants peuvent non seulement assez facilement élaborer de nouvelles situations, mais peuvent aussi s'échanger des situations. On voit ainsi se développer de plus en plus des banques de situations, que ce soit à l'échelon local, régional ou national, par exemple en Belgique, au Liban, au Québec..., ou encore au niveau international, grâce notamment à la Communauté des pratiques suscitée par le Bureau International de l'Éducation (BIE, UNESCO).

Malgré de tels échanges, l'évaluation des compétences – quelle que soit la modalité utilisée – restera vraisemblablement la phase la plus complexe pour ceux qui pratiquent l'approche par les compétences. Le recours à des situations complexes construites nous paraît indispensable. Cette modalité sera faisable pour les enseignants pour autant qu'un certain nombre de conditions soient remplies, que ce soit à l'initiative du système éducatif en tant que tel ou à celle des enseignants eux-mêmes :

- disposer d'un « référentiel de situations » décrivant les situations-cibles que l'élève doit être à même de résoudre au terme du niveau de formation concerné ;
- définir, pour chaque famille de situations, les paramètres permettant de caractériser les éléments invariants de toutes ces situations ;
- définir les critères d'évaluation qui seront utilisés, en limitant le nombre de ceux-ci tout en veillant à leur pertinence et à leur indépendance ;
- élaborer des situations qui correspondent aux paramètres de la famille de situations et qui permettent d'évaluer au moins trois fois et de manière indépendante chaque critère, notamment en proposant trois consignes ou trois tâches différentes mais aussi complexes l'une que l'autre à partir d'un même contexte et/ou de mêmes supports ;
- valider ces situations en les expérimentant, en les échangeant, en sollicitant l'avis de collègues... ;
- formaliser sur la base de cette expérimentation et/ou de ces échanges les indicateurs les plus pertinents par rapport à chaque occasion et à chaque critère, et clarifier le plus possible la manière de gérer ces indicateurs pour déterminer le niveau de maîtrise des critères et de la compétence évaluée.

Ces conditions ne pourront évidemment être remplies que dans le cadre d'une véritable professionnalisation des enseignants : ceux-ci ne peuvent plus être de simples « transmetteurs de savoirs » mais doivent devenir de véritables professionnels du processus, complexe, de l'apprentissage.

4. Bibliographie

- ALLAL, L., WEGMULLER, E., BONAITI-DUGERDIL, S. & COCHET KAESER, F. (1998). Le portfolio dans la dynamique de l'entretien tripartite. *Mesure et évaluation en éducation*, 20, 5-31.
- ASHBY, W.R. (1973). Some peculiarities of complex systems, *Cybernetic Medicine*, vol.9, 1973 pp.1-8.
- AVENIER, M.-J. (1997). La stratégie « chemin faisant », *Economica*, Paris, pp. 13-20.
- BÉLAIR, L. (1999). *L'évaluation dans l'école. Nouvelles pratiques*, Paris : ESF.
- CIES (Centre d'Initiation à l'Enseignement Supérieur) Provence, Côte d'Azur, Corse (2006). *Du tout complexe à la complexité du tout – Un essai interdisciplinaire de description de la notion de complexité*, <http://www.lsis.org/~estratm/cies/>
- CHAROLLES, M. (1988) *Les études sur la cohérence, la cohésion et la connexité textuelle depuis la fin des années soixante*, " Modèles linguistiques ", Lille : Presses Universitaires de Lille.
- CRAHAY, M. (2006). Dangers, incertitudes et incomplétude de la logique de la compétence en éducation, *Revue Française de Pédagogie*, n° 154, octobre 2006, 97-110.
- DE KETELE, J.-M. (2006). De la nécessité et de la relativité de la notion de seuil en évaluation. In BAILLÉ, J. (Éd.). *Du mot au concept : Seuil*, Grenoble : Presses universitaires de Grenoble, pp. 93-112.
- DE KETELE, J.-M. & GERARD, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences, *Mesure et évaluation en Éducation*, Volume 28, n°3, 1-26.
- DE KETELE, J.-M. & ROEGIERS, X. (1993, troisième édition 1996). *Méthodologie du recueil d'informations*, Paris-Bruxelles : De Boeck-Université.
- DELORME, R. (1998), *From first order to second order complexity in economic theorizing*, Paris : CEPREMAP, Mimeo.
- DELORME, R. (1999), De l'emprise à l'en-prise. Agir en situation complexe, In : GRASCE (Groupe de recherche en analyse de système et calcul économique) (Ed.). *Entre systémique et complexité, chemin faisant...*, *Mélanges en hommage à Jean-Louis le Moigne*, pp. 31-32, Paris : Presses universitaires de France.
- DELORME, R. (2000). Théorie de la complexité et institutions en économie, *Colloque « Organisations et institutions en économie »*, Amiens, 25-26 Mai 2000.
- GERARD, F.-M. (2002). L'indispensable subjectivité de l'évaluation, *Antipodes*, n°156, avril 2002, 26-34.
- GERARD, F.-M. (2005), L'évaluation des compétences par des situations complexes, *Actes du Colloque de l'Admee-Europe*, IUFM Champagne-Ardenne, Reims, 24-26 octobre 2005.
- GERARD, F.-M. (2006a). Les outils ouverts d'évaluation, ou la nécessité de clés de fermeture, In : FIGARI, G. & MOTTIER LOPEZ, L. (Éds). *Recherche sur l'évaluation en éducation - Problématiques, méthodologies et épistémologie*. Paris : L'Harmattan, pp. 147-154.

- GERARD, F.-M. (2006b). L'évaluation des acquis des élèves dans le cadre de la réforme éducative en Algérie, in *Réforme de l'éducation et innovation pédagogique en Algérie*, Ministère de l'éducation nationale, Programme d'appui de l'UNESCO à la réforme du système éducatif, UNESCO-ONPS ; 2006, pp. 85-124.
- GERARD, F.-M. & ROEGIERS, X., 2006. *Mise au point relative à la notion de paramètre d'une famille de situations*, Document inédit, BIEF.
- JONNAERT, Ph. (2002). *Compétences et socioconstructivisme - Un cadre théorique*, Bruxelles : De Boeck.
- JONNAERT, Ph., BARRETTE, J., MASCIOTRA, D. & YAYA, M. (2006). La compétence comme organisateur des programmes de formation revisitée, ou la nécessité de passer de ce concept à celui de « l'agir compétent », Montréal : ORÉ.
- LE MOIGNE, J.-L. (1990), *La modélisation des systèmes complexes*, Paris : Bordas.
- MORIN, E. & LE MOIGNE, J.-L. (1999). *L'intelligence de la complexité*, Paris : L'Harmattan.
- NUNZIATI, G. (1990). Pour construire un dispositif d'évaluation formatrice. *Cahiers pédagogiques*, 280, 47-64.
- 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.
- SCALLON, G. (2004). *L'évaluation des apprentissages dans une approche par les compétences*. Bruxelles : De Boeck Université.