

Chapitre 8

PBL Problem Based Learning ou APP Apprentissage Par Problèmes

Dieudonné LECLERCQ et Cees VAN DER VLEUTEN

A.	LA PEDAGOGIE MEDICALE S'AUTO-CRITIQUE	2
B.	L'APPRENTISSAGE PAR PROBLEMES DANS LE MONDE	2
C.	CURRICULUM EN SPIRALE	4
E.	PAS DE COURS <i>EX-CATHEDRA</i>	5
F.	CAS CONCRETS ET INTERDISCIPLINARITE	5
G.	LA METHODE DES "7 JUMPS"	6
H.	UN PAYSAGER D'ETUDE	6
I.	LE <i>SKILLSLAB</i> ET LES STAGES	6
J.	EQUILIBRE ENTRE L'OBLIGATOIRE ET L'OPTIONNEL	7
K.	LE ROLE DES ENSEIGNANTS EST MODIFIE	7
L.	LES TESTS DE BLOC	8
M.	LES TESTS "DE PROGRES"	8
N.	<i>FEEDBACKS</i> DETAILLES AUX TESTS DE PROGRES	9
O.	ETUDES COMPARATIVES DES RESULTATS DE L'APP EN MEDECINE	10
P.	DISCUSSION	

A. LA PEDAGOGIE MEDICALE S'AUTO-CRITIQUE

Dans les années '70 et '80, de nombreuses facultés de médecine ont remis en cause leur pédagogie, sous la pression de principes de l'OMS visant à donner priorité aux soins de santé primaires. On a investi aussi bien le préventif que le curatif. Dans le tiers monde¹. Notamment, on a voulu éviter des situations telles que

- "a[...] une relation inversement proportionnelle entre la fréquence d'une maladie sur le terrain et l'importance qui lui a été accordée durant les études de médecine;
- b[...] les facultés d'observation, de déduction et d'initiative des étudiants restent atrophiées par rapport aux facultés de mémorisation;
- c[...] les étudiants sont initiés à des techniques complexes de diagnostic et de traitement, alors qu'il est fréquent qu'un hôpital de district ne dispose que d'un seul médecin (secondé par une infirmière qualifiée et une poignée d'auxiliaires) pour plus de 100 000 personnes et ne possède ni appareil de radiologie, ni eau courante, ni salle d'opération et seulement un maigre stock de médicaments divers... »(Mc GAGHIE *et al.*, 1978, 12).

Des diagnostics d'insatisfaction ont été posés sur la pédagogie médicale de l'époque :

« ... les causes profondes {d'insatisfaction}, particulièrement dans les pays qui ont subi l'influence coloniales [sont les suivantes] :

- l'enseignement médical est généralement calqué sur un modèle étranger ...
- de nombreuses écoles de médecine sont isolées de la clientèle que leurs diplômés seront appelés à servir ...
- la préoccupation principale de certains enseignants éminents est plus de comprendre la maladie que de préserver la santé ...
- les étudiants entrent en contact avec les malades dans des hôpitaux d'enseignement ... généralement peuplés de cas complexes et non résolus, de sorte que l'enseignement clinique a tendance à mettre l'accent sur le diagnostic et le traitement de troubles rares, plutôt que de troubles fréquents" (Mc GAGHIE *et al.*, 1978, 12).

Aussi bizarre que cela paraisse, les pays développés ont parfois des problèmes semblables à ceux du Tiers Monde en ce qui concerne la couverture médicale. Ainsi, dans l'Etat d'Illinois, de vastes zones rurales sont dépourvues de médecin (il arrive qu'une femme enceinte soit à plus de 200 km de tout gynécologue !). La formation médicale "adaptée aux besoins locaux" est devenue une préoccupation mondiale (COLES, 1991).

B. L'APPRENTISSAGE PAR PROBLEMES DANS LE MONDE

La pédagogie par problèmes est proche, comme on va le voir, de la pédagogie par projets.

Deux références incontournables concernant cette dernière approche sont d'une part John DEWEY (1905), philosophe de l'éducation américain et Célestin FREINET, instituteur français (1896-1966). Au centre de la pédagogie par projets, quatre principes : (1) la définition du projet par l'apprenant lui-même ; (2) le décloisonnement des matières et l'approche interdisciplinaire ; (3) la responsabilité personnelle de l'apprenant dans sa propre formation et (4) la coopération entre apprenants.

Partir des problèmes n'équivaut pas totalement à cette approche, en tout cas quant au premier principe, dans la mesure où les problèmes ne sont pas choisis par les apprenants. Néanmoins,

¹ Sur ce sujet, on consultera, pour la méthodologie, GACHIE *et al* (1978) et, pour la description de diverses expériences dans le monde, RICHARDS *et FULOP* (1989) d'une part et KRANTOWITZ *et al.* (1989) d'autre part.

après l'analyse du problème, l'étudiant (ou le ici le groupe) devra se définir un projet d'apprentissage. On rejoint donc ainsi la pédagogie du projet.

SCHERLY (1997) fait remarquer que l'APP est aussi fondée sur les préoccupations de la « *Situated Learning* » (BILLETT, 1996) qui vise à rendre similaires les contextes de l'apprentissage et de la pratique du métier donc à partir de problèmes authentiques et à apprendre ce qui sera utile à leur solution. Il en va de même pour l'« *Enchored learning* » (CTGV, 1992) qui insiste sur l'ancrage de l'apprentissage dans des situations concrètes et le « *Cognitive Apprenticeship* » (COLLINS *et al.*, 1989) qui tire parti des échanges entre apprenants. Les théoriciens de l'APP les plus connus sont BARROWS et TAMBLYN (1977, 1980), de la faculté de médecine de la *Southern Illinois University*. Selon eux (1980, 163-164), un problème « didactique » devrait :

1. présenter les informations comme le praticien les reçoit habituellement du patient, et non pas sous la forme d'un résumé prédigéré ;
2. amener l'étudiant à prendre des décisions qui s'enchaînent, comme dans la vie réelle, les actions n'étant plus réversibles ;
3. pouvoir être abordé de différentes façons conduisant éventuellement à des résultats différents en fonction de la stratégie ou de la compétence des étudiants ;
4. permettre de mettre en pratique et d'évaluer les différentes phases du raisonnement clinique (1980, 19-36) :
 - (a) Perception et interprétation de l'information : se forger une « image » du patient ;
 - (b) Génération d'hypothèses ;
 - (c) Enquête (questions, examens, tests, prélèvements) pour nuancer, vérifier, éliminer certaines hypothèses (b), sur base de la compétence clinique ;
 - (d) Synthèse du problème ;
 - (e) Diagnostic et décision thérapeutique ;
 - (f) Communication(s) au patient ;
5. être supporté par des documents multimédias (photo, radiographie, vidéo, bande sonore, ...).

C'est à la Faculté de Médecine de l'université *Mac Masters* (Hamilton, Ontario) qu'eut lieu le premier lancement de cette approche sur le terrain. Rapidement, l'approche *Problem based Learning* fut adoptée par de nombreuses autres facultés de médecine : la moitié des canadiennes, aux USA celles de Harvard (STARR, 1982) et de Stanford, etc. Les Québécois ont traduit PBL par APP (Apprentissage Par Problèmes). On peut tout aussi bien traduire par ABP (Apprentissage Basé sur les Problèmes), traduction plus proche de l'original anglais.

Lors de sa fondation en 1974, la Faculté de Médecine de Maastricht, qui forme des médecins généralistes en six ans, a adopté l'approche APP mais en a fait une "version Maastricht" originale, la plus « en pointe » dans le monde. Ses conditions d'implantation à Maastricht ont bénéficié de deux circonstances favorables : d'une part une volonté politique du Sud Limbourg néerlandais d'avoir une faculté de médecine et de la supporter, même (et surtout) dans son originalité méthodologique par rapport aux facultés anciennes du pays, d'autre part l'installation de l'APP dès la création de la faculté (seuls des enseignants adhérant à ces principes ont été recrutés). Il est plus facile de naître que de renaître !

L'Université de Sherbrooke (Québec) qui, au milieu des années '80, a opté pour cette même approche a dû démolir ses installations pédagogiques au bulldozer (les amphithéâtres) pour les remplacer par des infrastructures appropriées (grand nombre de petites pièces pouvant contenir dix personnes).

Si c'est la faculté de médecine qui a lancé le mouvement à Maastricht, les autres facultés l'ont rejointe sur ce terrain : les facultés d'économie (GIJSELAERS et SCHMIDT, 1995), de droit, de sciences de la santé, de psychologie et de sciences pratiquent aussi l'APP.

On trouvera ci-après les grands constats, diagnostics et principes mis en oeuvre en faculté de médecine dans les 4 premières années (1 à 4), les deux dernières (5 et 6) étant réservées aux "rotations cliniques" dans les hôpitaux et chez les généralistes praticiens. L'entrée en première année se fait sur base d'un *numerus clausus* de 200 étudiants par an, sélectionnés au hasard parmi les candidats ayant satisfait aux critères (élevés) d'admission.

La description qui suit doit beaucoup aux écrits de VANDERVLEUTEN et WIJNEN (1990), NORMAN et SCHMIDT (1992), DOLMANS, VANDERVLEUTEN et WOLFHAGEN (1998), LECLERCQ, (1998).

C. CURRICULUM EN SPIRALE

Le curriculum des quatre premières années est découpé en blocs de 6 semaines, une année comportant 6 de ces blocs. Chaque bloc est centré sur une fonction et ses pathologies associées (respiration, système cardiaque, etc..). Ainsi, le thème des douleurs dans le dos permettra d'étudier le squelette et le passage des nerfs dans les vertèbres. Le **curriculum** est conçu "en spirale" : les mêmes notions sont revues au cours des années ultérieures, mais de façon plus approfondie (principe aussi appliqué dans l'approche LQRT-SAFE, qui a le souci de "Considérer la logique de l'apprenant").

D. DES GROUPES TUTORIELS

Les 200 étudiants de chacune de ces quatre années sont répartis en **petits groupes** de 8 étudiants dont la composition change, au hasard, à chaque bloc. Ces groupes se réunissent deux fois par semaine dans des « sessions tutorielles » de deux heures. Un *tutor* (professeur, assistant ou moniteur) assiste le groupe dans sa démarche mais n'enseigne pas ! A la fin de la *tutoring session*, il fait des remarques méthodologiques au groupe. Dans le groupe, un des étudiants est le *chairman* (qui anime et régule les débats) et un autre le *recorder* (qui note au tableau). Ces fonctions sont réparties selon une tournante de façon à être également exercées par tous les étudiants.

Le groupe travaille en deux temps :

- a) La première heure est consacrée à la mise en commun du travail. Chacun verbalise – dans ses propres termes – ce qu'il a appris à propos du cas examiné lors de la session tutorielle précédente. En quelque sorte, les étudiants jouent tour à tour le rôle d'enseignant pour les autres (sans que ce principe soit aussi poussé que dans les Projets d'Animations Réciproques).
- b) Lors de la deuxième heure, le groupe est confronté à un problème, un cas médical concret présenté par écrit (le patient ne sera physiquement présent qu'à partir de la 4^e année). Le cas est alors analysé selon la méthode des "7 jumps" (voir ci-après). Suite à une répartition du travail entre les membres du groupe, chacun ira consulter des ouvrages de référence.

SCHERLY (1997) signale que « Selon les écoles, on peut dégager deux attitudes face au rôle des tuteurs (SCHMIDT et MOUST, 1995). Une perspective met l'accent sur les qualités personnelles du tuteur : être capable de communiquer d'une manière informelle avec les étudiants, de les

encourager à apprendre par l'intermédiaire des échanges d'idées. Les partisans de cette attitude estiment qu'un tuteur expert peut court-circuiter l'auto-apprentissage et ainsi être un obstacle au processus même de l'APP. L'autre attitude considère l'expertise du tuteur dans le domaine touchant le problème proposé aux étudiants comme déterminante pour l'apprentissage. Ceci est en partie confirmé par les expériences de NEEDHAM et BEGG (1991) qui montrent qu'un *feedback* correctif rapide est très important dans l'apprentissage. L'étude de SCHMIDT & MOUST (1995) a montré qu'en fait le tuteur idéal devrait rassembler trois qualités : expertise dans la matière traitée, engagement personnel pour l'apprentissage des étudiants et capacité à s'exprimer en utilisant le langage des étudiants. »

E. PAS DE COURS *EX-CATHEDRA*

Traditionnellement, l'apprentissage autonome des étudiants est difficile à l'université parce que les meilleures heures de la journée sont occupées par les cours *ex-cathedra*. C'est pourquoi ceux-ci sont quasiment supprimés dans l'APP à Maastricht. Des « conférences » (en anglais *Lectures*) existent encore, mais elles sont assez rares (une par semaine) et servent de support à l'apprentissage. Elles ne sont pas obligatoires.

F. CAS CONCRETS ET INTERDISCIPLINARITE

Dans l'approche traditionnelle, beaucoup d'étudiants quittent la 1^{re} ou la 2^e année de médecine, démotivés, notamment parce qu'ils étudient la chimie, la physique, la biologie pendant des années avant de voir un patient ou un « cas », ce qui est démotivant.

Au contraire, dans les trois premières années de l'APP, on part du cas d'un patient, dans sa version écrite, avec ses plaintes, ses radiographies, ses analyses d'urine et de sang, etc., comme dans la réalité, c'est-à-dire toutes disciplines confondues.

A partir de la 4^e année, les patients « didactiques » peuvent être réels ou simulés (en anglais « *standardized patients* », c'est-à-dire rendus conformes à des normes) : ils ont alors appris leur rôle comme des acteurs. La faculté de médecine de Maastricht fonctionne avec un pool de 130 patients didactiques. Ces derniers, souvent des volontaires (pour aider les étudiants à devenir de bons médecins), sont entraînés à cette fonction.

Voici un tel « cas » (DOLMANS *et al.*, 1998, 2) :

Au petit matin, Paul rentre saoul à la maison, après une nuit de copieuses libations. Il essaye d'ouvrir la porte avec la clé, mais plus il rapproche sa main de la porte, plus sa main est agitée de tremblements. Quand, finalement, il se retrouve chez lui à l'intérieur, il a le sentiment que la pièce tourne, même quand il ferme les yeux. Ses jambes vacillent, ont du shimmy. Il essaye de se faire une tasse de café, mais après avoir craqué une série d'allumettes et s'être brûlé, il y renonce.

G. LA METHODE DES "7 JUMPS"

D'habitude, la majorité des étudiants ne savent pas comment apprendre seuls parce qu'on ne les y a pas entraînés systématiquement. C'est au contraire une approche méthodique qu'adopte l'APP via les « 7 jumps » : 7 étapes pour aborder un problème, gérer son ignorance (ou plutôt sa connaissance partielle) et "apprendre".

SCHMIDT (1983) définit comme suit ces *seven jumps* :

1. Identifier et clarifier les termes, concepts, valeurs numériques, etc., non compris 2. Définir le problème (quel phénomène s'agit-il d'expliquer ?)
3. Analyser le problème (hypothèses causales, créativité nécessaire pour générer des idées)
4. Inventaire systématique (coapparitions, connections, liens,....fixation de priorités dans la vérification)
5. Formulation des objectifs d'apprentissage et distribution des tâches 6. Etude individuelle
7. Synthèse (collective) et critique des informations apportées par chacun. Mise à l'épreuve de l'acquis en essayant de résoudre un (ou plusieurs) autre(s) cas dans le même thème.

H. UN PAYSAGER D'ETUDE

Traditionnellement, l'ambiance d'une université (avec ses amphithéâtres en pente, etc.) se prête mal au travail individuel de centaines d'étudiants simultanément. Si l'on veut favoriser ce travail de recherche et d'étude individuels, l'architecture et la logistique doivent être conçus en conséquence. C'est pourquoi le vaste « *Study landscape* » (paysager d'étude) combine deux types de bibliothèques :

- a) La bibliothèque habituelle où plusieurs milliers d'ouvrages se trouvent en un seul exemplaire.
- b) La bibliothèque "de travail" où les ouvrages fondamentaux se trouvent en plusieurs dizaines d'exemplaires ainsi qu'une importante documentation audio-visuelle (diapos, films, vidéos, CD Roms). C'est donc une **médiathèque spécialisée**. Un étudiant (ou un groupe) peut y rassembler des livres sur une table de travail et les y laisser toute une journée.

I. LE SKILLSLAB ET LES STAGES

Les habiletés pratiques ne s'acquièrent pas de la même façon que les théoriques. L'APP s'est dès lors doté d'un système propre à l'acquisition de ces techniques (VAN DALEN, 1990, 19) : le Skillslab ou Laboratoire d'entraînement pratique.

Des salles sont spécialisées en vue d'entraîner à des habiletés précises :

- a) des capacités d'examen physique (durant 50 % du temps passé au Skillslab) :
 - ausculter par stéthoscope, palper, examiner les yeux, les oreilles, le rectum, etc. ;
 - prendre la température, le pouls, faire des injections, des prélèvements;
- b) des habiletés thérapeutiques : sutures, bandages, intubations, etc. (15 % du temps de Skillslab) :

- c) des habiletés de laboratoire : prises de sang, analyses d'urine, d'expectorations, de biopsie (15 % du temps) etc.;
- d) des capacités relationnelles, des savoir-être : un entraînement à la relation médecin/patient (psychologie médicale) : mener une anamnèse, conclure un entretien, communiquer un diagnostic, prescrire, former le patient à sa maladie (20 % du temps).

Dans ce *Skillslab*, des patients, vrais ou simulés, sont souvent présents et ce durant les quatre premières années. En outre, des **stages sur le terrain** (en hôpital, au cabinet médical) se déroulent en compagnie de praticiens professionnels et permettent la mise en œuvre de techniques spécialisées (radiographies, endoscopie, ...).

J. EQUILIBRE ENTRE L'OBLIGATOIRE ET L'OPTIONNEL

L'auto-formation a besoin d'être stimulée. Un minimum d'activités organisées est nécessaire parce qu'en-deçà de 10 heures "obligatoires" (en gris dans l'horaire ci-dessous), la stimulation est insuffisante et au-delà de vingt heures, les étudiants n'ont plus le temps de s'auto-former (VAN DER DRIFT et VOS, 1987). Voici une semaine typique dans l'agenda d'un étudiant (d'après VAN DER VLEUTEN, 1996, 11), où chaque activité dure deux heures :

Jour	Matin	Après-midi
Lundi		<i>Skillslab</i>
Mardi	GROUPE TUTORIEL	Conférence (<i>lecture</i>)
Mercredi	<i>SkillsLab</i>	
Jeudi		Stage chez un généraliste
Vendredi	GROUPE TUTORIEL	

K. LE ROLE DES ENSEIGNANTS EST MODIFIE

Traditionnellement, les professeurs sont peu disponibles pour les étudiants bien qu'ils se sentent eux-mêmes écrasés par les charges de cours. En fait, ils sont accaparés par la transmission du contenu dont il faut les libérer pour les rendre disponibles pour d'autres tâches. Dans l'APP, leur **rôle** consiste désormais à :

- créer l'environnement d'apprentissage
- repérer des cas, tenir à jour l'information scientifique
- créer des questions
- assister les étudiants dans leur auto-apprentissage et les y former

Toutes ces tâches sont comptabilisées pour calculer la "charge" de l'enseignant. Par exemple, "*tutorer*" un groupe de huit étudiants vaut 30 heures, être membre d'un "*planning group*" (qui conçoit les cas, etc.) vaut 30 heures également, et être membre d'une commission de création de tests vaut 120 heures. Pour devenir *chairman* d'un *planning group*, il faut en avoir été membre, et pour en être membre, il faut avoir été tuteur.

Les tuteurs sont évalués par les étudiants. Si plusieurs rapports consécutifs sont négatifs, des dispositions sont prises pour changer la situation.

Le plus souvent possible, chaque enseignant, pour une matière donnée (ex : la cardiologie) assume SOIT une fonction d'évaluation, SOIT une fonction de formation. Dans ce dernier cas, il

est donc entièrement engagé aux côtés des étudiants vers leur réussite. Ainsi, le tuteur du groupe n'a pas de fonction d'évaluation. L'année suivante, le professeur peut changer de fonction dans chaque discipline.

Le professeur perd de son « pouvoir » traditionnel. Par exemple, il peut « soumettre » des questions à l'équipe d'évaluation, mais celle-ci n'est pas forcée de les accepter. Cette révision des questions par les pairs amène à en éliminer 15% (hors matière, ambiguës, trop difficiles...). Après que les questions aient été posées, 10 à 15 % d'entre elles sont encore éliminées sur base de leurs caractéristiques psychométriques et des critiques des étudiants, qui ont 15 jours, après l'affichage des résultats et des solutions correctes, pour contester les questions par écrit (PRINCE et VISSER, 1997).

L. LES TESTS DE BLOC

D'habitude, les étudiants ne se rendent compte que très (trop) tard des exigences des cours et de leur degré de maîtrise de la matière. Or ils devraient pouvoir se confronter, immédiatement après l'étude de chaque matière, à une évaluation formative, ce qui implique banque de questions, diagnostic, prescriptions et éventuellement remédiation.

Dans l'APP en médecine à Maastricht, des tests obligatoires sont administrés après l'étude de chaque thème (ou bloc de six semaines), avec conseil (éventuel) de cours de rattrapage organisés.

M. LES TESTS "DE PROGRES"

Traditionnellement, une faculté (et encore moins une université entière) est incapable de dire, preuves à l'appui, si les étudiants s'améliorent ou si leur qualité baisse d'une année à l'autre. La raison en est que les épreuves ne sont pas comparables d'une session à l'autre, puisqu'elles ne sont pas créées à partir de grilles d'objectifs de référence, avec accord de tous les professeurs concernés sur la pertinence et la qualité de chaque question de la banque (ce qui force à la concertation), avec échantillonnage rigoureux, avec vérification des qualités métriques de chaque question, etc.

Dans l'APP en médecine à Maastricht, chaque année, à trois mois d'intervalle, les étudiants ont quatre occasions d'évaluations "sanctionnées" via les Tests de Progrès. Ces tests comptent environ 250 questions à consigne « VRAI-FAUX GENERALISE ² ». Ils portent sur toute la matière de médecine.

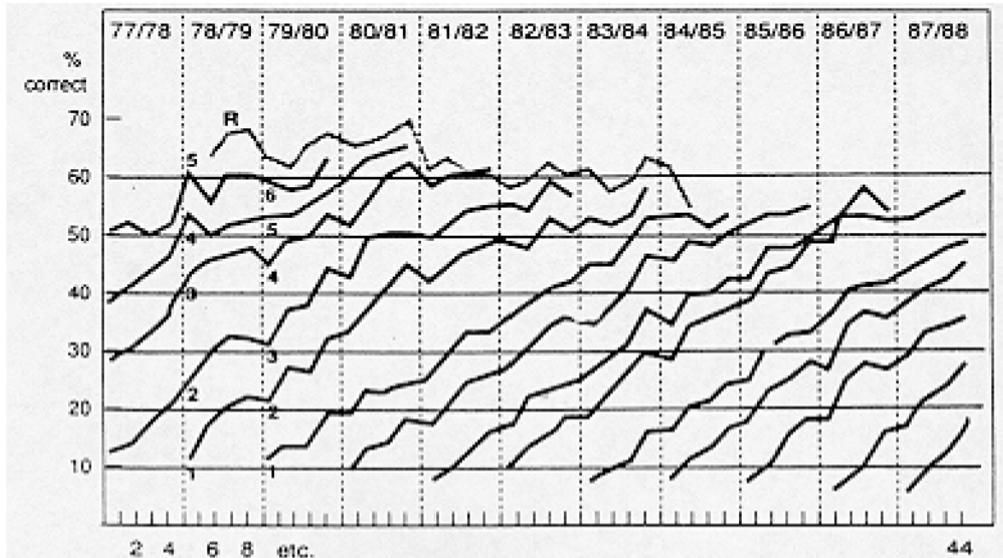
Cette formule présente un avantage pédagogique non négligeable : le stress des jours précédant l'examen est diminué, dans la mesure où il est impossible de se préparer à TOUTE la matière de médecine dans la semaine qui précède le test. La solution consiste à faire, au jour le jour, tout au long de l'année, le mieux possible son métier d'apprenant (NEWBLE et JAEGER, 1983 ; SCOTT et CHAFE, 1997). Ceci contraste avec les habitudes du classique "*Test directed studying*" : les apprenants étudient ce sur quoi ils seront testés ... pour l'oublier peu après (SEMB et ELLIS, 1994).

Les étudiants des 6 années passent ce même test ensemble dans un grand hall de la ville loué pour l'occasion. Chacun peut ainsi mesurer ses résultats, les progrès qu'il a réalisés depuis le test

²VFG = QCM où l'étudiant doit fournir une réponse VRAI-FAUX pour chaque solution proposée. En cas de réponse correcte, l'étudiant reçoit 1 point, en cas d'incorrecte, il en perd 1. C'est la « *correction for guessing* » classique (voir LECLERCQ, 1987).

précédent, parallèle, portant sur la même matière. Il est normal que les étudiants des années inférieures ignorent (en écrivant "?") de nombreuses réponses. En conséquence, le taux de réussite attendu diffère selon les années. Ainsi, on attend de l'étudiant de première année qu'il réussisse 10% des questions, de celui de deuxième année 20%, etc.. et de celui de sixième année 60%. S'il n'atteint pas le pourcentage fixé pour son année, l'étudiant repasse une seconde, voire une troisième fois ce test et c'est seulement après ces repêchages qu'il se voit contraint de doubler son année.

Voici le graphique des résultats (44 tests au total, comme l'indique l'axe des x) obtenus par des cohortes différentes et par les mêmes cohortes au cours des années (VANDERVLEUTEN et WIJNEN, 1990, 31) :



Comme comparaison inter-cohortes au cours de la même année de calendrier, prenons à titre d'exemple l'année 79/80. Les étudiants de première année (E1) ont obtenu un score d'un peu plus de 10% dès le premier test et de près de 20% lors du dernier test de l'année. Les E2 dépassaient 20% lors du premier test et même 30% lors de leur dernier (4ième) test de cette année. La moyenne des E6, cette année-là a atteint les 60% (score plancher de réussite) lors du premier et du dernier test, mais pas lors des tests 2 et 3.

En haut du graphique, les scores obtenus par les **R** ou **praticiens en fonction** (non formés à Maastricht) sont supérieurs aux scores des E6, ce qui conforte les responsables dans leur conviction que le système forme à « apprendre tout au long de sa vie professionnelle » (VANDERVLEUTEN et WIJNEN, 1990, 31).

N. FEEDBACKS DETAILLÉES AUX TESTS DE PROGRES

D'ordinaire, les étudiants reçoivent trop peu de précisions sur la qualité de leurs résultats. Or, chacun devrait pouvoir se situer au moins par rapport aux autres (référence normative). La solution dans l'APP en médecine à Maastricht consiste à ce que chaque étudiant reçoive un dossier individuel après chaque épreuve (voir *fac-similé* ci-dessous). En outre, les étudiants reçoivent le test et les réponses correctes, qu'ils peuvent critiquer sur base d'arguments tirés de la littérature (ils ont deux semaines pour le faire).

Résultats par discipline (en %)

Description	Résultats individuels				
	Nb Q	Cor.	Inc.	?	C-I
1. Anatomie/embryologie	13	54++	0--	46	54++
2. Biochimie	13	77+	8-	15	69++
3. Pharmacologie	7	57+	0-	43	57++
4. Physiologie	11	73	0--	27++	73++
5. Génétique – Biol. Cell.	8	50	25	25	25
6. Immunologie	4	75+	25+	0-	50
7. Microbiologie médicale	6	83	0-	17+	83
8. Pathologie	10	60	0-	40	60
Total des sciences de base	72	65+	6--	29	60++
9. Médecine du travail	1	0	0	100	0
10. Chirurgie générale	16	38	6--	56	31++.
11. Cardiologie	10	30--	10-	60++	20
12. Dermatologie	5	40	0	60	40
13. Gynécologie – Obstétrique	6	33	17	50	17
14. Pratique familiale	9	44	11--	44+	33+
15. Médecine interne	20	65++	10-	25-	55++
16. Pédiatrie	8	38	0--	63+	38+
17. Nez, gorge, oreilles	5	20	20+	60	0
18. Neurologie	2	50+	0	50-	50++
19. Ophtalmologie	3	33	33++	33-	0
20. Orthopédie	5	20+	0	80	20++
21. Pneumologie	5	40	0-	60+	40
22. Médecine de réadaptation	3	0-	33+	67	-33
23. Urologie	1	100+	0	0	100+
Total des sciences cliniques	99	40	9-	51	31+
24. Soins de santé	2	0	100++	0-	100-
25. Epidémiologie	12	75++	8--	17	67+
26. Législation sanitaire	2	50	50++	0--	0
27. Psychiatrie clinique	6	0-	0-	100++	0
28. Psychologie médicale	3	33+	33+	33-	0
29. Sociologie médicale	2	100++	0	0-	100++
Total sciences comptmentales	27	48	19	33	30
GRAND TOTAL	198	51+	9--	40	41++

Résultats collectifs (n = 54)							
Cor.	Ec.typ	Inc.	Ec.typ	?	Ec.typ	C-I	Ec.typ
39	14	20	16	40	19	19	23
68	13	15	11	17	11	52	21
38	19	14	17	48	25	24	27
67	16	19	11	14	12	48	24
57	16	22	10	21	14	36	23
65	17	15	15	19	20	50	25
85	12	6	8	10	13	79	16
59	14	7	10	34	16	52	18
58	9	15	7	26	11	43	11
13	34	7	26	80	40	6	45
33	13	18	11	48	19	15	14
49	17	21	14	29	21	28	22
31	33	7	15	62	34	25	39
29	21	13	13	57	21	16	29
40	14	27	13	33	17	13	21
49	13	17	8	35	15	22	16
36	19	16	14	48	16	20	29
22	18	8	15	70	22	14	25
19	31	3	11	79	34	16	32
37	28	9	17	54	29	28	37
10	15	3	7	87	19	6	13
42	22	10	13	48	22	33	28
15	22	12	23	73	26	3	37
74	44	15	36	11	31	59	73
37	10	15	7	48	14	21	10
6	16	42	42	53	46	-36	43
65	18	18	9	17	18	48	22
40	40	13	26	47	42	27	52
18	19	10	16	72	28	8	22
17	25	18	25	65	41	-1	27
52	37	7	18	41	41	44	42
42	13	17	8	41	20	26	10
45	9	16	6	39	13	30	9

Résultats par catégories (en %)													
DESCRIPTION	Résultats individuels					Résultats collectifs (N = 54)							
	NbQ	Corr	Incor.	?	C-I	Corr	Ec.typ	Incor.	Ec.typ	?	Ec.typ	C-I	Ec.typ
1. Système respiratoire	18	56++	6--	39	50++	44	11	23	10	33	14	22	16
2. Hématol. & syst. Lymphat.	15	73+	0-	27+	73++	65	12	17	9	18	11	48	19
3. Système musculo-osseux	19	37+	16	47-	21	30	9	12	10	58	16	18	11
4. Soins de santé mentale	9	11	11	78	0	18	18	13	15	70	28	5	16
5. Système reproductif	13	46+	15	38-	31+	35	13	16	9	49	15	19	17
6. Système cardio-vasculaire	23	43-	4--	52++	39	54	15	19	11	27	17	35	21
7. Syst. endocrinien + métabol.	14	57-	14	29+	43	65	13	14	9	22	11	51	18
8. Peau	9	22	11	67	11	27	21	14	12	59	26	14	21
9. Santé sociale	7	43+	43++	14--	0	30	20	19	18	52	32	11	19
10. Système digestif	19	58++	5-	37	53++	45	12	15	10	40	17	29	15
11. Système urinaire	12	50+	0--	50+	50++	44	11	16	13	40	18	28	17
12. Système nerveux	21	48++	10	43-	38++	31	12	12	10	57	18	18	14
13. Divers	9	78	0-	22+	78	82	13	6	6	12	13	76	16
14. Méthodologie et Philosophie	10	80+	10-	10	70+	67	20	19	12	14	19	49	26
GRAND TOTAL	198	51+	9--	40	41++	45	9	16	6	39	13	30	9

Légende :

Corr. = correct ; Inc.= Incorrect ; ? = réponses par le signe ? ; Ec.typ = Ecart-type. C-I = Différence entre Taux de Corrects © et Taux d'Incorrects (I)

Les croix (++ ou +) et les signes négatifs (- ou --) indiquent quand l'étudiant est supérieur ou inférieur à la moyenne collective.

O. ETUDES COMPARATIVES DES RESULTATS DE L'APP EN MEDECINE 1.

La définition des critères

De nombreuses études ont été consacrées à l'évaluation de l'efficacité de l'APP comparée à celle d'autres méthodes de formation médicale. Le taux d'abandons (*dropouts*) mérite d'être pris en considération en premier lieu : moins de 10 % à Maastricht alors qu'il est de 30 % en moyenne pour la Hollande. Le temps moyen d'études de médecine est aussi un indicateur très précieux : de 6,2 ans à Maastricht alors qu'il est de 7,5 ans en moyenne en Hollande.

2. Les principes des méta-analyses

Devant la diversité des études, de leurs objets, de leurs méthodes, de leurs instruments, de leurs résultats, s'est développée la nécessité d'en donner une vue synthétique, résultant de l'analyse secondaire des analyses de premier degré ou résultant de l'agrégation de plusieurs analyses, c'est-à-dire de méta-analyses. Dans son célèbre article « *Primary, Secondary and Meta-Analysis of Research* », GLASS (1976), partant du principe que « mille doctorats sont muets si personne ne les a lus » et que « les données sont dans les livres, mais la connaissance est dans les cerveaux » a développé des démarches et des formules devenues classiques.

Central dans son apport est le concept d'Ampleur de l'Effet (A.E.) en anglais *Effect Size (E.S.)*. Il s'agit de la différence des moyennes du groupe expérimental et du groupe contrôle, différence rapportée à (divisée par) l'écart-type du groupe contrôle.

Une variante a été proposée par HEDGES et OLKIN (1985) qui rapportent la différence à l'écart-type des deux distributions fusionnées (celle du groupe contrôle et celle du groupe expérimental).

Cette « standardisation » des résultats contourne la difficulté des différences de métrique d'une étude à l'autre.

D'habitude, une méta-analyse, présente ses données sous la forme d'un tableau comportant les colonnes suivantes : noms des auteurs de l'étude, date, lieu, nombre de sujets dans cette étude, nature de la méthode (ou variantes), nature des mesures d'efficacité (ex. : test QCM sur les connaissances de base ; capacité clinique jugée au lit du malade, motivation à poursuivre des études, degré de satisfaction mesuré par questionnaire, etc.), type de méthode de recherche (ex. : comparaison du groupe avec lui-même dans le temps, comparaison « statique » de deux groupes, etc.), Ampleur de l'Effet.

Les lignes du tableau sont constituées par les études et leurs données. Il apparaît immédiatement qu'un obstacle important à la comparaison et à la fusion des résultats entre plusieurs recherches est la diversité des objets étudiés et des méthodes de ces études. C'est pourquoi, dans une métaanalyse, on trouve plusieurs tableaux.

Ainsi, la méta-analyse de l'efficacité de l'APP réalisée par VERNON et BLAKE (1993) comporte 4 tableaux, c'est-à-dire 4 méta-analyses, des études sur l'efficacité de l'APP :

- a) Un tableau des comparaisons des évaluations (subjectives) du programme par les étudiants (ce tableau regroupe 15 études).
- b) Un tableau des comparaisons sur le rendement des étudiants aux tests de *National Boards of Medical Examinations* (NBME) comportant 8 études présentant 28 séries de données différentes.

- c) Un tableau des comparaisons de performances cliniques, comportant 12 études présentant 20 séries de données différentes.
- d) Un tableau des comparaisons des connaissances cliniques, comportant 4 études présentant 12 séries de données.

Deux autres méta-analyses, celle d'ALBANESE et MITCHELL (1993) et celle de BERKSON (1993) parues la même année dans la même revue (*Academic Medicine*) présentent elles aussi plusieurs tableaux récapitulatifs.

3. Résultats des méta-analyses

De ces trois méta-analyses, nous relevons que :

- a) Parmi les facultés pratiquant l'APP, on peut observer une **grande variété** de résultats par exemple entre les facultés de la *Michigan State University* ou du Colorado d'une part, et dans les facultés de Mac Master (Ontario) ou de New Mexico d'autre part. Les unes peuvent avoir une Ampleur d'Effet (*Effect Size*) positive (rendement supérieur au groupe contrôle) et les autres une Ampleur d'Effet négative.
- b) Pour les **connaissances** (souvent testées au NBME), l'APP est inférieur aux études traditionnelles. L'Ampleur de l'Effet va de de -1.00 à $+0,27$ dans la méta-analyse d'ALBANESE et MITCHELL et de $-0,24$ à $+0,06$ dans la méta-analyse de VERNON et BLAKE.
- c) Pour les **performances cliniques**, l'APP est supérieur aux études traditionnelles, l'A.E. variant de $+0,02$ à $+0,51$ chez ALBANESE et MITCHELL et de $+0,40$ à $+0,70$ ($dw = 0,55$)³ pour VERNON et BLAKE. Bien sûr, il faut ici se méfier d'un possible effet Hawthorne⁴.
- d) Pour le choix des spécialisations, en moyenne, l'APP est supérieur ($dw = +0,33$ avec E.S. allant de $-0,22$ à $+0,96$) chez ALBANESE et MITCHELL, avec un plus grand choix du métier de médecin de famille (généraliste) chez les étudiants formés par l'APP.
- e) Pour d'autres caractéristiques, des différences apparaissent aussi. Ainsi, ALBANESE et MITCHELL font apparaître que :
 - Le nombre de consultations par mois (moins pour les APP : $-0,28$). - Le revenu par mois (moins pour les APP : $-0,27$).
 - Les patients par mois (moins pour les APP : $-0,37$).
 - Les coûts à la consultation (plus pour les APP : $+0,35$), car ils demanderaient plus d'expertises secondaires.
 - Les consultations relevant de la psychologie (plus pour les APP : $+0,23$).

Il serait intéressant de mesurer aussi la capacité à (et la volonté de) s'auto-former, à consulter une bibliothèque ainsi que d'autres compétences transversales, mais cela n'a pas fait l'objet de méta-analyses, sauf dans l'étude de BERKSON (1993) selon laquelle les étudiants APP seraient plus réticents à s'engager dans des apprentissages de surface (pour le court terme) et plus enclins à étudier pour comprendre. Il serait aussi intéressant de disposer de comparaisons de résultats

³ $Dw = \text{« difference weighted »}$. Cet indice est une somme pondérée (par les nombres dans chaque échantillon) des « Effect Size » moyens dans diverses études.

⁴ Les étudiants sont stimulés par le sentiment de faire partie d'une expérience et leur jugement surévalue les bénéfices retirés. Ce phénomène a été appelé « Effet HAWTHORNE » du nom de la localité dans laquelle cette expérience a été menée pour la première fois, sur l'influence des intensités lumineuses sur le rendement au travail. On a ainsi montré que c'est quand ils croyaient que l'intensité lumineuse avait augmenté que les ouvriers amélioreraient leur production.

avec des méthodes qui mêleraient les deux approches, l'APP et la traditionnelle, où est plus présente la notion de soutien par les formateurs (d'étayage, selon l'expression de BRUNER).

4. Conclusions

Les résultats comparatifs sont encore assez flous, étant donné les difficultés méthodologiques. Globalement, on ne peut dire qu'un système méthodologique est nettement meilleur que l'autre. Par ailleurs, à l'intérieur de chaque système, des écarts importants existent. Enfin, on doit se demander quelle sera l'évolution des systèmes respectifs. Pour BERKSON, il faut s'attendre à ce que les différences de rendement s'amenuisent.

P. DISCUSSION

On peut se demander si un curriculum de médecine générale ne présente pas des caractéristiques très particulières par rapport aux autres formations universitaires qui expliqueraient le succès de l'APP en formation de médecins généralistes.

Tout d'abord, la cohorte d'étudiants qui s'engage dans ces études va faire l'objet d'un traitement uniforme pendant six ans, car il y a peu (ou même pas) d'options. Un système APP portant sur les six ans peut donc compter sur une cohérence d'ensemble de l'approche, ce qui pose le problème au départ de « s'adapter » au système, mais qui évite celui de « se réadapter » ou d'être « compatible » avec le système traditionnel qui ne réapparaît plus dans la formation.

Ensuite, la situation médicale est naturellement une situation de résolution de problèmes, plus particulièrement de « cas », la plupart du temps, individuels. Ce n'est pas le seul domaine de ce type. Le droit, l'économie se présentent aussi de la sorte. On ne voit pas pourquoi, cependant, la psychologie y échapperait, pas plus que les langues étrangères ! Dans tous ces domaines, il est possible de pratiquer les principes d'isomorphisme : faire vivre lors de la formation les expériences de la vie professionnelle.

Enfin la démarche médicale de résolution de problème est elle-même bien formalisée : Observations, Diagnostic, Communication et Perception, Traitement, Suivi. Mais les autres disciplines n'auraient-elles aucune démarche « rigoureuse » ?

Nous laissons ce débat ouvert.

*Je ne vois qu'un moyen de savoir jusqu'où on peut aller,
c'est de se mettre en route et de marcher.*

Henri BERGSON