

LE DÉVELOPPEMENT DU RAISONNEMENT BIOMÉDICAL ET CLINIQUE AU COURS DU CURSUS MÉDICAL

Anne Collard
Mars 2014

Thèse présentée en vue de l'obtention du grade de
Docteur en Sciences psychologiques et de l'éducation à l'Université de Liège

Sous la Direction de Jean-Pierre Bourguignon et Serge Brédart

Jury composé de Marie Claude Audetat (UdeM), Jacques Boniver (ULg),
Jean-Pierre Bourguignon (ULg), Serge Brédart (ULg), Dieudonné Leclercq (ULg)
et Dominique Vanpee (UCL)



Université
de Liège

La recherche a été menée à l'Université de Liège



Faculté
de Médecine

À la Faculté de Médecine



Et à l'Institut de Formation et de Recherche en Enseignement
Supérieur



fnrs
LA LIBERTÉ DE CHERCHER

Ce projet de recherche a été partiellement financé par le Fonds de la
Recherche Fondamentale et Collective, fonds associés du Fonds
National de la Recherche Scientifique

PROMOTEURS

Professeur Jean-Pierre Bourguignon, Faculté de Médecine, IFRES – Université de Liège

Professeur Serge Brédart, Faculté de Psychologie et Sciences de l'éducation – Université de Liège

COMITÉ D'ACCOMPAGNEMENT

Professeur Jacques Boniver, Faculté de Médecine – Université de Liège

Professeur Dieudonné Leclercq, Faculté de Psychologie et Sciences de l'éducation – Université de Liège

MEMBRES EXTERNES DU JURY

Professeur Marie-Claude Audetat, Faculté de Médecine – Université de Montréal

Professeur Dominique Vanpee, Faculté de Médecine et Médecine dentaire – Université Catholique de Louvain

REMERCIEMENTS

Je tiens tout d'abord à remercier Jean-Pierre Bourguignon, promoteur de ce travail, pour la confiance qu'il m'a accordée, pour son soutien, son accompagnement, son écoute, ses relectures attentives et la générosité dans le partage de ses idées ainsi que pour m'avoir incitée à aller toujours plus loin.

Je remercie également Serge Brédart, promoteur de ce travail, pour son soutien, sa collaboration, ainsi que pour m'avoir recentrée sur mon travail au moment nécessaire.

Je remercie Jacques Boniver, initiateur de ces travaux de recherche, par son implication et son souci pédagogique à la Faculté de médecine ainsi que pour son accompagnement et ses commentaires tout au long du travail. Je remercie également Dieudonné Leclercq pour son accompagnement et ses relectures et commentaires concernant les différents travaux de recherche.

Mes remerciements à Madame Audetat et Monsieur Vanpee pour avoir accepté de faire partie du Jury de cette thèse.

Je tiens également à remercier les membres du Bureau Pédagogique de la Faculté de médecine et en particulier France Mélot, pour sa disponibilité, son aide et sa collaboration dans la collecte des données mais aussi dans la rédaction des articles de recherche.

Mes remerciements au Collège d'enseignants de Bac 2 et Bac 3 de la Faculté de médecine pour m'avoir inspirée et guidée dans mes recherches. Je remercie tout particulièrement Pierre Bonnet, Ernst Heinen et Jean-Olivier Defraigne qui ont participé activement aux recherches que j'ai pu mener.

Je souhaite remercier tous mes collègues de l'IFRES pour leur soutien et disponibilité. Et tout particulièrement, Eliane Nivart, Pascal Detroz et Laurent Leduc pour leur écoute et soutien indéfectible ; Merci à Olivier Borsu pour son aide dans la mise en page de ce travail et pour avoir réalisé la couverture ; Merci aux collègues du SMART et plus précisément à Marc Pirson pour son temps et son aide dans l'analyse des données longitudinales ; Merci à Virginie Staelens,

ma collègue et amie pour sa présence, son écoute, l'environnement de travail stimulant et rassurant qu'elle parvient à créer ainsi que pour les moments de détente, parfois bien nécessaires.

Je remercie les tuteurs du module endocrino-métabolisme ainsi que tous les étudiants volontaires ayant accepté de participer à ces recherches. Je remercie également toutes les personnes ayant collaboré à la rédaction des articles de recherche : Jean-Pierre Bourguignon, Jacques Boniver, Serge Brédart, Jean-Olivier Defraigne, Sabine Gelaes, France Mélot et Sophie Vanbelle.

Mes remerciements aux Fonds de la Recherche Fondamentale et Collective qui par leur soutien financier ont permis d'initier ce travail.

Enfin, je souhaite remercier ma famille pour leur présence, leur soutien et leur disponibilité. En particulier, mon papa, Michel, pour ses relectures et le soutien en anglais ; ma maman, Louise, et mes sœurs pour leur disponibilité et leur écoute ; mon compagnon, Benoit, et mon fils Jules, personnes stimulantes dans ma vie.

TABLE DES MATIÈRES

INTRODUCTION.....	11
PARTIE 1 : LE RAISONNEMENT BIOMÉDICAL ET CLINIQUE	15
1. CADRE : LA FACULTÉ DE MÉDECINE DE L'UNIVERSITÉ DE LIÈGE	15
1.1. Les objectifs de la formation médicale	15
1.2. Structure des études médicales (2000 à 2011)	19
2. LE RAISONNEMENT BIOMÉDICAL ET CLINIQUE - DÉFINITIONS.....	20
3. LES PROCESSUS DE RAISONNEMENT	22
3.1. Les processus analytiques	22
3.1.1. Le raisonnement hypothético-déductif (backward reasoning)	22
3.1.2. L'application de règles de causalité ou conditionnelles (forward reasoning)	23
3.2. Les processus non analytiques	23
3.2.1. Pattern Recognition.....	23
3.2.2. Instance-based.....	24
3.3. Les processus mixtes analytiques et non analytiques	24
4. L'ORGANISATION DES CONNAISSANCES	26
4.1. Les architectures fondées sur les exemples concrets (instances)	27
4.2. Les architectures prototypiques	27
4.3. Les architectures complexes	28
4.3.1. Les réseaux sémantiques.....	28
4.3.2. Les scripts	28
4.3.3. Les schémas	29
5. LE DÉVELOPPEMENT DU RAISONNEMENT DU NOVICE À L'EXPERT	30
6. LA FORMATION AU RAISONNEMENT	36
6.1. Principes de base pour l'apprentissage du raisonnement clinique	37
6.1.1. Établir des liens avec les connaissances antérieures	37
6.1.2. Faciliter le raisonnement hypothético-déductif.....	37
6.1.3. Favoriser à la fois l'usage des processus analytiques et non analytiques de raisonnement clinique	38
6.1.4. Favoriser le transfert des connaissances.....	38
6.1.5. Favoriser l'organisation et l'activation des connaissances	39
6.1.6. Favoriser une récolte de données cliniques pertinentes et discriminantes	39

6.2.	La capacité de transfert des apprentissages.....	40
6.3.	Le PBL pour développer le raisonnement clinique ?	42
6.3.1.	La théorie éducationnelle du PBL.....	43
6.3.2.	Les étapes du PBL	44
6.4.	Les effets du PBL ?.....	45
6.5.	Le PBL à la Faculté de Médecine de l'Université de Liège	48
6.5.1.	Les AIC – Apprentissage de l'Intégration des connaissances	49
6.5.2.	Les APP – Apprentissage Par Problèmes	50
6.5.3.	Les ARC – Apprentissage du Raisonnement Clinique	52
6.5.4.	Les ARPC – Apprentissage de la Résolution de Problèmes complexes	53
7.	L'ÉVALUATION DU RAISONNEMENT	56
7.1.	Les modalités d'évaluation du raisonnement clinique	56
7.1.1.	Patient Management Problem (PMP)	56
7.1.2.	Les grilles d'évaluation globale.....	57
7.1.3.	Les Questions à Choix Multiples (QCM)	57
7.1.4.	L'oral	57
7.1.5.	La question rédactionnelle	58
7.1.6.	L'Examen Clinique Objectif et Structuré (ECOS).....	58
7.1.7.	Problems and Modified Essay Questions (MEQ).....	58
7.1.8.	Les questions à appariement étendu (EMQ, Extended Matching Questions)	59
7.1.9.	L'examen par éléments clés	59
7.1.10.	Le Test de Concordance de Scripts (TCS).....	59
7.1.11.	L'évaluation basée sur la performance : Le miniCEX (Clinical Examination Exercise)	60
7.1.12.	Diagnostic Thinking Inventory (DTI)	60
7.2.	Les évaluations en Faculté de Médecine	64
7.2.1.	Les évaluations privilégiant l'évaluation des connaissances	64
7.2.2.	Les évaluations privilégiant l'évaluation du raisonnement.....	65
7.2.3.	Les évaluations mixtes.....	66
7.3.	La capacité d'auto-évaluation de sa base de connaissances	67
PARTIE 2 : LES TRAVAUX DE RECHERCHE		
1.	INTRODUCTION	69
1.1.	Cadre des recherches	69
1.2.	Les questions de recherche	69

2.	RÉSUMÉ DES MATÉRIELS ET MÉTHODES.....	72
2.1.	Participants	72
2.2.	Matériels	72
3.	SYNTHÈSE DES RÉSULTATS.....	75
3.1.	La capacité de transfert des connaissances	75
3.2.	L'évolution de la capacité de raisonnement durant la formation	75
3.3.	La capacité de raisonnement et la rétention des connaissances factuelles	76
3.4.	La capacité de raisonnement et la capacité de délimitation du champ des connaissances	76
3.5.	L'évolution de la capacité de délimitation du champ des connaissances	77
3.6.	L'impact d'un exercice travaillant la capacité d'auto-validation des connaissances au cours d'une séance APP....	77
4.	ARTICLE 1 : CONTEXT IMPACT OF CLINICAL SCENARIO ON KNOWLEDGE TRANSFER AND REASONING CAPACITY IN A MEDICAL PBL CURRICULUM	80
4.1.	Introduction	81
4.2.	Methods	84
4.3.	Results	89
4.4.	Discussion	91
4.5.	References	94
5.	ARTICLE 2 : REASONING VERSUS KNOWLEDGE RETENTION AND ASCERTAINMENT THROUGHOUT A PROBLEM-BASED LEARNING CURRICULUM. 99	
5.1.	Introduction	100
5.2.	Methods	102
5.3.	Results	109
5.4.	Discussion	113
5.5.	References	119
6.	ARTICLE 3 : LONGITUDINAL ASSESSMENT OF PROGRESS IN REASONING CAPACITY AND RELATION WITH SELF-ESTIMATION OF KNOWLEDGE BASE	123
6.1.	Introduction	124
6.2.	Materials and Methods	126
6.3.	Results	129
6.4.	Discussion	136
6.5.	References	143
7.	ARTICLE 4 : EFFETS D'UN EXERCICE DE VALIDATION DES ÉLÉMENTS D'UN RAISONNEMENT BIOMÉDICAL SUR LES COMPOSANTES INDIVIDUELLES ET COLLECTIVES DE L'APPRENTISSAGE AU COURS D'UN SÉMINAIRE APP	147

7.1.	Introduction	148
7.2.	Méthodes Etude 1	152
7.3.	Résultats Etude 1	157
7.4.	Discussion Etude 1	166
7.5.	Méthodes Etude 2	169
7.6.	Résultats Etude 2	169
7.7.	Discussion Etude 2	171
7.8.	Conclusions Générales	172
7.9.	Références	173
PARTIE 3 : DISCUSSION		177
1.	FORCES ET LIMITES DE NOS TRAVAUX.....	177
1.1.	Les participants	177
1.2.	Les outils	178
1.3.	Le contexte de la formation	180
2.	ANALYSE DES RÉSULTATS PRINCIPAUX	182
2.1.	La capacité de transfert des connaissances	182
2.2.	L'évolution de la capacité de raisonnement durant la formation	183
2.3.	La capacité de raisonnement et la rétention des connaissances factuelles	184
2.4.	La capacité de raisonnement et la capacité de délimitation du champ des connaissances	185
2.5.	L'évolution de la capacité de délimitation du champ des connaissances	186
2.6.	L'impact d'un exercice travaillant la capacité d'auto-validation des connaissances au cours d'une séance APP. 187	
3.	LE PROCESSUS DE RAISONNEMENT EN LIEN AVEC LA BASE DE CONNAISSANCES ET SON UTILISATION : UN MODÈLE.....	189
3.1.	Les limites du modèle	198
4.	LES ERREURS DE RAISONNEMENT.....	200
5.	EVOLUTION DU PBL A LA FACULTÉ DE MÉDECINE DE L'ULG ET PISTES D'ADAPTATION DE LA FORMATION.....	206
5.1.	Le PBL à la Faculté de Médecine de l'ULg depuis 2011	206
5.2.	Propositions d'adaptations pour les modalités de formation PBL à la Faculté de Médecine	208
CONCLUSION ET PERSPECTIVES		212
RÉFÉRENCES.....		218

INTRODUCTION

Un objectif majeur de la formation médicale est le développement du raisonnement clinique, essentiel à l'exercice de la médecine. Bien que la recherche sur le raisonnement clinique soit très étendue, la plupart des études s'intéressent essentiellement au raisonnement diagnostique et à son développement au niveau des études de Master, le début de la formation clinique. Relativement moins d'études s'intéressent au développement du raisonnement clinique dès le début du cursus médical, au niveau des années précliniques.

Plusieurs recherches dans les années 90 ont démontré que les capacités de raisonnement clinique ne sont pas indépendantes des connaissances. Effectivement, elles suggèrent une théorie du développement de l'expertise médicale, dans laquelle l'acquisition des connaissances et du raisonnement clinique se fait conjointement (Boshuizen & Schmidt, 1992 ; Schmidt & Boshuizen, 1992a ; Schmidt, Boshuizen, & Norman, 1992). Traditionnellement, les premières années du cursus médical sont dédiées à l'acquisition des connaissances de base ; l'apprentissage des connaissances cliniques et du raisonnement intervient au niveau du Master dès les premiers stages cliniques. Une critique fréquente qui a été émise est que ces étudiants, après les années de baccalauréat, disposent bien de connaissances mais éprouvent des difficultés à les exploiter dans un contexte où elles peuvent être appliquées (Nendaz & Bordage, 2002). Ces observations ont été faites par un grand nombre d'enseignants et ont inspiré de nouvelles approches d'enseignement de la médecine, comme le Problem-Based Learning (Apprentissage Par Problèmes) qui a notamment pour objectif le développement du raisonnement clinique par l'apprentissage actif sur base d'un contexte concret (Norman & Schmidt, 1992).

Dans ce travail, nous aborderons le développement du raisonnement clinique dès la 2^{ème} année du cursus médical à l'Université de Liège. La Faculté de médecine de l'ULg encourage les étudiants à développer simultanément leur base de connaissances et leur capacité

de raisonnement clinique de façon progressive. Cette volonté est présente dans le cadre de l'enseignement par la mise en place de séminaires d'Apprentissage Par Problèmes (PBL) qui vise le développement précoce du raisonnement biomédical, prémisse du raisonnement clinique ; ainsi que dans le cadre de l'évaluation qui teste de façon spécifique les connaissances et/ou le raisonnement. Dès la 2^{ème} année, les étudiants sont encouragés à développer leur raisonnement hypothético-déductif à travers l'analyse de cas qui leur permet également d'utiliser et développer graduellement leur base de connaissances. Leur base de connaissances devient de plus en plus élaborée à travers l'application et l'utilisation de ces connaissances pour raisonner à partir de cas de plus en plus complexes. Ce cursus fournit donc un environnement de choix pour investiguer le processus de raisonnement et les facteurs qui impactent sur son développement.

Les objectifs de ce travail sont les suivants :

- étudier la maturation du raisonnement biomédical et clinique au cours du cursus PBL de la Faculté de Médecine de l'Université de Liège ;
- étudier dans quelle mesure les étudiants sont capables de transférer les connaissances acquises dans le cadre des séminaires PBL pour les utiliser dans une tâche de raisonnement inédite ;
- étudier la relation entre les performances en raisonnement et en connaissances et la capacité à auto-évaluer ses propres connaissances.

A travers ce travail, nous souhaitons inscrire et expliciter le processus de raisonnement biomédical et de raisonnement clinique dans un continuum commun, en lien avec la base de connaissances acquises par les étudiants et la façon dont cette base de connaissances est structurée en mémoire.

La première partie précisera le cadre de ce travail, la Faculté de médecine de l'Université de Liège, en définissant les objectifs et les modalités de la formation médicale qui y est proposée. Elle se centrera ensuite sur la définition du raisonnement clinique et son

développement. Nous aborderons les aspects cognitifs liés au développement du raisonnement, ses modalités d'enseignement et d'évaluation. Nous discuterons le lien étroit entre les capacités de raisonnement et la base de connaissances et l'importance de travailler précocement avec les étudiants la structuration et l'utilisation de cette base de connaissances.

La deuxième partie développera quatre articles. Le premier article abordera la capacité de transfert et d'utilisation des connaissances acquises dans le contexte de séminaires PBL à d'autres contextes inédits. Le second article présentera une étude expérimentale transversale évaluant pour des étudiants de la 3^{ème} à la 6^{ème} année, la maturation des performances en raisonnement mesurée par un Test de Concordance de Scripts et l'évolution des performances en connaissances ainsi que leur auto-évaluation. Ces évaluations portent exclusivement sur l'endocrinologie. Le troisième article évaluera les mêmes processus que le précédent mais dans un contexte longitudinal utilisant les données des évaluations certificatives des étudiants de l'année 2 à l'année 5 sur l'ensemble des systèmes anatomiques abordés. Enfin, le dernier article relatera une expérience menée avec une vingtaine d'étudiants mesurant l'impact d'un exercice de délimitation de leur champ de connaissances sur leur processus de raisonnement dans le cadre d'un séminaire APP expérimental.

Dans la dernière partie, nous rappellerons les résultats principaux de nos recherches et les discuterons à la lumière d'un modèle explicitant le lien entre raisonnement et base de connaissances, selon le niveau d'expertise. Sur base de ce modèle, nous discuterons alors les erreurs de raisonnement et les modalités de la formation au raisonnement. Ces réflexions et le cheminement de notre travail nous conduiront à identifier quelques nouveaux défis pédagogiques de la Faculté de médecine et à suggérer quelques propositions pour y répondre. Enfin, nous tirerons les conclusions de notre travail et en dégagerons quelques perspectives.

PARTIE 1 : LE RAISONNEMENT BIOMÉDICAL ET CLINIQUE

1. CADRE : LA FACULTÉ DE MÉDECINE DE L'UNIVERSITÉ DE LIÈGE

Pour répondre aux besoins actuels de la société, pour faire face à l'accroissement exponentiel des connaissances dans le domaine médical et s'adapter au nombre d'étudiants qui étaient alors en réduction, la Faculté de Médecine de l'Université de Liège a entamé dès 1998 une profonde réflexion sur ses modalités d'enseignement. Ces réflexions et les actions entreprises ont été pilotées notamment par un Collège d'enseignants-cliniciens et non-cliniciens de cette faculté. Les travaux de recherche exposés dans le cadre de cette thèse ont été développés dans la continuité de cette profonde réforme de l'enseignement et dans la ligne des réflexions pédagogiques menées et questions identifiées par ce Collège. Dès lors, nos travaux ont été largement suscités et guidés par les préoccupations pédagogiques de ces enseignants.

1.1. Les objectifs de la formation médicale

Début des années 2000, le Doyen de la Faculté de médecine de l'Université de Liège en fonction, Jacques Boniver, a rédigé un plan stratégique pour la Faculté, dans lequel il définissait les objectifs de la formation et les choix méthodologiques pour les modalités d'enseignement et d'évaluation de cette Faculté. Ce plan stratégique approuvé par la Commission Académique de l'ULg en 2001 définissait les objectifs de la formation médicale comme suit (extrait du plan stratégique mentionné dans un article de Boniver, 2004) :

« EN VUE DU SERVICE À OFFRIR AUX MALADES

- *les préparer à exercer la médecine, soit comme médecin généraliste, soit comme médecin spécialiste, en ayant acquis les connaissances nécessaires à une bonne pratique des actes médicaux dans le respect et au service des malades et de la société ;*
- *les former à exercer la médecine en connaissant la portée morale, physique et psychologique des actes qu'ils réalisent, en respectant et en soulageant la souffrance d'autrui et en connaissant les règles de l'éthique médicale.*

POUR UNE PRÉPARATION À LA MÉDECINE GÉNÉRALE OU À LA MÉDECINE SPÉCIALISÉE ET À LEUR INTÉGRATION

- *leur donner une formation de base, -1^{er} et 2^{ème} cycles, qui leur permette d'accéder, selon leur choix :*
 - *soit à la formation spécialisée de 3^{ème} cycle en médecine générale,*
 - *soit à la formation spécialisée de 3^{ème} cycle en une spécialisation médicale ;*
- *les informer précocement de la diversité de la pratique médicale, qu'il s'agisse de médecine générale ou de médecine spécialisée, qu'il s'agisse de médecine à l'hôpital ou à l'hôpital universitaire, ou de médecine extra-hospitalière, au cabinet, en pratique isolée ou en groupe ;*
- *à l'issue d'un DES en médecine générale, les rendre capables de pratiquer la médecine curative et la médecine préventive qui caractérisent l'exercice de la médecine générale ;*
- *à l'issue d'un DES en une spécialisation médicale, les rendre capables de postuler avec succès un poste hospitalier ou hospitalo-universitaire ou d'exercer la médecine spécialisée en cabinet à un haut niveau, et/ou de pratiquer la médecine préventive.*

POUR UNE COHÉRENCE ENTRE LA PRATIQUE MÉDICALE ET LA SOCIÉTÉ

- *au cours de la formation de base, leur donner un enseignement qui tienne compte de la prévalence des affections, de leur gravité et de leur exemplarité, en pondérant, en hiérarchisant les différentes disciplines, avec un souci d'intégration ;*
- *les former à situer la maladie et la médecine dans le contexte socio-économique de l'environnement où ils exerceront leur pratique; ceci comprend la prise en compte des facteurs socioéconomiques influençant la santé, les conséquences socio-économiques de la pratique médicale et, en particulier, l'impact de leur activité médicale sur le coût de la santé dans le cadre de la sécurité sociale ou au niveau du patient ;*
- *les sensibiliser à envisager la pratique médicale et leur rôle dans la société au-delà de la région dont ils sont issus, en étant instruits de la place et de l'aspect de l'exercice médical dans d'autres pays du monde, selon leur culture et leur développement.*

EN VUE DU DÉVELOPPEMENT, DE LA BONNE UTILISATION ET DE L'ENTRETIEN DES CONNAISSANCES

- *les aider à acquérir une vraie formation ou culture médicale, qui se caractérise par l'aptitude à bien utiliser leurs connaissances et à appréhender, à différencier, à intégrer, à hiérarchiser les éléments du diagnostic et de la décision thérapeutique, à en percevoir les limites et en évaluer les conséquences ;*
- *les convaincre que l'exercice de la pratique médicale requiert un entretien des connaissances et des aptitudes, et ce à travers la formation continuée et une démarche permanente d'amélioration; les étudiants en médecine doivent être préparés à mener cette formation continuée selon le mode de l'auto-apprentissage ;*
- *les intégrer dans des activités de recherche fondamentale et clinique, que l'Institution se doit de développer; l'intérêt et l'importance de la recherche dans la formation médicale doivent être mis en valeur ;*

- *les instruire des moyens modernes de communication et d'accès aux informations dans des banques de données, ... et en particulier à travers Internet.*

POUR UNE COMMUNICATION OPTIMALE AVEC LES PATIENTS ET LEURS AUTRES SOIGNANTS

- *les aider à acquérir et entretenir de grandes capacités de communication, qu'il s'agisse de la communication avec le patient ou avec les autres médecins ; cet apprentissage doit comprendre aussi la communication avec les autres prestataires de soins, avec le milieu hospitalier, avec l'Université, avec les administrations, avec le monde économique, avec la société ;*
- *les former à construire la relation médecin/malade, avec toutes ses facettes, au bénéfice du malade ;*
- *les préparer à interagir avec les autres médecins et autres prestataires de soins dans le cadre d'un respect mutuel. »*

Pour atteindre ces objectifs, l'enseignement doit :

- être recentré sur l'étudiant ;
- être basé sur des objectifs d'apprentissage définis par les enseignants travaillant en collège (apprentissage, c'est-à-dire non seulement l'acquisition de connaissances - le savoir- mais aussi de compétences - le savoir-faire et le savoir-être) ;
- privilégier une démarche active de l'étudiant.

Les moyens retenus par la Faculté pour suivre cette stratégie ont été notamment :

- l'organisation de l'enseignement sur un mode multidisciplinaire, par modules basés sur les systèmes anatomiques définis par les enseignants réunis en collège plutôt que sur les disciplines ;
- l'implantation de méthodes pédagogiques qui visent à stimuler la participation active de l'étudiant à son apprentissage, à savoir le «Problem Based Learning» (PBL), organisé sous forme de séminaires par petits groupes, remplaçant partiellement les cours ou cliniques ex cathedra ;

- l'organisation de l'évaluation des connaissances (examens) sur un mode multidisciplinaire.

Depuis ce Plan stratégique, deux nouveaux Doyens ont été en fonction à la Faculté de médecine. Le contexte a bien entendu évolué depuis 2000 mais la volonté de conserver un enseignement multidisciplinaire et une participation active des étudiants dans leur apprentissage reste une préoccupation des enseignants de la Faculté.

1.2. Structure des études médicales (2000 à 2011)

Chaque année académique est composée de 2 quadrimestres (Q) d'enseignement, dont chacune de ces périodes de cours est suivie par une session d'examen.

En première année (Q1 et Q2), les enseignements de chimie, physique, mathématiques, biologie, génétique, embryologie et anatomie sont dispensés. En deuxième année, le Q3 comprend l'enseignement des sciences médicales de base (biochimie, physiologie, histologie, immunologie, microbiologie générale). Le Q4, ainsi que les Q5 et Q6 de la troisième année, sont consacrés à l'étude de la structure et du fonctionnement de l'homme normal (anatomie, embryologie, histologie, biochimie et physiologie humaine) et des fondements des maladies (biochimie et physiologie pathologiques, anatomie pathologique générale). Au cours des 4^{ème} et 5^{ème} années (Q7 à Q10) la pathologie (les maladies) est enseignée. Enfin, les 6^{ème} et 7^{ème} années sont essentiellement consacrées à des stages à temps plein, à l'exception de deux journées mensuelles de cliniques, cours et séminaires.

Nous décrivons en fin de cette section les différents séminaires PBL proposés au cours des différentes années de la formation médicale à l'Université de Liège.

2. LE RAISONNEMENT BIOMÉDICAL ET CLINIQUE - DÉFINITIONS

Le raisonnement clinique est un processus complexe qui implique l'intégration et l'application de différents types de connaissances, l'utilisation de différentes formes de raisonnement, et une réflexion critique sur le processus utilisé pour arriver à une solution diagnostique (Anderson, 2006). Le raisonnement clinique ne se résume toutefois pas à la solution correcte du problème posé. L'essentiel du processus est le travail d'explication des signes, symptômes et résultats d'investigation qui sont présentés à l'étudiant.

La base de connaissances utilisée pour raisonner peut prendre la forme de connaissances en sciences de base, de la connaissance de la prévalence ou la probabilité d'une maladie dans un contexte donné, d'expérience antérieure similaire, etc. Selon le niveau d'élaboration de cette base de connaissances, différents types de raisonnement peuvent être appliqués pour expliquer les signes et symptômes d'un nouveau cas clinique (Anderson, 2006).

Le raisonnement clinique correspond à l'intrication de tous ces éléments mis ensemble et peut, en résumé, être défini comme un processus de résolution de problèmes de santé qui implique que le clinicien décode et synthétise les informations d'un cas clinique, les intègre aux connaissances et expériences antérieures et les utilisent pour prendre des décisions diagnostiques et de prise en charge (Boshuizen & Schmidt, 1992a).

Contrairement au raisonnement clinique, le raisonnement biomédical n'implique pas de prise de décision, ni de catégorisation diagnostique. Nous le définirons comme le processus de pensée qui permet l'identification d'éléments clés d'une situation clinique et la définition des liens qui les unit à la lumière des connaissances des sciences de base permettant la compréhension des mécanismes physiopathologiques sous-jacents et la représentation globale d'un problème clinique.

Ces deux formes de raisonnement impliquent l'intégration des connaissances et leur mise en relation, ce qui nous semble être au cœur du processus de raisonnement et nous incite à proposer de placer ces 2 formes de raisonnement sur un même continuum.

La majorité des études concernant le raisonnement dans le domaine de la santé s'intéressent exclusivement au raisonnement clinique qui est au cœur de l'exercice professionnel. Elles impliquent donc rarement des étudiants de niveau pré-clinique dont l'objectif de la formation est essentiellement l'acquisition des connaissances de base. Le raisonnement clinique a d'ailleurs longtemps été considéré comme une compétence intuitive qui ne peut être enseignée car elle émerge seulement après une considérable expérience clinique (Groves, 2008). Dans notre travail, nous mettrons en évidence le caractère évolutif du processus de raisonnement qui s'initie très tôt au sein du cursus en Faculté de médecine de l'Université de Liège.

Cette partie du mémoire présente essentiellement les données issues de la recherche sur le raisonnement clinique. L'étendue de la littérature et la diversité des approches rendent difficile une approche exhaustive du concept de raisonnement dans le domaine de la santé. Nous sommes conscients que le processus de raisonnement est plus complexe que l'utilisation efficace de ses connaissances. En effet, des dimensions contextuelles comme la relation avec le patient et la gestion émotionnelle interviennent pour une grande part dans ce processus. Toutefois, ces dimensions ne seront pas abordées dans ce travail. Ce choix a résulté de l'implication, dans nos études, d'étudiants qui, pour les plus jeunes, ne sont pas encore confrontés au patient et dont les dimensions émotionnelles sont notamment liées au contexte d'apprentissage et aux relations avec leurs pairs. Ces dimensions peuvent être assez différentes de celles vécues par les étudiants de master qui sont en contact avec le patient dès le début de ce cycle. Nous avons donc volontairement sélectionné la littérature en lien avec les aspects plus cognitifs du raisonnement, permettant une meilleure compréhension de l'aspect développemental de cette compétence clinique.

3. LES PROCESSUS DE RAISONNEMENT

Les problèmes médicaux auxquels doivent faire face les cliniciens sont habituellement « mal définis », toutes les données ne sont pas disponibles d'emblée, les objectifs poursuivis sont complexes et les résultats souvent ambigus et difficiles à prédire. Les informations concernant le problème du patient peuvent être fragmentaires, inconsistantes ou imprécises. Tout cela rend le processus de formulation du problème, son diagnostic et sa prise en charge complexes et difficiles (Kassirer & Kopelman, 1991). Différents travaux ont mis en évidence les différentes formes de raisonnement mises en œuvre par les cliniciens lorsqu'ils abordent un cas clinique. Dans la suite, nous aborderons ces différents types de raisonnement antérieurement décrits par Nendaz, Charlin, Leblanc et Bordage (2005).

3.1. Les processus analytiques

3.1.1. Le raisonnement hypothético-déductif (backward reasoning)

Le raisonnement hypothético-déductif a été mis en évidence par des études observant des experts cliniciens ou des études d'analyse d'enregistrement du raisonnement à voix hautes (Barrows & Bennett 1972, Elstein, Shulman, & Sprafka, 1978). Ces études ont démontré que le clinicien émet quelques hypothèses diagnostiques dès le début de l'entrevue médicale. Ces hypothèses servent alors de cadre au recueil d'informations supplémentaires par une démarche analytique délibérée. Les nouvelles données recueillies sont interprétées à la lumière de ces hypothèses émises. Les hypothèses sont elles-mêmes réévaluées pour être retenues ou écartées. Ce processus recommence avec de nouvelles hypothèses jusqu'à ce que le diagnostic soit atteint. Ce processus qui va de la solution entrevue vers les données implique une démarche analytique et consciente sur le mode déductif.

Lorsque le clinicien est face à un nouveau problème clinique, la solution est rarement univoque et il doit pouvoir faire le tri parmi la

quantité de données disponibles chez le patient. La génération précoce d'hypothèses guidant la recherche d'information permet de structurer le problème et rend la gestion des informations et la complexité des cas gérables pour la mémoire de travail dont les capacités sont limitées (Elstein, Shulman, & Sprafka, 1978, 1990). L'utilisation du raisonnement hypothético-déductif est une découverte importante pour l'apprentissage du raisonnement clinique car il correspond au mode de raisonnement utilisé par les cliniciens tout au long de leur carrière pour des problèmes inédits (Elstein et al., 1978 ; Hmelo, 1998).

3.1.2. L'application de règles de causalité ou conditionnelles (forward reasoning)

Dans le cadre de cas complexes, pour lesquels une hypothèse précoce pertinente ne peut être émise, il a été démontré que les cliniciens peuvent raisonner en appliquant des règles stockées en mémoire du type « si le symptôme x est présent, cela implique le diagnostic y » (Gilhooly, 1990 ; Patel & Groen, 1986). Le raisonnement est dans ce cas, un processus conscient et analytique qui part des données cliniques en suivant des règles causales ou conditionnelles, à l'inverse du processus hypothético-déductif qui part des hypothèses pour rechercher les données pertinentes.

3.2. Les processus non analytiques

Les processus non analytiques sont inconscients et automatiques, ils sont utilisés lorsque le clinicien est confronté à un cas clinique familier, non problématique. Il s'agit de faire une correspondance sur base du niveau de similarité entre le cas clinique présenté et un cas ou une configuration de signes similaires précédemment rencontré et stocké en mémoire.

3.2.1. Pattern Recognition

Ce type de raisonnement est utilisé pour le traitement de cas familier. Il est caractérisé par une récupération automatique de connaissances stockées et est déclenché par l'identification d'éléments

clés d'un cas (Arocha, Patel, & Patel, 1993). Cette forme de raisonnement nécessite une certaine expérience car il suppose que des cas similaires aient été rencontrés précédemment pour qu'ils soient considérés comme des cas familiers ou routiniers (Patel & Groen, 1986). Les caractéristiques de ce processus de raisonnement suggère une forte association entre les capacités de raisonnement et la base de connaissances, non seulement sa quantité mais aussi son stockage, sa récupération et son utilisation. En d'autres mots, avec l'expérience, le processus de raisonnement devient automatique. L'expertise en raisonnement clinique est caractérisée par une base de connaissances hautement conceptualisée et une capacité non analytique de reconnaître et traiter des situations familières.

3.2.2. Instance-based (Schmidt et coll. 1990)

Ce modèle suppose que des cas rencontrés précédemment peuvent être stockés, tels qu'ils ont été vécus, sans abstraction, en mémoire. Lorsqu'un clinicien rencontre un nouveau cas, il est alors comparé à ceux stockés en mémoire. S'il trouve un cas similaire, la réponse appliquée à ce cas stocké est transposée telle quelle au cas présent (ce cas ressemble à celui de Madame X qui avait le diagnostic Y ; donc c'est le diagnostic Y pour ce cas). L'utilisation de ce processus non analytique nécessite également une certaine expérience puisqu'il nécessite le stockage de cas précédemment rencontrés (Brooks, Norman, & Allen, 1991 ; Hatala, Norman, & Brooks, 1996).

3.3. Les processus mixtes analytiques et non analytiques

L'utilisation de processus mixtes lors de l'analyse d'un cas clinique semble le mécanisme le plus probable pour expliciter la démarche du clinicien : La première étape est la reconnaissance d'une similarité qui permettrait d'activer une hypothèse et le réseau de connaissances correspondant en mémoire ; la seconde étape est la vérification de l'hypothèse émise par un processus hypothético-déductif

(Par exemple : Le symptôme polyurie évoque un diabète. S'il s'agit bien de cela, on devrait trouver une glucosurie à l'examen d'urine).

L'utilisation de processus mixtes a été démontrée dans l'étude de Kulantunga, Brooks et Norman (2001) ainsi que dans les études de Norman (Norman, 2005 ; Norman, Brooks, Regher, Mariott, & Shali, 1996) et celles de Leblanc et al. (Leblanc, Norman, & Brooks, 2001 ; Leblanc, Dore, Norman, & Brooks, 2004). Ces études démontrent que la qualité du diagnostic considéré précocement par le clinicien est associée à la reconnaissance des signes cliniques pertinents d'un cas.

4. L'ORGANISATION DES CONNAISSANCES

Les performances en raisonnement ne peuvent pas s'expliquer par le type de processus de raisonnement utilisé. La génération précoce d'hypothèses ne garantit pas d'obtenir le bon diagnostic (Barrows, Norman, Neufeld, & Feightner, 1982). Le développement du raisonnement est intimement lié au développement des connaissances et leur structuration en mémoire. L'organisation des connaissances en mémoire représente l'ensemble des connexions et liens entre concepts et expériences vécues (Regehr & Norman, 1996). L'expertise en raisonnement se caractérise par des réseaux élaborés, c'est-à-dire, constitués de liens qui déterminent la force des associations entre les éléments de connaissances. Ces réseaux étant constitués sur base de l'expérience personnelle, ils sont toujours uniques et propres à une personne (Grant & Marsden, 1987). L'organisation des connaissances en mémoire permet aux cliniciens de traiter efficacement l'information lors de l'analyse d'un cas clinique (Charlin, Tardif, & Boshuizen, 2000 ; Regehr & Norman, 1996). *« Les signes cliniques d'une situation activent des connaissances, des concepts ou des exemples similaires au sein du réseau cognitif, qui à leur tour en appellent d'autres. Ce processus d'activation s'étend, de sorte que, très rapidement, le clinicien dispose d'un riche ensemble de connaissances adaptées à la situation clinique, ce qui lui permet de générer des hypothèses pertinentes et de rechercher des informations supplémentaires pour les confirmer ou infirmer. Selon cette conception, l'acte diagnostique d'un expert ne résulte pas de l'application d'une stratégie générale de résolution de problème, mais de l'activation en mémoire de réseaux de connaissances appropriés »* (Nendaz et al. 2005).

Ces représentations cognitives peuvent contenir différents types d'architecture décrites ci-dessous (Custers, Regehr, & Norman, 1996a).

4.1. Les architectures fondées sur les exemples concrets (instances)

Les connaissances correspondent à une collection de cas cliniques tels qu'ils ont été vécus, sans abstraction, c'est-à-dire que chaque élément clinique du cas garde ses caractéristiques propres à ce cas et n'est pas modifié. Les éléments du contexte d'apparition du cas peuvent être stockés, comme par exemple le fait que le patient ait 3 enfants ou qu'il habite la région de Liège. L'apprentissage consiste ainsi à l'addition successive d'exemplaires de cas clinique. Ces exemplaires sont comparés au nouveau cas clinique rencontré afin de reconnaître un cas similaire (comme décrit ci-dessus). Ce modèle met en évidence l'importance du contexte dans le processus non analytique de reconnaissance immédiate et de transfert qui peut dans certains cas, conduire à des erreurs diagnostiques potentielles (Allen, Norman, & Brooks, 1992).

4.2. Les architectures prototypiques

Les prototypes sont développés en utilisant un processus d'abstraction pour construire une trace mnésique commune de patients avec la même maladie (Neiser, 1986). Ils peuvent être considérés soit comme un ensemble de caractéristiques décontextualisées, une description sommaire d'un cas ou comme un exemple « idéal » de la maladie ou du patient (Bordage & Zacks, 1984). Selon ce modèle, les catégories diagnostiques sont constituées de prototypes, qui représentent l'abstraction d'un exemple typique d'une catégorie. Une catégorie de maladie est structurée autour des prototypes qui comprennent les caractéristiques les plus communément observées des maladies. Les membres les plus typiques d'une catégorie, c'est-à-dire, ceux qui partagent le plus de signes avec le prototype, sont ceux qui seront le plus facilement rappelés et identifiés par les cliniciens (Bordage & Zacks, 1984). En effet, il a été démontré que les médecins ont une meilleure compétence diagnostique lorsque le cas est typique que lorsqu'il ne l'est pas (Papa, Stone, & Aldrich, 1996).

4.3. Les architectures complexes

Dans ces types d'architectures, la connaissance est vue comme un réseau de concepts ou de nœuds, reliés entre eux par des liens indiquant la relation entre ces éléments. Ces nœuds sont activés par des stimuli avec une activation qui se répand sélectivement à travers les liens. Selon ces modèles, le diagnostic peut être considéré comme le processus qui permet de trouver le chemin à travers les réseaux de connaissances médicales aux nœuds qui représentent les catégories de maladies, utilisant les caractéristiques cliniques du cas donné comme guide.

4.3.1. Les réseaux sémantiques

Les réseaux sémantiques sont des réseaux de concepts reliés entre eux par des associations. Il a été démontré que face aux données d'un cas clinique, le clinicien transforme l'information pour se construire sa propre représentation du problème (Lemieux & Bordage, 1992). Cette représentation du problème permet, par une transformation sémantique, de donner du sens aux éléments, énoncer des hypothèses et activer les représentations cognitives pertinentes qui vont permettre d'aboutir à une solution diagnostique. Il s'agit donc d'un processus adaptatif plus complexe que la simple reconnaissance oui/non des architectures prototypiques ou basé sur les exemplaires. Ces réseaux sémantiques sont plus ou moins développés selon la richesse de l'organisation de leurs connaissances. Habituellement ces réseaux sémantiques sont considérés comme essentiellement l'application de connaissances biomédicales alors que les scripts et les schémas impliquent des connaissances cliniques, développées à travers l'expérience clinique (Schmidt, Norman, & Boshuizen, 1990 ; Schmidt & Boshuizen, 1993).

4.3.2. Les scripts

Les scripts sont des architectures de connaissances permettant de réaliser des actions spécifiques (Schank & Abelson, 1977 ; Nelson, 1986 ; Fayol & Monteil, 1988). Dans le domaine médical, cela signifie que les cliniciens disposent de connaissances spécifiquement organisées

pour être efficaces dans leurs tâches cliniques (Charlin, Tardif, & Boshuizen, 2000 ; Feltovich & Barrows, 1984 ; Schmidt, Norman, & Boshuizen, 1990). Les scripts de maladie contiennent les associations que le clinicien a établies entre l'entité pathologique, ses signes et symptômes et des cas similaires rencontrés précédemment. Un script de maladie comporte une série de caractéristiques cliniques, chacune ayant une valeur qui qualifie son attribut. Les caractéristiques contenues dans un script ne sont pas exclusives à ce script mais bien les associations entre elles. L'activation des scripts permet la sélection des hypothèses par la comparaison des informations contenues dans ce script et les données du patient. Les scripts guident également la recherche d'information. Ils évoluent avec l'expérience et deviennent de plus en plus fonctionnels.

4.3.3. Les schémas

La notion de schéma a été développée notamment par Rumelhart (1981) qui le définit comme une structure de données représentant des concepts génériques stockés en mémoire. La théorie des schémas propose que lorsqu'un individu acquiert de nouvelles connaissances, il tente de raccrocher cette nouvelle connaissance à une structure en mémoire afin de lui donner du sens. Cette théorie suppose que les individus regroupent les informations dans des concepts génériques qui sont stockés par catégorie facilitant un rappel ultérieur. En conclusion, les schémas sont des structures mentales qui permettent à l'apprenant de comprendre et associer les données auxquelles ils sont confrontés. Les schémas représentent les connaissances que l'on a d'un concept, ses propriétés générales : les objets et les relations qu'il a avec d'autres objets, situations, événements, etc.

Certains auteurs proposent une organisation de certaines connaissances médicales en schémas prenant une forme algorithmique dans laquelle un problème clinique spécifique est associé à l'information à rechercher et aux hypothèses diagnostiques à évoquer (Coderre, Mandin, Harasym, & Fick, 2003 ; Mandin, Jones, Woloschuk, & Harasym, 1997).

5. LE DÉVELOPPEMENT DU RAISONNEMENT DU NOVICE À L'EXPERT

La différence entre un expert et un novice ne repose pas tellement sur la nature du processus de raisonnement mais surtout sur son application (Groves, 2008). Dans le cadre de l'analyse d'un cas clinique, les experts vont plus aisément identifier les éléments clés du cas, générer des hypothèses plus précises et plus spécifiques, utiliser des questions plus focalisées pour vérifier ces hypothèses et fournir une explication plus cohérente du cas. Les experts possèdent généralement une base de connaissances plus large et plus intégrée.

Le raisonnement clinique se développe via l'interaction entre les connaissances et l'expérience clinique. Une composante essentielle de cette interaction est le niveau d'organisation des connaissances en mémoire à long terme (Bordage & Lemieux 1991 ; Bordage & Zacks, 1984 ; Norman, Rosenthal, Brooks, Allen, & Muzzin, 1989 ; Patel & Groen, 1986 ; Patel, Groen, & Arocha, 1990).

Schmidt et ses collaborateurs (Schmidt, Norman, & Boshuizen, 1990 ; Schmidt, Boshuizen, & Norman, 1992) ont formulé dans les années 90 une théorie du développement de l'expertise en raisonnement qui envisage le développement en parallèle du raisonnement clinique et de l'acquisition des connaissances. Cette théorie propose 4 étapes dans le développement de l'expertise qui résulte de changement progressif au niveau des structures de connaissances.

La première étape : Le développement de réseau élaboré de liens de causalité

Au début de la formation, les étudiants apprennent les informations biomédicales, et les organisent en réseau structurel de concepts simples ou d'unités de connaissances reliés entre eux par des liens de causalité. Le raisonnement biomédical est essentiellement basé sur des explications en termes de processus physiopathologique. Avec l'acquisition de nouvelles connaissances, le réseau va s'élargir, de nouveaux nœuds et liens vont se créer, les éléments existants vont se préciser et se renforcer.

A ce niveau, le raisonnement clinique correspond à une chaîne courte de concepts des sciences de base et à la capacité de générer des hypothèses explicatives, mais avec une intégration limitée des caractéristiques cliniques dans la représentation globale du problème (Joseph & Patel, 1990).

La deuxième étape : L'encapsulation des connaissances

Lorsque l'étudiant débute sa formation clinique, il est amené à appliquer ses connaissances aux cas cliniques rencontrés. Les concepts à appliquer ainsi que les liens entre ces différents concepts sont activés. Si un ensemble de concepts et de liens sont activés régulièrement en même temps, un regroupement de ces différents éléments peut s'opérer. Ces groupements consistent en l'encapsulation de concepts pertinents d'un point de vue diagnostique et peuvent résumer des processus physiopathologiques entiers. Ces regroupements qui correspondent à l'abstraction d'une plus grande quantité d'informations, permettent d'identifier plus rapidement et plus efficacement les éléments pertinents d'un cas. Selon Schmidt et al. (1990, 1992) l'encapsulation permet l'intégration des connaissances biomédicales et des connaissances cliniques. L'encapsulation des connaissances, par les raccourcis qui sont constitués, permet de faire un lien direct entre les caractéristiques cliniques du cas et des hypothèses diagnostiques. Les études de protocole de raisonnement, à voix haute notamment, ont permis de démontrer qu'à ce stade de développement, les cliniciens ne reviennent pas aux concepts des sciences de base pour l'explication d'un cas clinique mais peuvent y revenir si nécessaire, par exemple pour un cas difficile ou atypique.

La troisième étape : L'émergence des scripts

Peu à peu les réseaux de connaissances reliées par des liens de causalité vont évoluer vers des réseaux associatifs. Une transformation graduelle des connaissances encapsulées en « label diagnostique » ou « modèle mental abstrait » sera observée. Ces scripts qui utilisent la force des liens positifs et négatifs unissant signes et maladies, vont permettre d'expliquer les données cliniques et de prendre des décisions diagnostiques ou de traitement (Schmidt et al. 1990). Ce mode de

raisonnement est moins exigeant cognitivement que le raisonnement sur un mode de relations causales.

Les scripts de maladie sont des structures abstraites qui, par la présence de conditions d'application dans leur construction, impliquent des liens contextuels et temporels, en plus des liens de causalité (Feltovich & Barrows, 1984). Les scripts comportent une structure sérielle : les données cliniques d'un cas sont réorganisées pour correspondre au script activé qui se structure comme une histoire plutôt que dans un ordre strictement chronologique de présentation des éléments. Le développement des scripts implique à la fois, des connaissances conceptuelles de caractéristiques de maladie acquises formellement, et des connaissances perceptuelles acquises informellement. Ce dernier type de connaissances concerne principalement les conditions d'application et la variabilité des caractéristiques (Boshuizen, Schmidt, Custers, & van der Wiel, 1995). La composition et la qualité des scripts de maladies dépendent de la qualité ainsi que de la quantité de l'expérience clinique et de l'expérience personnelle du clinicien. Ils font partie de structures conceptuelles plus larges et sont reliés les uns aux autres par des éléments communs. Ces liens sont dynamiques et peuvent être mis à jour avec les nouvelles expériences (Hassebrock & Prietula, 1992). Ils existent à différents niveaux de généralité, allant de représentations de catégories de maladies générales à des exemples de maladies spécifiques d'un patient particulier. Les scripts sont des entités intégrées dont tous les éléments sont simultanément et automatiquement activés ensemble lors de la confrontation à un patient. Son activation entraîne ainsi l'anticipation d'autres caractéristiques cliniques chez le patient.

Le raisonnement clinique de l'expert serait donc un processus de recherche, sélection et vérification de scripts (Schmidt et al. 1990).

Quatrième étape : Les scripts exemplaires « Instance scripts »

Ces « instance scripts » sont le résultat de la confrontation répétée entre les scripts de maladie et les valeurs réelles chez les patients rencontrés. Le processus de vérification de la correspondance entre les 2 renforce les éléments et les associations récurrentes pour les résumer encore davantage dans un exemplaire de script. Ces

exemplaires sont stockés en mémoire à long terme, en supplément des scripts de maladie. Plus un clinicien a d'expérience clinique, plus il a développé d'« instance script » et plus il a de possibilités de rencontrer une correspondance parfaite entre un « instance script » et les données d'un patient particulier. Ce degré de correspondance influence la rapidité du diagnostic. Un cas typique ou routinier pour lequel un « instance script » est disponible, est diagnostiqué très rapidement (Belleza & Bower, 1981).

Cette théorie suppose donc que toutes les connaissances médicales soient organisées dans de vastes réseaux d'encapsulations dans lesquels les concepts biomédicaux et leurs liens de causalité sont inclus dans des propositions cliniques plus larges qui sont connectées entre elles par des mécanismes associatifs (Schmidt et al. 1990). Ce réseau de connaissances consiste en de nombreux sous réseaux de propositions cliniques qui comportent les caractéristiques cliniques et la physiopathologie sous-jacente, le tout relié par les liens de causalité ou associatifs appropriés.

Le Tableau 1 ci-dessous, adapté de l'article de Boshuizen et Schmidt (2008) synthétise les éléments ci-dessus selon le niveau de développement en intégrant le type de représentation des connaissances et leur organisation, le processus de raisonnement, son contrôle et sa charge cognitive ainsi que ses modalités d'application. Les novices et les étudiants de niveau intermédiaire possèdent des réseaux de connaissances moins élaborés qui nécessitent une charge cognitive supérieure lorsqu'ils analysent une situation clinique.

Tableau 1. Développement de l'expertise en raisonnement clinique en lien avec la base de connaissances (Tableau adapté de Boshuizen & Schmidt, 2008)

La restructuration des connaissances, le raisonnement Clinique et les niveaux d'expertise						
Niveau d'expertise	Représentation des connaissances	Acquisition des connaissances et (re)structuration	Raisonnement Clinique	Contrôle du raisonnement clinique	Charge cognitive	Clinical reasoning in action
Novice	Réseaux	Accroissement et validation des connaissances	Parcours long et détaillé à travers des réseaux encapsulés	Monitoring actif de chaque étape du raisonnement	Elevée	Difficulté de combiner la collecte et l'évaluation des données et le raisonnement clinique
Intermediaire	Réseaux	Encapsulation	Raisonnement à travers des réseaux encapsulés	Monitoring actif de chaque étape du raisonnement	Moyenne	...
Expert	Scripts de maladie	Formation de scripts de maladie	Activation de script de maladie et « instantiation »	Monitoring du niveau de « script instantiation »	Faible	Ajustement de la collecte des données selon le temps disponible et selon le niveau de vérification /faisification des hypothèses

Patel et Kaufman (2000) proposent un modèle différent dans lequel les connaissances biomédicales et les connaissances cliniques constitueraient « 2 mondes » différents. Selon ce modèle, il existerait un réseau composé de connaissances biomédicales dans lequel les concepts physiopathologiques sont connectés par des liens de causalité et un réseau distinct de connaissances cliniques dans lequel les concepts de maladie sont connectés aux signes et aux symptômes via des liens associatifs. Le réseau de connaissances biomédicales n'est activé que lorsque les liens associatifs du réseau de connaissances cliniques ne permettent pas le diagnostic.

Pour ces 2 modèles, les connaissances biomédicales ne sont activées que pour des cas atypiques ou rares (Groves, 2008). Van der wiel, Boshuizen et Schmidt en 2000 ont démontré que, sans contrainte de temps, les cliniciens experts sont capables d'expliquer les cas cliniques par les sciences de base. Ces données suggèrent un réseau où les connaissances cliniques et biomédicales sont interreliées. Il semblerait qu'avec le temps, les liens associatifs reliant les signes, les symptômes, les causes des maladies et les concepts cliniques de haut niveau deviennent plus fort que les liens de causalité entre les concepts cliniques et les mécanismes physiopathologiques. Ils deviennent effectivement les plus souvent activés et les plus utilisés. Le processus de raisonnement active tous les liens reliés à la maladie mais les liens de causalité entre les concepts biomédicaux exigent plus de temps pour être activés à un degré similaire aux liens associatifs (van der Wiel et al. 2000). Le concept de « chunking » permet de mettre tout le monde d'accord. Un « chunk » peut être considéré comme le résultat du développement d'un lien associatif fort dans le réseau de connaissances. Ce concept dépasse le problème du déploiement des connaissances biomédicales encapsulées tout en permettant la reconnaissance rapide et automatique caractérisant le raisonnement clinique expert.

6. LA FORMATION AU RAISONNEMENT

L'apprentissage du raisonnement clinique débute habituellement avec la formation clinique après les premières années réservées à l'acquisition des connaissances de base. Les méthodes d'apprentissage varient : lecture d'ouvrages cliniques sur les maladies, conférences de cas cliniques, observation d'experts cliniciens, séminaires d'Apprentissage du Raisonnement Clinique ou encore à l'occasion des stages par essais-erreurs. Chacune de ces méthodes comportent ses avantages et ses limites comme le décrit Nendaz (2005) dans son article. Il propose différents principes issus de la psychologie cognitive qui permettent de réfléchir à des activités d'enseignement efficaces pour l'apprentissage du raisonnement clinique (Regehr & Norman, 1996 ; Tardif, 1992). Ces principes sont :

- établir des liens avec les connaissances antérieures ;
- faciliter le raisonnement hypothético-déductif ;
- favoriser l'usage à la fois des processus analytiques et non analytiques de raisonnement clinique ;
- favoriser le transfert des connaissances ;
- favoriser l'organisation et l'activation des connaissances ;
- favoriser une récolte de données cliniques pertinentes et discriminantes.

Ces différents principes que nous allons expliciter ci-dessous nous indiquent que l'apprentissage du raisonnement pourrait s'initier très tôt dans le cursus, dès les premières années précliniques par un développement parallèle des connaissances et des capacités de raisonnement. Nous discuterons ensuite plus particulièrement le transfert des apprentissages et l'usage du Problem Based Learning comme méthode d'apprentissage du raisonnement biomédical et clinique.

6.1. Principes de base pour l'apprentissage du raisonnement clinique

6.1.1. Établir des liens avec les connaissances antérieures

Tout apprentissage est le résultat d'une démarche active et personnelle de confrontation et d'intégration de nouvelles informations aux connaissances antérieures. Lorsque nous sommes face à une nouvelle situation, nous tentons de lui donner du sens sur base de ce que nous connaissons déjà. La réactivation des connaissances antérieures dans toute situation d'apprentissage est donc essentielle. Les nouvelles connaissances ne se reçoivent pas telles quelles, elles sont transformées et mises en relation avec les connaissances antérieures afin de construire une nouvelle interprétation de la situation rencontrée en intégrant les nouveaux éléments.

La capacité de raisonnement est largement dépendante de la qualité de l'élaboration et de l'organisation de la base de connaissances. Afin de faciliter l'encodage et la récupération des connaissances, il est important d'inciter les étudiants à créer des liens entre les connaissances à acquérir et les connaissances déjà acquises et leur donner l'occasion d'organiser ces connaissances en fonction des tâches à réaliser. Nendaz (2005) souligne que *« chaque apprenant construit lui-même sa propre base de connaissances et qu'il est important de favoriser la validation des connaissances nouvellement acquises et des stratégies de raisonnement par une interaction avec les pairs étudiants et avec les cliniciens expérimentés »*.

6.1.2. Faciliter le raisonnement hypothético-déductif

L'enseignement traditionnel exigeait de l'étudiant de recueillir toutes les informations cliniques d'un patient avant d'émettre des hypothèses (forward reasoning). Or, les études sur le raisonnement de cliniciens plus expérimentés nous ont révélé qu'ils génèrent précocement des hypothèses qu'il cherche à valider ou invalider par la recherche d'informations chez le patient (backward reasoning). Ces études ont incité les encadrants à proposer de nouvelles modalités

d'apprentissage comme les séminaires d'apprentissage du raisonnement clinique développés sur base des méthodes du Problem Based Learning que nous décrivons ci-dessous (Kassirer, 1983 ; Chamberland, 1998).

Dans le cadre de ce type d'activité, les étudiants sont incités à poser des questions pour découvrir la situation ; chaque question doit être justifiée et argumentée et doit venir renforcer ou déforer une hypothèse. Cette activité permet aux étudiants de développer une stratégie efficace d'identification et d'interprétation des éléments pertinents d'un cas.

6.1.3. Favoriser à la fois l'usage des processus analytiques et non analytiques de raisonnement clinique

Comme décrit plus haut, il n'existe pas une seule forme de raisonnement, les approches analytiques (algorithme diagnostique, raisonnement hypothético-déductif) et non analytiques (pattern recognition) peuvent être utilisées seules ou conjointement pour résoudre un cas clinique particulier. La formation au raisonnement clinique devrait donc promouvoir ces 2 formes de raisonnement dès le début du cursus en permettant aux étudiants de développer plusieurs stratégies de résolution de problème. Notamment, Eva (2005) a suggéré que donner explicitement la consigne à des cliniciens de prendre en compte leur jugement de similarité initial (processus non analytique) tout en traitant de façon explicite les différentes données présentées (processus analytique) conduit à de meilleures performances que lorsque l'on ne tient compte que d'un seul de ces processus.

6.1.4. Favoriser le transfert des connaissances

Multiplier les exemples d'un même phénomène permet aux étudiants de créer un répertoire de cas et plusieurs voies d'accès en mémoire (Eva, Neville, & Norman, 1998). Hatala (2003) a démontré que cela renforçait l'apprentissage et permettait une meilleure reconnaissance lors de la présentation de nouveaux cas. Il est essentiel d'inciter les étudiants à rechercher les similarités et les différences entre divers problèmes afin qu'ils puissent en abstraire les concepts sous-jacents et se rendre compte que des problèmes en apparence similaire

peuvent être très différents et, inversement, des problèmes en apparence très différents peuvent concerner des concepts semblables. A chaque nouveau problème, il est important d'identifier les concepts sous-jacents et de les relier à ceux appris antérieurement.

Le concept de transfert est décrit plus précisément ci-dessous.

6.1.5. Favoriser l'organisation et l'activation des connaissances

Afin de faciliter l'activation des connaissances des étudiants lorsqu'ils sont face à une situation clinique, il est important qu'ils aient une compréhension en profondeur du problème pour en abstraire les phénomènes sous-jacents et qu'ils en construisent une représentation adéquate en mémoire. La présentation des connaissances dans le contexte de cas spécifiques favorise l'organisation des connaissances par la construction d'une représentation adaptée au contexte d'application de ces connaissances. Ainsi, différentes méthodes d'apprentissage utilisent la schématisation pour favoriser l'organisation des connaissances comme l'usage de schéma physiopathologique ou de cartes conceptuelles (Coté, Bellavance, Chamberland, & Graillon, 2005 ; Marchand & Gagnayre, 2005).

6.1.6. Favoriser une récolte de données cliniques pertinentes et discriminantes

Différentes études ont démontré l'importance d'orienter la recherche d'information selon les hypothèses précoces émises. La qualité des informations recueillies dépend de la pertinence des hypothèses émises. Ainsi Kassirer (1983) propose d'initier le recueil d'information de façon logique par rapport à la plainte du patient plutôt que favoriser la recherche exhaustive des données sans hiérarchisation. Bien qu'il soit important d'orienter la recherche d'information autour d'hypothèses, il faut veiller au biais de confirmation qui consiste à tenir compte de toutes les informations qui confirment notre hypothèse et négliger les autres (Klayman & Brown, 1993). Il convient dès lors d'inciter les étudiants à expliciter les liens entre une information particulière et une liste d'hypothèses et de les contraster afin d'aider l'étudiant à comprendre et retenir ce qui les différencie (Avrahami et al., 1997). Cet

exercice favorisera le développement des liens entre les différents concepts et les maladies en mémoire et en facilitera la récupération face à de nouveaux cas, par la création de différentes voies d'accès. Ils apprendront alors que certaines caractéristiques du cas peuvent être discriminantes alors que d'autres assez peu spécifiques. Reprenons l'exemple du diabète : perdre du poids peut en être un signe mais peu spécifique à l'inverse d'uriner beaucoup, qui orientera plus sûrement vers le diagnostic. Certaines études ont démontré l'effet bénéfique de la recherche d'éléments discriminants pour l'apprentissage, c'est-à-dire pour la rétention et l'utilisation ultérieure de ces connaissances (Wigton, Connor, & Centor, 1998).

Tous ces éléments relevés par Nendaz et al. (2005) mettent à nouveau en évidence la nécessité de développer simultanément la base de connaissances, la structuration de ces connaissances et les différentes formes de raisonnement. Bien que cette approche vise à développer le raisonnement diagnostique chez les étudiants de niveau de formation clinique, tous ces conseils sont tout à fait transférables à l'enseignement du raisonnement biomédical au niveau préclinique.

L'enseignant doit prendre conscience qu'il dispose d'un savoir « compilé » et qu'il a tendance à utiliser des raccourcis dans sa pensée lorsqu'il raisonne. Il est nécessaire de rendre explicite tous les liens établis entre les éléments de la situation, la physiopathologie, les connaissances et expériences antérieures afin de clarifier, développer et « rendre utilisables » les réseaux de connaissances.

6.2. La capacité de transfert des apprentissages

Toute formation et les différentes pratiques d'enseignement mises en œuvre ont pour objectif de développer la capacité des apprenants à mobiliser les connaissances acquises pour les utiliser dans des situations où elles sont nécessaires. Le processus de transfert est cette capacité à mobiliser et utiliser des connaissances antérieures pour résoudre une situation inédite (Frenay, 2004).

Bracke (2004) décrit le transfert comme un ensemble de processus dont le point de départ est la confrontation à un problème à résoudre. Ces processus nécessitent d'identifier les éléments pertinents de la situation afin de pouvoir construire une représentation mentale du problème. Cette représentation mentale du problème sera alors confrontée aux informations et connaissances stockées en mémoire afin d'identifier des éléments similaires entre le problème présenté et des expériences d'apprentissage antérieures. Quand l'étudiant est confronté à une nouvelle situation pour laquelle il ne dispose pas encore en mémoire de « modèle sur-mesure », il peut activer un problème similaire dont la solution est connue et ensuite apparier la solution à ce problème source avec la situation actuelle (Gick & Holyoak, 1987).

Carveni, Nguyen-Xuan, Hoc et Politzer (1990) identifient 2 difficultés associées au transfert : la difficulté à trouver une situation similaire en mémoire et la difficulté à trouver les similarités entre deux situations. Selon Ripoll (1992), ces difficultés renvoient à la représentation et l'organisation des connaissances en mémoire à long terme ainsi que des processus d'activation et de récupération de ces connaissances. Un réseau de connaissances bien organisé et une bonne intégration entre les nouveaux éléments et les connaissances antérieures sont des prédicteurs de la capacité de transfert (Dixon & Brown, 2012).

Les caractéristiques communes entre 2 situations peuvent être « de surface » (l'habillage de la situation, éléments non reliés directement aux objectifs et résultats de la situation) et « structurelles » (éléments reliés directement aux objectifs et résultats de la situation). Ces deux éléments vont affecter la similarité perçue entre 2 situations (Gick & Holyoak, 1987).

Dès 1988, Novick met en évidence une différence entre les experts et les novices d'un domaine spécifique dans la capacité à identifier la similarité entre 2 situations selon la représentation qu'ils peuvent se faire d'un problème. Les représentations des novices impliquent essentiellement des informations superficielles et de surface alors que les représentations des experts incluent des caractéristiques structurelles et de surface.

En conclusion, la capacité de transfert est la capacité à faire face aux nouvelles informations d'une situation inédite par l'activation en mémoire de connaissances antérieures pertinentes pour résoudre cette nouvelle situation (en percevant les similarités entre les tâches) et par l'assimilation des données nouvelles dans sa propre structure de connaissances.

Dans le domaine de la santé, il n'est pas rare d'observer des étudiants qui réussissent brillamment leurs examens de connaissances mais qui présentent des difficultés à utiliser ces connaissances pour résoudre des problèmes inédits (Norman, 2009). Ces observations ont conduit à réfléchir à de nouvelles modalités d'enseignement comme l'implémentation du PBL dont un objectif est une meilleure intégration entre les connaissances en sciences de base et les problèmes cliniques. L'impact positif de la contextualisation des apprentissages sur le stockage et la récupération des informations a été largement démontrée (Chi, Feltovich, & Glaser, 1981 ; Chi, Glaser, & Farr, 1988 ; Lesgold et al. 1988 ; Newell & Simon, 1972). Les étudiants exposés à de multiples exemples d'application d'un concept en contexte lors de leur apprentissage développent une représentation plus flexible des connaissances et ont plus de facilité à abstraire les éléments pertinents d'un problème (Gick & Holyoak, 1983 ; Spiro, Vispoel, Schmidt, Samarapungavan, & Boerger, 1987).

Pour entraîner la capacité de transfert, Norman (2009) suggère de relier les nouveaux concepts à apprendre aux connaissances antérieures et d'utiliser de multiples exemples pour identifier les structures communes et les différentes applications d'un même concept.

6.3. *Le PBL pour développer le raisonnement clinique ?*

Le Problem Based Learning fut introduit tout d'abord à la Faculté de médecine de l'Université de Mc Master au Canada en 1965 en réaction au niveau de connaissances dont faisaient preuve les étudiants diplômés après avoir suivi les cours basés sur les disciplines. La nouvelle

approche est dérivée de la méthode d'étude des cas. Cette modalité pédagogique fut initiée par Howard Barrows et fut ensuite largement développée comme stratégie d'éducation et comme modalité principale d'apprentissage.

6.3.1. La théorie éducationnelle du PBL

Le PBL est basé sur le modèle constructiviste de l'éducation bien que, lorsque celui-ci est appliqué à la médecine, il contienne également des éléments du socio-constructivisme. Selon Schmidt (1983), les principes cognitifs du PBL peuvent être résumés comme suit :

- La connaissance antérieure est le déterminant le plus important pour la quantité de nouvelles informations qui pourront être traitées et pour la qualité de ce traitement (niveau de compréhension).
- Pour que le processus commence, il faut que les connaissances antérieures soient activées par des indices du contexte dans lequel l'information a été étudiée.
- L'activation est dépendante de la façon dont les connaissances sont structurées en mémoire et cela détermine son accessibilité. Les connaissances sont organisées comme des réseaux sémantiques composés de concepts connectés entre eux par des liens qui spécifient la relation entre ces concepts (causal, temporel, conditionnel, d'attribution, ...) (Patel & Groen, 1986). La profondeur et la qualité de la compréhension augmentent lorsque le réseau se développe. Les réseaux sémantiques reflètent non seulement la connaissance formelle d'une personne mais aussi et surtout de son expérience et de ses idées. C'est cette unicité qui explique les différences entre les apprenants.
- Le développement des réseaux apparaît à travers un processus élaboratif qui prend part lors de l'apprentissage actif. L'élaboration est la construction de nouvelles associations ou liens entre des concepts là où il n'y en avait pas précédemment. Il en résulte une compréhension plus approfondie à travers la création de multiples chemins d'activation pour la récupération des informations.

- L'activation et l'accès aux connaissances stockées sont dépendants des indices contextuels. Les recherches sur les effets du contexte sur l'apprentissage et la récupération de connaissances ont montré que des informations contextuelles indicielles sont stockées en mémoire à long terme en même temps que les informations apprises intentionnellement, facilitant leur récupération future dans un même contexte.
- La motivation à apprendre augmente le temps d'étude et par là le résultat. La curiosité épistémique est plus favorable à la production d'un apprentissage en profondeur et est stimulée par une discussion en groupe dans laquelle les idées existantes sont clarifiées et mises à l'épreuve.

6.3.2. Les étapes du PBL

Le PBL implique que l'étudiant travaille à partir d'un problème clinique combinant des discussions en petits groupes généralement supervisées par un tuteur et de l'étude personnelle. La solution au problème exige d'être expliquée en termes de principes et mécanismes sous-jacents.

Les étapes proposées par Schmidt (1983) sont les suivantes :

- Définir et clarifier les termes et concepts ;
- Formuler le problème ;
- Générer des hypothèses pour expliquer le problème sur base des connaissances antérieures et de l'expérience ;
- Evaluer, affiner et hiérarchiser les hypothèses ;
- Identifier les manques en connaissances et générer des objectifs d'apprentissage pour expliquer les aspects inconnus du problème ;
- Etudier de façon indépendante entre les sessions de groupe ;
- Discuter des résultats trouvés et réévaluer le problème en appliquant les nouvelles informations apprises.

Selon Schmidt (1993), le suivi de ces étapes est cohérent avec la théorie cognitive sous-tendant le PBL et a plusieurs effets sur l'apprentissage de l'étudiant. L'analyse initiale du problème active les connaissances antérieures de l'étudiant. Ces connaissances sont activées

à travers la discussion du groupe et utilisées pour traiter activement les nouvelles informations acquises durant le travail personnel. La connaissance nouvellement acquise est intégrée aux connaissances existantes et est ensuite restructurée pour s'appliquer à un problème particulier. De plus, les étudiants développent un intérêt intrinsèque dans l'apprentissage car il prend place dans un contexte dans lequel ils devront l'utiliser. Un bénéfice supplémentaire est que la nature réaliste et intégrée du PBL prépare les étudiants à gérer l'incertitude et le besoin de baser leurs décisions sur des probabilités plutôt que sur des faits indiscutables, caractéristiques inhérentes à la pratique de la médecine.

6.4. Les effets du PBL ?

De nombreuses études ont été réalisées sur l'impact du PBL chez les étudiants en médecine. Les résultats restent mitigés. La plupart des études évaluant les effets du PBL concernent la comparaison entre des étudiants de cursus PBL à des étudiants de cursus plus traditionnel mais aucun effet spécifique positif ou négatif de l'un et l'autre cursus n'a pu être démontré incontestablement (Nouns et al. 2012).

Selon Barrows (Barrows, 1984 ; Barrows & Tamblyn, 1980), le PBL a des effets positifs sur la rétention des connaissances, sur l'intégration des sciences de base avec les connaissances cliniques, le développement des capacités cliniques, la motivation intrinsèque et le développement de la capacité d'apprentissage autodirigé. Les étudiants PBL sont caractérisés par une attitude plus positive face à l'apprentissage, un plus grand intérêt pour la compréhension conceptuelle que pour l'apprentissage par cœur (Coles, 1985 ; Newble & Clark, 1986) ; des habilités métacognitives supérieures (Blumberg & Michael, 1992 ; Glaser, 1991) ; ils ont une meilleure autogestion (Vernon & Blake, 1993) et sont plus enclins à la formation continuée (Shin, Haynes, & Johnston, 1993). Barrows et Tamblyn (1980) suggèrent que le PBL promeut la stimulation et le maintien d'un intérêt intrinsèque pour la matière car les étudiants identifient eux-mêmes leur manque de connaissances basé sur ce qu'ils perçoivent comme pertinent et nécessaire pour la solution au problème donné. Cependant, bien que De

Volder, Schmidt, Moust et De Grave (1986) aient montré que les discussions entreprises dans le cadre du séminaire PBL motivent les étudiants à entreprendre l'étude et à les intéresser davantage au sujet, cet intérêt ne se reflète pas dans les scores des étudiants sur cette même matière.

Les recherches sur les effets du PBL sur les résultats académiques sont, en effet, plus mitigées. Bien qu'il y ait un certain nombre d'études en faveur du PBL, une crainte souvent émise est que cette modalité d'apprentissage n'assure pas des connaissances solides en sciences de base (O'Neill, 2000). Une revue de la littérature réalisée par Koh, Khoo, Wong et Koh (2008) indique que les étudiants PBL évaluent eux-mêmes leurs connaissances médicales générales comme inférieures aux étudiants de cursus plus traditionnel. Nandi, Chan, Chan, Chan et Chan (2000) ont montré que bien que les étudiants PBL démontrent des habiletés inter-personnelles et des connaissances psychosociales supérieures ainsi qu'une meilleure attitude envers le patient ; les étudiants en cursus plus traditionnel obtiennent des scores supérieurs aux évaluations en sciences de base. Prince et al. (2003) ont analysé les performances en anatomie pour 8 écoles de médecine avec des cursus PBL et des cursus plus traditionnels. Ils n'ont pas trouvé des performances inférieures en connaissances en anatomie. La revue de la littérature réalisée par Vernon et Blake (1993) identifie des scores en sciences de base légèrement inférieurs et des scores en connaissances cliniques légèrement supérieurs pour les étudiants PBL comparés aux résultats obtenus par des étudiants ayant suivi un cursus plus conventionnel. Cependant, une étude plus récente de Blake, Hosoka & Riley (2000) montre une différence significative en faveur des étudiants PBL sur les 2 types de connaissances. Selon d'autres études, les étudiants PBL ont démontré une rétention des connaissances jusqu'à 60% plus grande, 2 à 4 années après l'apprentissage initial (Einsenstadt, Barry, & Glanz, 1990 ; Martenson, Eriksson, & Ingelman-Sundberg, 1985) ; ils produisent des explications pathophysiologiques plus cohérentes et détaillées (Hmelo, 1998) et sont plus capables d'intégrer les concepts des sciences de base aux aspects cliniques d'un cas. Ces caractéristiques ont été attribuées aux mécanismes d'apprentissage actif

en contexte et permettent en conséquence le développement de structures de connaissances élaborées. En d'autres mots, bien que les étudiants PBL peuvent initialement apprendre moins, ils investissent leur apprentissage plus intensément et le retiennent plus longtemps (Norman & Schmidt, 1992 ; Schmidt, De Grave, De Volder, Moust, & Patel, 1989). De plus, les étudiants PBL ont plus de facilité à transférer les connaissances acquises d'un contexte à un autre qui exige l'application des mêmes concepts pour sa résolution (Needham & Begg, 1991).

Le PBL promeut le modèle hypothético-déductif du raisonnement clinique à travers un cycle itératif de génération, testing et révision d'hypothèse selon les données collectées jusqu'à ce qu'une explication finale qui « colle » au problème soit atteinte.

Il existe certains éléments en faveur de l'idée que les étudiants PBL pensent effectivement différemment des étudiants ayant suivi un cursus conventionnel (Patel et al. 1991 ; Thomas, 1997). Comme attendu, les étudiants PBL utilisent le raisonnement hypothético-déductif alors que les autres utilisent plutôt des stratégies inférentielles (Arocha, Patel, & Patel, 1993 ; Patel, Groen, & Norman, 1991). Les étudiants PBL produisent plus d'explications possibles du problème et leurs explications montrent une cohérence globale, probablement car le PBL implique l'élaboration structurée des connaissances. Au contraire, les étudiants ayant suivi un cursus conventionnel produisent peu d'hypothèses et leurs explications sont moins élaborées : elles consistent davantage en une liste de diagnostics possibles plutôt que de se référer aux principes des sciences de base. Leurs explications montrent moins de cohérence globale que celles des étudiants PBL (Patel et al. 1991). Le PBL, à travers l'utilisation du raisonnement hypothético-déductif, renforce l'utilisation et l'application des principes des sciences de base, encourageant ainsi la construction de structures de connaissances élaborées exigée par le développement de l'expertise.

Les résultats parfois contradictoires observés dans les différentes études sont notamment dus aux différentes applications du PBL au sein des différentes écoles étudiées, aux différentes

méthodologies, ce qui implique les difficultés des mesures prises et leur variabilité.

6.5. Le PBL à la Faculté de Médecine de l'Université de Liège

L'utilisation du PBL a débuté à l'Université de Mc Master au Canada. Rapidement d'autres universités ont adopté ce nouveau mode d'apprentissage soit comme base de leurs cursus comme à l'Université de Maastricht, soit intégré à un cursus préexistant comme à l'Université de Newcastle en Australie. Différentes applications du PBL et de ses principes se sont donc développées à travers le monde. A l'Université de Liège, différentes mises en application du PBL ont été développées selon le niveau du curriculum mais toujours autour de principes communs. Dans le cadre d'un enseignement de type PBL, l'apprentissage est basé sur un problème concret, contextualisé dans la pratique professionnelle future sous forme de cas cliniques (vignettes cliniques), ce qui donne un sens concret aux sciences de base. Habituellement le cas clinique est intégratif, interdisciplinaire ; il fait appel à des notions structurelles (anatomie, histologie, embryologie) et fonctionnelles (biochimie, physiologie). L'enjeu est d'ancrer les questions pratiques posées dans les fondements théoriques qui les sous-tendent, nécessitant une démarche active de l'étudiant. L'analyse de ce problème clinique par l'étudiant l'amène à découvrir lui-même ses objectifs d'apprentissage soit, concrètement, la matière qu'il doit comprendre et apprendre sur base de ses acquisitions antérieures. Le PBL se pratique sous forme de séminaires rassemblant un petit nombre d'étudiants (habituellement 8 à 10) qui effectuent une démarche d'auto-apprentissage accompagnée par un enseignant. Celui-ci n'est plus un dispensateur de savoir mais bien un facilitateur de la démarche d'auto-apprentissage de l'étudiant.

A l'Université de Liège, dès la deuxième année et jusqu'à la fin de leur parcours de formation les étudiants sont confrontés à ce type de séminaire. Le Tableau 2 présente les différentes formes de PBL selon le niveau du cursus.

Tableau 2. Structure des études médicales 2000-2011

Année du cursus	Quadrimestre	Etape de la formation	Séminaire PBL
BAC 1	Q1 et Q2	Sciences fondamentales	
BAC 2	Q3	Sciences biomédicales générales	AIC Apprentissage à l'Intégration des Connaissances
	Q4	Homme normal et maladies (principes généraux)	APP Apprentissage Par Problèmes
BAC 3	Q5 et Q6		
Master 1	Q7 et Q8	Les maladies et le malade	ARC Apprentissage du Raisonnement Clinique
Master 2	Q9 et Q10		
Master 3	Q11 et Q12	Le malade	ARPC Apprentissage à la Résolution de Problèmes Complexes
Master 4	Q13 et Q14		

6.5.1. Les AIC – Apprentissage de l'Intégration des connaissances

Dès le début de la 2^{ème} année, ces séminaires sont organisés à 5 reprises. Les étudiants sont répartis en groupes de 10, encadrés par un tuteur (premier assistant, chef de travaux, chercheur permanent). Chacun de ces séminaires porte sur un thème précis proposé par un enseignant. L'objectif est d'amener l'étudiant à réactiver dans sa mémoire des connaissances acquises antérieurement dans les différentes disciplines enseignées et d'établir des réseaux entre elles afin de les utiliser dans un contexte nouveau (le thème du séminaire). L'étudiant intégrera ainsi des éléments déjà connus dispersés dans divers enseignements. Concrètement ces séminaires se déroulent sur une journée, débutant par la présentation du thème par un enseignant en amphithéâtre. Après 1h30 de travail, chaque groupe écrit au tableau les différents éléments d'intérêt en rapport avec le thème. A partir de cette liste, une ébauche de schéma explicatif est dessinée au tableau.

L'enseignant commente les résultats de cette première phase, puis, les groupes s'installent dans des salles de séminaires où ils travaillent en utilisant les documents de référence (essentiellement les syllabi et livres de cours) pendant 2h30. Les groupes se rassemblent alors dans l'amphithéâtre et mettent en commun les résultats de leurs recherches. Une discussion aboutit à une synthèse sous forme de schéma définitif. La séance se termine par une évaluation du fonctionnement du séminaire par chaque étudiant.

6.5.2. Les APP – Apprentissage Par Problèmes (Figure 1)

Lors du 2^{ème} quadrimestre de la 2^{ème} année, l'enseignement est organisé sur un mode multidisciplinaire par modules autour d'un système anatomique particulier (cardiovasculaire, respiratoire, néphro-urinaire, ...). A ce stade de la formation, l'étudiant apprend la structure et le fonctionnement de l'organisme normal (anatomie, embryologie, histologie, biochimie, physiologie humaine) ainsi que les principes généraux qui sous-tendent les maladies (biochimie et physiologie pathologiques, anatomie pathologique générale, sémiologie, ...). Les étudiants sont répartis en groupes de 8 à 10 encadrés par un tuteur clinicien. Certains étudiants sont amenés à remplir un rôle particulier pour la durée du séminaire. Ainsi, un étudiant est l'animateur : il va assurer le bon fonctionnement du groupe ; un autre étudiant est secrétaire : il écrit au tableau la production du groupe. Chaque étudiant reçoit une vignette clinique, le problème, point de départ de l'apprentissage. La première phase de l'exercice, « la phase aller » dure environ 1h30. Ils commencent par lire et résumer le problème. Ensuite, ils identifient dans la vignette les «phénomènes à expliquer». Ils cherchent alors des «hypothèses explicatives» à chacun de ces phénomènes sur base de leurs connaissances antérieures. Ne possédant pas encore toutes les connaissances nécessaires pour obtenir une compréhension complète de la situation, ils identifient leurs limites en connaissances et se fixent des objectifs d'apprentissage. Ils terminent cette première phase de l'exercice par une synthèse des hypothèses explicatives en reliant les différents phénomènes entre eux et en établissant une ébauche de schéma physiopathologique. Les étudiants disposent alors de deux ou trois demi-journées de travail personnel, à la

bibliothèque de la Faculté de médecine ou à domicile, pendant lesquelles ils recherchent, dans des livres de référence concernant les différentes disciplines, les matières correspondant aux objectifs d'apprentissage et les réponses aux questions posées lors de la séance « aller ». Ils corrigent et complètent alors leur propre schéma physiopathologique. Pour la dernière partie de ce séminaire, « la phase retour », les étudiants se réunissent à nouveau en groupe. Ils remettent au tuteur une copie du schéma physiopathologique qu'ils ont élaboré pendant la phase de travail personnel. En groupe, ils vérifient les hypothèses explicatives et en font la synthèse en reconstruisant un schéma physiopathologique commun. Ces séminaires d'APP constituent l'essentiel de la formation à ce stade. Toutefois, chaque module de séminaire d'APP est encadré par 4 à 6 heures de conférences d'introduction et de synthèse. Elles sont données par les enseignants, le plus souvent sous forme intégrée. Au seuil du premier quadrimestre qui comprend les séminaires APP, les étudiants bénéficient d'une journée de formation au cours de laquelle les principes du PBL et les modalités des séminaires APP sont expliquées et illustrées.

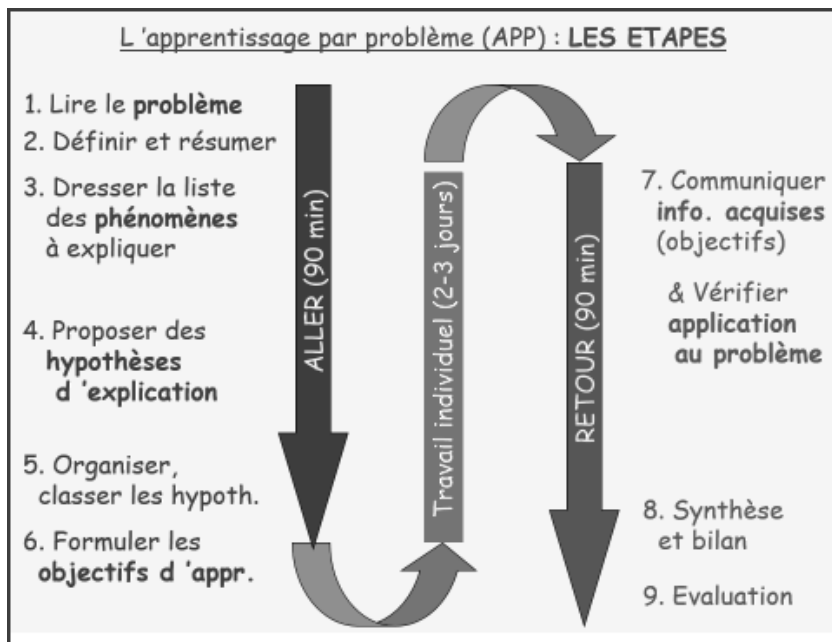


Figure 1. Les étapes du séminaire d'Apprentissage Par Problèmes

6.5.3. Les ARC – Apprentissage du Raisonnement Clinique

En master soit à partir de la 4^{ème} année, les enseignements sont toujours organisés autour de modules portant sur les systèmes anatomiques. Néanmoins, les enseignements ex cathedra reprennent une place plus importante. Les séminaires ARC sont organisés à nouveau en petits groupes de 8 à 10 étudiants encadrés par un tuteur clinicien expérimenté. Un des étudiants, le dispensateur de données, détient un document qui contient toutes les données anamnestiques, cliniques et para-cliniques d'un cas. Cet étudiant se présente avec une plainte initiale. Chaque étudiant du groupe va alors avoir l'occasion de poser ses questions afin d'obtenir les différentes informations concernant ce patient. Un étudiant commence par recueillir des données anamnestiques; lorsqu'il est au bout de son interrogatoire, il ébauche des hypothèses diagnostiques précoces. Si nécessaire, les autres étudiants complètent par d'autres questions. Le tuteur stimule le groupe pour que l'anamnèse conduise à la génération d'hypothèses diagnostiques (précoces). Ensuite, le groupe continue de la même façon pour les données de l'examen clinique, et enfin para-cliniques. L'objectif étant que chaque étudiant puisse aller au bout de son raisonnement en obtenant toutes les informations qui lui sont nécessaires. Les hypothèses précoces sont réévaluées et renforcées ou déforcées voire exclues ; d'autres hypothèses sont éventuellement avancées. Le tuteur veille à ce que chaque question soit justifiée et mise en relation avec une hypothèse diagnostique. L'exercice vise à se rapprocher progressivement du mode de raisonnement de l'expert clinique : à partir de symptômes et signes, aboutir au plus vite et au mieux à une hypothèse diagnostique plausible et exclure les autres hypothèses. Le séminaire dure environ 1 heure et se termine par une évaluation de la démarche effectuée et une brève discussion de la stratégie thérapeutique.

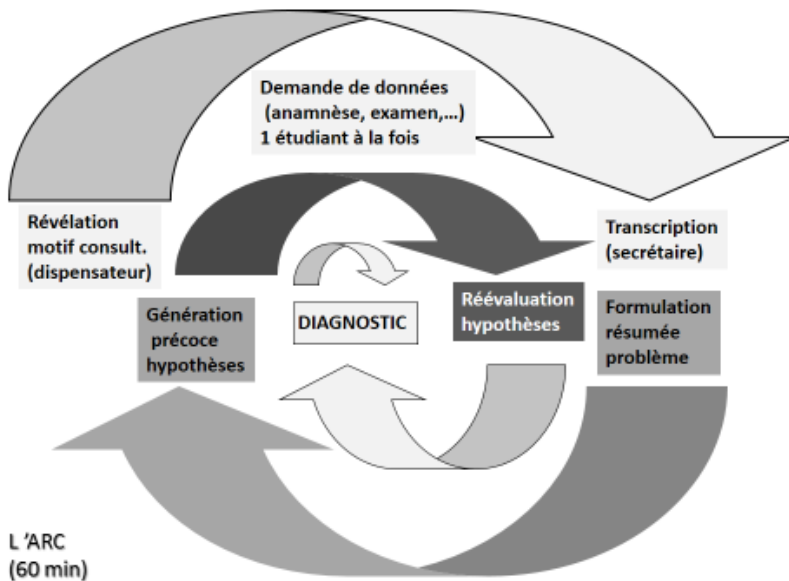


Figure 2. Les étapes du séminaire d'Apprentissage du Raisonnement Clinique

6.5.4. Les ARPC – Apprentissage de la Résolution de Problèmes complexes

A ce niveau de la formation, les étudiants sont essentiellement en stage. Ils suivent 2 journées de cours par mois. Durant ces 2 jours, ils participent aux ARPC. Pour ces séminaires qui s'organisent en 2 temps, les étudiants sont également répartis en petits groupes encadrés par un tuteur clinicien. Ces séminaires portent sur des situations cliniques complexes telles qu'on en retrouve fréquemment dans la vie professionnelle. Ils sont confrontés à des situations où le patient présente, en plus d'un problème clinique, des problèmes familiaux, professionnels, sociaux, économiques, etc. Ces situations, pour être résolues, doivent, en général, faire appel à d'autres compétences que celles d'ordre strictement médical. Au cours de la première phase, la séance «aller», les étudiants analysent le cas qui leur est soumis par le tuteur sous la forme d'un jeu de rôle à partir de la situation. Ils identifient les problèmes médicaux et non médicaux qu'il pose; ils répartissent toutes les questions restées en suspens entre les membres

d'un panel d'experts mis à leur disposition le lendemain (médecins, avocats, psychologues, assureurs, etc.). Le lendemain matin, les étudiants consultent les experts, bénéficiant des questions posées par les étudiants des autres groupes, puis se rassemblent à nouveau pour mettre en commun et construire ensemble la ou les solution(s) au problème complexe qui leur a été soumis.

Tableau 3. Les étapes du séminaire d'Apprentissage de la Résolution de Problèmes complexes (Giet et al. 2005).

Déroulement d'une séance d'apprentissage à la résolution de problèmes complexes (ARPC)	
Etapes de l'ARPC	Description
JOUR 1	
Durée : 20 min	<p>Séquence 1. Mise en situation (de type « jeux de rôles ») (10 groupes de 8 étudiants, sous le tutorat d'un clinicien spécialiste de la discipline concernée)</p> <ul style="list-style-type: none"> - Le tuteur ou un étudiant incarne un patient et introduit le cas auprès des étudiants par une phrase de départ. Ceux-ci interrogent le malade, explorent ses plaintes et le contexte dans lequel elles s'inscrivent.
Durée : 1h	<p>Séquence 2. (mêmes groupes et même tuteur que pour la séquence 1. Un étudiant-animateur est désigné)</p> <ul style="list-style-type: none"> - Le groupe lit la description complète du cas et découvre éventuellement des éléments qu'il n'a pas soulevés lors de la mise en situation. - Les problèmes qui restent sans réponse après débat au sein du groupe sont listés. - Les étudiants formulent des questions à soumettre à des experts de spécialités multiples (avocat, infirmier, médecin spécialiste, psychologue, etc.). Des documents de référence liés aux disciplines concernées sont, si cela est nécessaire, proposés aux étudiants. - La spécialité des 4 ou 5 experts présents le lendemain (voir jour 2 ci-dessous) est révélée. - Les questions en suspens sont réparties entre ces 4 ou 5 domaines d'expertise et les étudiants du groupe se subdivisent en 4 ou 5 sous-groupes, chacun étant chargé de rencontrer un expert et d'investiguer un domaine ainsi défini.
JOUR 2	
Durée : 1h	<p>Séquence 3. Consultation des experts (1 ou 2 représentants de chaque groupe d'étudiants à la rencontre d'un expert)</p> <ul style="list-style-type: none"> - Le ou les représentant(s) de chaque groupe d'étudiants interagissent avec les experts en leur posant tour à tour des questions précises.
Durée : 1h30	<p>Séquence 4. Synthèse (mêmes groupes et même tuteur que pour les séquences 1 et 2)</p> <ul style="list-style-type: none"> - Les mêmes groupes d'étudiants se réunissent en vue de la synthèse des informations recueillies auprès des experts et élaborent un consensus qui apporte des réponses aux questions et problèmes restés en suspens pour une prise en charge globale du patient.

7. L'ÉVALUATION DU RAISONNEMENT

Ces dernières années, différentes méthodes de mesure du raisonnement clinique ont été développées. Néanmoins aucune de ces modalités d'évaluation ne rencontre une approbation unanime et ne prétend mesurer le processus de raisonnement clinique dans toute sa complexité et ses dimensions.

Les difficultés liées à l'évaluation du raisonnement clinique sont à la fois dues à la complexité du processus lui-même, comme la diversité des cheminements de raisonnement ou la spécificité de contenu mais aussi au processus d'évaluation lui-même qui doit être valide, fidèle et praticable.

7.1. *Les modalités d'évaluation du raisonnement clinique*

Charlin, Bordage et Van der Vleuten (2003) présentent quelques outils d'évaluation du raisonnement clinique dont l'essentiel est décrit ci-dessous.

7.1.1. Patient Management Problem (PMP)

L'idée de cette évaluation est de simuler le processus de collecte d'informations et de prise de décisions du clinicien (Mc Guire & Babbott, 1967). La procédure de base implique la présentation d'une amorce de cas clinique. L'étudiant reçoit alors une série d'options possibles, il doit choisir parmi elles toute donnée supplémentaire qui serait nécessaire concernant l'histoire du patient, les examens et les investigations afin de lui permettre de sélectionner un diagnostic et des décisions de prise en charge sur base de l'interprétation des données collectées. Le parcours du raisonnement de l'étudiant est comparé à un groupe de référence, une note est attribuée en fonction du caractère complet de la collecte de données, de son efficacité et de sa pertinence (Van der Vleuten & Newble, 1995).

Les limites de cette évaluation sont tout d'abord liées au format des questions qui par les indices proposés peuvent améliorer artificiellement le score. Ensuite, le système de cotation lui-même pourrait évaluer la capacité à rassembler des données plutôt que la compétence diagnostique.

7.1.2. Les grilles d'évaluation globale

Ces grilles comportant une liste de critères à évaluer sont essentiellement utilisées dans le cadre des stages. Elles sont complétées par une ou plusieurs personnes amenées à être en contact avec l'étudiant pendant quelques semaines à quelques mois. La diversité des critères à évaluer devrait permettre d'aborder tous les aspects de la compétence clinique mais seuls un petit nombre de maîtres de stages évaluent spécifiquement le raisonnement clinique (Norman, 1993). Bien que l'utilisation de ces grilles soit très répandue, plusieurs études ont montré une fidélité relativement faible de ces grilles (Streiner, 1985).

7.1.3. Les Questions à Choix Multiple (QCM)

Les QCM sont très largement utilisées dans le cadre des évaluations certificatives avec un grand nombre d'étudiants à évaluer. Ces évaluations standardisées permettent un haut degré de fidélité et d'objectivité. Elles offrent la possibilité d'évaluer un large spectre de connaissances ainsi que de grandes populations d'étudiants. Bien que le QCM ne soit pas l'outil de choix pour l'évaluation du raisonnement clinique car il est reconnu comme outil d'évaluation des connaissances factuelles, il est possible d'approcher l'étude du raisonnement par la construction de QCM autour de présentations de cas cliniques (Jolly & Grant, 1997).

7.1.4. L'oral

L'utilisation de l'examen oral est largement répandue car il permet l'évaluation conjointe des connaissances, de la capacité de résolution de problème ainsi que des attitudes et caractéristiques propres à l'étudiant interrogé. Les biais les plus souvent rapportés pour ce type d'évaluation sont soit un excès de clémence, soit un excès de sévérité ou encore une tendance à donner à tous un score moyen (Jean,

DesMarchais, & Delorme, 1993). Malgré ces biais, l'examen oral est très utilisé pour sa facilité d'organisation (bien que coûteux en temps) et parce qu'il permet effectivement de se rendre compte de la qualité du raisonnement.

Si cet examen oral est articulé autour d'un cas clinique réel, il peut aussi comporter un biais lié au(x) domaine(s) limité(s) au(x)quel(s) se rapporte le cas, à l'exclusion des autres domaines.

7.1.5. La question rédactionnelle

L'usage de la question rédactionnelle dans le cadre d'évaluation certificative est à nouveau très fréquent malgré une fidélité inter-juge habituellement assez basse (Norcini et al. 1990). Il est préférable d'utiliser un plus grand nombre de questions à réponse ouverte courte. Ce type de question permet d'évaluer le raisonnement sous-jacent à une prise de décision.

7.1.6. L'Examen Clinique Objectif et Structuré (ECOS)

Cet examen est basé sur la mise en situation clinique des étudiants. L'évaluation correspond à l'observation du comportement et du discours de l'étudiant sur base d'une grille prédéfinie des éléments attendus. L'examen consiste en une série de « stations » qui, chacune, évalue un comportement clinique dans une situation particulière, en général simulée (Harden & Gleeson, 1979). L'étudiant dispose d'un temps fixe dans chaque station. La structure standardisée de cet examen permet d'obtenir une bonne fidélité et différents domaines sont explorés ainsi que les compétences communicationnelles en situation simulée.

7.1.7. Problems and Modified Essay Questions (MEQ)

Cette modalité d'évaluation s'apparente au PMP (Norman & Feightner, 1981). Il s'agit de l'introduction séquentielle d'un cas qui commence par une vignette clinique. Sur la base des informations contenues dans la vignette, on demande à l'étudiant d'identifier l'information additionnelle nécessaire. Les réponses sont collectées et les informations transmises. Des questions sont alors posées sur l'interprétation des données, sur les données supplémentaires

nécessaires ou le diagnostic et la prise en charge. Le MEQ nécessite que l'étudiant rédige de courtes réponses aux questions liées aux aspects spécifiques de chaque élément d'information obtenue séquentiellement (Feletti, 1980).

Difficulté « pratique » de cette évaluation : nécessité de collecter les réponses à chaque étape avant de fournir la série de données suivantes.

7.1.8. Les questions à appariement étendu (EMQ, Extended Matching Questions)

Il s'agit d'une variante de la modalité QCM. L'amorce de la question constitue un motif de consultation associé à un ensemble de signes. Une liste de diagnostics possibles est alors proposée à l'étudiant (Case, Swanson, & Stillman, 1988) parmi lesquels il doit déterminer le plus pertinent face à la situation présentée. Cette modalité standardisée permet l'évaluation d'un grand nombre d'étudiants par une correction automatisée.

7.1.9. L'examen par éléments clés

Cet examen est constitué de scénarios cliniques suivis par une série de questions de différentes formes (QCM, QROC, choix parmi une longue liste) conçues pour évaluer les éléments clés du cas (Page & Bordage, 1995). L'accent étant mis sur les éléments clés de résolution du problème concerné, cette méthode permet de mesurer la capacité de prise de décision sur base de ces éléments (Hatala & Norman, 2002). Habituellement 20 à 40 cas sont nécessaires pour obtenir une fidélité suffisante (Jolly & Grant, 1997). Comme il est nécessaire d'obtenir une seule bonne réponse possible, les situations univoques sont privilégiées.

7.1.10. Le Test de Concordance de Scripts (TCS)

Ce test mesure la capacité à raisonner en contexte d'incertitude et à résoudre des problèmes mal définis c'est-à-dire un problème dont les données, les buts et les solutions ne sont pas univoques. Cette capacité renvoie à l'expertise professionnelle (Schön, 1983). Il est basé sur la théorie des scripts, réseaux de connaissances mobilisés en situation de résolution de problèmes. Une série de problèmes cliniques brefs, mal définis sont présentés à l'étudiant. Il lui est alors demandé de

prendre des décisions diagnostiques, d'investigation ou de traitements lorsque de nouveaux éléments lui sont présentés. Concrètement, une hypothèse est proposée et l'étudiant doit déterminer quel effet (positif, négatif ou neutre) la découverte d'une donnée complémentaire a sur le statut de cette hypothèse. Le TCS est construit pour vérifier si l'organisation des connaissances cliniques (scripts) de l'évalué se rapproche (concorde) des décisions prises par un panel de référence (experts) (Charlin, Gagnon, Sibert, & Van Der Vleuten, 2002a). En effet, le score obtenu par l'étudiant pour chaque réponse est fonction du nombre d'experts ayant répondu comme lui. Cet outil présente une bonne fidélité et une bonne validité (Charlin, Desaulniers, Gagnon, Blouin, & Van der Vleuten, 2002b). Comme le précise Charlin, Bordage et Van Der Vleuten (2003), la démarche de construction d'outil d'évaluation du raisonnement clinique a été jusqu'alors une tentative de simulation de la réalité. Le TCS, quant à lui, est issu d'une théorie du raisonnement clinique et vise à mesurer un processus essentiel du raisonnement plutôt que l'issue du raisonnement face à une situation qui mime la réalité.

7.1.11. L'évaluation basée sur la performance : Le miniCEX (Clinical Examination Exercise)

Cette évaluation se déroule sur un lieu de pratique professionnelle (salle d'urgence, clinique, ...). Il est demandé à l'étudiant de réaliser un examen physique et raisonner à voix haute sur un cas particulier (Holmboe & Hawkins, 1998). L'évaluateur donne ensuite un feedback à l'étudiant et complète une grille d'observation. *L'observation directe de l'étudiant dans son interaction avec le patient est essentielle et semble apporter une valeur ajoutée à l'évaluation du raisonnement clinique* (Wass & Jolly, 2001).

7.1.12. Diagnostic Thinking Inventory (DTI)

Le DTI a été développé par Bordage, Grant et Marsden (1990) pour l'évaluation du raisonnement diagnostique. Il est constitué d'une série d'items, chacun comprenant une amorce suivie par 2 affirmations opposées séparées par une échelle de likert en 6 points représentant un continuum entre ces 2 affirmations. Cet outil évalue 2 aspects différents

du raisonnement : la flexibilité du raisonnement et l'organisation des connaissances. Cet instrument est basé sur les recherches sur le développement du raisonnement clinique du novice à l'expert. La flexibilité du raisonnement évalue dans quelle mesure les processus de réflexion peuvent être appliqués durant la démarche diagnostique (Exemples issus de l'article Beulens, Struyf, & van Damme, 2006: « *When I am interviewing a patient : I often seem to get an idea stuck in my mind about what might be wrong versus I usually find it easy to explore various possible diagnoses* »). L'organisation des connaissances concerne la disponibilité des connaissances stockées en mémoire durant le processus diagnostique (« *When the patient presents his or her symptoms: I think of the symptoms in the precise words used by the patient versus I think of the symptoms in more abstract terms than the expressions actually used (e.g. '4-day duration' becomes 'acute', '2 hands' becomes 'bilateral')* »). Selon Bordage et al. (1990) le DTI est un outil valide et fidèle.

Le raisonnement clinique est un processus multidimensionnel, il implique notamment la capacité à intégrer différentes données, à générer des hypothèses pertinentes, à évaluer ces hypothèses à la lumière des différentes informations obtenues, à prendre des décisions diagnostiques, et de prise en charge (Charlin et al. 2003). Aucun instrument ne permet l'évaluation de toutes ces dimensions. En général, les outils qui visent l'évaluation du résultat du processus de raisonnement (le bon diagnostic par exemple) sont privilégiés aux outils d'évaluation du processus de raisonnement lui-même. Cette préférence s'explique notamment par la meilleure fidélité de ces outils ainsi que par leur plus grande faisabilité. Les processus de raisonnement étant beaucoup plus variables que les résultats de ces processus (Charlin et al. 2003). Or en cours de formation, il est essentiel de pouvoir aider les étudiants à identifier les différentes étapes et leur rendre un feedback sur leur processus afin également de leur proposer les remédiations appropriées.

Tableau 4. Forces et limites des outils d'évaluation du raisonnement clinique

Outils d'évaluation	Validité	Fidélité	Faisabilité
Patient Management Problems	Evalue la capacité à résoudre des problèmes ; Evalue la capacité à rassembler les données (≠ la compétence diagnostique) ; Difficultés liées à l'effet d'indices et à la multiplicité des cheminements possibles.	Difficulté liée à la spécificité de contenu, il est nécessaire d'impliquer un grand nombre de cas pour obtenir une fidélité acceptable.	Coûteux en temps pour la préparation du matériel.
Les grilles d'évaluation globale	Peu d'éléments de ces grilles évaluent le raisonnement.	Faible (difficultés liées à la longue durée, effet de halo, prise en compte du caractère de l'étudiant, tendance à la surévaluation, subjectivité).	Acceptable.
Les questions à choix multiple (QCM)	Généralement faible, évalue plutôt l'application des connaissances que le processus de raisonnement.	Elevée (standardisation, objectivité).	Acceptable (nécessite un investissement pour la création de questions de qualité).
L'oral	Elevée (permet de se rendre compte du processus de raisonnement de l'étudiant). Biais si basé sur un seul cas (domaines restreints explorés).	Faible (biais liés aux caractéristiques des évaluateurs et des évalués).	Coûteux en temps pour l'administration de l'examen.
La question rédactionnelle	Permet d'évaluer le processus de raisonnement sous-jacent à une prise de décision.	Faible (difficultés liées à la qualité de la copie au niveau de la calligraphie, à la grammaire ou l'orthographe notamment).	Coûteux en temps pour la correction.

L'examen clinique objectif et structuré (ECOS)	N'évalue qu'une part du processus de raisonnement et uniquement ce qui est observable. Compétences communicationnelles évaluées par la mise en situation.	Elevée (standardisation et objectivité).	Coûteux en ressources matérielles et humaines.
Modified essay questions	Evalue la capacité à déterminer un diagnostic plutôt que le processus de raisonnement.	Acceptable avec un nombre de cas élevé.	Coûteux en temps pour la préparation du matériel ; Difficulté liée à la nécessité de récolte des réponses à chaque étape.
Les questions à appariement étendu	Evalue le processus de raisonnement non analytique « pattern recognition ».	Elevée (standardisation et objectivité).	Acceptable (nécessite un investissement pour la création de questions de qualité).
L'examen par éléments clés	Evalue la capacité de prise de décisions sur base de la prise en compte d'éléments clés ; Porte sur des situations proches de la vie professionnelle mais à réponse univoque.	Acceptable avec un nombre élevé de cas (40).	Coûteux en temps pour la préparation du matériel.
Le test de concordance de scripts (TCS)	Evalue le processus d'interprétation des informations cliniques dans un contexte mal défini.	Elevée (standardisation et objectivité).	Acceptable (nécessite un investissement pour la création de questions de qualité).
Le miniCEX	Evalue en contexte authentique essentiellement les comportements observables.	Faible (spécificité de contenu, biais lié à l'évaluateur).	Coûteux en temps pour l'administration.
Diagnostic Thinking Inventory	Evalue essentiellement l'auto-évaluation de son processus de raisonnement et de l'utilisation de ses connaissances.	Acceptable.	Elevée, ne nécessite que peu de ressources.

7.2. Les évaluations en Faculté de Médecine

La plupart des modalités d'évaluation à la Faculté de médecine de l'ULg se veulent multidisciplinaires. Elles sont organisées à 3 moments dans l'année. Une première période d'évaluation se déroule en janvier. Elle concerne les évaluations des cours du premier quadrimestre. Une deuxième période d'évaluation a lieu en juin, concernant les évaluations du deuxième quadrimestre ou l'évaluation des cours qui s'étalent sur les 2 quadrimestres. Une seconde session est prévue en aout pour les étudiants ayant échoué à leurs premières évaluations.

7.2.1. Les évaluations privilégiant l'évaluation des connaissances

- Les Questions à Choix Multiples avec Degré de Certitude (QCM DC)

Dès la première année, les étudiants sont confrontés à des questionnaires à choix multiples évaluant essentiellement leurs connaissances factuelles. Dans certains cas, ces QCMs peuvent évaluer l'application des connaissances comme par exemple lorsqu'elles portent sur l'analyse d'un schéma. Pour les années 2 à 5, ces questions sont principalement constituées d'une amorce suivie de 4 affirmations dont les possibilités de réponses sont : toutes vraies, toutes fausses, une vraie.

Pour chaque réponse fournie, l'étudiant doit déterminer son degré de certitude, c'est-à-dire, dans quelle mesure il est sûr de sa réponse, sur une échelle de likert à 6 niveaux (1: 0-25% ; 2: 25-50% ; 3: 50-70% ; 4: 70-85% ; 5: 85-95% ; 6: 95-100%). En 6^{ème} et 7^{ème} années, cette modalité d'évaluation est fortement réduite. Elle ne concerne plus qu'un seul cours.

- Les Questions à Réponse Ouverte Courte (QROC)

Cette modalité d'évaluation se pratique à nouveau tout au long du cursus. Elle concerne principalement l'évaluation de l'application des connaissances.

7.2.2. Les évaluations privilégiant l'évaluation du raisonnement

- Discussion Multidisciplinaire d'une Vignette Clinique (DMVC)

Cette évaluation, qui se pratique dès la deuxième année, consiste en l'analyse d'une vignette clinique et la présentation d'un schéma reliant les différents éléments de cette vignette, permettant d'expliquer la situation clinique. Cette analyse est ensuite présentée et discutée avec un jury d'enseignants. Pour les étudiants des années 2 et 3, il s'agit d'identifier les éléments pertinents et de les relier par des relations de cause à effet afin d'expliquer et intégrer les phénomènes physiopathologiques sous-jacents à la situation. Pour les années 4 et 5, les étudiants identifient les données cliniques et proposent un diagnostic et une prise en charge thérapeutique. Pour les années 6 et 7, ce type d'évaluation du raisonnement persiste mais les modalités de mise en œuvre peuvent varier légèrement selon le cours évalué.

- Examen Clinique Objectif Structuré (ECOS)

Un ECOS est organisé en 6^{ème} année dans le cadre de l'examen de Pédiatrie. L'examen est constitué de 7 à 8 stations, parcourues en 1 heure. L'étudiant est mis en situation de 2 façons : les situations simulées où l'interrogateur joue par exemple le rôle d'un parent ; et les situations discutées où l'interrogateur invite l'étudiant à expliciter certains aspects de la prise en charge. La tâche porte par exemple sur l'anamnèse, l'interprétation de résultats d'examens complémentaires, la prescription ou l'explication d'un traitement, etc. Les éléments de réponse attendus sont repris dans une grille préétablie et cochés par un examinateur. Les séances d'examen ont lieu de préférence après que les étudiants aient presté au moins un mois de stage de pédiatrie.

Dans le cadre de l'examen de sélection en Médecine Générale, un ECOS est également organisé fin de la 7^{ème} année. Il est constitué de 17 stations. La réussite de cette évaluation conditionne la possibilité de poursuivre le Master Complémentaire en Médecine Générale pour 70 % de la note globale à l'examen de sélection.

- Test de Concordance de Scripts (TCS)

Un Test de Concordance de Scripts est également intégré à l'examen de sélection en Médecine Générale. Cet examen s'organise à la fin de la 7^{ème} année pour les étudiants inscrits à l'année préparatoire en Médecine Générale. Ce TCS est constitué d'une cinquantaine de questions diagnostiques, d'investigations ou thérapeutiques.

7.2.3. Les évaluations mixtes

- Les Questions à Réponse Ouverte Longue (QROL)

Cette modalité d'évaluation permet de prendre en considération la capacité d'intégration, de synthèse ou d'analyse des étudiants. Elles sont pratiquées tout au long du cursus.

Au-delà de ces différentes modalités d'évaluation, les étudiants sont évalués lors de leurs travaux pratiques, sous forme de QCM, QROC ou QROL ainsi que lors de mises en situation plus procédurales. Les étudiants sont également évalués en fin de cursus sur leur performance en stage et pour la dernière année sur l'élaboration d'un travail de fin d'études. Ce travail peut concerner des thématiques diverses (la clinique, la recherche, la médecine du travail, la médecine légale, etc.) et est défendu devant un jury de 5 enseignants.

Il est à noter que la pratique réflexive des étudiants est évaluée de façon explicite et pour un quart de la note globale du cours de gynécologie-obstétrique de fin de cursus (6^{ème} année).

Tableau 5. Les évaluations en Faculté de Médecine

Année du cursus	1	2	3	4	5	6	7
Evaluation privilégiant les connaissances	QCM DC QROC	QCM DC QROC	QCM DC QROC	QCM DC QROC	QCM DC QROC	/	QCM QROC
Evaluation privilégiant le raisonnement	/	DMVC	DMVC	DMVC	DMVC	DMVC ECOS	DMVC ECOS TCS
Evaluations mixtes	QROL	QROL	QROL	QROL	QROL	Stages	QROL Stages TFE

DMVC : Discussion Multidisciplinaire d'une Vignette Clinique ; ECOS : Examen Clinique Objectif et Structuré ; QCM DC ; Questions à Choix Multiples avec Degrés de Certitude ; QROC : Questions à Réponse Ouverte Courte ; QROL : Questions à Réponse Ouverte Longue ; TCS : Test de Concordance de Scripts ; TFE : Travail de Fin d'Etudes.

7.3. La capacité d'auto-évaluation de sa base de connaissances

La formation au raisonnement dès le début du cursus médical a pour objectif de promouvoir le développement du raisonnement en parallèle au développement du réseau de connaissances. A ce stade précoce de la formation, l'important est de raisonner à partir des connaissances en sciences de base et non pas de chercher à poser un diagnostic. Bien que la confrontation précoce à des situations cliniques ait démontré un certain nombre de bénéfices (Bell, Boshuizen, Scherpbier, & Dornan, 2009 ; Diemers, Dolmans, Verwijnen, Heineman, & Scherpbier, 2008 ; Dorman & Bundy, 2004 ; Prince et al. 2000 ; Seabrook, 2004), le risque d'un contact précoce avec la clinique pourrait être d'inciter les étudiants à se focaliser sur les connaissances cliniques et la compétence diagnostique et à négliger les mécanismes des sciences de base qui sont sous-jacents (Diemers, Van De Wiel, Scherpbier, Heineman, & Dolmans, 2011). Le travail en groupe amène également certains étudiants à se fixer comme objectif l'obtention du schéma explicatif établi de manière coopérative en négligeant ou suivant passivement le travail de raisonnement qui y conduit. Ce phénomène a

notamment pu être mis en évidence lors des évaluations orales de discussion multidisciplinaire de vignette clinique. Les enseignants du jury ont pu quelquefois observer des étudiants qui présentent un schéma expliquant la situation de manière complète et correcte mais qui ont des difficultés à expliquer les phénomènes physiopathologiques sous-jacents. Or la validation des connaissances de base sous-jacentes au raisonnement est essentielle pour la validation du processus de raisonnement (Langendyk, 2006). Nendaz (2005) précise qu'il est important de favoriser la validation des connaissances nouvellement acquises. La qualité du processus de raisonnement dépend notamment de la capacité de contrôle que l'on peut avoir sur notre processus de raisonnement (Mamede, Schmidt, & Rikers, 2007). La capacité de délimiter son champ de connaissances peut ainsi être une part importante du processus de validation du raisonnement.

Cette capacité d'auto-évaluation des connaissances sera étudiée notamment par l'utilisation des degrés de certitude aux tests de connaissances. Comme décrit plus haut, les degrés de certitude évaluent dans quelle mesure un étudiant est sûr de sa réponse à une question, ici de connaissances. Il s'agit d'établir un jugement sur sa réponse sur base de l'état de ses connaissances (A-t-il pu récupérer distinctement les connaissances nécessaires pour répondre à la question ? ; Le niveau de maîtrise de ses connaissances est-il suffisant pour répondre avec certitude à la question ?).

PARTIE 2 : LES TRAVAUX DE RECHERCHE

1. INTRODUCTION

1.1. *Cadre des recherches*

Comme décrit plus haut, les études médicales à l'Université de Liège comportent, depuis 2000, des séminaires « PBL » ou Problem Based Learning qui s'échelonnent de la deuxième à la dernière année du cursus (Boniver, 2004). Les méthodes d'évaluation ont également été révisées afin de garantir une cohérence entre les objectifs de la formation médicale, ses méthodes d'enseignement et son évaluation. Le Tableau 1 ci-dessous résume la structure des études médicales. Ce cursus « PBL » est en réalité un dispositif mixte comportant des séminaires PBL (qui concernent l'ensemble de la matière) ainsi que des activités classiques dont les cours magistraux et les travaux pratiques.

Un des effets du PBL a été d'établir une continuité entre la formation de base (baccalauréat) et la formation clinique (master). Nos travaux de recherche s'inscrivent dans le cadre de ce dispositif pédagogique de la Faculté de médecine de l'ULg et des questions et réflexions qu'il a suscité chez les enseignants impliqués. La mise en place de cette méthode pédagogique a en effet fait émerger un questionnement sur les modalités d'apprentissage des étudiants en médecine et en particulier l'apprentissage du raisonnement biomédical et clinique. De fait, une originalité de notre recherche a été d'impliquer les étudiants dès la 2^{ème} année du cursus, soit à un stade précoce et avec un suivi transversal ou longitudinal de la maturation de ces processus pendant les années suivantes.

1.2. *Les questions de recherche*

Nos différents travaux de recherche abordent donc la capacité de raisonnement des étudiants dans différents contextes ainsi que les facteurs pouvant influencer sur cette capacité, parmi lesquels les

connaissances et leur auto-évaluation. Les questions de recherche posées à travers les 4 articles sont les suivantes :

- a) Les étudiants bénéficiant d'un apprentissage contextualisé (un cas concret) dans le cadre des séminaires PBL sont-ils capables de transférer les connaissances acquises pour les appliquer lors d'une tâche de raisonnement à appliquer à un autre cas, dans un nouveau contexte? Cette capacité varie-t-elle selon le niveau d'expérience ? (Article 1)
- b) La capacité de raisonnement des étudiants augmente-t-elle avec l'expérience au cours du cursus médical PBL ? (Article 2 et 3)
- c) La capacité de raisonnement et sa maturation sont-elles liées à la rétention des connaissances et son évolution ? (Article 2 et 3)
- d) La capacité de raisonnement et sa maturation sont-elles liées à la capacité d'auto-évaluation des connaissances ? (Article 2 et 3)
- e) La capacité d'auto-évaluation des connaissances augmente-t-elle avec l'expérience au cours du cursus médical PBL ? (Article 3)
- f) Travailler explicitement la capacité d'auto-évaluation des connaissances lors d'une séance PBL permet-il de renforcer l'apprentissage et la validation du raisonnement chez les étudiants ? (Article 4)

Tableau 1. Structure des études médicales à l'ULg et activités PBL qui y sont menées

Année du cursus	1		2		3		4		5		6		7	
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Formation	Sciences fondamentales		Sc. Biomed	Homme normal et maladies (principes généraux)			Les maladies et le malade				Le malade			
PBL			AIC	APP			ARC				ARPC			
Evaluation des connaissances	QCM avec degrés de certitude QROC QROL													
Evaluation du raisonnement	/	/	Discussion Multidisciplinaire d'un cas clinique avec un jury de 3 à 5 enseignants (discussion sur les aspects anatomiques, histologiques, biochimiques et physiopathologiques)	Discussion Multidisciplinaire d'un cas clinique avec un jury de 3 enseignants (discussion sur le diagnostic, l'anamnèse, l'examen physique et les examens complémentaires)										Analyse de cas cliniques
Stages					120 h		144h	144h	168h	168h				21 à 24 mois

AIC : Apprentissage de l'intégration des Connaissances ; APP : Apprentissage Par Problèmes ; ARC : Apprentissage du Raisonnement Clinique ; ARPC : Apprentissage de la Résolution de Problèmes complexes ; QCM : Questionnaire à Choix Multiple ; QROC : Question à Réponse Ouverte Courte ; QROL : Question à Réponse Ouverte Longue

2. RÉSUMÉ DES MATÉRIELS ET MÉTHODES

2.1. *Participants*

La participation aux recherches expérimentales des articles 1, 2 et 4 s'est déroulée sur base volontaire. Les étudiants et tuteurs recrutés pour ces travaux étaient ceux ayant participé aux séminaires PBL concernés, les séminaires portant sur la matière de l'endocrinologie.

Par ailleurs, une recherche longitudinale (article 3) a exploité les données des évaluations certificatives de 3 cohortes d'étudiants au cours de leur cursus pour l'ensemble des matières traitées dans les séminaires APP (Apprentissage Par Problèmes – Années 2 et 3) et ARC (Apprentissage du Raisonnement Clinique – Années 4 et 5).

Le Tableau 2 décrit plus précisément les populations de participants selon les différentes recherches.

2.2. *Matériels*

L'évaluation du raisonnement

Le Test de Concordance de Scripts (TCS) est une méthode développée pour évaluer précisément le raisonnement clinique (Charlin et al. 2000, 2002, 2004). Un TCS (articles 1 et 2) a été créé pour aborder des questions de raisonnement biomédical pouvant être gérées dès le baccalauréat. Le TCS est un outil d'évaluation des capacités de raisonnement de problèmes partiellement définis, c'est-à-dire en contexte d'incertitude. Ce choix délibéré vise à la cohérence avec les problèmes rencontrés sur le terrain professionnel et pour lesquels une partie seulement des paramètres utiles au raisonnement sont disponibles. La construction des items pour ce test s'est basée sur les problèmes traités par les étudiants lors des séances APP et ARC du module endocrinologie.

Les données de l'examen oral intégratif de discussion multidisciplinaire d'un cas clinique des étudiants ont également été exploitées comme mesure des capacités de raisonnement (article 3). Cet

examen implique d'identifier les éléments pertinents d'un cas (signes, symptômes, éléments contextuels) et de les mettre en relation les uns avec les autres afin de présenter une représentation globale du cas et une compréhension des phénomènes impliqués dans ce cas. Pour les années 2 et 3 (baccalauréat), il s'agira d'expliquer et d'intégrer les mécanismes physiopathologiques ; pour les années 4 et 5 (masters) il s'agira d'expliquer l'utilisation des données cliniques dans la démarche diagnostique ainsi que les choix d'examen complémentaires et de traitement. La capacité de raisonnement est ici évaluée à travers l'analyse du problème. Elle mesure la capacité de l'étudiant à sélectionner et transformer les éléments factuels dans une construction schématique et sémantique qui permet d'évaluer la progression vers la structuration, l'analyse et la compréhension du problème. Cette capacité est essentielle au raisonnement clinique (Chang, Bordage, & Connell, 1998 ; Nendaz et al. 2005).

Les schémas physiopathologiques créés par les étudiants lors des séances APP ont également été utilisés comme mesure de l'évolution du processus de raisonnement en cours de séance APP (article 4). Le schéma de l'étudiant contient les éléments qu'il a pu identifier comme pertinents et les relations qu'il a établies entre ces différents éléments aux différentes étapes d'une séance APP. Dans le cadre de notre travail, l'évolution en termes de nombre d'éléments et de liens ainsi que le niveau de certitude et de compréhension de ces éléments et liens ont été mesurés. L'analyse de ces schémas était accompagnée d'un questionnaire spécifiquement créé et a été suivie d'un débriefing. Ces 2 dernières modalités d'évaluation concernaient le fonctionnement des étudiants en APP.

L'évaluation des connaissances

L'évaluation des connaissances a été réalisée dans les différents travaux par des questionnaires vrai-faux avec degrés de certitude. Chaque question était constituée d'une amorce suivie de 4 items pour lesquels il fallait déterminer si l'affirmation était vraie ou fausse. Dans le cas des évaluations certificatives, les options de réponses étaient : toutes vraies, toutes fausses, une vraie et 3 fausses.

L'évaluation de la capacité d'auto-évaluation des connaissances

Pour chacune des questions de connaissances, les participants étaient amenés à déterminer leur degré de certitude par rapport à la réponse fournie. Il s'agit d'indiquer dans quelle mesure le participant est sûr de sa réponse, à l'aide d'une échelle de Likert à 6 niveaux (0-25% ; 25-50% ; 50-70% ; 70-85% ; 85-95% ; 95-100%). Il faut noter que ces degrés de certitude font partie des éléments fournis par l'étudiant lors des évaluations certificatives dès le début du cursus. Les étudiants sont donc familiers avec la procédure et entraînés à l'utiliser. Nous l'avons exploitée dans le cadre des questions vrai-faux des évaluations certificatives (article 3) aussi bien qu'expérimentales (article 2). Une capacité d'auto-évaluation optimale devrait se traduire par un degré de certitude élevé pour les réponses correctes et un degré de certitude faible pour les réponses incorrectes et donc une différence maximale entre les deux. Ainsi, un indice de la capacité de discriminer le connu de l'ignoré a été calculé sur base de la différence entre les degrés de certitude pour les réponses correctes et les degrés de certitude pour les réponses incorrectes.

3. SYNTHÈSE DES RÉSULTATS

3.1. *La capacité de transfert des connaissances*

Les résultats ont démontré, dès la 3^{ème} année du cursus PBL en Faculté de médecine, une capacité des étudiants à utiliser les connaissances acquises dans le contexte d'une situation clinique donnée pour les appliquer lors d'une tâche de raisonnement sur une nouvelle situation (article 1). Lorsque les questions portent sur des situations inspirées des séances d'Apprentissage du Raisonnement Clinique (ARC) mais n'impliquent aucun élément contextuel rencontré dans ces séances PBL, elles sont les mieux réussies par les étudiants. Donc, un contexte neuf favorise la mobilisation des capacités de raisonnement. Par ailleurs, les cliniciens expérimentés présentent des performances moins dispersées et plus élevées que les étudiants pour l'ensemble des catégories de questions TCS.

3.2. *L'évolution de la capacité de raisonnement durant la formation*

Nos différents travaux ont évalué comment la capacité de raisonnement varie au cours du cursus. Nous avons pu observer que le raisonnement biomédical fait partie des capacités de l'étudiant dès les 2^{ème} et 3^{ème} années du cursus médical à l'Université de Liège. Nous avons également démontré une augmentation des performances en raisonnement durant ce cursus, aussi bien à travers une étude expérimentale transversale mesurant les capacités de raisonnement via un test spécifiquement créé (article 2) qu'à travers une étude longitudinale exploitant les données des évaluations certificatives (article 3).

3.3. *La capacité de raisonnement et la rétention des connaissances factuelles*

Les performances aux tests de rétention des connaissances (vrai-faux) n'évoluent pas parallèlement aux performances aux tests de raisonnement au cours du cursus médical. Notre étude expérimentale transversale (article 2) a démontré une diminution des performances en connaissances et une augmentation des performances en raisonnement lorsque nous évaluons une même matière sur 4 années du cursus (années 3 à 6) avec un délai qui augmente entre l'acquisition et l'évaluation de la rétention. A l'inverse, ce délai est court et constant lorsque nous avons exploité les données longitudinales des évaluations certificatives pour l'ensemble des matières abordées (article 3). Dans ce cas, nous observons un gain global entre les années 2 et 5 avec une certaine réduction entre l'année 3 (dernière année dans les sciences de base) et l'année 4 (première année en sciences cliniques). Par contre, nous observons une augmentation linéaire pour les performances en raisonnement.

En ce qui concerne la corrélation entre performances au test de connaissances et au test de raisonnement, la première étude (article 2) qui utilise un Test de Concordance de Scripts pour l'évaluation du raisonnement, a montré une corrélation positive pour les premières années (3 et 4) mais absente pour les années suivantes (5 et 6). La seconde étude (article 3) basée sur l'examen oral de discussion multidisciplinaire d'un cas clinique impliquant une discussion sur les connaissances sous-jacentes au raisonnement, a montré une corrélation positive de l'année 2 à l'année 5.

3.4. *La capacité de raisonnement et la capacité de délimitation du champ des connaissances*

Nos études (articles 2 et 3) ont montré une corrélation positive entre les performances en raisonnement évaluées par un Test de Concordance de Scripts ainsi que par une discussion multidisciplinaire

sur un cas clinique et la capacité de délimiter son propre champ des connaissances. Celle-ci a été évaluée par la différence entre les degrés de certitude aux réponses correctes et les degrés de certitude aux réponses incorrectes.

3.5. L'évolution de la capacité de délimitation du champ des connaissances

L'étude expérimentale transversale n'a pas pu mettre en évidence une augmentation de la capacité à discriminer le connu de l'ignoré au cours du cursus médical (article 2). Par contre, lorsque nous étudions cette capacité chez un même étudiant de l'année 2 à l'année 5, nous observons une augmentation significative (article 3).

La caractéristique développementale et la signification de cette capacité sont confirmées par la comparaison de la performance chez les étudiants selon qu'ils échouent ou réussissent lors des examens de la 2^{ème} année du cursus. Nous observons une capacité de délimitation du champ des connaissances supérieure pour les étudiants qui réussissent leur année. Parmi ceux qui échouent une première fois leur année mais qui la réussissent après avoir recommencé cette même année, ils améliorent le score et atteignent une capacité de délimitation des connaissances équivalente à ceux ayant réussi dès la première tentative (article 3).

3.6. L'impact d'un exercice travaillant la capacité d'auto-validation des connaissances au cours d'une séance APP

Il s'agit d'un APP expérimental avec 3 groupes d'étudiants auxquels des consignes de validation du schéma explicatif sont soit données à la fin de la séance aller, soit durant la séance retour, soit après la séance retour (article 4). La majorité des étudiants qui participent à cet APP expérimental déclarent à posteriori discerner de

manière implicite les éléments qu'ils connaissent, ceux dont ils doutent et les éléments incomplets lors de leur participation aux séminaires d'APP. Néanmoins, les exercices proposés lors de l'APP expérimental et qui les forcent à faire ce travail de façon explicite, leur permettent de significativement mieux se rendre compte de leur évolution en identifiant ces différents éléments. La façon dont la mise en commun coopérative a été menée au retour a eu également un impact positif sur leur fonctionnement ; toutefois, aucun effet significatif de cette mise en commun sur la validation du raisonnement n'a pu être démontré.

Bien qu'on ait pu observer une évolution significative des schémas entre le début et la fin de la séance retour, les exercices de délimitation du champ des connaissances ne semblent pas avoir modifié de manière significative le contenu des schémas individuels. Les schémas finaux étaient en effet comparables pour les différents groupes quelque soit le moment où les consignes de validation ont été données. L'analyse des schémas produits avant la mise en commun permet, toutefois, de dégager des différences dans le processus individuel de correction du schéma selon les consignes données aux étudiants. En effet, lorsque celles-ci prennent place avant le travail de mise en commun autour du schéma coopératif, cela entraîne une réduction d'environ 60% du nombre d'éléments manquants identifiés après le schéma coopératif et renforce donc l'apprentissage individuel. Dès lors, il semblerait que la validation « forcée » de son raisonnement avant la construction du schéma coopératif et la confrontation aux connaissances sous-jacentes favorise la démarche personnelle de l'étudiant dans la prise de conscience de ses acquis et de ses limites, ce qui pourrait renforcer l'apprentissage.

Tableau 2. Résumé des matériels, méthodes et résultats des articles de recherche

Article	Participants	Matériels et Méthodes	Résultats
1. Context impact of clinical scenario on knowledge transfer and reasoning capacity in a medical PBL curriculum	- 104 étudiants volontaires - année 3 à 6 - 9 tuteurs	- TCS - Endocrinologie	- Augmentation des performances au TCS avec l'expérience - Performances aux questions ARC dans un contexte inédit les mieux réussies dès la 3 ^{ème} année
2. Reasoning versus knowledge retention and ascertainment throughout a problem-based learning curriculum	- 104 étudiants volontaires - année 3 à 6 - 9 tuteurs	- TCS - Vrai/Faux avec degrés de certitude - Endocrinologie	- Augmentation des performances au TCS avec l'expérience - Diminution des performances au test de connaissances après la 3 ^{ème} année - Corrélation positive entre les performances au TCS et au test de connaissances uniquement pour les années 3 et 4 - Corrélation positive entre les performances au TCS et la capacité de discerner le connu de l'ignoré - Pas d'augmentation de la capacité à discerner le connu de l'ignoré
3. Longitudinal assessment of progress in reasoning capacity and relation with self-estimation of knowledge base	- 213 étudiants - année 2 à 5 (+ 66 étudiants échouant année 2)	- 7 DMVC - 7 QCM vrai/faux avec degrés de certitude - Toutes matières confondues	- Augmentation des performances aux MDCC avec l'expérience - Augmentation des performances aux QCMs entre les années 2 et 5 avec une certaine réduction entre l'année 3 et l'année 4 - Augmentation de la capacité à discerner le connu de l'ignoré avec l'expérience - Corrélation positive entre les performances au TCS et les performances aux QCMs - Corrélation positive entre les performances au TCS et la capacité de discerner le connu de l'ignoré - Lien significatif de la capacité à discerner le connu de l'ignoré avec la réussite en 2 ^{ème} année
4. Effets d'un exercice de validation des éléments d'un raisonnement biomédical sur les composantes individuelles et collectives de l'apprentissage au cours d'un séminaire APP	APP expérimental : - 20 étudiants volontaires - année 3 2009-2010 répartis en 3 groupes - 3 tuteurs Questionnaires : + 43 étudiants non participants année 3 2009-2010 - Groupe contrôle : 23 étudiants année 3 2010-2011	- Analyse d'une vignette APP - Endocrinologie - Consignes de validation des connaissances (3 moments) - Questionnaire - Fonctionnement en APP - Débriefing	- Les exercices proposés rendent explicite le travail de validation des connaissances - Les exercices proposés permettent à l'étudiant d'identifier l'évolution de ses capacités - Les exercices proposés avant la mise en commun renforce le processus de correction individuel lors de la mise en commun. Cet effet se traduit par une réduction du nombre d'éléments manquants identifiés après la mise en commun.

4. ARTICLE 1 : CONTEXT IMPACT OF CLINICAL SCENARIO ON KNOWLEDGE TRANSFER AND REASONING CAPACITY IN A MEDICAL PBL CURRICULUM

A. Collard, S. Brédart and J-P. Bourguignon

Soumis à la revue « Higher Education Research and Development »

ABSTRACT

Background: Since 2000, the Faculty of Medicine at the University of Liège has integrated Problem-Based Learning seminars from year 2 to 7 in its 7-year curriculum. The PBL approach has been developed to facilitate student's acquisition of reasoning capacity. This contextualized learning raises the question of the de- and re-contextualization process and its development during medical education.

Methods: A Script Concordance Test (SCT) to assess clinical reasoning was administered to 108 volunteer students from year 3 to 6 of the Faculty of Medicine and to a reference panel. The SCT has been created to measure, in a standardized way, reasoning capacity for ill-defined clinical problem. We used clinical cases from PBL seminars related to endocrinology from 3rd and 4th Year. Based on each clinical case used for APP and ARC seminars, 2 SCT scenarios aiming at the same learning objectives were created: one in the original learning context and one in a different, new context.

Results: Irrespective of the learning context, the reference panel showed higher scores than all groups of students. Students in year 5 and 6 obtained higher scores than younger students. A significant difference was also found in relation to the context (Original/New) of the question. Higher scores were unexpectedly obtained for questions related to ARC cases than APP cases and in new contexts versus original ones, irrespective of the study level. The higher score were obtained for 4th year cases in new context.

Conclusion: This study shows that since third year, students are able to transfer the knowledge acquired throughout PBL seminars for application to questions with information out of the learning context. Moreover, reasoning capacity is growing with experience, irrespective of the process of transfer.

4.1. Introduction

One of the most important learning aim is to develop the capacity of using knowledge acquired in one context to solve a new dissimilar problem in another context. This phenomenon is identified in cognitive psychology as transfer capacity. It is a real difficulty to use basic knowledge acquired during the learning process to solve or explain a new problem. *“Typically, students who have learned a concept in one problem context will only have a 10-30% success rate in applying the concept to solve a new problem”* (Norman, 2009; Norman, Dore, Krebs, & Neville, 2007). Johnson, Dixon, Daugherty, and Lawanto (2011) have reported the inability of students to recognize the transferability of concepts learned from solving well-structured problems in classroom to ill-structured problems faced outside of the classroom. This observation raises the question as to whether the lack of transfer capacity results from absent or insufficient training or inadequacy of learning conditions as compared to clinical problem solving conditions. The latter hypothesis is supported by research on the capacity of generalization or transfer where the issue of factors influencing access to memory has been addressed. This has lead to show that the environment in which learning occurs (the room; the time of the day; the surrounding noise, ...) (McGeoch, 1932; Smith, 1988; Smith & Vela, 2001) or the characteristics of the learners like experience or knowledge (Eva, Neville & Norman, 1998) as well as emotions can influence the retrieval of information from memory. This phenomenon has been identified during the 70’s by Tulving as “context specificity” (Tulving, 1973; Smith & Vela, 2001).

In medical education, it is not uncommon that students, after passing exams successfully, are still unable to apply the acquired knowledge to solve a new problem (Norman, 2009). Such observations have led to revised educational strategies like the implementation of Problem-Based Learning (PBL) throughout medical curriculum with the aim of better integration of basic science with clinical problems. The specificity of PBL is learning from clinical cases, including basic sciences, in other words contextualization of learning. Using clinical cases in PBL gives students the opportunity to be closer to the real professional world

and therefore can enhance student's general learning transfer and problem-solving skills (Hmelo-Silver, 2004). The cognitive approach has largely demonstrated the positive impact of integrated learning on storage and retrieval of information and also on reasoning and problem solving capacity (Chi, Feltovich, & Glaser, 1981; Chi, Glaser, & Farr, 1988; DeGroot, 1965; Lesgold et al. 1988; Newell & Simon, 1972).

PBL however still faces the question of the difficulty of knowledge transfer after being acquired in a specific context. Thus, it is up to the individual student to extract knowledge from clinical cases and transfer that knowledge to new contexts. The student has to retrieve a learned concept to apply that concept in understanding and solving a new problem. Thus, the difficulty not only involves extraction and transfer of information but also lies in the mental representation, the abstraction and organization of the knowledge base. To use a concept for solving a problem, the common features that the new problem shares with a concept acquired must be recognized (Norman, 2009). Expert problem solvers remember the underlying conceptual structure that makes two problems similar although they have different surface features. By contrast, novice problem solvers tend to prioritize surface similarities between problems (Sutton, 2003). This ability makes easier the transfer of concept because of the conscious effort to abstract knowledge and concepts from one context for application to another (Johnson, 1995). Sometimes the new problem context has clues that do exactly match problem already solved but more frequently, it doesn't look analogous on the surface. Then, the transfer task is more difficult because in order to identify the analogies, the process must take place at the level of the deep structure of the problem (Norman, 2009). As an example, a systolic murmur at the aortic valves will be interpreted by the novice as a turbulence of blood flow generating a murmur like in stenosis of other heart valves. In contrast, the expert will rapidly think in terms of consequences of malfunction of this particular valve for pressure upstream in the left heart and flow rate and blood supply downstream in organs including heart.

In the cognitive perspective of learning process, it is proposed that learning progresses through interpretation of new knowledge in the light of what is already known (Norman, 2009). A well-organized knowledge base and efficient integration of any new learning to what a student already knows are predictors of the transfer capacity (Dixon & Brown, 2012). Without an extensive and organized knowledge base, transfer strategies are less likely to operate outside the instructional context since there is nothing or too few to connect together (Haskell, 2001). Experts not only have an appropriate knowledge base in their long-term memory, they also have the ability to access and use that knowledge. An important difference between experts and novices is how their knowledge base is organized (Glaser, 1987). A structured knowledge enables experts to “see into” a problem more quickly than novices. Further, this difference enables them to more precisely encode and store information. As a consequence, access to and retrieval of relevant information is more efficient. Haskell considers that knowledge base is an absolute requirement not only for transfer but also for thinking and reasoning. Along the same line, Galloti (1989) concluded *“it appears that some of what predicts good everyday reasoning is the breadth and depth of the knowledge base. The knowledge that one possesses affects the type of mental models that one can construct as well as the type of problems that can even be recognized.”*

In the medical domain, Regher and Norman (1996) suggested that expertise is characterized by “scripts” i.e. construction of knowledge networks that connect specific experiences (contextualized information), and/or abstract concepts (de-contextualized information). The cognitive approach identifies major differences between experts’ scripts and novices’ scripts. With experience, physicians develop more extensive and better organized scripts (Custers, Boshuizen & Schmidt, 1998; Feltovich & Barrows, 1984; Schmidt, Norman & Boshuizen, 1990). Expertise also enhances pattern of information perception, and then, activation of scripts (Custers, Boshuizen & Schmidt, 1996). Some argued that experienced physicians show superior performance in relevant script activation from prior information available at the first step of clinical reasoning (Elstein, Shulman & Sprafka, 1978), even with limited

information (Hobus, Schmidt, Boshuizen & Patel, 1987; Hobus, Hofstra, Boshuizen, & Schmidt, 1988; Hobus, Boshuizen, & Schmidt, 1990; Hofstra, Hobus, Boshuizen, & Schmidt, 1988). Finally, with experience, the capacity of information processing progresses, in rapidity and accuracy. According to knowledge and experience, experts are able to filter relevant information and then prioritize them.

Since 2000, the Faculty of Medicine at the University of Liège has integrated PBL seminars from year 2 to 7, in its 7-year curriculum (Boniver, 2004). The Problem-Based Learning (PBL) approach has been developed to facilitate student's process of acquisition, organization and retrieval of knowledge (Norman & Schmidt, 1992; Regher & Norman, 1996; Schmidt, 1983; Van Gessel, Nendaz, Vermeulen, & Junod, 2003). One aim of PBL is to help the student integrating new information in a rich and connected knowledge network that would be later activated (illness scripts) (Barrows, 1986). PBL promotes transfer skills, on top of clinical reasoning skills (Dixon & Brown, 2012; Hmelo-Silver, 2004).

The present study addressed, in a PBL setting, the de- and re-contextualization process and its development during medical education. The investigation tools involved a Script Concordance Test (SCT) (Caire, Sol, Charlin, Isodori, & Moreau, 2004; Charlin, Roy, Brailovsky, Goulet, & Van Der Vleuten, 2000; Charlin, Gagnon, Sibert, & van Der Vleuten, 2002; Charlin & StJean, 2002; Charlin & Van Der Vleuten, 2004) to assess clinical reasoning throughout the curriculum (year 2 to 6). The Script concordance test has been created to measure, in a standardized way, reasoning capacity for ill-defined clinical problem.

4.2. Methods

Participants

The study was based on 108 volunteer students from year 3 to 6 of the Faculty of Medicine at the University of Liège (with a minimum of 20 students for each study year). Endocrinology was the selected topic of the investigations. The students differed according to the PBL

seminars and the clerkship period they had attended (Table 1). In the second and third years, the students learned the structure and normal functioning of the human body systems, through APP (Apprentissage Par Problèmes) seminars. In the clinical cases used for APP seminars, reasoning aimed at identifying among the observations (history, examination, complementary assessment, treatment) those that needed to be explained and providing relevant explanation. In the clinical cases used for ARC seminars, the aim of reasoning was much more focused: obtaining information (history, examination, complementary assessment) to raise diagnostic hypotheses as early as possible and prioritize one hypothesis as proposed diagnosis. In year 4 and 5, the diagnostic and therapeutic bases of diseases of the different systems were learned through lectures and ARC (Apprentissage du Raisonnement Clinique) seminars. Written informed consent was obtained from the participants who responded to the test questions anonymously. Among the tutors handling the PBL seminars, nine served as reference panel for SCT scoring (see below). Two teachers wrote the test questions.

Table 1. Characteristics of the participants

Year of the curriculum	3rd	4 th	5th	6 th
Training completed (endocrine system) at the time of study	APP seminars	APP seminars	APP and ARC seminars	APP and ARC seminars
Lectures on diagnosis and treatment of endocrine disorders	-	-	done	Done
N participants / N Total	35/114	20/100	25/96	24/92
Reference Panel	9 tutors involved in APP and ARC seminars about the endocrine system			

Testing procedure

Originally, a SCT question consists in a short and incomplete clinical scenario which are challenging even for an expert. Following the scenario, a diagnostic or management hypothesis is proposed and then, a new information was provided. The responders have to assess the impact of this new information on the initial diagnostic or management hypothesis. The response is given on a 5-point Likert scale. Each individual response is credited by a score calculated on the basis of the distribution of responses given by a reference panel (Charlin et al. 2000, 2002; Charlin, Tardif & Boshuizen, 2000).

While keeping the original SCT format, we adapted the test to assess skills of clinical reasoning and knowledge transfer in young students not yet familiar with the diagnostic process. We used clinical cases from PBL seminars related to endocrinology: 5 cases from 3rd year seminars (APP) and 5 cases from 4th year seminars (ARC). For each clinical case, 2 SCT scenarios aiming at the same learning objectives were created: one in the original learning context and one in a different, new context. The “Original Context” scenarios involved information coming only from the clinical cases of PBL seminars used for learning purposes, without any irrelevant information. In addition, the physiological function was disturbed in the same way as in the learning case. The “New Context” scenarios involved information that was new (different from the learning clinical cases) and mixed with some irrelevant information (i.e. that was not usable for reasoning in this particular case). The physiological function was disturbed differently from the learning case. Tables 2a and 2b give an example of these 2 series of questions. For each scenario, three independent hypotheses were proposed, each followed by additional new information.

Table 2. Examples of SCT questions in “Original Context” (a) and in “New Context” (b) with the element of the original context in *italic*

(a) A young women with Hashimoto’s thyroiditis treated by 75 µg/day of L-thyroxin complains about tiredness, loss of hair and intolerance to coldness			
If the hypothesis is	And you are informed that	The hypothesis becomes	Answer of refer.
The dose of thyroxine is insufficient	Reverse T3 serum levels are in the normal range	Much less likely	-
		Less likely	3
		Not affected	6
		More likely	-
		Much more likely	-

(b) A young man is admitted in the emergency room for drug intoxication. He could have taken 12 tablets (6g as a whole) of aspirin and 10 tablets (2mg as a whole) of L-thyroxin. His heart rate is 72/min.			
If the hypothesis is	And you are informed that	The hypothesis becomes	Answer of refer.
The patient took less medication than suspected	The medications were taken 1 hour before admission	Much less likely	1
		Less likely	7
		Not affected	1
		More likely	-
		Much more likely	-

The present study was based on 48 SCT questions including, 22 APP and 26 ARC questions. Those 48 questions were validated by the reference panel whereas 12 questions were discarded based on two exclusion criteria (ambiguous or confusing writing and/or content out of the learning objectives).

Scoring of SCT

The original scoring method developed by Charlin and collaborators was “the aggregate scoring process” (Charlin et al. 2002; Charlin & Van Der Vleuten, 2004). Participants ‘response to each item, is compared with the response of the panel members. The score was attributed to each responder based on how many of panel members have chosen the same response. A maximum score of 1 was given for the modal answer. Other responses are given a partial credit depending on the number of panel members who provided that answer. For example, on a given item, if 8 (out of 10) members of the reference

panel responded “much less likely” on the Likert scale, this option was credited by the maximum score of 1 point (8/8, i.e. the modal response). Then, if 2 members selected “no effect” as response, this option was credited 0.25 point (2/8). The answers not selected by any of the panel members received no credit whether they were close to or far from the modal answer.

That scoring process was adapted for this study using the kappa coefficients methodology (Vanbelle, Massart, Giet, & Albert, 2007). Weighted kappa coefficient (Cohen, 1968) quantified the agreement between 2 raters on an ordinal scale and corrected for agreement due to chance. This methodology was extended to quantify the agreement between the reference panel and a responder. The weights permitted to assign higher scores to responses closer to the modal response(s) given by the reference panel. The weights were derived from the reference panel responses by computing the frequency distribution of the number of category differences separating two experts’ responses of all possible pairs of experts. This led to a weight equal to 1 was given for a category of the Likert scale consistent with the modal response of the reference panel, a weight of 0.96 for the adjacent categories and weights of 0.19, 0.06 and 0 at a distance of 2, 3 and 4 categories, respectively. Cohen’s kappa coefficient in the reference panel was also calculated (Collard et al. 2009).

Statistical analysis

For each category of SCT question (APP Original Context, APP New Context, ARC Original Context, and ARC New Context), the performances were determined using a modified scoring method (Vanbelle et al. 2007). The results were expressed as mean (\pm SD). The effect of experience (curriculum level) of the responders was tested using a two-way ANOVA, with activity (APP vs. ARC) and context (Original vs. New) as within factors. Planned comparisons were used to determine significant differences. The results were considered to be significant at the 5% probability level ($p < 0.05$).

4.3. Results

SCT scores

The mean scores of the five groups (students of year 3 to 6 and referents) for the 4 categories of questions are given in Table 3. The mean scores were significantly different among the 5 groups ($F(4,108) = 12.56$; $p < 0.0001$). More specifically, the reference panel showed higher scores ($p < 0.05$) than all groups of students; students in year 5 and 6 obtained higher scores ($p < 0.05$) than younger students (year 3 and 4). Significant differences were found based on the level in the curriculum and the type of seminar (APP/ARC) to which the question was related ($F(1,108) = 5.28$; $p < 0.05$). A significant difference was also found in relation to the context (Original/New) of the question ($F(1,108) = 7.97$; $p < 0.001$). Higher scores were unexpectedly obtained at ARC than at APP and in new contexts *versus* original ones, irrespective of the study level (Figure 1). Also, statistical analysis revealed a significant relation between the type of seminar (APP/ARC) and the context of the question (Original/New) ($F(1,112) = 5.65$; $p < 0.05$). Planned comparisons showed a statistically significant difference between ARC related questions in new context and the three other categories of questions ($p < 0.05$).

Table 3. Mean Score and Standard deviation (SD) in the different groups at the 4 categories of SCT questions

Group	APP Questions		ARC Questions		All questions
	Original Context	New Context	Original Context	New Context	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Year 3 (n=35)	0.55 (0.25)	0.58 (0.16)	0.54 (0.28)	0.61 (0.21)	0.57 (0.16)
Year 4 (n=20)	0.51 (0.23)	0.55 (0.17)	0.49 (0.24)	0.67 (0.17)	0.55 (0.15)
Year 5 (n=25)	0.61 (0.17)	0.64 (0.18)	0.69 (0.19)	0.73 (0.12)	0.67 (0.10)
Year 6 (n=24)	0.63 (0.18)	0.61 (0.17)	0.57 (0.27)	0.74 (0.14)	0.64 (0.11)
Reference Panel (n=9)	0.76 (0.12)	0.76 (0.11)	0.79 (0.10)	0.84 (0.08)	0.79 (0.05)

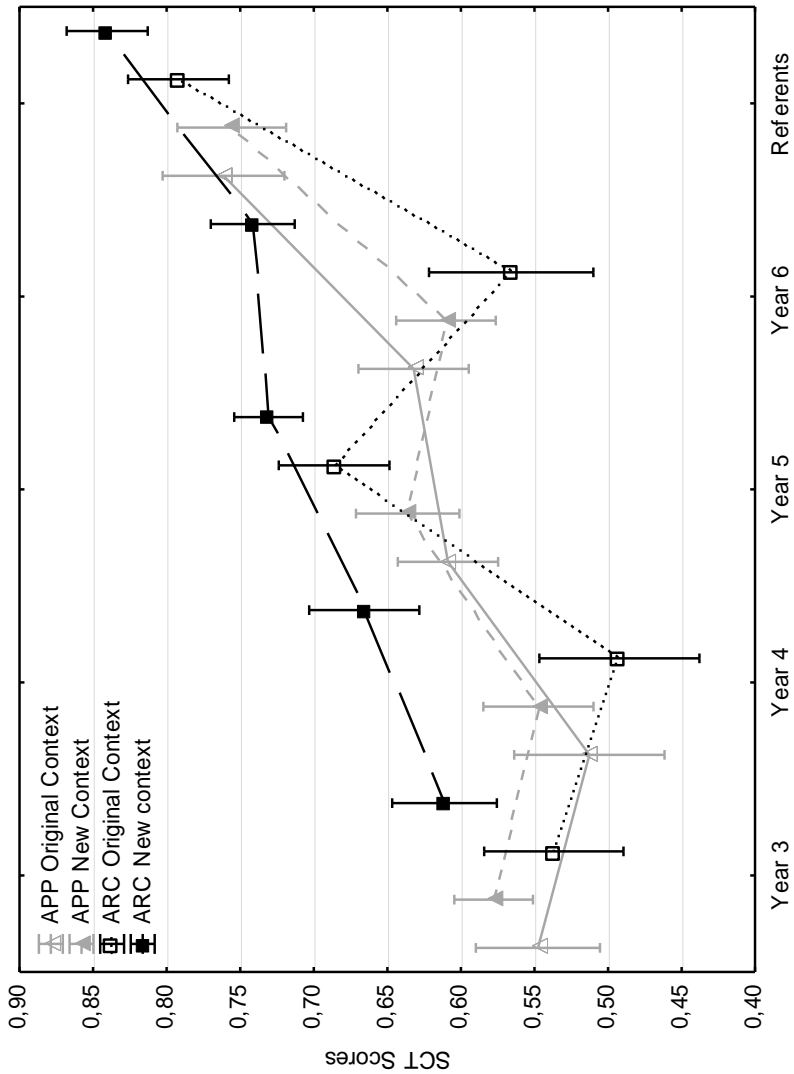


Figure 1. Changes in SCT Score (Mean \pm Standard Error) throughout the curriculum. The data are shown for APP and ARC questions and in each category, the scores are compared for questions in the original learning context and in a new context

4.4. Discussion

This study shows that as early as in the third year of the curriculum, the students who have experienced PBL seminars for one and a half year, have already good reasoning and transfer capacity. During the subsequent years, the SCT scores increase throughout the curriculum. The latter findings are consistent with the observations made later during training and showing an increase in SCT scores between post-graduate students and expert physicians (Caire et al. 2004; Charlin et al. 2000, 2002; Charlin & StJean, 2002; Charlin & Van Der Vleuten, 2004). We have reported previously the increase in reasoning maturation throughout the curriculum that was found to be correlated with knowledge base as well as the capacity of core knowledge self-estimation (Collard et al. 2009).

At each level in the curriculum, the students are able to transfer knowledge acquired during learning into a new context. A specific effect of PBL seminars could not be determined since we could not compare with students trained without PBL seminars. Norman (2009) has argued that substantial gain can be observed when students are engaging in multiple teaching examples. The students are exposed to multiple contexts in learning when it includes examples demonstrating wide application of what is being taught, like in a PBL curriculum. In such instances, they develop a flexible representation of knowledge and are likely to abstract the relevant features of concepts that make two unique problem scenarios similar (Gick & Holyoak, 1983; Spiro, Vispoel, Schmidt, Samarapungavan, & Boerger, 1987). To train transfer capacity, PBL could be a relevant method because like Norman stated: *to improve transfer, you have to link the concept learned to what they already know, use multiple examples to identify common deep structure, and use of mixed practice with multiple examples to focus on identifying when a concept applies*. According to Spiro, Feltovich, Jackson, & Coulton (1991) the difficulty to transfer concepts that students learn in schools to the real world is due to the over simplification of education material to make it easier to teach. Problem-based learning do not suffer from a lack of context or an oversimplification of content. Learning throughout

problem-based activities can enhance students' general learning transfer and problem-solving skills (Hmelo-Silver, 2004; Dixon & Brown, 2012).

Besides the expected findings that were described here above, some unexpected observations were made. Firstly, the scores at ARC questions were higher than at APP questions for the younger students (year 3 and 4) though they had never attended ARC seminars. Secondly, the scores obtained at questions in new context were higher than at questions in the original learning context. The highest scores were obtained for new context ARC questions whatever the group. Several explanations could be proposed. The ARC scenarios of SCT were written on the basis of clinical cases discussed in the ARC seminars centred on the diagnostic process while APP cases were constructed to meet the learning objectives of basic sciences. So, the realistic nature of ARC cases might result in greater activation of reasoning than APP cases.

The hypotheses proposed in SCT questions were, most of the time, different from the hypotheses discussed in PBL seminars. Moreover, the additional information to be evaluated was out of the original learning context for all questions. Thus, the lower scores obtained for original context questions could also involve possible interference created by new information in SCT questions. A proactive interference could be created by the first learning experience and entail difficulties for students when they analyse a familiar case with discordant elements compared to the learning context (Bernstein, Penner, Clarke-Stewart, & Roy, 2007). Transfer between tasks is a function of the similarity between the tasks and learning experiences (Dixon & Brown, 2012). Transfer is therefore affected by the context of the original learning; so processing new hypotheses and new additional information in an already known context (i.e. the original learning context) could be more difficult for students than treating new information in a new context without interference. With experience, the clinicians develop a more flexible representation of knowledge and are likely to abstract the relevant features of concepts. Thus they are less influenced with contextual surface element of a case. The interference phenomenon is therefore not so important for the tutor group.

In addition, even in a new context, if the responders generated a given hypothesis based on the SCT scenario while the SCT proposed another hypothesis less likely for the student, it could be difficult for them to ignore the hypothesis they generated when processing the new information. Considering these possible biases in writing test material, the SCT questions used in further research could be selected on the basis of realism or plausibility of both the SCT scenarios and hypotheses. A research design with think-aloud protocol or a writing justification of each response could provide an explanation of the features that influence reasoning and transfer capacity.

This study shows that since third year, students are able to transfer the knowledge acquired throughout PBL seminars for application to questions with information out of the learning context. Moreover, reasoning capacity is growing with experience, irrespective of the process of transfer.

4.5. References

- Barrows, H. (1986). The scope of clinical education. *Journal of Medical Education*, 61, 23-33.
- Bernstein, D., Penner, L.A., Clarke-Stewart, A., & Roy, E.J. (2007). *Psychology* (8th ed.). Boston: Cengage Learning. pp262-263.
- Boniver, J. (2004). Les études de médecine à l'Université de Liège : Le renouveau pédagogique de la Faculté de Médecine. *Revue Médicale de Liège*, 59, 717-730.
- Caire, F., Sol, J.C., Charlin, B., Isodori, P., & Moreau, J.J. (2004). Le test de concordance de script (TCS) comme outil d'évaluation formative des internes en neurochirurgie: Implantation du test sur Internet à l'échelle nationale. *Pédagogie Médicale*, 5, 87-94.
- Charlin, B., Gagnon, R., Sibert, R., Van der Vleuten, C. (2002). Le test de concordance de script, un instrument d'évaluation du raisonnement clinique. *Pédagogie Médicale*, 3, 135-144.
- Charlin, B., Roy, L., Brailovsky, C., Goulet, F., & Van der Vleuten, C. (2000). The script concordance test, a tool to assess the reflective clinician. *Teaching and Learning in Medicine*, 12, 189-195.
- Charlin, B., & St-Jean, M. (2002, Mai). Le test de concordance de script: Un outil pour évaluer le jugement en médecine. *Bulletin du CEFES, Université de Montréal*, 6, 4-5.
- Charlin, B., Tardif, J., & Boshuizen, P. (2000). Scripts and medical diagnosis knowledge: Theory and applications for clinical reasoning instruction and research. *Academic Medicine*, 75, 182-190.
- Charlin, B., & Van der Vleuten, C. Standardized assessment in context of uncertainty: The script concordance approach. *Evaluation in the Health Profession*, 27, 304-319.

- Chi, M., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121-152.
- Chi, M., Glaser, R., & Farr, M. (1988). *The nature of expertise*. Hillsdale, J.J.: Erlbaum.
- Cohen, J. (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement of partial credit. *Psychological bulletin*, *70*, 213-220.
- Collard, A., Gelaes, S., Vanbelle, S., Bredart, S., Defraigne, J.O., Boniver, J., & Bourguignon, J.P. (2009). Reasoning versus knowledge retention and ascertainment throughout a problem-based learning curriculum. *Medical Education*, *43*, 854-865.
- Custers, E., Boshuizen, H., & Schmidt, H. (1998). The role of illness scripts in the development of medical diagnostic expertise: Results from an interview study. *Cognition and Instruction*, *16*, 367-398.
- Custers, E., Boshuizen, H., & Schmidt, H. (1996). The influence of medical expertise, case typicality, and illness scripts component on case processing and disease probability estimates. *Memory and Cognition*, *24*, 384-399.
- DeGroot, A. (1965). *Tought and action in chess*. New York: Basic Books.
- Dixon, R.A., & Brown, R.A. (2012). Transfer of learning: connecting concepts during problem solving. *Journal of technology education*, *24*, 2-17.
- Elstein, A.S., Shulman, L.S., & Sprafka, S. (1978). *Medical problem solving*. Cambridge: Harvard University Press.
- Eva, K., Neville, A., & Norman, G. (1998). Exploring the aetiology of content specificity: Factors influencing analogic transfer and problem solving. *Academic Medicine*, *73*, 1-5.

Feltovitch, P., & Barrows, H. (1984). Issues of generality in medical problem solving. In H. Schmidt & M. De Volder (Eds.), *Tutorials in problem-based learning: A new direction in teaching the health Professions* (pp.128-142). Assen: Van Gorcum.

Galloti, K.M. (1989). Approaches to studying formal and everyday reasoning. *Psychological Bulletin*, *105*, 331-351.

Gick, M., & Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1-38.

Glaser, R. (1987). Learning theories and theories of knowledge. In E. de Corte, H. Lodewijks, R. Parmentier, & P. Span (Eds.), *Learning and Instruction vol. 1* (pp. 397-413). Oxford: Pergamon Press.

Haskell, R.E. (2001). *Transfer of Learning: Cognition, Instruction and Reasoning*. San Diego: Academic Press.

Hmelo-Silver, C.E. (2004). Problem-Based Learning: What and How Do Students Learn? *Educational Psychology Review*, *16*, 235-266.

Hobus, P., Boshuizen, H., & Schmidt, H. (1990). *Expert-novice differences in the mental representations of patient prototypes*. Paper presented at the 2nd International Symposium on the Problem-based Learning. Yogyakarta, Indonesia.

Hobus, P., Hofstra, M., Boshuizen, H., & Schmidt, H. (1988). De context van de klacht als diagnosticum (The context of the complaint as a diagnostic tool). *Huisarts en Wetenschap*, *31*, 261-264.

Hobus, P., Schmidt, H., Boshuizen, H., & Patel, V. (1987). Contextual factors in the activation of the first diagnostic hypotheses: Expert-novice differences. *Medical Education*, *21*, 471-476.

Hofstra, M., Hobus, P., Boshuizen, H., & Schmidt, H. (1988). De invloed van ervaring op diagnostische prestaties van huisartsen (The influence of experience on diagnostic accuracy of family physicians). *Huisarts en Wetenschap*, *31*, 282-284.

Johnson, S.D. (1995). Transfer of learning. *The technology teacher*, 54, 33-35.

Johnson, S.D., Dixon, R., Daugherty, J., & Lawanto, O. (2011). General versus specific intellectual competencies: the question of learning transfer. In M. Barack & M. Hacker (Eds.), *Fostering human development through engineering and technology education* (pp. 55-74). Netherlands: Sense Publishers.

Lesgold, A., Rubinson, H., Feltovitch, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures. In M. Chi, R. Glaser, & M. Farr, (Eds.), *The Nature of Expertise* (pp. 311-342). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

McGeoch, J. (1932). Forgetting and the law of disuse. *Psychological Review*, 39, 352-370.

Newell, A., & Simon, H. (1972). *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall.

Norman, G. (2009). Teaching basic science to optimize transfer. *Medical teacher*, 31, 807-811.

Norman, G., Dore, K., Krebs, J., & Neville, A.J. (2007). The power of the plural: effect of conceptual analogies on successful transfer. *Academic Medicine*, 82, s16-s18.

Norman, G.R., & Schmidt, H.G. (1992). The psychological basis of Problem-based Learning: a review of the evidence. *Academic Medicine*, 67, 557-565.

Regehr, G., & Norman, G.R. (1996). Issues in cognitive psychology: implications for professional education. *Academic medicine*, 71, 988-1001.

Schmidt, H.G. (1983). Problem-Based Learning: rationale and description. *Medical Education*, 17, 11-16.

Schmidt, H., Norman, G., & Boshuizen, H. (1990). A cognitive perspective on medical expertise: Theory and implication. *Academic Medicine*, 65, 611-621.

Smith, S. (1988). Environmental context-dependent memory. In G. Davies & D. Thomson (Eds.), *Memory in context: Context in memory* (pp. 13-43). New York: John Wiley & Sons.

Smith, S., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin and Review*, 8, 203-220.

Spiro, R.J., Feltovitch, P.L., Jackson, M.L., & Coulton, R.L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domain. *Educational Technology*, 31(5), 24-33.

Spiro, R.J., Vispoel, W.P., Schmitz, J.P., Samarapungavan, A., & Boerger, A.E. (1987). Knowledge acquisition for application: cognitive flexibility and transfer in complex domains. In B.K. Britton & S.M. Glynn (Eds.), *Executive control processes in reading* (pp.177-199). Hillsdale, N.J.: Erlbaum.

Sutton, M.J. (2003). Problem representation, understanding, and learning transfer: implications for technology education research. *Journal of Industrial Teacher Education*, 40, 47-61.

Tulving, E., & Thomsons, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), 352-373

Vanbelle, S., Massart, V., Giet, D., & Albert, A. (2007). Test de concordance de script: un nouveau mode d'établissement des scores limitant l'effet du hasard. *Pédagogie médicale*, 8, 71-80.

Van Gessel, E., Nendaz, M.R., Vermeulen, B., Junod, A., & Vu, N.V. (2003). Development of clinical reasoning from the basic sciences to the clerkships: a longitudinal assessment of medical student's needs and self-perception after a transitional learning unit. *Medical Education*, 37, 966-974.

5. ARTICLE 2 : REASONING VERSUS KNOWLEDGE RETENTION AND ASCERTAINMENT THROUGHOUT A PROBLEM-BASED LEARNING CURRICULUM

Anne Collard, Sabine Gelaes, Sophie Vanbelle, Serge Bredart, Jean-Olivier Defraigne, Jacques Boniver & Jean-Pierre Bourguignon

Publié dans *Medical Education* 2009: 43: 854–865

ABSTRACT

Context: Since 2000, problem-based learning (PBL) seminars have been introduced into the curriculum of medical studies at the University of Liège. We aimed to carry out a cross-sectional investigation of the maturational increase in biomedical reasoning capacity in comparison with factual knowledge retention throughout the curriculum.

Methods: We administered a factual knowledge test (i.e. a true/false test with ascertainment degree) and a biomedical reasoning test (i.e. an adapted script concordance test [SCT]) to 104 students (Years 3–6) and a reference panel. The selected topic was endocrinology.

Results: On the SCT, the students obtained higher scores in Years 5 and 6 than in Years 3 and 4. In Year 3, the scores obtained on SCT questions in a new context indicated transfer of reasoning skills. On the true/false test, the scores of Year 3 students were significantly higher than those of students in the other three year groups. A positive correlation between SCT scores and true/false test scores was observed only for students in Years 3 and 4. In each group, the ascertainment degree scores were higher for correct than for incorrect responses and the difference was calculated as an index of self-estimation of core knowledge. This index was found to be positively correlated to SCT scores in the four year groups studied.

Conclusions: Biomedical reasoning skills are evidenced early in a curriculum involving PBL and further increase during training. This is accompanied by a decrease in factual knowledge retention. The self-estimation of core knowledge appears to be related to reasoning capacity, which suggests there is a link between the two processes.

5.1. Introduction

In addition to their necessary learning of a large body of knowledge, medical students must develop clinical reasoning skills. Clinical reasoning has mainly been studied in the context of medical diagnosis, which is considered to involve a problem-solving process. The mental representation of a problem is a key issue in that process. For Feltovich and Barrows (1984), the mental representation of a diagnostic problem takes the form of an (instantiated) “illness script” into which patient characteristics, signs and symptoms of the disease, and knowledge on the underlying links are organized. Illness scripts may also contain memories of previous patient experiences which help in the diagnosis of new cases (Schmidt, Norman & Boshuizen, 1990). According to Schmidt et al. (1990), illness scripts are developed through continuous exposure to patients and thus result from experience. These scripts are used by experienced doctors when diagnosing routine cases. Other authors claim that the first step in the process of clinical reasoning is the early generation of hypotheses, which are then tested and validated through an active process of acquisition of relevant data (Barrows, Norman, Neufeld & Feghter, 1982; Bordage, 1994; Elstein, 1994; Schmidt & Boshuizen, 1993; Van Gessel, Nendaz, Vermeulen, Junod, & Vu 2003). Although these mechanisms of clinical reasoning are thought to operate mainly by the end of medical studies, when they have been evidenced, we hypothesized that reasoning may be active earlier in the curriculum and may mature subsequently.

Findings from studies on clinical reasoning have influenced educational strategies in order to promote reasoning and its assessment. Problem-based learning (PBL) was developed to facilitate the student’s process of acquisition, organization and retrieval of knowledge (Norman & Schmidt, 1992; Regher & Norman, 1996; Schmidt, 1983; Van Gessel et al. 2003). One aim of PBL is to help the student integrate new information in a rich and connected knowledge network that will be activated later (illness scripts) (Barrows, 1986). Through such mechanisms, PBL promotes clinical reasoning skills (Barrows, 1986). Since 2000, the Faculty of Medicine at the University of Liège has

integrated PBL seminars into Years 2–7 of a 7-year curriculum (Boniver, 2004). The present study aimed to investigate the maturation of reasoning capacity throughout a medical curriculum.

Some authors have attempted to analyze the relationship between the knowledge base and the maturation of clinical reasoning at different levels of expertise by using a think-aloud protocol applied during clinical problem-solving tasks (Boshuizen & Schmidt, 1992; Patel & Groen, 1986). Patel and Groen (1986) have shown that experts use less basic science in their reasoning explanations than novices. Boshuizen and Schmidt (1992) demonstrated that the application of biomedical knowledge (use of basic knowledge to explain the process underlying the case) decreased with increasing levels of expertise. They proposed a three-stage model of expertise development in medicine, consisting of the acquisition of biomedical knowledge, practical experience and the integration of both theoretical and experiential knowledge resulting in knowledge encapsulation (Schmidt & Boshuizen, 1993; Schmidt & Rikers, 2007). There are, however, few studies addressing the retention of factual knowledge over time. Some authors have reported decreasing scores for recognition and recall tasks, together with increasing scores for application and interpretation tasks (Semb & Ellis, 1994).

In the present study, investigation tools involved a Script Concordance Test (SCT) (Caire, Sol, Charlin, Isodori & Moreau, 2004; Charlin, Gagnon, Sibert & Van Der Vleuten, 2002; Charlin, Roy, Bralovsky, Goulet, & Van Der Vleuten, 2000; Charlin & StJean, 2002; Charlin & Van Der Vleuten, 2004) to assess biomedical and clinical reasoning and a true / false test to assess factual knowledge. The maturation of reasoning capacity was compared with knowledge retention. We hypothesized that retention of factual knowledge in the pre-clinical sciences might be highest in Year 3 and might decrease afterwards, whereas reasoning capacity might be lowest in Year 3 and might further increase throughout the curriculum. Degree of ascertainment is used as part of routine procedure for multiple choice question and true/false evaluations in our institution (Leclercq, 1975, 1982; Leclercq, Boxus, De

Brogniez, Lambert, & Wuidar, 1993). The ascertainment degree measures how sure the responder is about each of his or her answers and, thus, gives an indication of the subject's self-estimation of what he or she knows and does not know (Leclercq, 1975, 1982; Leclercq et al. 1993). The relationship between self-estimation of core knowledge (through the ascertainment degree on the true/false test) and reasoning was investigated.

5.2. Methods

Participants

During the academic year 2006–2007, 104 volunteer students from Years 3–6 in the Faculty of Medicine at the University of Liège participated in the study (a minimum of 20 students in each study year). Endocrinology was the selected topic of the investigations. The students differed in terms of the PBL seminars and clerkship periods they had undertaken (Table 1). In Years 2 and 3, the majority of learning occurs in PBL seminars (named APP, *Apprentissage par Problèmes*) (Barrows & Tamblyn, 1980; Boniver, 2004), which focus on the structure and functioning of the human body systems. In Years 4 and 5, the diagnostic and therapeutic bases of diseases of the different systems are learned through lectures and PBL seminars (named ARC, *Apprentissage du Raisonnement Clinique*) (Barrows & Tamblyn, 1980; Boniver, 2004). There was no difference among the groups of volunteer students in terms of their global academic achievement (assessed on the basis of Year 3 notes). Likewise, volunteers did not differ from non-participants, except for in Year 5, in which volunteer notes were of significantly better quality than those of the rest of the Year 5 cohort. Written informed consent was obtained from the participants and tests were completed anonymously. Confidentiality and anonymity were guaranteed. No student names were used in the statistical analysis. Each participant was assigned a number with which he or she was able to check his or her scores online and compare them with the whole-group scores.

Two teachers wrote the testing materials. Nine tutors (three in academic appointments and six in clinical appointments at the university

hospital), who handled the PBL seminars on the endocrine system, served as a reference panel for SCT scoring (see below). They also provided answers to the true/false test. Written informed consent was obtained from the tutors, who completed the tests anonymously.

Table 1. Characteristics of the participants

Year of the curriculum	3rd		4th		5 th		6 th	
	Volunteers	Non Participants	Volunteers	Non Participants	Volunteers	Non Participants	Volunteers	Non Participants
N	35	79	20	77	25	71	24	68
Gender F/M	24/11	51/28	10/10	55/22	13/12	46/25	13/11	33/55
Mean score Year 3	13.9	13.4	14.4	13.7	14.5*	13.3	14.1	13.5
Completed seminars (endocrine system) at the time of study	APP		APP		APP and ARC		APP and ARC	
Lectures on diagnosis and treatment of endocrine disorders	-		-		Done		Done	
Completed clerkship (all disciplines) at the time of study	-		Half-time 4 months		Half-time 12 months		Half-time 12 months and full time 4 months	
Reference Panel	9 tutors involved in APP and ARC seminars about the endocrine system							

* p<0.05 versus non participants year 5

Test materials

Clinical reasoning test: the script concordance test

Originally, the SCT (Caire et al. 2004; Charlin, Gagnon, Sibert & Van Der Vleuten, 2002; Charlin et al. 2000; Charlin & StJean, 2002; Charlin & Van Der Vleuten, 2004) proposed a series of short and incomplete clinical scenarios, each of which was followed by either a diagnostic hypothesis or a proposal for the therapeutic management of the patient. When new information was provided, the responders had to assess the impact of this information on the basis of the initial diagnostic or management hypothesis. Responses were given on a 5-point Likert scale. Because the incomplete clinical scenario allowed different answer options, the response was rarely unequivocal. Each individual response was credited by a score determined on the basis of the distribution of responses given by a reference panel (Charlin, Gagnon, Sibert & Van Der Vleuten, 2002; Charlin et al. 2000; Charlin, Tardif & Boshuizen, 2000).

Starting from the initial SCT format, we adapted the test to assess skills of biomedical reasoning in students not yet familiar with the process of clinical diagnosis. The SCT had not previously been used for this purpose. The present study was based on 48 SCT questions with scenarios inspired by the clinical vignettes used in the PBL seminars (APP and ARC). The SCT scenarios were based on clinical cases taken from PBL seminars on endocrinology: five cases were sourced from Year 3 seminars (APP) and five cases from Year 4 seminars (ARC). For each clinical case, two SCT scenarios which aimed at the same learning objectives were created, either in the original learning context or in a new context (in which case characteristics differed from those of the learning clinical cases). For each scenario, three independent hypotheses were proposed, each followed by newly given information. A single hypothesis could be used three times, as in the example provided in Table 2. The task was to evaluate the impact of new information on the initial hypothesis. From a total of 60 administered questions, 12 were discarded by the reference panel because the scenario was misunderstood as a result of unclear writing.

Table 2. Examples of SCT questions in a new context

A 50 years old premenopausal woman shows up at a routine visit in the department of occupational medicine. Her BMI is 28; she is sedentary. Glycosuria is found at screening urine analysis.			
<i>If the hypothesis is</i>	<i>And you know that</i>	<i>The hypothesis becomes</i>	<i>Answer of refer.</i>
Glycosuria results from early stage type 2 diabetes	The patient takes medication for hypercholesterolemia	Much less likely	
		Less likely	3
		Not affected	6
		More likely	
		Much more likely	
Glycosuria results from early stage type 2 diabetes	Together with a fasting blood sugar of 126 mg/dl, plasma insulin level was high	Much less likely	
		Less likely	
		Not affected	
		More likely	2
		Much more likely	7
Glycosuria results from early stage type 2 diabetes	She was fasting at the time of urine collection	Much less likely	
		Less likely	1
		Not affected	1
		More likely	3
		Much more likely	4

Factual knowledge test: true / false test

A 120-question true/false test was elaborated. No clinical scenario was presented in vignette in this test format. The questions explored factual knowledge through statements. They were related to the same topics as the SCT, which enabled the comparison between the maturation of biomedical reasoning and the retention of factual knowledge during medical studies. Sample questions are provided in Table 3.

For each true/false test answer, an ascertainment degree (Leclercq, 1975, 1982; Leclercq et al. 1993) was requested to measure how sure the participant was of his or her response. The ascertainment degree was given on a 6-point Likert scale (Table 3) ranging from 0 (the response has a probability of 0–25% of being correct) to 5 (the response has a probability of 95–100% of being correct). This validated scale (Leclercq, 1975, 1982; Leclercq et al. 1993) has been used for several years in the certificating evaluation at the Faculty of Medicine at Liège.

For each responder, the mean ascertainment degree was calculated separately for correct answers versus incorrect answers. The difference between the mean ascertainment degrees (correct–incorrect answers) was used as an estimate of the capacity for self-discrimination between known and unknown material.

Table 3. Examples of True/False Test questions

Questions :	T/F	Ascertainment Degree												
1. In the adipose tissue, insulin activates fatty acid synthesis		<table border="1"> <tr> <td>0</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>0%</td> <td>25%</td> <td>50%</td> <td>70%</td> <td>85%</td> <td>95%100%</td> </tr> </table>	0	1	2	3	4	5	0%	25%	50%	70%	85%	95%100%
0	1	2	3	4	5									
0%	25%	50%	70%	85%	95%100%									
2. Insulin causes increase in plasma free fatty acid concentrations		<table border="1"> <tr> <td>0</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>0%</td> <td>25%</td> <td>50%</td> <td>70%</td> <td>85%</td> <td>95%100%</td> </tr> </table>	0	1	2	3	4	5	0%	25%	50%	70%	85%	95%100%
0	1	2	3	4	5									
0%	25%	50%	70%	85%	95%100%									
3. The insulin receptor has enhanced affinity for insulin in type 2 diabetes		<table border="1"> <tr> <td>0</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>0%</td> <td>25%</td> <td>50%</td> <td>70%</td> <td>85%</td> <td>95%100%</td> </tr> </table>	0	1	2	3	4	5	0%	25%	50%	70%	85%	95%100%
0	1	2	3	4	5									
0%	25%	50%	70%	85%	95%100%									

Test administration

The tests were administered at the beginning of the second quarter of the academic year (February 2007). All the students were gathered to complete the tests at the same time. The reference panel completed the tests before the students. Each student and tutor started with the SCT, which was followed by the true/false test. There was no time limit. The tests were completed individually. We asked the participants to answer all questions on both tests. We created three different versions of each test in which the questions appeared in a pseudo-random order.

Scoring process on the SCT

Responses to SCT questions were marked using a modified scoring method (Vanbelle, Massart, Giet, & Albert, 2007). The original scoring process used by Charlin and collaborators was “the aggregate scoring process” (Charlin, Gagnon, Sibert, & Van Der Vleuten, 2002; Charlin & Van Der Vleuten, 2004). The score attributed to the responder for each item of a question was based on the proportion of panel

members who had selected the same answer. The maximum score was 1 for the modal answer. The number of panel members who provided an answer on the Likert scale was divided by the number of experts in the modal value for the item. To give an example, we refer to the referee responses to the sample question illustrated in Table 2. On the last item, the highest number (four of nine) of referees responded with “much more likely” on the Likert scale. This option was thus credited with the maximum score of 1 point (4/4 [i.e. the modal response]). Then, as three referees had selected “more likely” as a response, this option was credited with 0.75 points (3/4). The answers “not affected” and “less likely” were credited with 0.25 points (1/4) each. The answer(s) not selected by any of the panel members (“much less likely” in this example) received no credit, whether they were close to or far from the modal answer.

Here, the scoring method used by Charlin et al. was adapted using the kappa coefficients methodology (Vanbelle, Massart, Giet, & Albert, 2007) to account for scores obtained fortuitously. A weighted kappa coefficient (Cohen, 1968) quantified the agreement between two raters on an ordinal scale and corrected for agreement resulting from chance. This methodology was extended to quantify the agreement between the reference panel and a responder. The weights permitted higher scores to be assigned to responses closer to the modal response(s) given by the reference panel. The weights were derived from the reference panel responses by computing the frequency distribution of the number of category differences separating two experts’ responses for all possible pairs of experts. This led to a weighting of 1.00 for a category of the Likert scale consistent with the modal response of the reference panel, a weighting of 0.96 for the adjacent categories and weightings of 0.19, 0.06 and 0.00 at distances of two, three and four categories, respectively. Cohen’s kappa coefficient for the reference panel was also calculated.

Statistical analysis

The results on the SCT and true/false tests were expressed as mean (\pm standard deviation [SD]). On the SCT, the effect of experience (curriculum level) of the responders was tested using a one-way ANOVA.

On the true/false test, Cronbach's alpha was used to assess reliability and performances throughout the curriculum were analyzed using a one-way ANOVA. The ascertainment degree was analyzed using a repeated-measures ANOVA with the correctness of answer as a within-subjects factor. Subsequent a priori planned comparisons were used to determine significant differences. Pearson's correlations were used to characterize the association between SCT and true/false test scores and between SCT and the self-discrimination of knowledge (estimated by the difference between ascertainment degrees for correct and incorrect answers on the true/false test). Missing data were kept as such. The results were considered to be significant at the 5% probability level ($p < 0.05$).

5.3. Results

SCT scores

The mean scores of the five groups are given in Table 4. The mean scores differed significantly among the five groups ($F(4,108) = 12.56, p < 0.0001$). More specifically, the reference panel showed higher scores than all groups of students; students in Years 5 and 6 obtained higher scores than students in earlier years (Years 3 and 4), with a reduced scattering of the data as indicated by the SDs. As shown in Figure 1, the maturational increase in SCT scores was observed similarly using questions related to PBL seminars run in Year 3 (APP) and in Years 4 and 5 (ARC). It was noteworthy that students in Years 3 and 4 performed better on ARC than APP questions, although the former had not yet attended any ARC seminars and the latter had not yet attended endocrinology ARC seminars.

Table 4. Script Concordance Test score, True/False Test score and Ascertainment Degree in relation to study level in a medical curriculum

Group (N)	SCT		True/False Test					
	Score		Score		Ascertainment degree (%)			
	Mean	SD	Mean	SD	Correct answer		Incorrect answer	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Year 3 (35)	0.57	0.16	75.8 ^c	8.22	71.1 ^d	11.9	60.3	14.9
Year 4 (20)	0.55	0.15	65.6	6.16	57.8 ^d	15.1	49.8	13.7
Year 5 (25)	0.67 ^b	0.10	68.6	3.90	68.2 ^d	10.1	58.4	11.3
Year 6 (24)	0.64 ^b	0.11	66.3	6.29	58.2 ^d	12.6	48.6	13.4
Tutors (9)	0.79 ^a	0.05	72.6	7.46	68.7 ^d	15.6	55.1	17.1

SCT score are Cohen's kappa scoring the degree of agreement between the individual responder and the whole panel of tutors in the different groups. They vary between -1 and +1.

T/F data included the test score and the ascertainment degree for correct and incorrect answers. True/false Test score vary between 0 and 100.

^a p<0.05 versus year 3 to 6

^b p<0.05 versus year 3 and 4

^c p<0.05 versus year 4 to 6

^d p<0.05 versus ascertainment degree for incorrect answer

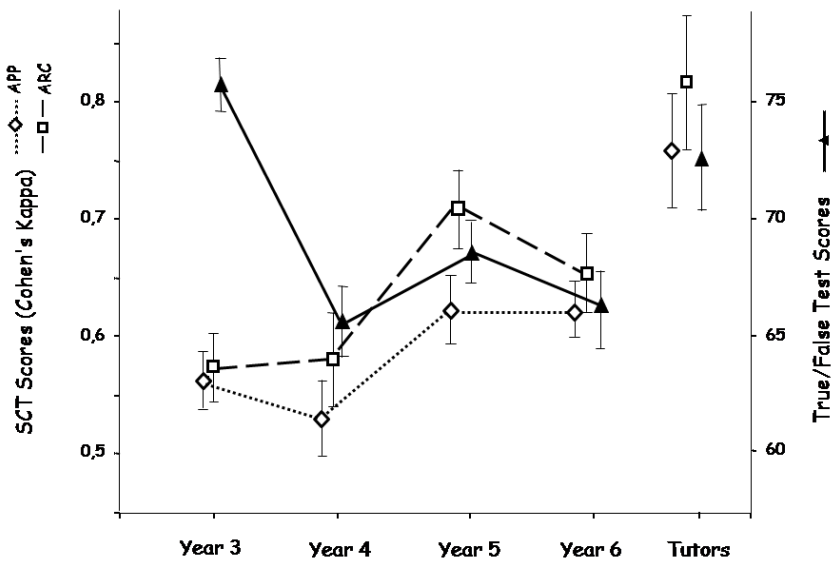


Figure 1. Changes in reasoning skills (evaluated through Kappa of Cohen at Script Concordance Test, SCT) and knowledge retention (evaluated through scores at True/False test) throughout a medical curriculum (3rd – 6th year) in comparison with a panel of tutors. The SCT data are shown separately for questions related to APP seminars (run in 3rd year) and ARC seminars (run in 4th and 5th years)

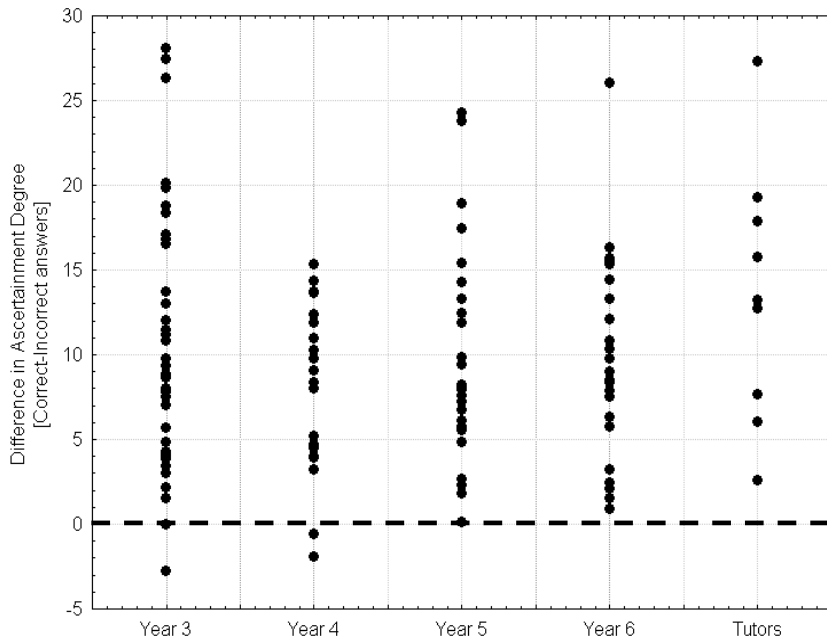


Figure 2. Individual differences between Ascertainment Degrees provided for correct answers and incorrect answers at four stages in the medical curriculum and in tutor

True/false test scores and ascertainment degree

Cronbach’s alpha coefficient was 0.79 for the True/false test, indicating acceptable reliability.

The mean scores of the five groups are given in Table 4 and shown in Figure 1. Results of the true/false test differed among the groups ($F(4,108) = 10.54, p < 0.0001$). The scores of Year 3 students were significantly higher ($p < 0.05$) than the scores obtained by all other student groups (Years 4–6). Moreover, the scores of the reference panel did not significantly differ from those of Year 3 students, but were higher than those of students in Years 4 and 6.

Ascertainment degree analysis showed significantly higher scores for correct than for incorrect responses ($F(1,108) = 224.11, p < 0.0001$), whatever the group of participants. The individual differences in ascertainment degree between correct and incorrect answers are shown in Figure 2. The data show similar values and distributions in the four student groups, as well as among the tutors. No significant correlation

was found in any group between ascertainment degree for correct answers and the difference between ascertainment degrees for correct and incorrect answers (data not shown).

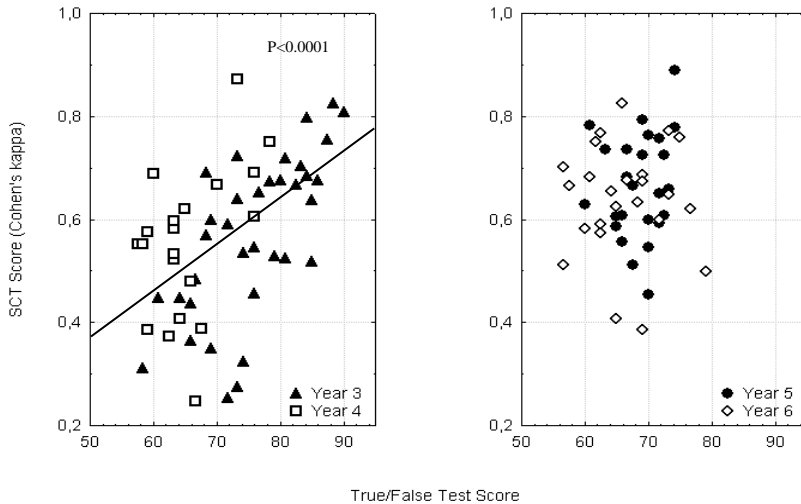


Figure 3. Correlation between True/False Test scores and SCT scores compared in two groups of medical students at different stages in the curriculum

SCT scores versus true/false test scores and ascertainment degree

Although the SCT scores increased during the medical education curriculum, the true/false test scores decreased after Year 3 (Table 4, Figure 1). Because the differences in SCT scores were not significant between Years 3 and 4, and Years 5 and 6, respectively, these groups were combined for the correlation analysis. A significant positive correlation (Figure 3) between SCT scores and true/false test scores was observed for students in Years 3 and 4 ($r = 0.53$, $p < 0.0001$), but not for those in Years 5 and 6 ($r = 0.07$, $p = 0.64$). A significant positive correlation between SCT scores and differences in ascertainment degrees for correct and incorrect answers (Figure 4) was observed for all students (Years 3 and 4: $r = 0.37$, $p = 0.006$; Years 5 and 6: $r = 0.32$, $p = 0.03$).

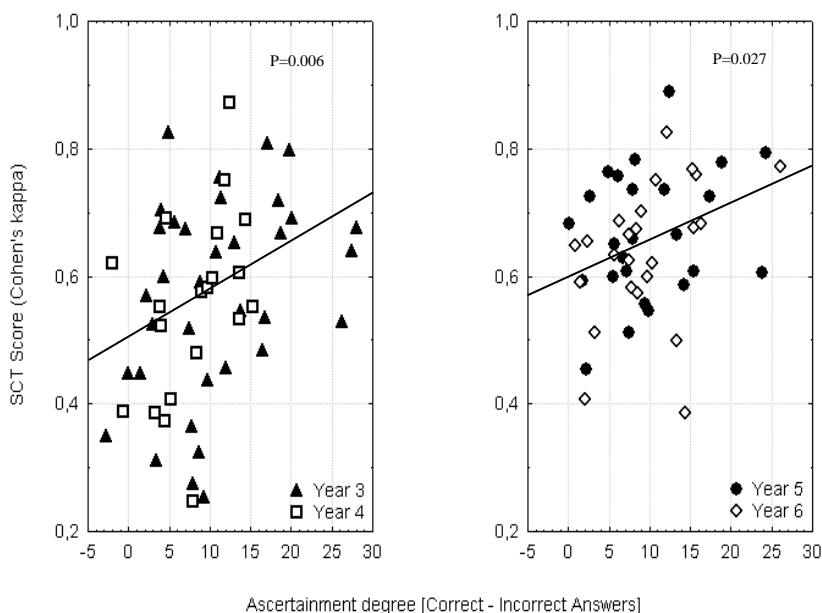


Figure 4. Correlation between SCT scores and differences between Ascertainment Degrees for correct answers - incorrect answers at True/False Test compared in two groups of medical students at different stages in the curriculum

5.4. Discussion

It appears from the present study that the SCT can be adapted to assess biomedical and clinical reasoning at different stages in a medical curriculum. There are several validation issues that must be addressed. Validity of the content refers to consistency with teaching content, which was ensured by involving two teachers in the writing of the test. Validation for clarity of writing would have required pre-validation of the SCT questions with tutors and students. This, however, would have resulted in the exclusion of those test persons from an already limited group of participants. Therefore, we elected to first administer the whole test to both the students and tutors and then, in a second process, to review the questions for clarity and consistency with the panel of tutors. Following this process, 12 of the 60 questions were

invalidated and were not included in the analysis. The significance of our adapted SCT for assessment of reasoning skills was unfortunately not validated directly. We thought about making a possible validation through correlating the SCT scores with notes obtained at a certifying examination which involved a pathophysiological discussion of a short clinical scenario. However, this was incompatible with the anonymous and non-certifying nature of our study.

Although this is the first time that SCT scores have been shown to increase in the course of a PBL medical education, our findings are consistent with observations made later during training, which show that SCT scores of expert doctors are increased over those of postgraduate students (Caire et al. 2004; Charlin, Gagnon, Sibert & Van Der Vleuten, 2002; Charlin et al. 2000; Charlin & StJean, 2002; Charlin & Van Der Vleuten, 2004). It is of note that the reasoning skills of Year 3 students were already obvious and could be applied in the context of ARC related scenarios to which those students had not been previously exposed. Such a finding suggests that reasoning skills can be transferred to a context that differs from that in which they were learned. The SCT results show a significant maturational increase in terms of biomedical and clinical reasoning capacity between the two earlier and two later year groups studied. These outcomes confirm previous findings showing that gaining experience and being exposed to patients contributes to the development of clinical reasoning (Schmidt, Norman, & Boshuizen, 1990). This is consistent with the substantial increase in the time periods our students spend in clerkships during Years 5 and 6 in comparison with Years 3 and 4. However, evaluation of the possible contributions of PBL seminars or clerkships or any other learning activity in this maturation process was not the aim of the present study. Because our training curriculum uses a mixed system which combines PBL seminars and other modalities, including formal lectures, it would be almost impossible to establish whether the skills observed on the SCT are related to a specific part of or change in training. Rather, it is fair to state that the introduction of PBL seminars into the curriculum was a reason for teachers and students to promote reasoning skills and for us, as investigators, to attempt an evaluation of those skills.

A strength of this study concerns the original application of a single reasoning test throughout the medical curriculum in a cross-sectional sample of students, all of whom experienced similar learning conditions, including PBL seminars. A longitudinal prospective study would carry some added value as it would result in increased homogeneity and comparability within the student sample. However, in such conditions, different tests bearing on different topics would be administered, thus diminishing their comparability. In addition, enhancements in student capacity as the curriculum progresses, which reflect students' increasing familiarity with a new testing procedure, may bias a longitudinal study. Finally, the number of student volunteers may decrease with time and invalidate the study through an insufficient number of participants.

Results on the true/false test revealed that student scores decreased in Years 4–6 of the curriculum, compared with those of Year 3 students, which were not significantly different from those of the reference panel. The higher scores obtained by Year 3 students may be explained by the fact that these students had just sat an examination on the topic involved in the test and were still easily able to retrieve the knowledge acquired from their memory. For students in the later year groups, this knowledge might be more difficult to retrieve because a longer period had elapsed since it had been learned and other unrelated knowledge had been processed in the interim. Moreover, organization of the knowledge network may mature less rapidly than acquisition of knowledge, which would account for a lower yield of reactivation, particularly in a test in which contextualized information is not triggered. The knowledge networks of the reference panel may have been better organized and more easily activated because the panel consisted of specialists in the study domain who processed the information in their daily practice. Expert involvement in the clinical practice of endocrinology was more likely to account for the higher scores on the true/false test than global expertise independent of the study domain. These observations are consistent with the data reported by Boshuizen and Schmidt (1992) and with their encapsulation theory, which proposes that basic knowledge is still available even if experts do not process basic

knowledge in itself when they make a diagnosis (Boshuizen & Schmidt, 1992; Rikers, Loyens, Winkel, Schmidt, & Sins, 2005; Rikers, Schmidt, & Moulaert, 2005; Schmidt & Rikers, 2007). In addition, some members of the reference panel taught endocrinology and therefore regularly reactivated and manipulated knowledge in this particular domain. Presumably, the true/false test study of performances of expert doctors in a domain other than endocrinology would result in a performance level similar to or lower than that of Year 6 students.

The increasing scores on the biomedical reasoning test and the decreasing scores on the factual knowledge test throughout the medical curriculum were in line with results of studies comparing performance in remembering tasks and application tasks (Semb & Ellis, 1994). In some studies evaluating the maturation of clinical reasoning and knowledge, both reasoning and knowledge were found to increase during the curriculum. In this particular case, however, knowledge was evaluated in a contextualized diagnostic processing (Boshuizen, Van Der Vleuten, Schmidt & Machiel-Bongaerts, 1997; Verhoeven, Verwijn, Scherpbier, & Van Der Vleuten, 2002) (based on clinical scenarios). By contrast, the retention of knowledge acquired in the first years of training was studied here through non-contextualized questions. Thus, the discrepancy between the two studies might suggest the facilitating impact of context on knowledge retrieval. This hypothesis could be substantiated in a study addressing this issue specifically, bearing in mind that the effects of context on knowledge retrieval from the memory network should be differentiated from the possible role played by context at the time of knowledge acquisition (i.e. during a learning process, such as PBL seminars).

A positive correlation between scores on the SCT and the true/false test was observed for students in the earlier years (Years 3 and 4), but not for those in the later years. Such correlations suggest that, in students at the beginning of a PBL curriculum, clinical reasoning skills are directly linked with level of background knowledge. The absence of any significant correlation in students in the later years may indicate that a relative independence of factual knowledge and clinical

reasoning has developed with experience. Throughout maturation, reasoning may become more related to experience and matching with known cases, whereas knowledge is not influenced in the same manner. According to Patel, Glaser, and Arocha (2000), expert performance is not a result of generally superior memory skills. It is, rather, a function of a well-organized knowledge base equipped to recognize a familiar configuration of stimuli. This superior quality of experts' mental representations allows them to rapidly select and use relevant information, and to adapt the processing of information to changing circumstances. As practitioners gain experience in executing a task, their performance becomes increasingly smooth, efficient and automatic (Boshuizen, Van Der Vleuten, Schmidt, & Machiels-Bongaerts, 1997). The difference between expert and less skilled individuals is not merely a matter of the amount and complexity of accumulated knowledge; it mostly reflects qualitative differences in cognitive strategies, knowledge organization and problem representations (Patel, Glaser, & Arocha, 2000). This concept is consistent with our finding that reasoning capacity increases throughout the medical curriculum. It is, however, interesting to observe that Year 3 medical students show obvious reasoning skills. Further studies are warranted to delineate the possible determinants of such early development of reasoning capacity, especially the possible contribution of PBL seminars.

The difference between ascertainment degrees provided for correct and incorrect answers on the true/false test has been used here to indicate the capacity of self-estimation of core knowledge on the part of the participants. It was anticipated that, when maximal, such a capacity would result in degrees of ascertainment that were high for correct answers and low for incorrect answers, and that this would be translated into a marked difference between the two degrees. By contrast, a minimal self-estimation capacity would be translated into a small difference. This capacity, however, was not related to experience as there were no differences among the four groups of students and the tutors. Furthermore, the capacity to discriminate correct and incorrect answers was not related to the global level of ascertainment for correct answers. A positive correlation between SCT scores and the differences

in ascertainment for correct and incorrect answers on the true/false test was observed in all student groups. These results suggest that either the ability to discriminate what is known and what is unknown has common determinants with reasoning capacity, or that the two parameters can influence each other. The higher the difference between the ascertainment degrees provided for correct and incorrect answers, the better the reasoning capacity. This finding is crucial because the effectiveness of reasoning depends on the correctness of the knowledge retrieved during the reasoning process. Such an issue is worthy of further study, especially in a PBL curriculum where reasoning capacity is stimulated early in the training.

In summary, in a 7-year medical curriculum in which PBL seminars are integrated with other learning modalities, Year 3 students demonstrated acquired reasoning skills (assessed through an adapted SCT) that they were able to transfer into a new context. These skills increased throughout the curriculum, whereas knowledge retention showed a rapid reduction.

A direct correlation between knowledge and reasoning skills was observed initially and disappeared with maturation. A new finding, however, showed that student capacity to self-estimate core knowledge was directly related to reasoning skills in the four year groups studied. In further research, it would be interesting to compare these results with those of other tests assessing the same processes of knowledge retention and biomedical reasoning that we studied here. In addition, the relationship between reasoning, and knowledge and its self assessment, should be investigated in more depth. From an educational perspective, the extent to which the individual student is stimulated to develop reasoning skills and to validate his or her core knowledge in the learning activities appears to be important. It may be that greater emphasis is more commonly placed on the former aspect in PBL and on the latter in formal lectures. In further research, we will investigate the possible impact of activities integrated into APP seminars and directed towards self-estimation of core knowledge.

5.5. References

- Barrows, H. (1986). The scope of clinical education. *Journal of Medical Education*, 61, 23–33.
- Barrows, H.S., Norman, G.R., Neufeld, V.R., & Feightner, J.W. (1982). The clinical reasoning of randomly selected physicians in general medical practice. *Clinical and Investigative Medicine*, 5, 49–55.
- Barrows, H.S., & Tamblyn, R. (1980). *Problem-based Learning: an Approach to Medical Education*. New York, NY: Springer.
- Boniver, J. (2004). Les études de médecine à l'Université de Liège: le renouveau pédagogique de la Faculté de Médecine. *Revue Médicale de Liège*, 59(12), 717–730.
- Bordage, G. (1994). Elaborated knowledge: a key to successful diagnostic thinking. *Academic Medicine*, 69, 883–885.
- Boshuizen, H.P.A., & Schmidt, H.G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, 16, 153–184.
- Boshuizen, H.P.A., van der Vleuten, C., Schmidt, H.G., & Machiels-Bongaerts, M. (1997). Measuring knowledge and clinical reasoning skills in a problem-based curriculum. *Medical Education*, 31, 115–121.
- Caire, F., Sol, J.C., Charlin, B., Isodori, P., & Moreau, J.J. (2004). Le test de concordance de script (TCS) comme outil d'évaluation formative des internes en neurochirurgie: implantation du test sur Internet à l'échelle nationale. *Pédagogie Médicale*, 5, 87–94.
- Charlin, B., Gagnon, R., Sibert, R., & van der Vleuten, C. (2002). Le test de concordance de script, un instrument d'évaluation du raisonnement clinique. *Pédagogie Médicale*, 3, 135–144.

Charlin, B., Roy, L., Brailovsky, C., Goulet, F., & van der Vleuten, C. (2000). The script concordance test, a tool to assess the reflective clinician. *Teaching and Learning in Medicine*, 12, 189–195.

Charlin, B., & St-Jean, M. (2002, Mai). Le test de concordance de script: un outil pour évaluer le jugement en médecine. *Bulletin du CEFES, Université de Montréal*, 6, 4–5.

Charlin, B., Tardif, J., & Boshuizen, P. (2000). Scripts and medical diagnosis knowledge: theory and applications for clinical reasoning instruction and research. *Academic Medicine*, 75, 182–190.

Charlin, B., & van der Vleuten, C. (2004). Standardized assessment in the context of uncertainty: the script concordance approach. *Evaluation & the Health Professions*, 27, 304–319.

Cohen, J. (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement of partial credit. *Psychological Bulletin*, 70, 213–220.

Elstein, A.S. (1994). What goes around comes around: return of hypothetico-deductive strategy. *Teaching and Learning in Medicine*, 6, 121–123.

Feltovich, P.J., & Barrows, H.S. (1984). Issues of generality in medical problem solving. In H.G. Schmidt, & M.L. De Volder (Eds.), *Tutorials in Problem-based Learning: a New Direction in Teaching the Health Professions* (pp.128-142). Assen: Van Gorcum.

Leclercq, D. (1975). *L'évaluation subjective de la probabilité d'exactitude des réponses en situation pédagogique* (Thèse de Doctorat en Pédagogie), Université de Liège, Liège, Belgique.

Leclercq, D. (1982). Confidence marking, its use in testing. In B. Choppin, & N. Postlethwaite (Eds.), *Evaluation in Education* (pp.161-287). Oxford: Pergamon Press.

Leclercq, D., Boxus, E., de Brogniez, P., Lambert, F., & Wuidar, H. (1993). The TASTE Approach: General implicit solutions in MCQ, open books

exams and interactive testing and self-assessment. In D. Leclercq, & J. Bruno (Eds.), *Item Banking, Interactive Testing and Self-Assessment* (pp.210-232). Berlin: Springer Verlag.

Norman, G.R., & Schmidt, H.G. (1992). The psychological basis of problem-based learning: a review of the evidence. *Academic Medicine*, *67*, 557–565.

Patel, V.L., Glaser, R., & Arocha, J.F. (2000). Cognition and expertise: acquisition of medical competence. *Clinical and Investigative Medicine*, *23*(4), 256–260.

Patel, V.L., & Groen, G.J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science*, *10*, 91–110.

Regehr, G., Norman, G.R. (1996). Issues in cognitive psychology: implications for professional education. *Academic Medicine*, *71*, 988–1001.

Rikers, R., Loyens, S., Winkel, W., Schmidt, H.G., & Sins, P.H.M. (2005). The role of biomedical knowledge in clinical reasoning: a lexical decision study. *Academic Medicine*, *80*(10), 945–949.

Rikers, R., Schmidt, H.G., & Moulaert, V. (2005). Biomedical knowledge: encapsulated or two worlds apart? *Applied Cognitive Psychology*, *19*(2), 223–231.

Schmidt, H.G. (1983). Problem-based learning: rationale and description. *Medical Education*, *17*, 11-16.

Schmidt, H.G., & Boshuizen, H.P.A. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, *5*, 205–221.

Schmidt, H.G., Norman, G.R., & Boshuizen, H.P.A. (1990). A cognitive perspective on medical expertise: theory and implication. *Academic Medicine*, *65*, 611–621.

Schmidt, H.G., & Rikers, R. (2007). How expertise develops in medicine: knowledge encapsulation and illness script formation. *Medical Education*, *41*, 1133–1139.

Semb, G.B., & Ellis, J.A. (1994). Knowledge taught in school: what is remembered? *Review of Educational Research*, *64*(2), 253–286.

Vanbelle, S., Massart, V., Giet, D., & Albert, A. (2007) Test de concordance de script: un nouveau mode d'établissement des scores limitant l'effet du hasard. *Pédagogie Médicale*, *8*, 71–80.

van Gessel, E., Nendaz, M.R., Vermeulen, B., Junod, A., & Vu, N.V. (2003). Development of clinical reasoning from the basic sciences to the clerkships: a longitudinal assessment of medical students' needs and self-perception after a transitional learning unit. *Medical Education*, *37*, 966–974.

Verhoeven, B., Verwijnen, G., Scherpbier, A., & van der Vleuten, C. (2002). Growth of medical knowledge. *Medical Education*, *36*, 711–717.

6. ARTICLE 3 : LONGITUDINAL ASSESSMENT OF PROGRESS IN REASONING CAPACITY AND RELATION WITH SELF- ESTIMATION OF KNOWLEDGE BASE

A. Collard, F. Mélot and J-P. Bourguignon

Sous Presse dans la revue « Assessment and Evaluation in Higher Education »

ABSTRACT

The aim of the study was to investigate progress in reasoning capacity and knowledge base appraisal in a longitudinal analysis of data from summative evaluation throughout a medical Problem-Based Learning curriculum.

The scores at Multidisciplinary Discussion of a Clinical Case and at Multiple Choice Questionnaires (MCQs) with ascertainment degree were studied longitudinally for 213 students from Year 2 to 5. The capacity of core knowledge delimitation was calculated as the difference between the levels of average ascertainment degrees given for correct and incorrect answers at MCQ.

For both Multidisciplinary Discussion of a Clinical Case evaluation and self-estimation of core knowledge, the capacity increases throughout the curriculum. The reasoning capacity assessed through Multidisciplinary Discussion of a Clinical Case is positively correlated with MCQ scores and the capacity to discriminate the mastered core knowledge.

In conclusion, this study indicates that self-estimation of core knowledge is associated with increase in reasoning performance through a well-organized knowledge base. Since that ability is related to success or failure, it is suggested that student awareness about delimitation of mastered core knowledge is considered as part of learning.

6.1. Introduction

Because clinical reasoning is a central process in professional practice in health sciences, developing such skills is an important goal in a medical curriculum. Problem-based learning (PBL) was developed to facilitate the student's process of acquisition, organization and retrieval of knowledge (Norman & Schmidt, 1992; Regehr & Norman, 1996; Schmidt 1983; van Gessel, Nendaz, Vermeulen, Junod, & Vu, 2003). Through such processes, PBL promotes reasoning skills (Barrows 1986). Since 2000, the Faculty of Medicine at the University of Liège has integrated PBL seminars into year 2–7 of a 7-year curriculum (Boniver, 2004). The curriculum is built of learning units, each related to a body system (e.g. cardiovascular, respiratory, ... see Table 1). Each units involved both plenary lectures and PBL activities with clinical case discussion as early as in the 2nd year of the curriculum. To evaluate biomedical and clinical reasoning developed through PBL, oral multidisciplinary discussion of a clinical case (MDCC) with a jury of 3-5 teachers has been used for 13 years in addition to classical written Multiple Choice Questionnaires (MCQs).

Since clinical reasoning is thought to operate mainly by the end of the medical curriculum, most studies on clinical reasoning have been performed at that late stage of training. More recently, researchers were interested in the mechanism of reasoning capacity at early stages in the curriculum (Collard et al. 2009; Diemers, Van De Wiel, Scherpbier, Heineman, & Dolmans, 2011; Hoff, Bestawros, Kassis, & Charlin, 2010; Humbert et al. 2011). An original feature in our conditions is the early introduction of clinical cases used throughout the entire curriculum both for learning and evaluation, enabling to assess maturation of student reasoning. In a previous cross-sectional study in non-summative conditions, we provided evidence that biomedical reasoning was already effective early in the curriculum (year 3) and showed subsequent maturational increase that was in contrast to a decrease in factual knowledge retention (Collard et al. 2009).

Another feature in our curriculum has been that, in MCQ evaluation, each question is complemented by certainty degrees provided by the students. These degrees aim at giving more credit to correct answers with high certainty and less penalty to wrong answers with low certainty (Leclercq 1982, 1993; Leclercq, Boxus, De Brogniez, Lambert, & Wuidar, 1993). In a former study, we developed the concept that those certainty degrees could enable to estimate the student capacity of delimitation of oneself knowledge base. This issue can be important since students could neglect or ignore the importance of knowledge that underpins reasoning and thus develop reasoning using wrong or inaccurate knowledge. Such an issue is consistent with the importance for a student to be capable to delimitate oneself knowledge base for reasoning validation that was emphasized by some authors (Jones, 1992; Langendyk, 2006). We reported in our previous study that such self-estimation of core knowledge delimitation was significantly correlated with reasoning capacity (Collard et al. 2009). According to Patel and Groen (1996), reasoning performance does not result from superior memory skills but is a function of a well-organized knowledge base to recognize a familiar configuration of stimuli. Along the same line, Jones (1992) considered that clinical reasoning errors could be due to incapacity to retrieve the relevant information in memory. Audetat, Laurin, and Sanche (2011) also identified that insufficient knowledge base and organization could explain clinical reasoning difficulties. An optimal knowledge organization allows rapid selection and use of relevant information (Jones, 1992; Langendyk, 2006), presumably implying that the student discriminates the known and the unknown in core knowledge. This hypothesis is consistent with our previous findings that the capacity of self-estimation of knowledge base is correlated with reasoning performance.

The aim of the present study was to investigate longitudinally in summative conditions the change in student reasoning performance and knowledge base retrieval as well as the relation with the student capacity of knowledge base delimitation. We also explored the possible association between the capacity of knowledge base self-estimation and the outcome of examinations.

6.2. Materials and Methods

Participants

This study involved a total of 213 students belonging to three cohorts beginning year 2 of the medical curriculum in 2002, 2003 and 2004 and followed for four years till year 5 of the curriculum. Besides the longitudinal analysis of data from 213 students who succeeded four years without repeating grade, the data from 66 students who failed year 2 and, among them, 48 who repeated year 2 were included in an analysis to compare the capacity of delimitation of knowledge base in relation to passing or failing.

Materials

The scores obtained at the tests administered twice a year from year 2 to 5 were used in this study. The topics addressed in each of the four years and the modalities of learning and evaluation (two examinations each year) are summarized in Table 1. We used the mean yearly scores of individual students at multidisciplinary MCQs (7 true/false tests in a four year period, each including a total of 80 to 200 questions) as a reflection of knowledge retrieval. Each question addressed four true or false items with 3 possible conditions of responses: either all true, either all false or one true and three false. Ascertainment degrees were provided through six levels corresponding with the following ranges of certainty: level 1 (0-25%); level 2 (25-50%); level 3 (50-70%); level 4 (70-85%); level 5 (85-95%); level 6 (95-100%) at each MCQ question. To encourage the students towards critical self-estimation, there was a bonus when high or low degree of certainty was given to a correct or wrong answer, respectively. Conversely there was a penalty when low or high degree of certainty was given to a correct or wrong answer, respectively. Based on those degrees, the capacity to estimate core knowledge delimitation was calculated as the difference [average ascertainment degree for correct answers – average ascertainment degree for incorrect answers] (Collard et al. 2009).

The oral Multidisciplinary Discussion of a Clinical Case consisted of a task similar to PBL seminars, i.e. analysis of the content of a short

clinical case and preparation of a diagram relating the different phenomena in order to explain the case (45 min). That analysis was then presented and discussed with a multidisciplinary team of 3 to 5 teachers (20-30 min). This task required that students identified relevant elements (signs, symptoms, contextual elements) of the clinical case and connected those elements together with causal relation to explain and integrate the pathophysiological mechanisms (year 2 and 3 students) or the clinical findings and the diagnostic and therapeutic management of a case (year 4 and 5 students). The reasoning capacity was evaluated through the problem analysis. It involved the student capacity to select and transform factual elements in a schematic and semantic construction that permitted to evaluate the progression towards structuring, analysing and understanding the problem. This capacity is essential for clinical reasoning (Chang, Bordage, & Connell 1998; Nendaz, Charlin, Leblanc, & Bordage, 2005). Such a procedure was also thought to evaluate knowledge retrieval and delimitation though, in the assessment, we could not separate what was related to these components from reasoning. Following the Multidisciplinary Discussion of a Clinical Case examination, a global score (average of the 3-5 teachers) was given and took into consideration knowledge retrieved for reasoning and reasoning capacity together. The qualitative scale used (/20) was 18-20: outstanding, 16-17: excellent, 14-15: very good, 12-13: good, 10-11: passable, 8-9: insufficient and <7: very insufficient.

Statistical Analysis

The data based on test scores were calculated as mean (\pm SD). Normality of data distribution was confirmed using Kolmogorov-Smirnov test. The significance of maturational change in performance throughout the curriculum was investigated using repeated-measure Anova. Moreover the relation between average ascertainment degree and the success or failure at year 2 was evaluated through a one-way Anova for each year. Post-Hoc analyses (Fisher's test) were used to determine significant differences. Pearson's correlations were used to characterize the association between our different performance scores (knowledge, reasoning and ascertainment). The results were considered to be significant at the 5% probability level ($p < 0.05$).

Table 1. Topics addressed, learning and evaluation modalities from year 2 to 5 of medical curriculum

Year of the curriculum	2		3		4		5	
	June 2003, 2004, 2005	January 2004, 2005, 2006	June 2004, 2005, 2006	January 2005, 2006, 2007	June 2005, 2006, 2007	January 2006, 2007, 2008	June 2006, 2007, 2008	
Time of cohort exams								
Endpoints	Basic sciences, normal and abnormal structures and functions of systems		Basic sciences, normal and abnormal structures and functions of systems		Clinical sciences, disease diagnosis and management		Clinical sciences, disease diagnosis and management	
Topics	Cardiovascular Respiratory Renal Blood	Digestive Endocrine Female reproductive	Nervous Musculo-skeletal Immune Homeostasis control	Cardiology Pulmonology General medicine	Obstetric & neonatology Hemato-oncology Immunorheumatology Endocrinology Nephrology Dermatology	Gastroenterology Urology Obstetrics & neonatology	Neurology Orthopedics Ophthalmology General medicine	
Knowledge evaluation	Multiple Choice Questionnaires with ascertainment degree (all topics)							
Reasoning evaluation	x	x	x	x	x	x	x	x
	Multidisciplinary Discussion of a Clinical Case with 3 teachers							
	(discussion about diagnosis, anamnesis, physical examination and complementary examination)							
N Teachers	20	20	20	13	17	13	18	
N hours Lectures	~ 40	~ 40	~ 40	~ 150	~ 70	~ 125	~ 150	
N hours Practical Exercises	~ 60	~ 40	~ 65	~ 30	~ 30	~ 30	~ 30	
N PBL seminars	~ 20 (60h)	~ 20 (60h)	~ 25 (75h)	~ 22	~ 30	~ 22	~ 30	
N hours Clerkship	/	120		144	144	168	168	

6.3. Results

Maturational changes in reasoning capacity, knowledge retention, and delimitation of core knowledge

In Figure 1A are shown the 4-year changes in average scores obtained at the Multidisciplinary Discussion of a Clinical Case (MDCC). The average score of 13.7 (./20) in year 2 suggested that the students could demonstrate early reasoning capacity. Such a capacity increased throughout the curriculum as confirmed by statistical analysis ($F(3,636) = 34.85$; $p < 0.001$). Post Hoc analysis showed a significant difference between year 2 and the following years as well as between year 3 and year 4. When the average scores at MCQ were studied over the same periods (Figure 1B), an overall gain appeared between year 2 and year 5 with some reduction in the meantime between year 3 (i.e. the last year in basic sciences) and year 4 (i.e. the first year in clinical sciences). The changes were significant throughout the whole period as well as among individual years ($F(3,636) = 203.39$; $p < 0.001$). The capacity of core knowledge delimitation using the difference in ascertainment degree for correct and incorrect answers showed a steady and significant ($F(3,636) = 11.85$; $p < 0.001$) maturational increase throughout the 4-year study period (Figure 1C). Post hoc analysis showed a significant difference between year 5 and the previous years as well as between year 2 and 4.

The capacity of core knowledge delimitation in relation to success or failure at the beginning of the curriculum

Table 2 presents the mean values and standard deviation for ascertainment degree, Multidisciplinary Discussion of a Clinical Case and MCQ at Year 2. The results of students who failed and repeated year 2 are presented (Year 2 B).

As shown in Figure 2, the average capacity of core knowledge delimitation as estimated through the difference in ascertainment degree for correct and incorrect answers at MCQ was significantly different among the student groups at year 2 A (June session) ($F(4,274) = 6.83$; $p < 0.001$). Post Hoc analysis showed a significantly higher performance in the students passing after year 2 A examinations than in all the other students failing in year 2A. Interestingly, the students who

passed at year 2 B, i.e. after repeating grade, experienced a catch up in capacity of core knowledge delimitation, reaching ultimately an average level similar to that of the students having passed one year earlier. Moreover, they attained a significantly higher level than the students failing year 2B or leaving the medical curriculum ($F(2,44) = 4.03$; $p = 0.02$).

Table 2. Mean (SD) Multiple Choice Questionnaire (MCQ) and Multidisciplinary Discussion of a Clinical Case (MDCC) scores and difference in Ascertainment Degree (AD) at Year 2 A (first attempt) and 2B (attempt after repeating grade) for June session

	N	MDCC (./20)		MCQ (./20)		AD (%)	
		Year 2A	Year 2B	Year 2A	Year 2B	Year 2A	Year 2B
Success 2A	213	13.7 (2.5)		12.2 (2.5)		14.3 (5.7)	
Failure 2A Success 2B June	28	10.8 (1.8)	12.1 (1.3)	9.7 (1.2)	11.2 (1.3)	11.6 (3.8)	14.2 (5.9)
Failure 2A Success 2B September	12	8.1 (4.6)	10.5 (3.9)	8.9 (1.2)	10.2 (1.3)	10.7 (5.9)	9.9 (5.4)
Failure 2A Failure 2B	8	8.1 (2.4)	7.1 (4.5)	7.9 (1.8)	8.6 (2)	9.6 (5.3)	9 (4.6)
Failure 2A Stop ME	18	7.8 (3.8)		7.6 (1.5)		8.8 (4.9)	

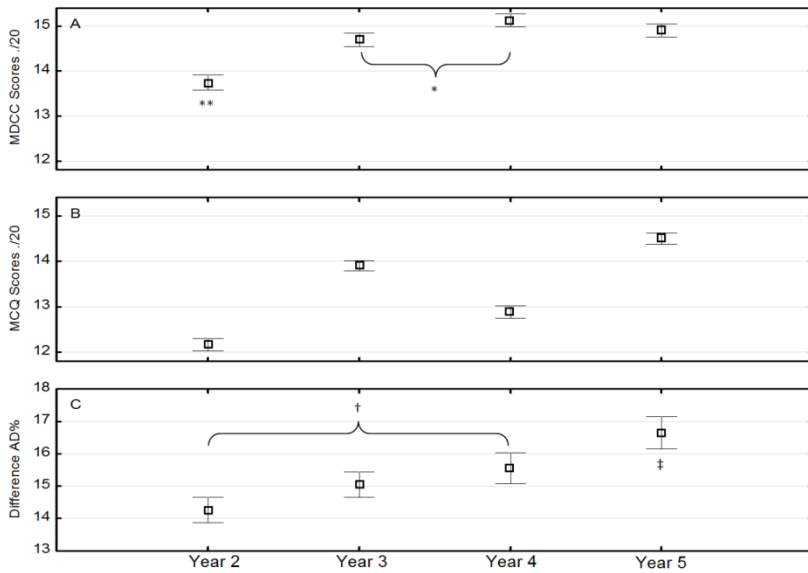


Figure 1. Change in mean \pm Standard error scores obtained at multidisciplinary discussion of a clinical case - MDCC (panel A), multiple choice questionnaire - MCQ (panel B) and difference in ascertainment degree – AD between correct and incorrect answers (panel C) in 213 medical students throughout 4 years of the curriculum. A significant increase is observed for each of the three studied parameters.

* significant difference between the year 3 and year 4 $p < 0.001$

** significant difference between year 2 and all others years

† significant difference between year 2 and year 4

‡ significant difference between year 5 and all others years

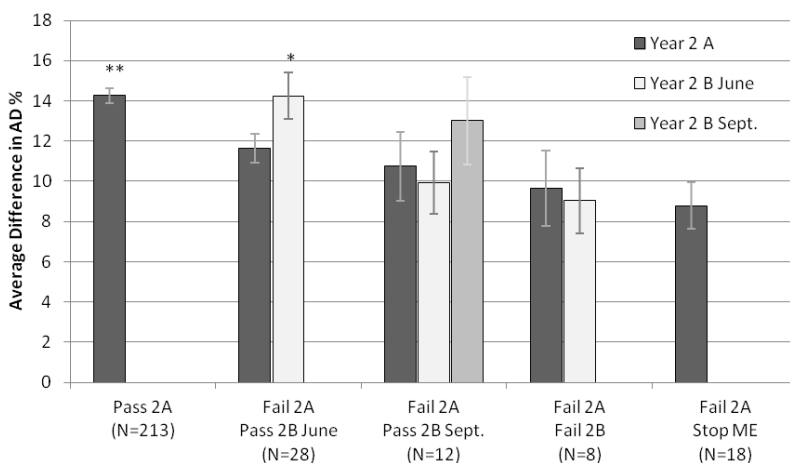


Figure 2. The average difference (+/- standard error) between ascertainment degree for correct and incorrect answers is shown for different student groups in 2nd year; 213 students who succeeded and entered the 3rd year and 66 students who failed. In these failing students, 28 were successful at year 2B June session, 12 succeeded at Year 2B September Session, 18 stopped medical studies.

** significant difference between the students passing year 2 A than in all other students groups failing in year 2A

* significant difference between the students passing year 2 B in June than the other students failing year 2B in June and September

The relation between Multidisciplinary Discussion of a Clinical Case scores, MCQ scores and core knowledge delimitation

A highly significant correlation between Multidisciplinary Discussion of a Clinical Case scores and MCQ scores is illustrated in Figure 3. Such a positive correlation was observed both for years 2 and 3 ($r = 0.71$; $P < 0.05$) as well as year 4 and 5 ($r = 0.68$; $p < 0.05$). With progression throughout the curriculum, the performance of students at MCQ increased significantly between years 2-3 and years 4-5 since the mode of score distribution was 12.5 and 14.6, respectively. However, analysis year by year (Figure 1) showed that years 3 and 5 resulted in average scores higher than years 2 and 4, possibly following student adaptation to the learning strategies (relative number of lectures) and teachers who changed between years 3 and 4. To a lesser extent, a similar finding was obtained with the Multidisciplinary Discussion of a

Clinical Case score where the mode was 14.5 and 15.5, respectively, the difference being significant.

As shown in Figure 4, a weak but significant positive correlation between the capacity of core knowledge delimitation and oral Multidisciplinary Discussion of a Clinical Case examination performance was observed both for the periods year 2-3 ($r = 0.18$; $p < 0.05$) and year 4-5 ($r = 0.24$; $p < 0.05$) of the curriculum. Noteworthy, the individual values of difference in ascertainment degree for correct and incorrect answers showed more scattering and the correlation coefficients were higher in year 4 and 5 than in year 2 and 3. Also shown in Figure 4, the capacity of core knowledge delimitation appeared to be significantly correlated with the MCQ scores in a similar manner, raising the question as to whether the correlation observed with the oral examination scores could be explained by the knowledge component of that evaluation. In an attempt to clarify this, we separated the year 4 and 5 cohort in 2 subgroups based on the correlation between Multidisciplinary Discussion of a Clinical Case scores and MCQ scores (Figure 3). On the one hand the students achieving an oral examination score greater than that calculated for their MCQ score based on the equation of the regression line and, on the other hand those achieving a lower oral examination score than predicted from MCQ score. Thus, the former and the latter could have high or low reasoning capacity relatively to core knowledge. In the subgroup with relatively high reasoning capacity, the Multidisciplinary Discussion of a Clinical Case scores were expectedly all in the upper range without significant correlation with the capacity of core knowledge delimitation (Figure 5). In the group with relatively low reasoning capacity, however, a greater capacity of core knowledge delimitation predicted a greater score at Multidisciplinary Discussion of a Clinical Case.

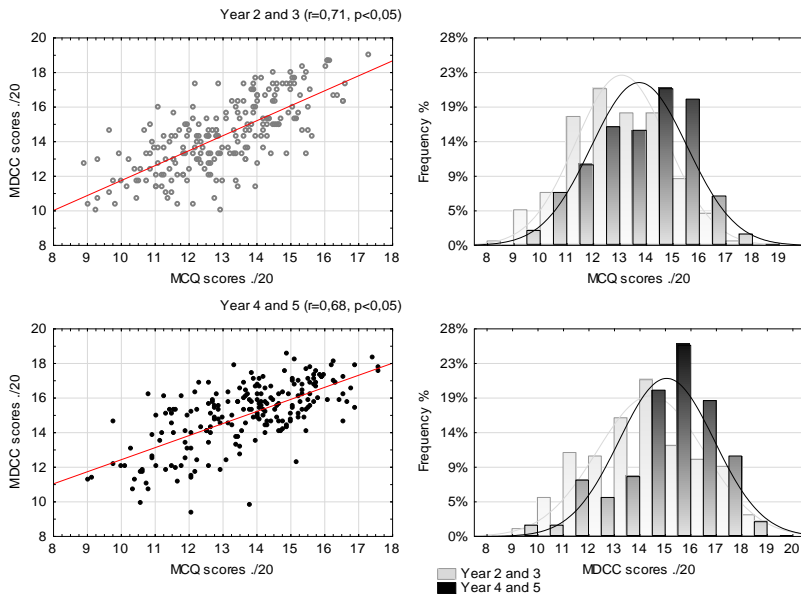


Figure 3. Left panels : Correlation between the individual average scores obtained by 213 medical students at Multidisciplinary Discussion of a Clinical Case (MDCC) in 2nd and 3rd year (upper) and 4th and 5th year (lower) and the average scores obtained at Multiple Choice Questionnaires (MCQ) administered in the same periods of the curriculum. Right panels: Frequency distribution (%) of the scores obtained by 213 medical students at MCQ (upper) and MDCC (lower)

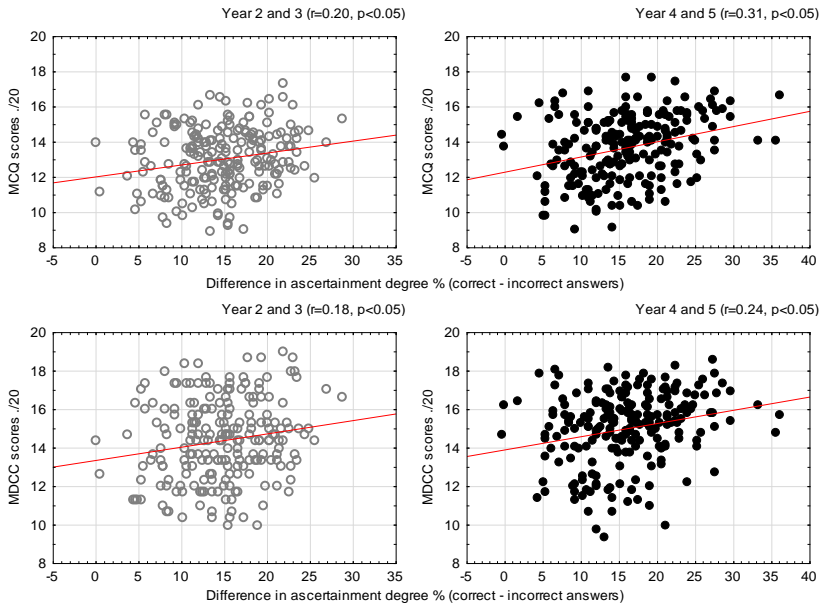


Figure 4. Correlation between the individual average scores obtained by 213 medical students at Multiple Choice Questionnaire (MCQ) (upper panels) and at Multidisciplinary Discussion of a Clinical Case (MDCC) (lower panels) in 2nd and 3rd years (left panels) and 4th and 5th year (right panels) and the average difference between ascertainment degree for correct and incorrect answers at MCQ (an appraisal of core knowledge delimitation) administered in the same periods of the curriculum

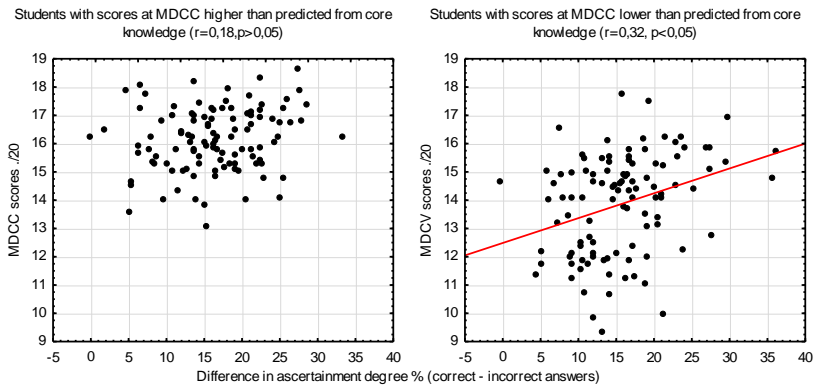


Figure 5. Correlation between the individual average scores obtained by 4th and 5th year medical students at Multidisciplinary Discussion of a Clinical Case (MDCC) and the average difference between ascertainment degree for correct and incorrect answers at Multiple Choice Questionnaire (MCQ) (an appraisal of core knowledge delimitation) administered in the same periods of the curriculum. The 213 students are separated into 2 groups; left panel: 108 students achieving higher MDCC scores than predicted from their MCQ scores; right panel: 105 students achieving lower MDCC scores than predicted from their MCQ scores

6.4. Discussion

The present longitudinal study provides evidence that reasoning capacity is an early ability of medical students and increases throughout the curriculum. This reasoning capacity is correlated with the capacity of core knowledge retrieval and is also associated with the student ability of core knowledge delimitation, particularly for students achieving Multidisciplinary Discussion of a Clinical Case scores lower than predicted from their core knowledge.

The maturational changes in reasoning and knowledge retrieval

The Multidisciplinary Discussion of a Clinical Case results enable to show a significant maturational increase in reasoning capacity that is similar to the data reported in the previous study using the Script Concordance Test (SCT). A strength of the previous study was the administration of the same test designed specifically for reasoning evaluation in students of 4 different years in the curriculum. A strength

of the present study is the longitudinal aspect and we used data from an examination that involves assessment of reasoning capacity. The similar global maturational increase in performance at both Multidisciplinary Discussion of a Clinical Case and Scripts Concordance Test could be related with the fact that they evaluate a similar capacity, that is biomedical and clinical reasoning though this was not specifically demonstrated. The gain in reasoning capacity throughout the curriculum is consistent with findings showing that increasing experience and exposure to clinical cases contribute to the development of clinical reasoning (Schmidt, Norman, & Boshuizen, 1990).

The maturational increase in reasoning capacity is not paralleled by similar changes in knowledge performance, although Multidisciplinary Discussion of a Clinical Case and MCQ which were administered twice a year assessed the same topics at the same time. This difference suggests that these two tests explore different abilities and, indirectly, further supports our above hypothesis that Multidisciplinary Discussion of a Clinical Case does assess other abilities than core knowledge.

The oral examination data however should be interpreted with caution due to possible biases related to the non-standardization of the examination modalities: juries, questions, topics, ... as opposed to MCQs and Scripts Concordance Tests which are standardized tests administered to all students. In addition, enhancement in student capacity throughout the curriculum could involve student's increasing familiarity with the testing procedure that can influence the outcome of evaluation. While we showed in the previous study that reasoning capacity was effective in the third year of the curriculum, it is remarkable that, here, 2nd year medical students already prove to be capable of biomedical reasoning after having participated in PBL seminars for half a year. Confirmation of the impact of those seminars, however, would require Scripts Concordance Test (or another reasoning test) administration before and after the first series of PBL seminars. Despite the biases mentioned, the present longitudinal study in summative conditions exploring different topics suggests that the medical students experience reasoning maturation comparable with that evidenced in an experimental cross-sectional study exploring a single topic.

It appears in this study that the capacity of knowledge retrieval increases somewhat during the curriculum whereas a developmental decrease in knowledge retention was observed in our previous cross-sectional study. Importantly however, the MCQs administered in the present study were different each half year and all related to the learning content of the past 4-6 months. Such conditions do not allow investigation of the time related change in knowledge retention as opposed to the MCQ administered in our previous study. In that study, the same test related to the learning content of year 3 was used across four years and similarly answered by students of year 3-6 (Collard et al. 2009). Therefore the maturational differences in performance at MCQs in this study could involve other factors. The memorization skills can be influenced by the circumstances of knowledge acquisition: quantity and complexity of knowledge for a specific topic (for instance nervous system learning in year 3) or the learning methods used by the teachers. The conditions of knowledge restitution could play a role as well: presence or absence of clue, congruence with context of knowledge acquisition,... . In addition, increasing experience can be associated with more efficient organization or elaboration at encoding as well as more efficient retrieval strategies and such developmental change may be assessed in the conditions of the present study.

The maturational changes in core knowledge delimitation

As proposed in our previous work (Collard et al. 2009), the difference between ascertainment degrees provided for correct and incorrect answers at MCQs was used here to assess the capacity of self-delimitation of core knowledge. It was anticipated that, when maximal, such a capacity would result in degrees of ascertainment that were high for correct answers and low for incorrect answers, thus resulting in a greater difference between the two degrees. By contrast, a low self-estimation capacity would be translated into a small difference. We consider that this capacity could depend on proper organization of knowledge in memory. The more the knowledge is well-structured in memory, the more it is easy to access and then easy to determine the mastered and not mastered knowledge. Such a concept is consistent with the cognitive theories about the reasoning process (Groves, 2008,

2012; Groves, O'Rourke, & Alexander 2003; Schmidt & Rikers 2007). The reasoning process is influenced by the knowledge data base, the capacity of data analysis and synthesis (Elstein, Shulman, & Sprafka 1978, 1990; Gale, 1982) and the awareness and monitoring of the thinking process (Mamede, Schmidt, & Rikers, 2007; Nickerson, Perkind, & Smith, 1985). To facilitate the development of reasoning process for students, it is important to make them aware of both the needed knowledge data and the reasoning process. This requires that the students access and select the accurate knowledge base to identify and classify the different information about a case. So, the capacity of core knowledge delimitation could be an important part of the reasoning process validation by students.

The mean values observed for this index could be felt to be low knowing that the scale of confidence could range from 0% to 100%. However, the results obtained here are congruent with the values generally obtained with this measure (Collard et al. 2009; Leclercq, 2003). Those low scores are partly explained by a student tendency to select intermediate levels of certainty where they reduced both the risk of penalty (low certainty for a right answer) and the chance of bonus (high certainty for a right answer), especially in summative conditions. In our previous cross-sectional non-summative study (Collard et al. 2009), there were no changes in core knowledge delimitation capacity throughout the curriculum suggesting no effect of experience. This conclusion however could have been biased by the reduction in knowledge retention that occurred with time. This could affect ascertainment and interfere with the ability to demonstrate changes in core knowledge delimitation. Such a bias is unlikely in the present longitudinal study since the MCQs were all administered at a similar time after learning. In these conditions, we demonstrate a maturational increase in the capacity of core knowledge delimitation. Moreover, our results provide evidence of a link between this process and the success in year 2. The enhancement of the ability of core knowledge delimitation in students failing first and then passing after repeating year 2 further supports a maturational process. These outcomes, the best capacity of core knowledge delimitation for successful students and the

maturational dimension of this process, are consistent with the link between this capacity and the organization of knowledge in memory. As a result of experience, the knowledge organization is more complete, complex, structured and ordered (Boshuizen & Schmidt, 2008; Patel, Glaser, & Arocha, 2000; Schmidt & Rikers, 2007). With such a memory organization, it is easier to access and retrieve knowledge and, thus, to substantiate efficient reasoning (Boshuizen, Van Der Vleuten, Schmidt, & Machiels-Bongaerts, 1997).

The relation between data obtained at Multidisciplinary Discussion of a Clinical Case, MCQ and ascertainment degrees

A positive correlation between Multidisciplinary Discussion of a Clinical Case and MCQ was observed in all the 4 years studied in the medical curriculum. In our former research, a significant correlation between Scripts Concordance Test and MCQ scores was observed in year 3-4 only and was no longer seen in year 5-6. A positive correlation in the first years makes sense because the reasoning process for novices is strongly dependent of the access of knowledge base that is not yet structured and encapsulated. For the more experienced students, with higher performing knowledge network that created automatism, a relative independence between reasoning capacity and factual knowledge base can be observed (like in our first study). Noteworthy, the knowledge base involved in that study was rather restricted compared to the present one. Here, for Multidisciplinary Discussion of a Clinical Case the task implied that the student explain and validate his reasoning justified by the underlying knowledge base. Since the topics were just learned during the 4-month period preceding examination and thus easier to access, it is suggested that, independently of knowledge retention, the contribution of knowledge is greater in oral examination than in Scripts Concordance Test. Furthermore, the differences in the enhancement scores for Multidisciplinary Discussion of a Clinical Case and MCQ throughout the curriculum enforce the statement of a maturational process different for Multidisciplinary Discussion of a Clinical Case.

The distribution of the differences in ascertainment degree for correct and incorrect answers on MCQ in this study in summative

conditions is quantitatively similar to the distribution observed in our previous study in non-summative conditions. This suggests that the impact of ascertainment degrees on the final note did not influence how certain or uncertain the students self-estimated their MCQ responses. Like in the previous research, a positive correlation, although weaker, was observed between reasoning capacity (MDCC scores) and the differences in ascertainment degree for correct and incorrect answers on MCQ. However, as opposed to the previous study where a specific test (Script concordance test) was used for assessment of reasoning, the Multidisciplinary Discussion of a Clinical Case exam involved both reasoning and knowledge retrieval. This is the reason why we also studied the correlation between the differences in ascertainment degree for correct and incorrect answers on MCQ and the MCQ scores. Since this correlation is similar to that found with Multidisciplinary Discussion of a Clinical Case scores, it is suggested that the core knowledge could be a key element and common determinants in both correlations. Then, we have postulated that reasoning capacity was relatively higher or lower in the students achieving oral examination scores higher than predicted for MCQ scores based on the equation of linear regression between the two scores. We used data from years 4 and 5 when capacity of both reasoning and core knowledge delimitation had attained increased levels. It appears that only in students achieving lower Multidisciplinary Discussion of a Clinical Case scores than expected for to their core knowledge (suggesting a difficulty to use their core knowledge for reasoning in a clinical context) the capacity of core knowledge delimitation is correlated with the Multidisciplinary Discussion of a Clinical Case scores. In students achieving higher oral examination scores than expected for to their core knowledge, the less scattered Multidisciplinary Discussion of a Clinical Case scores could account for the absence of correlation. These observations strengthen our suggestion of the importance of core knowledge organization and discrimination that appear to be more important than the quantity of knowledge for the effectiveness of reasoning. Also these results suggest that it is particularly important for students with reasoning difficulty to learn how to access the appropriate knowledge and to select them in the analysis of a given clinical case. Through such a process, the capacity of

core knowledge delimitation could be essential for validation of reasoning. Further research is necessary to better delineate the role of core knowledge delimitation and validation in the reasoning process and its determinants.

In summary, longitudinal data from the summative examinations of 3 cohorts of students throughout four years of curriculum enabled to show early reasoning capacity that progressed throughout the curriculum. In addition, the capacity of core knowledge delimitation also increased in parallel and was linked with the pass or fail outcome of examinations. Compared with our previous cross-sectional study, a weakness was that Multidisciplinary Discussion of a Clinical Case evaluated reasoning capacity among other abilities while a strength was the constant time conditions of MCQ versus learning, thus allowing no impact of knowledge retention. Our observations suggest that self-estimation of core knowledge is related to reasoning capacity through a well-organized knowledge base. Moreover, the link between that ability and the outcome of examination could be possibly explained by effective retrieval of information required in a given context. Further research is needed to better understand the reasoning process and how to efficiently stimulate the development of these skills through acting on the identified determinants in an educational perspective. It is suggested that core knowledge delimitation and validation of the learning process by the learner should be integrated among the student tasks.

6.5. References

- Audetat, M.C., Laurin, S., & Sanche, G. (2011). Aborder le raisonnement clinique du point de vue pédagogique, I. Un cadre conceptuel pour identifier les problèmes de raisonnement clinique. *Pédagogie Médicale*, 12, 223-229.
- Barrows, H. (1986). The scope of clinical education. *Journal of Medical Education*, 61, 23–33.
- Boniver, J. (2004). Les études de médecine à l'Université de Liège: Le renouveau pédagogique de la Faculté de Médecine. *Revue Médicale de Liège*, 59(12), 717–730.
- Boshuizen, H.P., & Schmidt, H.G. (2008). The development of clinical reasoning expertise. J. Higgs, M.A. Jones, S. Loftus, N. Christensen (Eds.), *Clinical reasoning in the Health professions* (pp.113-121). Amsterdam: Elsevier.
- Boshuizen, H.P.A., van der Vleuten, C., Schmidt, H.G., & Machiels-Bongaerts, M. (1997). Measuring knowledge and clinical reasoning skills in a problem-based curriculum. *Medical Education*, 31, 115–121.
- Chang, R.W., Bordage, G., & Connell, K.J. (1998). The importance of early problem representation during case presentations. *Academic Medicine*, 73(10 Suppl), S109-111.
- Collard, A., Gelaes, S., Vanbelle, S., Bredart, S., Defraigne, J.O., Boniver, J., & Bourguignon, J.P. (2009). Reasoning versus knowledge retention and ascertainment throughout a problem-based learning curriculum. *Medical Education*, 43, 854-865.
- Diemers, A.D., van de Wiel, M.W.J., Scherpbier, A.J., Heineman, E., & Dolmans, D.H. (2011). Pre-clinical patient contacts and the application of biomedical and clinical knowledge. *Medical Education*, 45, 280-288.

Elstein, A.S., Shulman, I.S., & Sprafka, S. (1978). *Medical Problem Solving: An Analysis of Clinical Reasoning*. Cambridge: Harvard University Press.

Elstein, A.S., Shulman, L.S., & Sprafka, S. (1990). Medical problem solving: a ten-year retrospective. *Evaluation & The Health Profession*, 13, 5-36.

Gale, J. (1982). Some cognitive components of the diagnostic thinking process. *British Journal of Educational Psychology*, 52, 64-76.

Groves, M. (2008). *Clinical reasoning and diagnostic expertise*. Saarbrücken: VDM Verlag Dr Müller.

Groves, M. (2012). Understanding clinical reasoning: The next step in Working out how it really works. *Medical Education*, 46, 440-446.

Groves, M., O'Rourke, P., Alexander, H. (2003). The clinical reasoning characteristics of diagnostic experts. *Medical Teacher*, 25(3), 308-313.

Hoff, L., Bestawros, A., Kassis, J., & Charlin, B. (2010). Le test de concordance de script comme outil d'enseignement et d'apprentissage : un projet-pilote pour les étudiants de première année de médecine. *Pédagogie Médicale*, 11(1), 51-56.

Humbert, A.J., Johnson, M.T., Miech, E., Friedberg, F., Grackin, J.A., & Seidman, P.A. (2011). Assessment of clinical reasoning: A Script Concordance test designed for pre-clinical medical students. *Medical Teacher*, 33, 472-477.

Jones, M.A. (1992). Clinical reasoning in manual therapy. *Physical Therapy*, 72, 875-884.

Langendyk, V. (2006). Not knowing that they do not know: self assessment accuracy of third-year medical students. *Medical Education*, 40, 173-179.

Leclercq, D. (1982). Confidence marking, its use in testing. In B. Choppin, & N. Postlethwaite (Eds.), *Evaluation in Education* (pp.161-287). Oxford: Pergamon Press.

Leclercq, D. (1993). Validity, Reliability and Acuity of Self-Assessment in Educational Testing. In D. Leclercq, & J. Bruno (Eds.), *Item Banking, Interactive Testing and Self-Assessment* (pp.114-131). Berlin: Springer Verlag.

Leclercq, D. (2003). *Diagnostic cognitif et métacognitif au seuil de l'Université. Le projet MOHICAN mené par les 9 universités de la Communauté française Wallonie Bruxelles*. Liège : Editions de l'université de Liège.

Leclercq, D., Boxus, E., de Brogniez, P., Lambert, F., & Wuidar, H. (1993). The TASTE Approach: General implicit solutions in MCQ, open books exams and interactive testing and self-assessment. In D. Leclercq, & J. Bruno (Eds.), *Item Banking, Interactive Testing and Self-Assessment* (pp.210-232). Berlin: Springer Verlag.

Mamede, S., Schmidt, H.G., & Rikers, R. (2007). Diagnostic errors and reflective practice in medicine. *Journal of Evaluation in Clinical Practice*, 13(1), 138–45.

Nendaz, M., Charlin, B., Leblanc, V., & Bordage, G. (2005). Le raisonnement clinique : données issues de la recherche et implications pour l'enseignement. *Pédagogie Médicale*, 6, 235-254.

Nickerson, R.S., Perkind, D.N., & Smith, E.E. (1985). *The Teaching of Thinking*. Hillsdale: Lawrence Erlbaum Associates INC.

Norman, G.R., & Schmidt, H.G. (1992). The psychological basis of problem-based learning: a review of the evidence. *Academic Medicine*, 67, 557–565.

Patel, V.L., & Groen, G.J. (1996). Knowledge-based solution strategies in medical reasoning. *Cognitive Science*, *10*, 91–110.

Patel, V.L., Glaser, R., & Arocha, J.F. (2000). Cognition and expertise: acquisition of medical competence. *Clinical and Investigative Medicine*, *23*(4), 256–260.

Regehr, G., & Norman, G.R. (1996). Issues in cognitive psychology: implications for professional education. *Academic Medicine*, *71*, 988–1001.

Schmidt, H.G. (1983). Problem-based learning: rationale and description. *Medical Education*, *17*, 11–16.

Schmidt, H.G., Norman, G.R., & Boshuizen, H.P.A. (1990). A cognitive perspective on medical expertise: theory and implication. *Academic Medicine*, *65*, 611–621.

Schmidt, H.G., & Rikers, R.M. (2007). How expertise develops in medicine: knowledge encapsulation and illness script formation. *Medical Education*, *41*(12), 1133–1139.

van Gessel, E., Nendaz, M.R., Vermeulen, B., Junod, A., & Vu, N.V. (2003). Development of clinical reasoning from the basic sciences to the clerkships: A longitudinal assessment of medical students' needs and self-perception after a transitional learning unit. *Medical Education*, *37*, 966–974.

7. ARTICLE 4 : EFFETS D'UN EXERCICE DE VALIDATION DES ÉLÉMENTS D'UN RAISONNEMENT BIOMÉDICAL SUR LES COMPOSANTES INDIVIDUELLES ET COLLECTIVES DE L'APPRENTISSAGE AU COURS D'UN SÉMINAIRE APP

A. Collard, F. Mélot et J-P. Bourguignon

Accepté sous réserve de modifications mineures dans la revue « Pédagogie Médicale »

ABSTRACT

Contexte: L'Apprentissage Par Problème exige que l'étudiant sélectionne les informations pertinentes d'un cas clinique et les relie par des liens de causalité expliquant la physiopathologie sous-jacente. Dans le cadre de ces séances en petits groupes, l'interaction avec les pairs et l'encadrant force l'étudiant à organiser et structurer ses pensées et idées, utilisant ses connaissances et expliquant les interrelations entre elles. A travers cet exercice, l'APP renforce le développement du raisonnement biomédical.

Buts: L'objectif est d'évaluer l'effet d'un travail de délimitation du champ des connaissances et de validation de son raisonnement par l'étudiant sur son apprentissage, à savoir la gestion du schéma de raisonnement biomédical.

Méthodes: Un exercice de délimitation du champ des connaissances a été intégré à différents moments d'une séance APP. Le nombre d'items et de liens sur les schémas physiopathologiques proposés par les étudiants aux différentes étapes de la séance ont été mesurés. Les données d'un questionnaire sur le fonctionnement en APP et un débriefing ont également été recueillies.

Résultats et Conclusion : L'exercice de validation et délimitation du champ des connaissances, s'il prend place avant le schéma coopératif, entraîne une réduction d'environ 60% du nombre d'éléments manquants identifiés après le schéma coopératif et renforce donc l'apprentissage individuel. Dès lors, il semblerait que la validation « forcée » de son raisonnement avant la construction du schéma coopératif et la confrontation aux connaissances sous-jacentes favorise la démarche personnelle de l'étudiant dans la prise de conscience de ses acquis et ses limites, ce qui pourrait renforcer l'apprentissage.

7.1. Introduction

Un des objectifs fondamentaux de la formation médicale est l'apprentissage du raisonnement clinique. Un certain nombre de dispositifs pédagogiques ont été mis en place dans le but de promouvoir le développement de cette compétence. Dans ce contexte, la Faculté de médecine de l'Université de Liège a, depuis 2000, mis en place des séminaires d'Apprentissage Par Problèmes (APP) tout au long de son cursus de formation (Boniver, 2004). L'Apprentissage Par Problèmes (APP) a été développé pour promouvoir un traitement actif des connaissances, faciliter la réactivation des connaissances antérieures, ainsi que pour favoriser le processus d'acquisition, organisation et récupération des connaissances chez les étudiants (Norman & Schmidt, 1992 ; Regher & Norman, 1996 ; Schmidt, 1983 ; Van Gessel, Nendaz, Vermeulen, Junod, & Vu 2003). En outre, l'APP permet de créer des liens avec le contexte professionnel futur où ces connaissances devront être réactivées (Davis & Harden, 1999).

L'APP implique donc un apprentissage en profondeur et facilite le développement de l'apprentissage auto-dirigé (Davis & Harden, 1999 ; Newble & Clarke, 1986). L'APP entraîne le raisonnement hypothético-déductif, un ensemble de processus de génération, testing, révision et synthèse d'hypothèse qui est utilisé par les étudiants et les cliniciens confrontés à des cas non familiers ou atypiques, pour lesquels des processus « automatisés » n'ont pas encore été développés par l'expérience (Barrows & Pickell, 1991). A travers ces différents mécanismes, l'APP renforce le développement de la capacité de raisonnement (Barrows, 1986).

Schwartz, Burgett, Blue, Donnelly et Sloan (1997) précisent que les étudiants ayant bénéficié d'un curriculum APP ont démontré de meilleures performances en raisonnement ainsi qu'en connaissances médicales et factuelles. Friedman et al. (1990) ont également montré que les étudiants APP semblent atteindre des habiletés interpersonnelles supérieures, une plus grande pratique de l'apprentissage continué, et un plus haut niveau de satisfaction professionnelle. Plus récemment, Haak, HilleRisLambers, Pitre et

Freeman (2011) ont montré que l'APP appliqué en biologie facilitait l'apprentissage et diminuait la disparité entre étudiants avec des connaissances initiales inégales. Dans le cadre de ces séances d'APP en petits groupes, l'interaction avec les pairs et l'encadrant permet un élargissement de la base de connaissances et le développement d'habiletés d'auto-gestion. L'encadrant et les pairs aident l'étudiant à explorer les phénomènes identifiées dans le problème et le poussent à verbaliser ses pensées. Cela force l'étudiant à organiser ses pensées et idées et à expliquer leurs interrelations, utilisant et remaniant ainsi sa base de connaissances (Yeung, AU-Yeung, Chiu, Mok, & Lai, 1999).

Une recherche antérieure a montré l'accroissement de la capacité de raisonnement et son lien avec les connaissances factuelles dès les années précliniques au cours du curriculum impliquant l'APP à la Faculté de médecine de l'Université de Liège (Collard et al. 2009). En outre, la capacité individuelle de raisonnement s'est avérée liée à la capacité de délimiter le champ de ses connaissances, à tous les stades étudiés de la formation. Ce constat est important dans la mesure où parfois, certains étudiants privilégient le raisonnement et négligent, voire ignorent l'importance des connaissances sous-tendant ce raisonnement. Un tel questionnement souligne l'importance, pour un étudiant, d'être capable de délimiter ses connaissances pour valider son raisonnement (Jones, 1992 ; Langendyk, 2006).

Le processus de raisonnement dépend de la base de connaissances, de la capacité d'analyser et de synthétiser des données (Elstein, Shulman, & Sprafka, 1978, 1990 ; Gale, 1982) ainsi que de la conscience et capacité de contrôle de ce processus de raisonnement (Mamede, Shmidt, & Rikers, 2007 ; Nickerson, Perkind, & Smith, 1985). Ces facteurs influencent tous les aspects du raisonnement biomédical et clinique et peuvent eux-mêmes être améliorés lorsque l'étudiant ou le clinicien réfléchit consciemment aux informations récoltées (supportant ou infirmant ses hypothèses) sur lesquelles ses demandes de renseignements et ses décisions cliniques sont basées. Afin de faciliter le processus de raisonnement des étudiants, il est donc important de les rendre conscients des différents facteurs qu'il implique. Cela suppose,

comme l'illustre la Figure 1, que les étudiants puissent disposer des connaissances nécessaires et les sélectionnent de façon pertinente pour appréhender les données d'un cas clinique donné. En effet, lorsque le praticien est confronté à un nouveau cas clinique, les éléments de ce cas vont activer des informations en mémoire. Sur base de ses connaissances et du contexte du cas clinique, il va pouvoir appréhender le cas en sélectionnant les informations qui lui apparaissent les plus utiles. L'abstraction/décontextualisation de ces éléments va lui permettre d'identifier et d'activer rapidement le schème à appliquer à cette situation. Lorsqu'il s'agit d'un nouveau cas, qui comprend des éléments que le praticien n'a jamais rencontrés jusque-là, le schème qui comprend le plus d'éléments concordants sera activé et identifié puis adapté à la nouvelle situation rencontrée. Plus cette base de connaissances est structurée, plus il sera facile d'identifier et d'activer le bon schème et venir y intégrer les nouvelles connaissances et de nouveaux liens. Identifier les connaissances maîtrisées et celles qui ne le sont pas apparaît donc essentiel pour avancer de façon efficace dans le processus de raisonnement. La capacité de délimiter son champ de connaissance peut ainsi être une part importante du processus de validation du raisonnement. Mamede, Schmidt et Rikers (2007) précisent que la qualité du diagnostic dépend de comment le clinicien gère la coordination et l'intégration de ces différentes composantes utilisant la métacognition.

La présente étude vise à intégrer à un séminaire APP une démarche de validation qui renforce la conscience que les étudiants peuvent avoir de leur processus d'apprentissage aux différentes étapes de l'analyse d'un cas clinique. En effet, être conscient de son processus de raisonnement améliore sa propre organisation des connaissances (Mamede, Schmidt, & Rikers, 2007 ; Nickerson, Perkind & Smith, 1985). À travers la pratique et le feedback, l'objectif est de faciliter le processus de métacognition qui semble être un élément central dans la capacité de raisonnement clinique (Yeung et al. 1999). Plus spécifiquement, cette étude évalue de quelle manière et dans quelle mesure un travail de l'étudiant sur sa capacité de délimiter le champ de ses connaissances lui

permet de valider son raisonnement et contribue ainsi à son apprentissage.

Processus de raisonnement		Base de connaissances et utilisation	
		Novices	Experts
1	Confrontation: nouveau cas clinique	<i>Cas simple</i>	<i>Cas complexe</i>
2	Activation: connaissances en MLT	<i>* principalement biomédicales * peu intégrées * reliées en fonction de la causalité</i>	<i>* biomédicales, Cliniques et expérimentelles * intégrées * liens de causalité et associatifs</i>
3	Sélection: informations à utiliser	Hiérarchisation: <i>Niveau faible</i>	<i>Niveau élevé</i>
4	Distinction: éléments connus vs douteux ou ignorés, à clarifier Abstraction du connu	Discrimination: Compréhension: Abstraction: <i>Niveaux faibles</i>	<i>Niveaux élevés</i>
5	Représentation du problème	<i>Analytique et Contextualisée Catégorisation: <i>rigide</i></i>	<i>Globale et abstraite <i>souple</i></i>
6	Identification du (des) script(s)/réseau(x)	<i>Processus lent</i>	<i>Processus rapide</i>
7	Sélection du script	<i>Analyse Séquentielle</i>	<i>Analyse Parallèle</i>
8	Adaptation du script avec nouvelles données		
Validation du raisonnement appliqué au cas clinique		Métacognition: <i>Niveau faible</i>	<i>Niveau élevé</i>

Figure 1. Traitement des informations de la base de connaissances lors du processus de raisonnement

ETUDE 1

7.2. Méthodes Etude 1

Participants

20 étudiants de 3^{ème} Baccalauréat ont participé volontairement à un séminaire APP intégratif dans le cadre de la recherche. Sur base de leur note globale de 3^{ème} année, les étudiants de cet échantillon (Moyenne : 13,4/20 ; Ecart-type : 2,1) n'étaient pas significativement différents des autres étudiants de la même cohorte (Moyenne : 13/20 ; Ecart-type : 2,3).

Nous avons formé 3 groupes de 6-7 étudiants encadrés chacun par un tuteur enseignant. Ces 20 étudiants et 43 autres étudiants de la même cohorte ont rempli un questionnaire d'évaluation de leur fonctionnement habituel en séminaire APP.

Matériel

Dès la 2^{ème} année du curriculum, les enseignements sont organisés en modules autour des grands systèmes anatomiques (Tableau 1). Le module « endocrinologie » se déroule au premier quadrimestre de la 3^{ème} année du cursus. L'APP expérimental s'est déroulé en mars 2009, après l'évaluation certificative sur la matière d'endocrinologie concernée par l'APP expérimental.

Une vignette clinique portant sur les objectifs de ce module a été rédigée. De manière à éviter toute discrimination négative vis-à-vis des étudiants non participants, la vignette portait sur les objectifs d'apprentissage déjà envisagés lors des séminaires APP suivis par les étudiants mais en mêlant ceux relatifs à différentes vignettes et en les proposant dans un contexte inédit.

Un questionnaire évaluant la qualité de l'apprentissage et du raisonnement a été spécifiquement créé. Ce questionnaire distinguait le fonctionnement lors de cet APP expérimental du fonctionnement habituel en APP (questionnaire *in extenso* dans les Tableaux 4 et 5).

Tableau 1. Organisation du cursus médical à l'Université de Liège

Années du cursus	2	3	4	5
Objectifs de formation	Sciences de base, structure et fonctions normales et pathologiques des systèmes anatomiques	Sciences cliniques, diagnostic des maladies et prise en charge		
Modules	Cardiaque Respiratoire Rénal Hématologie	Digestif Endocrinométabolisme Génital	Nerveux Locomoteur Immunitaire Homéostasie	Cardiologie Pneumologie Médecine Générale
			Obstétrique et néonatalogie Héματο-oncologie Immuno-rhumatologie Endocrinologie Néphrologie Dermatologie	Gastroentérologie Urologie Obstétrique & néonatalogie
				Neurologie Orthopédie Ophtalmologie Médecine Générale

Procédure

La procédure habituelle des séminaires APP de 3^{ème} année se déroule en 3 temps : 1) une phase ALLER au cours de laquelle les étudiants encadrés par un tuteur prennent connaissance de la vignette clinique, mettent en évidence les phénomènes à expliquer, formulent des hypothèses d'explication et identifient des objectifs d'apprentissages; 2) Une phase de TRAVAIL PERSONNEL en l'espace de 3-5 jours, au cours de laquelle les étudiants tentent de répondre aux objectifs d'apprentissage par leur recherche dans les ouvrages de référence et réalisent un schéma physiopathologique (schéma individuel); 3) Une phase RETOUR, durant laquelle les étudiants construisent un schéma coopératif au tableau, sur base de leur travail personnel. Le plus souvent, il est de la responsabilité de l'étudiant d'intégrer son schéma individuel avec le schéma collaboratif. Parfois, le tuteur passe en revue certains schémas individuels et les commente.

Certains étudiants ont construit des schémas élaborés mettant en relation de nombreux éléments sans pour autant les maîtriser tous et sans qu'ils n'identifient ceux qu'ils maîtrisent mieux ou moins bien. D'autres se présentent avec des schémas simplifiés et lacunaires, comptant sur le travail coopératif pour disposer d'un schéma collectif final comme point de référence. L'exercice proposé dans le cadre de cette recherche vise donc à impliquer les étudiants dans un travail individuel de validation ou invalidation/correction (avec justifications) des éléments de leur propre schéma physiopathologique et, par là même, de délimitation des champs du connu et de l'ignoré.

Les participants ont été répartis aléatoirement en 3 groupes (procédure décrite Tableau 2). Le groupe A a suivi la procédure habituelle des séminaires PBL décrite ci-dessus. Comme d'habitude, au début de la séance retour, avant la mise en commun dans un schéma construit en coopération, le tuteur a récolté une photocopie de chaque schéma individuel élaboré durant la phase de travail personnel et identifié avec un code pour chaque étudiant. Durant la mise en commun, chaque étudiant a annoté et amendé ou non son propre schéma selon sa stratégie habituelle, aucune instruction n'étant donnée par le tuteur. A la fin du séminaire, le tuteur a récolté les schémas individuels

éventuellement annotés ainsi que la copie du schéma coopératif transcrite par chaque étudiant. Ces documents ont été photocopiés. Finalement, un exercice de délimitation de leur connaissance sur base de leur schéma physiopathologique personnel leur a été demandé. L'instruction était de souligner ou d'entourer (3 couleurs différentes) dans leur schéma les composants (éléments mentionnés et flèches de mise en relation) qu'ils avaient identifiés comme douteux (1), ceux qu'ils avaient invalidés car incorrects (2) et ceux qu'ils avaient ajoutés car manquants (3), sachant que les éléments non marqués étaient *de facto* considérés comme validés - corrects.

Le groupe B a suivi la procédure habituelle jusqu'au début de la phase retour. A ce moment, AVANT la construction d'un schéma physiopathologique coopératif, les étudiants ont été invités à marquer, sur une photocopie de leur schéma initial, les composants dont ils doutaient (1) et les parties de leur schéma où ils pensaient que des informations pouvaient manquer (3). Les éléments non marqués étant considérés comme « certains ». Ces photocopies étaient alors récoltées. Ensuite le schéma coopératif était construit, les étudiants ayant été au préalable invités, durant cette étape, à marquer dans leur schéma initial (3 couleurs différentes) les composants (éléments mentionnés et flèches de mise en relation) qu'ils identifiaient comme douteux et qui s'avéraient corrects (1), ceux qu'ils invalidaient car incorrects (2) et ceux qu'ils ajoutaient car manquants (3), sachant que les éléments non marqués étaient *de facto* validés comme corrects. Tous les schémas individuels annotés ont alors été récoltés ainsi que le schéma éventuellement transcrit à partir du travail collaboratif.

Le groupe C a suivi la même procédure que le groupe B mais ils ont été avertis à la fin de la séance ALLER qu'ils auraient à réaliser un exercice de délimitation de leur champ de connaissances basé sur leur schéma en début de phase retour.

A la fin de la séance RETOUR, chacun des 3 groupes a rempli un questionnaire les interrogeant sur leur fonctionnement habituel en APP et sur leur fonctionnement lors de cet APP intégratif particulier. Ce questionnaire n'a pas été administré préalablement à l'APP expérimental de manière à ne pas biaiser celui-ci par le contenu du

questionnaire qui aurait indiqué aux étudiants les aspects envisagés par le travail. Ces séminaires se sont achevés par un débriefing enregistré.

Tableau 2. Procédure étude 1

Groupes	A (n=7)	B (n=7)	C (n=6)
Phase Aller	x	x	x
Information sur les consignes de délimitation des connaissances			x
Phase de travail personnel + schéma personnel (1)	x	x	x
Phase retour :			
- Récolte du schéma personnel (1)	x	x	x
- Annotation du schéma individuel (orange=douteux, vert=manques) + récolte du schéma (2)		x	x
- Construction du schéma coopératif (+ schéma individuel éventuellement annoté 1') + récolte des schémas	x	x	x
- Annotation du schéma individuel (orange= douteux, vert=manques, rouge= incorrects) + récolte du schéma (3)	x	x	x
Questionnaire	x	x	x
Debriefing enregistré	x	x	x

Analyses statistiques

Etant donné le faible échantillon, nous avons réalisé des analyses statistiques non paramétriques.

Pour les schémas

Les moyennes (\pm ET) du nombre d'items et de liens contenus dans les schémas individuels des 3 groupes ont été calculées. Les différences entre les groupes en termes d'items et de liens ont été évaluées par une Anova de Kruskal Wallis. L'évolution des schémas en termes d'items et de liens au cours de la séance retour pour chacun des groupes a été évaluée par une Anova de Friedman.

Pour le questionnaire

Le pourcentage des réponses choisies aux différentes questions du questionnaire ont été calculées. Les différences entre les groupes ont été évaluées par une Anova de Kruskal Wallis. Les différences de fonctionnement à l'APP expérimental et lors des APP intégrés au cursus de formation ont été évaluées par une Anova de Friedman.

Le seuil de significativité a été fixé à $p \leq 0.05$.

7.3. Résultats Etude 1

1. L'analyse des schémas :

Lorsque nous regardons le nombre d'items et de liens présents sur les schémas physiopathologiques individuels aux différentes étapes de la séance retour, nous observons une différence significative pour les étudiants du groupe A qui présentent un nombre d'items et de liens plus importants que le groupe B après la construction du schéma coopératif ($p = 0.03$). Toutefois, une telle différence n'est pas observée avec le groupe C ce qui renvoie à un impact possible des différences de groupe d'étudiant et/ou de tuteurs. Ceci est renforcé par le maintien d'une différence dans le même sens au niveau du schéma coopératif, même si celle-ci n'est plus significative.

Exercice 1 de marquage des schémas (groupe B et C) :

Le Tableau 3 montre que le groupe C, le seul informé de la démarche de validation avant la séance retour, a marqué un nombre significativement plus grand d'items pour lesquels ils avaient un doute ou une hésitation par rapport au groupe B ($p = 0.047$). Aucune différence n'est observée pour le nombre de liens douteux ni sur les items et liens insuffisamment détaillés ou développés.

Exercice 2 de marquage du schéma (Tableau 3) :

Les étudiants du groupe A identifient significativement plus d'items ($p = 0.005$) et liens ($p = 0.01$) manquants ou incomplets par rapport aux groupes B et C. Aucune différence n'est observée pour les items et liens douteux ou incorrects.

Evolution des schémas au cours de la séance retour (Tableau 3) :

Les 3 groupes présentent une augmentation significative en termes de nombre d'items et de liens présents dans les schémas individuels entre le début et la fin de la séance retour (A: $p < 0.01$; B: $p < 0.001$; C : $p < 0.01$).

Tableau 3. Moyennes (E-T) de nombre d'items et de liens des schémas aux différentes étapes de la séance retour ainsi que nombre d'items et de liens marqués lors des exercices de marquages

Evolution des schémas individuels au cours de la séance retour												Schéma coopératif collectif au tableau	
Groupes	Schémas individuels en début séance retour		Schémas individuels après l'exercice de marquage 1		Schémas individuels après la construction du schémas coopératif		Schémas individuels après l'exercice de marquage 2		Nbre d'items	Nbre de liens	Nbre d'items	Nbre de liens	
	Nbre d'items	Nbre de liens	Nbre d'items	Nbre de liens	Nbre d'items Δ	Nbre de liens Δ	Nbre d'items	Nbre de liens					
A (N=7)*	42,7 (9,5)	48,1 (13,2)			56,3 (8,2)	65,3 (13,5)	57,6 (7,5)	67 (12,3)	56	66			
B (N=7)*	39,4 (4,3)	43,3 (5,2)	40,1 (3,7)	44,7 (4,5)	44,6 (4,1)	49,6 (4,1)	45,6 (4,1)	51,4 (5)	38	49			
C (N=6)*	44,2 (7,4)	50,8 (8,9)	44,2 (7,4)	50,8 (8,9)	50,7 (8,7)	59,3 (10,2)	51,3 (8,3)	60,8 (8,7)	39	50			
Exercices de marquage des schémas													
Exercice de marquage 1						Exercice de marquage 2							
Groupes	Eléments douteux		Eléments manquants		Eléments douteux		Eléments incorrects		Eléments manquants		Nbre d'items	Nbre de liens	
	Nbre d'items ‡	Nbre de liens	Nbre d'items	Nbre de liens	Nbre d'items	Nbre de liens	Nbre d'items	Nbre de liens	Nbre d'items †	Nbre de liens †			
A (N=7)					3 (2,2)	2,5 (1)		1 (0)	14,6 (6,3)	13,9 (5,5)			
B (N=7)	2,7 (1,7)	3,3 (1,9)	5,2 (2,4)	2,5 (1,4)			1 (0)	1 (0)	5,1 (1,8)	5,3 (2,7)			
C (N=6)	4,5 (1,4)	3,3 (1)	4,7 (2,2)	3 (1)			2 (1,4)	1,3 (0,6)	5,3 (4,7)	6,7 (4,9)			

Δ différence significative entre le groupe A et le groupe B

‡ différence significative entre le groupe B et C

† différence significative entre le groupe A et les autres groupes

* différence significative entre les schémas individuels en début de séance retour et en fin de séance retour

2. Les résultats du questionnaire :

Leur fonctionnement lors de l'APP expérimental (Tableau 5) :

Les 3 groupes semblent avoir travaillé de manière équivalente lors de la phase de travail personnel (même le groupe C). En ce qui concerne les quelques différences observées lors de la séance retour (items 32, 34 et 36), elle semblerait plutôt due à une différence de fonctionnement et de dynamique des différents groupes et tuteurs. Toutefois, le groupe C averti de la procédure avant la phase de travail personnel donne des réponses compatibles avec moins d'investissement de la phase retour ce qui peut suggérer que le travail de validation a été en partie anticipé durant la phase de travail personnel.

Leur fonctionnement habituel en APP (Tableau 4) :

Il faut préalablement souligner que, ce questionnaire étant rempli après l'APP expérimental, les réponses des étudiants participants peuvent avoir été influencées par la procédure expérimentale même si les questions sont explicitement posées à propos des autres séminaires. De manière générale, les étudiants ayant participé cernent d'avantage l'importance de distinguer les éléments qu'ils connaissent de ceux qu'ils ne connaissent pas (items 2 et 6). En ce qui concerne la construction de leur schéma lors de la phase de travail personnel, les étudiants participants déclarent mieux identifier les éléments connus, douteux et insuffisamment détaillés par rapport aux non participants (items 7, 8, 9, 13 et 15). Pour la séance retour, les participants confirment les éléments connus et complètent les éléments manquants d'avantage que les non participants (items 17 et 19). Ils corrigent également plus vite leur schéma après la séance retour que par rapport aux non participants (items 26 et 27).

La différence entre APP expérimental et APP habituel (Tableau 6):

Il semblerait que la plupart des étudiants participants cerneraient déjà de manière implicite les éléments qu'ils connaissent, ceux dont ils doutent et les éléments incomplets. Néanmoins, les exercices proposés les forçant à faire ce travail de façon explicite, permet de significativement mieux se rendre compte de leur évolution

en identifiant ces différents éléments (items 17-32; 18-33; 19-34; 21-36; 22-37). La façon dont la mise en commun coopérative a été menée semble avoir eu un impact positif sur leur fonctionnement ; néanmoins, aucun effet significatif de cette mise en commun sur la validation du raisonnement n'a pu être démontré.

3. Les résultats du débriefing :

De manière générale, les étudiants apprécient le côté intégratif de l'APP sur une matière déjà étudiée. Ils considèrent le travail rentable et permettant une bonne intégration parce que la matière a déjà été abordée. A noter toutefois que nous n'avons pas comparé avec un APP expérimental similaire qui aurait porté sur de nouveaux objectifs d'apprentissage. Ils ont également apprécié l'implication des tuteurs dans la mise en commun coopérative qui a permis une analyse en profondeur de la situation clinique. En ce qui concerne les exercices de délimitation des connaissances, il semblerait que les étudiants le pratiquent déjà habituellement. Néanmoins, les exercices proposés les forçant à faire ce travail de façon explicite, ceci a permis à certains de mieux se rendre compte de leur évolution. Les étudiants semblent cerner l'importance de distinguer ce que l'on connaît de ce que l'on ne connaît pas et de la capacité de s'auto-évaluer correctement. Certains parlent de ce processus d'auto-évaluation comme un raisonnement en soi.

Tableau 4. Résultats du questionnaire sur le fonctionnement en APP comparant les réponses des 20 participants à l'APP expérimental aux réponses de 43 autres étudiants de la même cohorte n'ayant pas participé à cet APP particulier

Items	p	Participants						Non participants					
		+++	++	+	-	---	---	---	-	+	++	+++	
De manière générale,													
1. Je suis capable d'identifier ce que je sais et les limites de mes connaissances	NS		10	15	65	10		5	2	42	44	7	
2. Cerner ce que je sais et les limites de mes connaissances est important pour la qualité de mon apprentissage	*			10	45	45		2	5	28	53	12	
3. La validité de mon raisonnement repose sur la quantité de mes connaissances	NS		15	45	40			2	30	28	30	10	
4. La validité de mon raisonnement repose sur la <u>solidité</u> (certitude) de mes connaissances	NS			30	35	35		2	5	28	51	14	
5. La quantité de mes connaissances est plus importante que la certitude de mes connaissances pour la validité de mon raisonnement	NS		15	60	25			12	53	23	7	5	
6. Cerner ce que je sais et les limites de mes connaissances est important pour la qualité de mon raisonnement clinique	*			11	63	26		2	12	42	3	7	
Habituellement, au terme de la phase de travail personnel,													
7. J'ai confirmé les éléments (items et liens) qui me paraissent connus	*	5		10	30	40	15	2	7	21	42	28	
8. J'ai identifié les éléments (items et liens) sur lesquels persistent un doute ou une hésitation	*				25	65	10	2	12	9	40	35	2
9. J'ai identifié les parties qui restent insuffisamment détaillées ou développées	*				15	65	10	2	7	14	23	51	2
10. Mon schéma reflète mon travail personnel de manière exhaustive	NS	15	15	20	25	15	10	7	17	29	21	21	5
11. Mon schéma reflète mon travail personnel de manière synthétique	NS	5	5	5	50	25	10	9	14	17	38	21	2
12. Mon schéma comporte uniquement les éléments (items et liens) dont je suis certain	NS	5	20	55	15		5	10	17	33	21	14	5
13. Mon schéma comporte des éléments (items et liens) sur lesquels j'ai un doute ou une hésitation	*			5	60	30	5	5	9	21	41	24	

14. Mon schéma m'a servi à identifier les limites de mes connaissances	NS			15	50	30	5	7	14	19	29	26	5
15. Mon schéma m'a servi à identifier les éléments (items et liens) sur lesquels j'ai un doute ou une hésitation	*			5	55	35	5	7	5	26	33	29	
16. Mon schéma m'a servi à identifier les parties insuffisamment détaillées ou développées	NS		5	5	45	35	10	7	2	19	41	26	5
Habituellement, au terme de la phase retour,													
17. J'ai confirmé les éléments (items et liens) qui me paraissent connus	*			5	20	60	15		2	5	48	43	2
18. J'ai clarifié les éléments (items et liens) sur lesquels j'avais un doute ou une hésitation	NS			10	45	35	10	2	5	9	53	28	
19. J'ai complété les parties insuffisamment détaillées ou développées *	*			10	35	45	10		7	28	37	23	5
20. Le schéma coopératif m'a servi à identifier les limites de mes connaissances	NS			25	40	25	10	9		14	42	28	7
21. Le schéma coopératif m'a servi à clarifier les éléments (items et liens) sur lesquels j'avais un doute ou une hésitation	NS	5		15	45	25	10	5	5	14	39	35	2
22. Le schéma coopératif m'a servi à compléter les parties insuffisamment détaillées ou développées	NS	5		5	60	25	5	2	2	31	31	29	5
23. Le schéma coopératif est un schéma de référence pour corriger mon schéma individuel	NS	20		10	20	35	15	9	2	16	30	30	12
24. Le schéma coopératif est un schéma de référence pour préparer mes examens	NS	5	25	20	15	30	5	9	7	21	23	33	7
25. Le schéma coopératif permet de valider mon raisonnement	NS			5	40	45	10	5	5	9	35	42	5
Habituellement, après la phase retour													
26. J'apporte rapidement les correctifs nécessaires à mes notes et mon schéma	*	5,3	10,5	5,3	15,8	47,4	15,8	9	7	14	42	21	7
27. J'attends la bloque pour apporter les corrections nécessaires à mes notes et à mon schéma	*		35	25	5			9	39	28	12	5	7
28. Je ne reviens pas sur mes notes et mon schéma	NS	15	35	40		10		14	35	16	26	7	2

* p<0.05

NS non significatif

Tableau 5. Résultats des réponses au questionnaire sur le fonctionnement lors de cet APP expérimental par les 20 étudiants participants répartis en 3 groupes A (n=7), B (n=7) et C (n=6)

Items	p	Participants					
		---	--	-	+	++	+++
Au terme de la phase de travail personnel,							
29. Mon travail était conforme à ce que je fais habituellement	NS		14	14	29	29	14.
					14	29	57
					83		17
30. Je me suis plus investi que lors d'une séance APP habituelle	NS	14	29	43	14		
			14	43	43		
			33	50	17		
31. Je me suis moins investi que lors d'une séance APP habituelle	NS	14	57	29			
		29	14	43	14		
			17	50	17	17	
Au cours de cette phase retour particulière,							
32. J'ai confirmé les éléments (items et liens) qui me paraissent connus	*				14	43	43
							100
					17	67	17
33. J'ai clarifié les éléments (items et liens) sur lesquels j'avais un doute ou une hésitation	NS				29	14	57
					14	14	72
					33	67	
34. J'ai complété les parties qui me semblaient insuffisamment détaillées ou développées	*				14	29	57
			17		14	29	57
					33	50	
35. Le schéma coopératif m'a servi à identifier les limites de mes connaissances	NS		14		29	43	14
				16	43	14	43
					17	50	17

36. Le schéma coopératif m'a servi à clarifier les éléments (items et liens) sur lesquels j'avais un doute ou une hésitation						14	72	14
						43	67	57
						33	67	33
37. Le schéma coopératif m'a servi à compléter les parties insuffisamment détaillées ou développées						14	29	57
				17		33	33	17
				17		33	17	33
38. Le schéma coopératif est un schéma de référence pour corriger mon schéma individuel						14	14	72
				17		33	33	17
						28	43	29
39. Le schéma coopératif permet de valider mon raisonnement						29	71	71
						50	33	17
						29	71	
40. Cerner ce que je sais et les limites de mes connaissances m'a permis une meilleure intégration de mes connaissances						14	57	29
						50	50	
						17	50	33
41. Cerner ce que je sais et les limites de mes connaissances m'a forcé à un apprentissage plus en profondeur						14	57	29
						50	17	33
						17	83	
42. Cerner ce que je sais et les limites de mes connaissances m'a permis d'identifier les limites de mon raisonnement						29	43	28
						67	17	16

* $p < 0.05$, différence significative entre les 3 groupes

NS non significatif

Tableau 6. Résultats du questionnaire comparant les réponses des 20 participants répartis en 3 groupes (A, B et C) sur leur fonctionnement en APP habituel par rapport à leur fonctionnement lors de cet APP expérimental

Items	p	APP habituel					APP expérimental							
		---	--	-	+	++	+++	---	--	-	+	++	+++	
17-32. J'ai confirmé les éléments (items et liens) qui me paraissent connus	*			14	43	43	57	43			14	43	43	100
18-33. J'ai clarifié les éléments (items et liens) sur lesquels j'avais un doute ou une hésitation	*			28	43	29	17	83			29	14	17	57
19-34. J'ai complété les parties insuffisamment détaillées ou développées	*				33	67					33	67		72
20-35. Le schéma coopératif m'a servi à identifier les limites de mes connaissances	NS			14	43	29	14	14			14	29	57	57
21-36. Le schéma coopératif m'a servi à clarifier les éléments (items et liens) sur lesquels j'avais un doute ou une hésitation	*		14	13	58	14	43	43	14		14	72	14	57
22-37. Le schéma coopératif m'a servi à compléter les parties insuffisamment détaillées ou développées	*			33	33	17	17				33	67		33
23-38. Le schéma coopératif est un schéma de référence pour corriger mon schéma individuel	NS			17	67	43					14	29	57	33
25-39. Le schéma coopératif permet de valider mon raisonnement	NS			14	29	43	14	14	29	14	29	43	29	71
					33	67					50	33		17

*p<0.05

NS non significatif

7.4. Discussion Etude 1

Si nous regardons le nombre d'items et de liens des différents groupes aux différents moments de la séance retour, nous n'observons pas de différence avant la mise en commun coopérative. En ce qui concerne le schéma coopératif construit par les groupes et les schémas individuels récoltés après cette mise en commun, il apparaît que les schémas du groupe A comprennent plus d'items et de liens que les autres groupes. Nous pourrions faire l'hypothèse que ce groupe ne sachant toujours pas à ce stade du séminaire l'objectif de la recherche mais étant par ailleurs informé du caractère expérimental de celui-ci, il était plus attentif à un schéma complet afin de pouvoir répondre à d'éventuelles questions à la suite de la mise en commun.

Les exercices de délimitation du champ des connaissances ne semblent pas avoir modifié de manière significative les schémas individuels des différents groupes. En effet, nous ne pouvons pas attribuer aux exercices de délimitation l'évolution des schémas individuels entre le début et la fin de la séance retour car aucune différence significative n'est observée entre le groupe A et les 2 autres groupes à ces différentes étapes.

En ce qui concerne la façon dont les différents groupes ont marqués leur schéma durant les exercices de délimitation, quelques différences sont apparues. Tout d'abord, les étudiants du groupe C ont marqués d'avantage d'éléments douteux par rapport aux étudiants du groupe B. Nous pourrions faire l'hypothèse que ces étudiants ayant été prévenus qu'ils auraient à mettre en évidence sur leur schéma les éléments dont ils étaient sûrs et ceux dont ils doutaient, ont été incités à noter plus d'éléments même s'ils en doutaient. Ensuite, nous observons une différence lors du 2^{ème} exercice de validation au cours duquel les étudiants du groupe A ont marqué beaucoup plus d'éléments manquants ou incomplets que les groupes B et C. Ces derniers avaient réalisé avant le schéma coopératif l'exercice 1 au cours duquel ils avaient dû marquer les éléments insuffisamment détaillés ou développés. Il faut d'ailleurs remarquer que l'observation est similaire si l'injonction « validation des éléments du schéma » est donnée avant ou après la phase de travail personnel. Logiquement, ces deux groupes

avaient également ajouté moins d'éléments lors de la phase de mise en commun. Lorsque nous comparons l'ensemble des éléments marqués au cours des 2 exercices par les 3 groupes, nous n'observons plus de différences entre les groupes. Si le résultat final est donc comparable, il résulte davantage d'un processus individuel chez les étudiants interpellés pour valider leur schéma avant la mise en commun. On peut s'interroger sur l'impact de cette différence à plus long terme.

Nous pouvons conclure que la mise en commun autour du schéma coopératif semble avoir un impact positif pour la validation/correction du schéma et des apprentissages par les étudiants. Toutefois, ce processus de validation/correction concrétisé dans le schéma peut avoir impliqué l'étudiant plus ou moins activement et n'exclut pas qu'il retranscrive passivement la production des autres étudiants. Nous ne sommes donc pas à même d'identifier un impact direct des exercices de délimitation du champ des connaissances sur la production finale. Par contre, du fait de l'analyse de schémas produits avant la mise en commun de la séance retour, nous pouvons dégager, selon les consignes données aux étudiants, des différences dans le processus individuel de correction du schéma avec un effet possible sur le processus individuel d'apprentissage. Si cette consigne explicite de validation de sa propre production était itérative, ce qui inciterait les étudiants à travailler de manière personnelle leur propre base de connaissances plutôt que de compter essentiellement voire exclusivement sur le travail collectif, nous pourrions supposer un renforcement des apprentissages plus en profondeur.

A la suite de cette recherche, nous pouvons nous interroger sur la nature du processus de délimitation du champ des connaissances. Est-ce un processus complexe déjà présent chez l'étudiant qui l'applique de manière implicite dans le cadre de ses APP quelle que soit la consigne ou au contraire, ce processus peut-il être suscité, « boosté » par une consigne explicite de validation qui conduira à la délimitation du champ des connaissances ?

La vignette utilisée pour cette séance d'APP expérimental intégratif a été utilisée pour la formation de la cohorte suivante dans le cadre du module endocrinologie d'octobre 2009. Une deuxième étude a

été menée avec pour objectif de comparer l'évolution des schémas individuels au cours de la séance retour sans consignes particulières transmises aux étudiants.

ETUDE 2

7.5. Méthodes Etude 2

Participants

Un échantillon de 23 étudiants de 3^{ème} année répartis dans 3 groupes encadrés par un tuteur.

Procédure

Les étudiants ont vécu une séance APP traditionnelle comme décrit plus haut. Les schémas physiopathologiques remis aux tuteurs en début de séance ont été récoltés et photocopiés. Les schémas individuels ont été également récoltés en fin de séance retour afin de comparer si le nombre d'items et de liens présents dans ces schémas augmentaient de la même façon après le schéma coopératif lorsqu'il n'y a pas de consigne particulière qui est donnée.

Statistiques

La différence entre les 2 cohortes en termes de nombre d'items et de liens présents sur les schémas en début de séance retour a été évaluée par une Anova à un facteur. L'évolution des schémas entre le début et la fin de la séance retour en termes de nombre d'items et de liens selon les cohortes a été évaluée par une Anova en mesures répétées. Le seuil de significativité a été fixé à $p \leq 0.05$.

7.6. Résultats Etude 2

La Figure 2 illustre l'évolution du nombre d'items et de liens présents sur les schémas physiopathologiques entre le début et la fin de la séance retour pour notre groupe d'étudiants ayant participé à l'APP expérimental en mars 2009 (cohorte 1) et la cohorte d'étudiants suivante ayant suivi le séminaire APP en octobre 2009 (cohorte 2) dans le cadre de leur module endocrinologie. Ces résultats indiquent un nombre d'items et de liens significativement plus importants pour les étudiants ayant participé à l'APP expérimental dès le début de la séance

retour ($p < 0.00001$ pour le nombre d'items et de liens) et un gain nettement plus important pour ces mêmes étudiants ayant bénéficié des exercices de validation et délimitation du champ de leurs connaissances au cours de cette séance ($p = 0.021$ pour le nombre d'items ; $p < 0.001$ pour le nombre de liens).

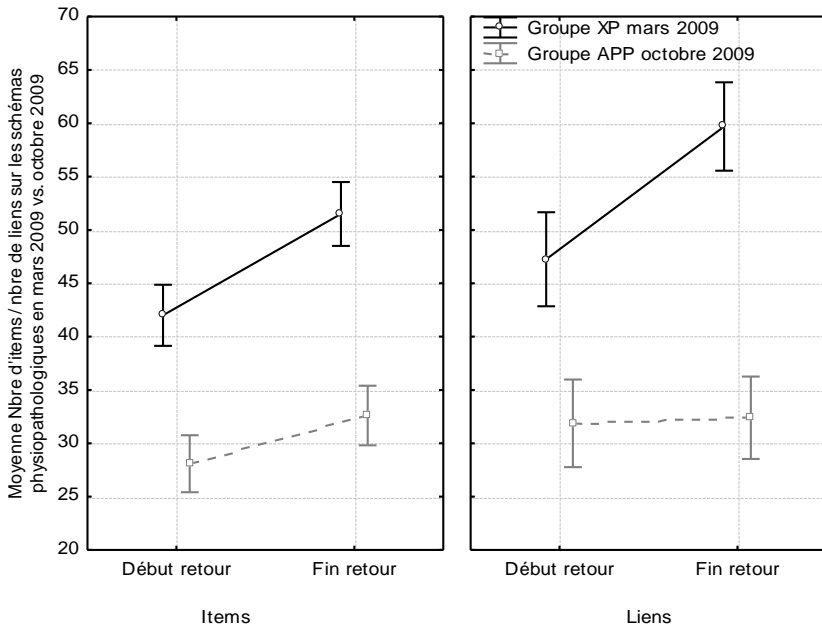


Figure 2. Moyenne ($\pm ET$) du nombre d'items (à gauche) et de liens (à droite) présents dans les schémas individuels du groupe ayant participé à l'étude 1 en mars 2009 et de la cohorte suivante ayant suivi la séance APP en octobre 2009 en début et en fin de séance retour

7.7. Discussion Etude 2

La différence de niveau de base du schéma de début de séance retour entre les 2 cohortes d'étudiants peut s'expliquer, du moins en partie, par les périodes de séance APP différentes. Effectivement, le groupe expérimental de la première cohorte avait déjà suivi l'ensemble du module endocrino-métabolisme le quadrimestre précédent l'expérimentation et avait également déjà subi l'évaluation de cette matière. Par contre, la cohorte 2 était en cours de module et avait donc investi moins de travail dans la matière dont elle n'avait pas encore une vision d'ensemble à ce stade. Une différence de connaissances de base en endocrinologie mais aussi d'expérience en termes de nombres de séminaires APP suivis peut contribuer à cette différence. Un facteur supplémentaire peut être lié à la motivation, les étudiants de l'étude 1 sachant qu'ils participaient à un séminaire APP dans lequel une procédure nouvelle serait introduite alors que les étudiants de l'étude 2 avaient simplement donné leur aval pour la communication et l'utilisation de leurs schémas.

Le gain en cours de séance qui est plus important pour le groupe expérimental de la cohorte 1 peut s'expliquer par différents facteurs. Tout d'abord, comme nous avons pu le voir dans l'étude 1, les participants à l'APP expérimental étaient des étudiants déjà plus impliqués dans la méthode APP que les étudiants non participants. Donc, un facteur motivationnel peut à nouveau jouer un rôle. Ensuite, la consigne de l'étude 1 incitait à corriger le schéma et à se poser des (bonnes) questions avant la mise en commun plutôt que simplement recopier le schéma du groupe. Enfin, les étudiants ayant déjà abordé l'ensemble de la matière et étudié ont créé un réseau de connaissance plus complet mais aussi plus structuré qui permet une meilleure appréhension du problème et une confrontation facilitée au réseau de connaissance créé.

7.8. Conclusions Générales

Cette étude avait pour objectif d'évaluer l'effet d'un travail de délimitation du champ des connaissances de l'étudiant sur son apprentissage et sur la validation de son raisonnement. La démarche suivie visait à renforcer la conscience que les étudiants peuvent avoir de leur processus d'apprentissage aux différentes étapes de l'analyse d'un cas clinique.

Un exercice de validation et délimitation du champ des connaissances avant la mise en commun collective force l'étudiant à évaluer sa propre production et par là interroge ses propres connaissances. La confrontation de son propre raisonnement à la construction du schéma coopératif et la validation « forcée » de ce raisonnement et des connaissances sous-jacentes pourraient renforcer l'apprentissage des étudiants par la prise de conscience de ses acquis et ses limites. Cet exercice forcé pourrait s'avérer particulièrement utile pour les étudiants qui sont moins actifs et participatifs lors des séances d'APP. Cet exercice individuel oblige à un traitement actif des connaissances et à la réactivation des connaissances antérieures sur lesquelles les nouvelles connaissances générées par le groupe lors de la mise en commun collective peuvent venir se greffer. La discussion collective amène alors l'étudiant à situer davantage son propre raisonnement par rapport à la démarche du groupe et, sans doute à verbaliser davantage son raisonnement et échanger avec ses condisciples. Nous n'avons pas étudié ce dernier aspect. En d'autres termes, une démarche individuelle renforcée qui précède la démarche en groupe peut catalyser à la fois l'apprentissage individuel en profondeur et pourrait contribuer à la quantité sinon la qualité des échanges en groupe.

7.9. Références

Barrows, H. (1986). The scope of clinical education. *Journal of Medical Education*, 61, 23–33.

Barrows, H., & Pickell, G. (1991). *Developing Clinical Problem Solving Skills: A Guide to More Effective Diagnosis and Treatment*. New York: Norton.

Boniver, J. (2004). Les études de médecine à l'Université de Liège: le renouveau pédagogique de la Faculté de Médecine. *Revue Médicale de Liège*, 59(12), 717–30.

Collard, A., Gelaes, S., Vanbelle, S., Bredart, S., Defraigne, J.O., Boniver, J., & Bourguignon, J.P. (2009). Reasoning versus knowledge retention and ascertainment throughout a problem-based learning curriculum. *Medical Education*, 43, 854-865.

Davis, M., & Harden, R. (1999). AMEE Medical Education Guide No. 15: Problem-based learning: a practical guide. *Medical Teacher*, 21(2), 130–140.

Elstein, A.S., Shulman, I.S., Sprafka, S.S. (1978). *Medical Problem Solving: An Analysis of Clinical Reasoning*. Cambridge, Mass: Harvard University Press.

Elstein, A.S., Shulman, L.S., & Sprafka, S.A. (1990) Medical problem solving: a ten-year retrospective. *Evaluation and the Health Professions*, 13, 5-36.

Friedman, C.P., deBliek, R., Greer, D., Mennin, S., Norman, G., Sheps, C., et al. (1990). Charting the winds of change: evaluating innovative medical curricula. *Academic Medicine*, 65, 8-14.

Gale, J. (1982). Some cognitive components of the diagnostic thinking process. *British Journal of Educational Psychology*, 52, 64-76.

Haak, D.C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased Structure and Active Learning Reduce the Achievement Gap in Introductory Biology. *Science*, *332*, 1213-1216.

Jones, M.A. (1992). Clinical reasoning in manual therapy. *Physical Therapy*, *72*, 875-884.

Langendyk, V. (2006). Not knowing that they do not know: self assessment accuracy of third-year medical students. *Medical Education*, *40*(2), 173-179.

Mamede, S., Schmidt, H.G., & Rikers, R. (2007). Diagnostic errors and reflective practice in medicine. *Journal of Evaluation in Clinical Practice*, *13*(1), 138-145.

Newble, D.I., & Clarke, R.M. (1986). The approaches to learning of students in a traditional and in an innovative problem-based medical school. *Medical Education*, *20*, 267-273.

Nickerson, R.S, Perkind, D.N., & Smith, E.E. (1985). *The Teaching of Thinking*. Hillsdale, NJ: Lawrence Erlbaum Associates INC.

Norman, G.R., & Schmidt, H.G. (1992). The psychological basis of problem-based learning: a review of the evidence. *Academic Medicine*, *67*, 557-565.

Regehr, G., & Norman, G.R. (1996). Issues in cognitive psychology: implications for professional education. *Academic Medicine*, *71*, 988-1001.

Schmidt, H.G. (1983). Problem-based learning: rationale and description. *Medical Education*, *17*, 11-16.

Schwartz, R.W., Burgett, J.E., Blue, A.V., Donnelly, M.B., & Sloan, D.A. (1997). Problem-Based Learning and Performance-Based Testing: Effective Alternatives for Undergraduate Surgical Education and Assessment of Student Performance. *Medical Teacher*, *19*, 19-23.

van Gessel, E., Nendaz, M.R., Vermeulen, B., Junod, A., & Vu, N.V. (2003). Development of clinical reasoning from the basic sciences to the clerkships: a longitudinal assessment of medical students' needs and self-perception after a transitional learning unit. *Medical Education*, 37, 966–974.

Yeung, E., Au-Yeung, S., Chiu, T., Mok, N., Lai, P. (1999). Application of Problem Based Learning Strategies to Enhance Clinical Reasoning and Self-Directed Learning Skills in a University Physiotherapy Program. In J. Marsh (Ed.), *Implementing Problem Based Learning Project: Proceedings of the First Asia Pacific Conference on Problem Based Learning* (pp.95-110). Hong Kong: The University Grants Committee of Hong Kong, Teaching Development Project.

PARTIE 3 : DISCUSSION

1. FORCES ET LIMITES DE NOS TRAVAUX

1.1. *Les participants*

Outre l'analyse longitudinale, la participation à nos différentes études sur base volontaire a pu entraîner certains biais dans les résultats obtenus. Ces étudiants, bien que représentatifs de leur cohorte sur base de leur note globale, pouvaient être animés par des éléments d'ordre motivationnel ou d'engagement dans leur apprentissage susceptibles d'avoir influencé les résultats obtenus. Toutefois, pour des raisons éthiques, il n'apparaissait pas possible de travailler sur une autre base qu'une participation volontaire. Celle-ci a été éclairée par une présentation des objectifs du travail (et, ultérieurement, des résultats) aux étudiants.

Par ailleurs, notre étude sur l'exercice de validation du raisonnement n'implique qu'un petit nombre d'étudiants. Dès lors, si nos résultats conduisent à un certain nombre de réflexions et ouvrent des pistes intéressantes, les mêmes réserves que celles formulées ci-dessus sont à garder à l'esprit et se doublent d'autres réserves liées au petit nombre.

Nos travaux ont permis d'étudier les performances pour un même outil d'évaluation du raisonnement (TCS) et un même outil d'évaluation des connaissances (QCM) sur 4 années du cursus de formation. Ces comparaisons transversales ont permis d'identifier l'évolution de ces performances pour une même matière, dans un contexte non-certificatif. De plus, ces mêmes performances ont également été évaluées longitudinalement et, cette fois, dans un contexte certificatif. Cette dernière évaluation, bien qu'elle permette de suivre un même étudiant sur 4 années de son parcours de formation, porte sur des matières différentes. Ces deux approches sont complémentaires et mènent à des conclusions cohérentes.

Nos analyses portent sur des étudiants de cohortes différentes allant de 2002 à 2009. Au fil de ces années et pour ces différentes cohortes, le nombre d'étudiants a fortement augmenté (voir plus loin) et il a donc été nécessaire d'augmenter également le nombre de tuteurs qui les encadrent, voire d'accroître le nombre d'étudiants par groupe. Lors des premières années de mise en place du PBL au sein de la Faculté de médecine, une formation systématique était suivie par chaque tuteur concernant l'accompagnement de petits groupes et les objectifs et modalités du tutorat. La vision qui prédominait était de promouvoir des tuteurs spécialistes de la méthode. Au fil des années, le recrutement des tuteurs s'est avéré plus difficile. L'équipe tutorale a comporté de plus en plus de spécialistes des disciplines et la formation des tuteurs a été de moins en moins formalisée. Ces évolutions dans la pratique du tutorat et donc dans la pratique du PBL à travers les différentes cohortes étudiées a pu interférer dans les résultats obtenus concernant les comparaisons de nos participants étudiants.

1.2. Les outils

L'évaluation du raisonnement

Dans le cadre de nos recherches, le raisonnement biomédical a été évalué à travers un Test de Concordance de Scripts et des examens oraux multidisciplinaires. Le Test de Concordance de Script est un outil spécifiquement créé pour l'évaluation du raisonnement et dont la validité a été démontrée à travers plusieurs études (Charlin et al. 2000, 2002a, 2002b ; Charlin & Van Der Vleuten, 2004). L'examen oral portant sur la discussion multidisciplinaire d'une vignette clinique fait partie des évaluations certificatives des étudiants. Cette évaluation vise principalement la mise en évidence du raisonnement de l'étudiant mais explore également les connaissances sous-jacentes à ce raisonnement. La mesure obtenue reflète donc une part de raisonnement et une part de connaissances. Par ailleurs, les modalités de mise en œuvre de cette évaluation ne sont pas toujours uniformisées et la note globale obtenue n'implique pas une proportion invariable de points accordés pour les connaissances et pour le raisonnement. Malgré ces différences, ces 2

outils d'évaluation du raisonnement mènent à une évolution comparable des performances.

Ces modalités d'évaluation du raisonnement mériteraient d'être étendues et comparées à d'autres tests de raisonnement.

L'évaluation de la capacité de délimitation du champ des connaissances

La capacité d'auto-évaluation des connaissances a été essentiellement mesurée à travers l'utilisation des degrés de certitude pour les réponses correctes et incorrectes aux tests de connaissances. Comme mentionné plus haut, dans le cadre des évaluations certificatives, pour encourager les étudiants à déterminer leur réelle auto-estimation, un bonus est accordé lorsque l'étudiant mentionne un degré de certitude élevé pour les réponses correctes et faible pour les réponses incorrectes. Inversement, il reçoit une pénalité pour un degré de certitude élevé pour une réponse incorrecte ou une certitude faible pour une réponse correcte. Le Tableau 1 ci-dessous détaille cette procédure. Ce système de cotation incite dans certains cas les étudiants à sélectionner un niveau intermédiaire de certitude afin de réduire le risque de pénalité tout en conservant un bonus suffisant (en sélectionnant systématiquement le degré 3 par exemple). Malgré cette limite, nous observons une différence significative entre les degrés de certitude pour les réponses correctes et les réponses incorrectes.

D'autre part, la confiance en soi qu'un étudiant peut avoir, varie considérablement d'un étudiant à un autre et peut avoir une influence dans la façon d'utiliser les degrés de certitude. Par exemple, un étudiant relativement peu sûr de lui pourrait ne jamais dépasser le niveau 3 de certitude alors qu'un étudiant très sûr de lui pourrait lui ne jamais descendre en-dessous du niveau 3. Malgré ces variations dans le niveau de base de confiance en soi qui impacte le choix des degrés de certitude, tous les étudiants distinguent les réponses correctes des réponses incorrectes par une sélection de degrés de certitude différents dans les 2 situations. C'est cette différence que nous avons choisi d'utiliser.

Tableau 1. Bonus et pénalité accordés selon l'utilisation des degrés de certitude

Degrés de Certitude	0 (0-25%)	1 (25-50%)	2 (50-70%)	3 (70-85%)	4 (85-95%)	5 (95-100%)
Scores/20 pour une réponse correcte	13	16	17	18	19	20
Scores/20 pour une réponse incorrecte	4	3	2	0	- 6	- 20

Il faut souligner que l'usage des degrés de certitude étant systématique dès la 1ère année pour tous les QCMs, l'exploitation de ceux-ci dans un but de recherche n'a pas été entachée par des aspects de familiarisation qui auraient pu influencer les réponses différemment entre les étudiants en début et en fin de cursus.

1.3. Le contexte de la formation

Le contexte de la réforme PBL dans lequel ces recherches ont été initiées et menées ont pu, comme déjà mentionné, guider et orienter les questions de recherche de ce travail. Notamment, l'article 4 est issu de la préoccupation des enseignants concernant l'observation à l'examen oral multidisciplinaire d'étudiants présentant un raisonnement tout à fait correct mais dont les connaissances sous-jacentes étaient lacunaires. Par contre la préoccupation du développement du raisonnement biomédical et clinique ainsi que le choix des outils de mesure sont indépendants de ce contexte. En tant qu'accompagnatrice pédagogique des réflexions concernant la mise en place des différentes modalités pédagogiques à la Faculté de médecine (notamment dans la mise en place et l'évolution des dispositifs PBL), j'ai évidemment développé un intérêt et un enthousiasme pour le dispositif mis en place. Néanmoins, ma formation de psychologue de la cognition et ma fonction de recherche en sciences de l'éducation m'ont incitée et encouragée à l'analyse et la remise en question de ces différents dispositifs. C'est avec ces 2 fonctions complémentaires et différentes que j'ai pu aborder ces travaux de recherche.

Les effets potentiellement favorables du PBL discutés dans ce travail sont essentiellement basés sur des arguments théoriques. Les effets du cursus PBL de l'Université de Liège sur l'apprentissage ne peuvent être démontrés directement car il n'y a pas d'études empiriques réalisées comparant ce cursus où PBL et enseignements magistraux sont intriqués avec d'autres cursus, notamment ceux qui sont univoques (uniquement PBL ou cours magistraux). Le point sans doute le plus important pour les recherches rapportées ici est qu'elles ont pu être menées en toute indépendance, sans attente des protagonistes enseignants que les travaux leur apportent des arguments pour ou contre le PBL ou une validation de curriculum.

2. ANALYSE DES RÉSULTATS PRINCIPAUX

2.1. *La capacité de transfert des connaissances*

Après un an et demi de cursus PBL, les étudiants sont capables de transférer les connaissances acquises dans le cadre de ces séminaires pour les appliquer dans une tâche de raisonnement proposant un contexte inédit (article 1). Un effet spécifique des séminaires PBL sur la capacité de transfert ne peut être démontré en l'absence de groupe témoin. Néanmoins, les séminaires PBL qui proposent divers exemples contextualisés d'application des connaissances, incitent les étudiants à lier les concepts qu'ils acquièrent avec leurs connaissances antérieures. Cette démarche nécessite de mettre les connaissances en relation les unes avec les autres et contribue ainsi à structurer le raisonnement. De cette façon, ces séminaires renforcent la compréhension en profondeur des connaissances. Tous ces éléments entraînent la capacité de transfert des connaissances (Norman, 2009) et la capacité de résolution de problèmes (Dixon & Brown, 2012 ; Hmelo-silver, 2004).

La capacité de transfert est fortement liée à la capacité de percevoir la similarité entre la tâche à réaliser et les expériences d'apprentissage antérieures (Dixon & Brown, 2012). Il est dès lors remarquable d'observer que cette capacité peut s'exercer tôt dans le cursus, alors que les expériences d'apprentissage antérieures ont été peu nombreuses. La capacité de transfert est aussi influencée par le contexte d'apprentissage (Smith & Vela, 2001 ; Tulving & Thomson, 1973) et a un impact sur le rappel des connaissances. Avec l'expérience, les cliniciens développent une représentation plus flexible des connaissances et peuvent abstraire les caractéristiques pertinentes des concepts. Ils sont ainsi moins influencés par les éléments contextuels de surface des différentes situations rencontrées. Les étudiants plus jeunes présentent quant à eux plus de difficultés à traiter des cas familiers (reprenant tels quels les éléments des séminaires APP et ARC pour l'amorce de la question) pour lesquels des éléments sont discordants avec le contexte d'encodage, c'est-à-dire une hypothèse et une information à évaluer parfois différentes de celles traitées lors de l'encodage créant ainsi une interférence proactive (Bernstein, Penner, Clarke-Stewart, & Roy, 2007).

Un tel phénomène pourrait expliquer les scores inférieurs obtenus face à des situations déjà rencontrées.

2.2. *L'évolution de la capacité de raisonnement durant la formation*

Une capacité de raisonnement présente dès la deuxième année et se développant avec l'expérience a pu être démontrée dans nos recherches, avec une augmentation significative de cette performance liée à l'exposition à des situations cliniques (articles 2 et 3). Nos résultats sont cohérents avec les différentes études sur le développement du raisonnement. Ces études démontrent notamment que l'application répétée des connaissances (comme à travers un enseignement de type PBL) et la confrontation à des patients réels à travers les stages notamment, soutiennent la formation des scripts et donc le développement du raisonnement (Charlin, Boshuizen, Custers, & Feltovich, 2007 ; Schmidt & Boshuizen, 1993 ; Schmidt, Norman & Boshuizen, 1990 ; Schmidt & Rikers, 2007). Plus récemment, d'autres études se sont intéressées au développement du raisonnement clinique dès le début du cursus médical. Hoff, Bestawros, Kassis et Charlin (2010) ont développé un Test de Concordance de Scripts pour favoriser l'apprentissage du raisonnement clinique dès la première année du cursus médical à l'Université de Montréal (cursus de 4 années). Selon ces auteurs, les réflexions et discussions induites par ces questions TCS formatives favoriseraient l'intégration des connaissances et développeraient le raisonnement clinique. Humbert, Johnson, Miech, Friedberg, Grackin et Seidman (2011), ont également administré un TCS à des étudiants de 2^{ème} (pré-clinique) et 4^{ème} année (fin de cursus) ainsi qu'à un panel d'experts et ont observé des résultats comparables aux nôtres. Ils concluent que les capacités de raisonnement clinique peuvent être évaluées très tôt dans le cursus médical. Nous pensons néanmoins qu'à ce stade de la formation, l'évaluation du raisonnement ne peut prétendre à l'implication du raisonnement diagnostique mais porte sur le raisonnement biomédical qui correspond mieux au niveau de développement du raisonnement à ce stade de la formation.

2.3. *La capacité de raisonnement et la rétention des connaissances factuelles*

Nous avons pu démontrer à travers nos résultats que, bien que les connaissances et le raisonnement soient intimement liés, les performances aux tests de rétention des connaissances n'évoluent pas parallèlement aux performances aux tests de raisonnement au cours du cursus médical. Dans notre étude transversale où le délai entre apprentissage et test de rétention varie entre moins d'un an et 4 années, il apparaît effectivement que la rétention diminue après un an et se stabilise ensuite (article 2). Dans notre étude longitudinale où le délai apprentissage – test de rétention est court et constant, la capacité de restitution des connaissances a tendance à augmenter mais non significativement (article 3). Bien que globalement notre quantité de connaissances augmente, c'est surtout la façon dont elles sont structurées et organisées qui évolue avec l'expérience (Boshuizen & Schmidt, 2012; Patel, Glaser, & Arocha, 2000 ; Schmidt & Rikers, 2007). Ainsi, lorsque nous évaluons les mêmes connaissances factuelles sur une même matière en dehors de tout contexte, nous pouvons observer une diminution de la rétention de ces connaissances. La trace en mémoire de ces connaissances devient de moins en moins accessible si celles-ci ne sont pas utilisées régulièrement (Bernstein, Penner, Clarke-Stewart et Roy, 2007 ; Custers, 2010). Ce qui va permettre de renforcer et restructurer les connaissances, ce sont leur occurrence d'utilisation en contexte qui permet de créer des liens avec d'autres connaissances. Ceci explique que des connaissances apprises depuis un délai plus long sans réactivation récente soient plus difficiles à récupérer que des connaissances acquises récemment ou utilisées dans la pratique quotidienne pour lesquelles des associations fortes existent (Bernstein, Penner, Clarke-Stewart, & Roy, 2007 ; Custers, 2010). Ceci explique aussi la tendance à l'augmentation de la récupération des connaissances avec la progression dans la formation qui confronte de plus en plus l'étudiant à des situations contextualisées, lors des séminaires et des stages (article 3).

Avec un test ciblant l'évaluation du raisonnement comme le Test de Concordance de Scripts (Dory, Gagnon, Vanpee, & Charlin, 2012), une corrélation entre capacité de raisonnement et performances en connaissances est observée uniquement pour les étudiants en début de cursus. Ces résultats sont cohérents avec l'évolution de la structuration des connaissances. Un étudiant novice possède des connaissances peu intégrées et un niveau d'abstraction relativement faible ; il base son raisonnement essentiellement sur les connaissances biomédicales auxquelles il peut avoir accès dans sa mémoire et les liens de causalité qu'il peut établir entre elles. Un étudiant en fin de cursus, plus expérimenté, possédant des connaissances plus intégrées et avec une capacité de généralisation plus grande, base son raisonnement sur des concepts plus abstraits, des connaissances cliniques et son expérience clinique (Jones, 1992). La capacité de raisonnement devient donc de moins en moins dépendante de la récupération consciente des connaissances factuelles biomédicales de base. L'examen oral de discussion multidisciplinaire d'un cas clinique évalue la capacité de raisonnement à travers la capacité d'identifier et de mettre en relation les éléments clés d'une situation clinique pour proposer une compréhension globale du problème. Cet examen implique également une discussion sur les connaissances sous-jacentes à ce raisonnement et exige de la part de l'étudiant de revenir aux connaissances biomédicales de base pour fonder leur raisonnement. Cette modalité d'évaluation présente donc une corrélation plus grande avec les performances aux tests de rétention des connaissances.

2.4. La capacité de raisonnement et la capacité de délimitation du champ des connaissances

Nos études ont démontré une corrélation entre les performances en raisonnement et la capacité à autoévaluer ses propres connaissances (articles 2 et 3). L'importance de s'interroger sur son propre processus de raisonnement et ses composantes a été mise en évidence notamment par Mamede et al. (Mamede, Schmidt, & Rikers, 2007 ; Mamede et al. 2012) dans le cadre du raisonnement

diagnostique. La métacognition a aussi été identifiée comme caractérisant l'expertise en raisonnement (Jones, 1992). Alors que le contrôle et la réflexion sur le processus de raisonnement dans son ensemble ont été abordés dans ces études, relativement peu d'études s'intéressent à cette capacité en lien avec les connaissances utilisées pour raisonner. Or un bon raisonnement, s'il est basé sur des connaissances erronées, ne peut conduire qu'à de mauvais résultats. La validation des connaissances est essentielle à la validation du raisonnement.

Le modèle constructiviste de l'éducation considère que l'apprenant construit ses propres connaissances à travers un apprentissage auto-régulé (Zimmerman, 2002). Le PBL s'ancre dans cette perspective en engageant l'étudiant dans un apprentissage auto-dirigé. Une habilité centrale dans ce processus d'apprentissage auto-dirigé nécessaire à l'étudiant est la capacité à s'auto-évaluer : discriminer ce qu'il connaît et ce qu'il ne connaît pas (Langendyk, 2006).

Par ailleurs, dans le cadre de l'apprentissage auto-dirigé suscité par la formation PBL, nos résultats nous ont permis de mettre en évidence l'importance du processus individuel d'auto-évaluation (article 4). Si un étudiant ne prend pas le temps d'interroger ses propres connaissances et laisse le groupe mener le travail en se contentant de baser son apprentissage sur le produit fini, validé par l'ensemble, il passe à côté du processus individuel qui lui permettra d'ancrer les nouvelles connaissances et de consolider celles-ci par une validation individuelle.

2.5. L'évolution de la capacité de délimitation du champ des connaissances

L'évolution de cette capacité avec l'expérience, chez un même étudiant, a pu être démontrée ainsi que son lien avec la réussite (article 3). Ces résultats démontrent l'importance de pouvoir distinguer ce que l'on connaît de ce que l'on ne connaît pas pour l'acquisition de nouveaux apprentissages qui sont à ancrer sur des bases de connaissances solides et pour la réussite des études.

Sans cette capacité à juger de leurs propres limites, les étudiants trouvent difficile de se donner des objectifs appropriés, d'adopter des stratégies pour atteindre ces objectifs et d'évaluer l'atteinte de ces objectifs (Fazey, 1993). Un objectif de la formation médicale devrait être de s'assurer que les étudiants développent cette capacité d'auto-évaluation qui sera essentielle pour le développement de leur apprentissage ainsi que tout au long de leur vie professionnelle (Langedyk, 2006).

2.6. L'impact d'un exercice travaillant la capacité d'auto-validation des connaissances au cours d'une séance APP

La méthodologie APP prévoit que l'étudiant remette en question ses acquis afin de venir y greffer les nouvelles connaissances à acquérir pour avoir une compréhension globale du cas clinique présenté. Cet exercice se déroulant en groupe, il n'est pas toujours évident pour les étudiants, en particulier les moins investis, de prendre le temps de s'interroger de façon individuelle sur ses propres limites de compréhension des concepts clés du cas. Habituellement, il n'y a pas d'incitation systématique de la part des enseignants à faire entrer l'étudiant dans cette démarche. L'ajout d'un exercice qui demande de faire ce travail de façon explicite et écrite, permet tout d'abord de s'assurer que chaque étudiant prend le temps de s'interroger sur ses connaissances. En outre, un tel exercice permet aux étudiants d'avoir une trace de leur évolution au cours d'une séance et oriente et structure le travail du groupe autour de ces besoins et limites perçues individuellement (article 4).

Dans le cadre d'une étude sur le raisonnement diagnostique au début des années de formation clinique, Mamede et al. (2012) ont pu démontrer l'effet bénéfique sur l'apprentissage d'une réflexion structurée attirant l'attention des étudiants de façon explicite sur les données d'une situation clinique qui vont dans le sens ou non de diagnostics. Cette étude suggère que cet exercice conduit à une compréhension en profondeur et une représentation mentale plus riche

des scripts de maladies. Selon ces auteurs, l'avantage de tels exercices est leur capacité à renforcer la restructuration des connaissances existantes. Cette restructuration des connaissances est capitale dans le développement du raisonnement clinique (Schmidt & Boshuizen, 1993 ; Schmidt, Norman & Boshuizen, 1990).

3. LE PROCESSUS DE RAISONNEMENT EN LIEN AVEC LA BASE DE CONNAISSANCES ET SON UTILISATION : UN MODÈLE

Les différents résultats discutés plus haut mettent en évidence le lien étroit entre le développement des capacités de raisonnement et l'organisation et la structuration des connaissances en mémoire. Ces éléments nous paraissent essentiels dans le cadre de la formation au raisonnement au cours du cursus médical dès les premières années précliniques. Un modèle qui met en perspective les différents facteurs impliqués dans le processus d'apprentissage du raisonnement en lien avec la base de connaissances et son utilisation nous apparaît être un outil nécessaire pour la réflexion sur les modalités de formation au raisonnement médical. Cet outil devrait être utile comme cadre de réflexion pour l'apprentissage du raisonnement lors des premières années du cursus. Il permet de constituer un socle solide pour le développement du raisonnement diagnostique et de prise en charge qui est enseigné plus tard. A ce stade plus avancé de la formation, d'autres modèles plus complexes comme celui développé par Charlin et al. (2012), prenant en compte la relation au patient et les éléments émotionnels notamment, permettent d'appréhender le raisonnement clinique en y intégrant ces dimensions essentielles à l'exercice de leur métier en contexte professionnel.

Notre modèle (Figure 1) présente le processus de raisonnement comme une série d'étapes successives, ceci afin d'en faciliter la lecture. Dans la réalité du processus, ces différentes étapes peuvent se chevaucher et se combiner.

Processus de raisonnement		Base de connaissances et utilisation	
		Novices	Experts
1	Prendre connaissance du nouveau cas clinique	<i>Cas simple</i>	<i>Cas complexe</i>
2	Activer les connaissances en MLT	<i>* principalement biomédicales</i> <i>* peu intégrées</i> <i>* reliées en fonction de la causalité</i>	<i>* biomédicales, Cliniques et expérientielles</i> <i>* intégrées</i> <i>* liens de causalité et associatifs</i>
3	Sélectionner les informations à utiliser	Hierarchisation: <i>Niveau faible</i>	<i>Niveau élevé</i>
4	Discriminer les éléments connus vs douteux ou ignorés, à clarifier Abstraire le connu	Discrimination: Compréhension: Abstraction: <i>Niveau faible</i>	<i>Niveau élevé</i>
5	Se Représenter le problème	<i>Analytique et Contextualisée</i> Catégorisation: <i>rigide</i>	<i>Globale et abstraite</i> <i>souple</i>
6	Identifier le (les) script(s)/réseau(x)	<i>Processus lent</i>	<i>Processus rapide</i>
7	Sélectionner le script	<i>Analyse Séquentielle</i>	<i>Analyse Parallèle</i>
8	Adapter le script avec les nouvelles données		
Valider le raisonnement appliqué au cas clinique		Métacognition: <i>Niveau faible</i>	<i>Niveau élevé</i>

Figure 1. Modèle du processus de raisonnement en lien avec l'utilisation de la base de connaissances

Etape 1 : La *Prise de connaissance et le décodage initial* du nouveau cas clinique

Le point de départ du processus est la confrontation à une situation clinique. En début de cursus, des cas simples sont généralement proposés aux étudiants novices ; ils permettent avant tout aux étudiants de se familiariser avec les éléments contextuels qui traduisent en situations pratiques les connaissances des sciences de base et biomédicales acquises jusque-là de façon linéaire, essentiellement à travers des cours disciplinaires juxtaposés. Ces cas ou vignettes cliniques, bien que correspondant à des situations réalistes, ne prêtent généralement pas à discussion diagnostique; ils ne contiennent pas d'ambiguïté et sont relativement bien définis. Au fur et à mesure, on confrontera l'étudiant à des situations de plus en plus complexes, impliquant des éléments contextuels moins univoques, des éléments dont la pertinence peut être ambiguë, donc, des problèmes généralement mal définis. Dans ce sens, La Rochelle et al. (2012) recommande d'introduire petit à petit des éléments contextuels plus authentiques (comme des vidéos de patients réels) impliquant la gestion et l'assimilation d'informations complexes.

La mise en place des séminaires PBL vise à confronter très tôt les étudiants aux problèmes contextualisés et les entraînent ainsi au processus de raisonnement par une augmentation progressive de la complexité des cas traités partant des séances d'AIC (Apprentissage de l'Intégration des Connaissances) allant jusqu'aux séances d'ARPC (Apprentissage de la Résolution de Problèmes complexes).

Etape 2 : L'*activation des connaissances en mémoire à long terme*

Les connaissances des jeunes étudiants en médecine sont essentiellement constituées de connaissances en sciences de base, de connaissances biomédicales et de leur expérience personnelle avec la médecine au sein de leur famille. Ces connaissances sont relativement peu intégrées les unes avec les autres mais plutôt juxtaposées ou reliées

par des liens de cause à effet (Groves, 2008 ; Higgs, Jones, Loftus, & Christensen, 2008).

Les connaissances augmentent avec l'expérience mais surtout elles se diversifient. Aux connaissances de base et biomédicales acquises dans les livres, viennent s'ajouter les connaissances cliniques et leur première expérience clinique de stage puis toute leur expérience professionnelle (Diemers et al. 2011). Ces connaissances sont intégrées les unes aux autres et reliées par des liens plus complexes que les liens de cause à effet ; des relations associatives selon le contexte peuvent aussi être présentes, par exemple. Avec l'expérience, un réseau de connaissances élaboré, hiérarchisé et de mieux en mieux organisé se développe (Eylon & Reif, 1984 ; La Rochelle et al. 2012).

Les résultats de nos études qui démontrent une augmentation des performances en raisonnement avec l'expérience (articles 2 et 3) renforcent ce modèle. Les corrélations observées entre performances en raisonnement et rétention des connaissances, principalement en début de cursus, sont également cohérentes avec une structuration et une utilisation de la base de connaissances qui évolue avec l'expérience. Le processus d'élaboration sur base des informations présentées dans les situations cliniques et la recherche de signification aident à activer les connaissances pertinentes supplémentaires et renforcent l'intégration et la réorganisation des connaissances préexistantes (Mamede et al. 2012).

Etape 3 : La *sélection des informations*

Dans une situation notamment clinique, toutes les informations n'ont pas la même valeur. Certaines sont essentielles, d'autres plus accessoires ou encore non pertinentes. L'identification des éléments clés du cas détermine la qualité du raisonnement. Alors que les cliniciens expérimentés sélectionnent d'emblée les informations pertinentes qui leur permettront de renforcer ou éliminer leurs hypothèses, les jeunes étudiants analysent les informations de façon séquentielle sans hiérarchisation préalable (Groves, 2008).

Nous n'avons pas directement travaillé cette étape dans nos travaux. Néanmoins, il a été démontré dans différentes études que la qualité du raisonnement dépend de la reconnaissance précoce des éléments (vignette) ou signes cliniques (anamnèse et examen) pertinents (Kalantunga, 2001 ; Leblanc, Norman, & Brooks, 2001 ; Leblanc, Dore, Norman, & Brooks, 2004 ; Norman, 2005). Kassirer (1983) propose d'initier le recueil d'information de façon logique par rapport à la plainte du patient plutôt que favoriser la recherche exhaustive des données sans hiérarchisation. Dans ce sens, une des premières consignes de l'APP à l'ULg après la lecture de la vignette est de proposer d'emblée un résumé du problème présentant les données qui apparaissent comme essentielles à la situation.

Etape 4 : La *distinction*: éléments connus vs douteux ou ignorés, à clarifier et l'*abstraction* du connu

Cette étape est probablement celle dont l'importance a pu être le plus largement démontrée par les différentes études qui ont été menées dans ce travail. Le clinicien va devoir confronter les éléments de la situation clinique à ses connaissances en mémoire pour en abstraire les concepts essentiels à la résolution du problème. Le niveau d'abstraction et de structuration des concepts en mémoire détermine le niveau de compréhension de ce même concept (Groves, 2008). Cette étape est cruciale car elle va déterminer la base sur laquelle le clinicien va se fonder pour sa recherche de solution. Si le clinicien ne peut pas distinguer les éléments qu'il maîtrise et dont il est sûr par rapport aux connaissances qu'il ne maîtrise pas, il risque d'appuyer son raisonnement sur des connaissances erronées qui peuvent conduire à des erreurs diagnostiques ou de prise en charge qui pourraient avoir de graves conséquences. Un clinicien qui sait qu'il ne maîtrise pas l'un ou l'autre élément de la situation pourra être prudent dans sa recherche de solution et envisagera la recherche d'informations auprès des ressources disponibles (livres de référence, collègues plus expérimentés, etc.). Un jeune étudiant aura plus de difficulté à discriminer ce qu'il connaît de ce qu'il ne connaît pas avec un niveau plus faible d'abstraction et de compréhension des concepts sous-jacents aux situations cliniques.

Nos études ont mis en évidence le lien entre la capacité de distinguer les connaissances maîtrisées de celles qui ne le sont pas et les performances en raisonnement (articles 2, 3 et 4). Une de nos études a montré l'importance d'inciter les étudiants à avoir une réflexion sur la maîtrise des connaissances lors d'une séance PBL et l'influence de cet exercice sur la validation du processus de raisonnement. Comme nous l'avons déjà évoqué, très tôt au sein d'un cursus PBL, les étudiants sont capables de transférer les connaissances acquises dans le cadre d'une mise en situation contextualisée. Toutefois, nos travaux ont permis de démontrer que les étudiants plus jeunes sont davantage influencés par le contexte d'apprentissage que les cliniciens expérimentés (article 1).

Etape 5 : La *représentation* du problème

Le jeune étudiant en médecine aura une représentation du problème qui comprend les éléments de contextualisation du cas présenté et une vision très analytique du problème qui juxtapose les différents éléments constitutifs du cas. Le clinicien expérimenté, quant à lui, aura une représentation globale du problème avec une vue abstraite et hiérarchisée des concepts sous-jacents à la situation. Les novices catégorisent donc le problème à un niveau superficiel alors que les experts sont capables de fouiller à un niveau bien plus profond de la structure du problème (La Rochelle et al. 2012 ; Sweller, 1988). Une représentation analytique et contextualisée du problème va rendre plus difficile le travail de recherche de solution correspondante en mémoire, par rapport à une représentation globale et abstraite qui permettra une catégorisation plus souple pour la recherche de correspondance. Dans la même ligne, Askew, Manthey et Mahler (2012) identifient les problèmes de représentation d'une situation clinique comme source d'erreur diagnostique.

Dans le cadre des séminaires PBL à la Faculté de médecine de Liège, les consignes de schématisation des vignettes cliniques renvoient à cette idée de représentation du problème sous une forme hiérarchisée. Cet exercice entraîne l'étudiant à représenter les données d'une situation clinique sous la forme d'un réseau reliant les éléments

du cas par des liens de cause à effet, notamment dans les premières années du cursus. Habituellement, il est conseillé aux étudiants d'utiliser les termes médicaux appropriés ou le concept abstrait sous-jacent à un ensemble d'observations afin d'obtenir une représentation du problème qui permet de mettre en avant les structures profondes de la situation. Cela leur permet de donner plus de sens à ces éléments et une compréhension plus en profondeur des concepts impliqués. Ceci facilite la mise en correspondance avec une éventuelle situation comparable stockée en mémoire (Ark, Brooks et Eva, 2007).

Etape 6 : L'identification du (des) scripts / réseau(x)

La représentation du problème va permettre d'activer un ou plusieurs scripts (ou réseaux plus simples chez les novices) de situations similaires rencontrées précédemment et pour laquelle une solution ou une explication avait déjà pu être envisagée. La rapidité d'identification dépendra de la qualité de la représentation du problème comme décrit ci-dessus ainsi que de la façon dont ces scripts ou réseaux de connaissances sont structurés, hiérarchisés et organisés en mémoire (Groves, 2008). Plusieurs scripts ou réseaux peuvent être activés par une même situation clinique car un recouvrement important entre la représentation du problème et les éléments des scripts et réseaux sont identifiés. Des éléments peuvent être communs à plusieurs scripts et ne pas être exclusifs à une seule pathologie. Il devient alors nécessaire d'identifier dans quelle mesure les éléments du script correspondent à l'histoire du patient (Charlin, Boshuizen, Custers, & Feltovich, 2007). Une recherche d'information supplémentaire pourra être nécessaire pour sélectionner le script le plus pertinent voire l'adapter si celui-ci ne contient pas encore toutes les variantes possibles de situations et solutions.

Nous avons pu démontrer à travers nos études (article 2) que les scripts des étudiants en fin de cursus concordent davantage aux scripts des experts que ceux des étudiants en début de cursus.

Etape 7 : La sélection du script

Comme le présente la Figure 2 ci-dessous, la recherche et la sélection du script se font de manière séquentielle pour les novices qui traitent une seule hypothèse à la fois et les renforcent ou les éliminent les unes après les autres jusqu'à les avoir toutes envisagées pour faire leur sélection. Les experts, quant à eux, vont aborder plusieurs scripts en parallèle et éliminer et renforcer les uns et les autres de façon concomitante jusqu'à aboutir plus rapidement à une solution qui se dégage même si toutes les possibilités n'ont pas encore été traitées jusqu'au bout.

Le raisonnement clinique

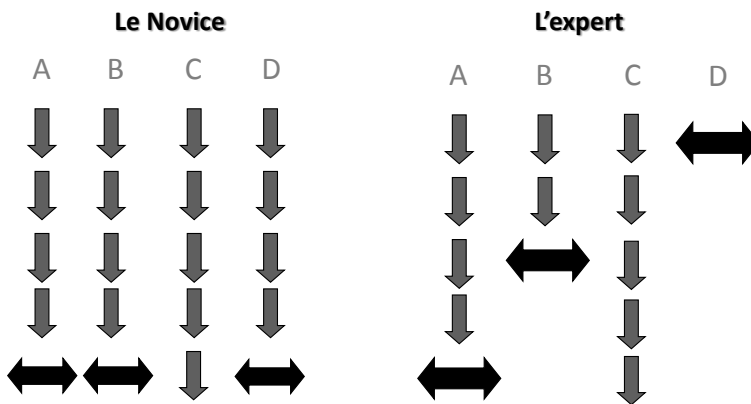


Figure 2. Le raisonnement clinique : comparaison du processus du novice et de l'expert

Etape 8 : L'adaptation du Script

Il est rare, en particulier pour les novices, d'avoir accès d'emblée à la solution pour un problème en tout point semblable à une situation précédemment rencontrée. Lorsque nous ne possédons pas toutes les connaissances nécessaires ou lorsque celles-ci sont encore floues, les recherches effectuées dans le cadre de l'analyse du nouveau cas

peuvent venir compléter ou renforcer l'état de nos connaissances. Dans ces situations, les connaissances antérieures disponibles permettent de faire face aux nouvelles informations issues de la situation inédite à laquelle on est confronté, par l'activation de scripts plus ou moins similaires. Les nouvelles données peuvent ensuite être assimilables dans la structure de connaissances existantes. Ces nouvelles connaissances intégrées aux connaissances antérieures viennent provoquer des réorganisations internes. Le script ou réseau existant est donc adapté ou complété pour venir y intégrer les nouvelles connaissances acquises. Le réseau de connaissances est alors enrichi et restructuré pour former un script plus complet.

L'enseignement PBL est fondé sur ce principe d'activation des connaissances acquises pour venir y greffer les nouvelles connaissances (Schmidt, 1983, 1993). Dans la perspective constructiviste du PBL, l'essentiel de l'apprentissage est la confrontation à la dissonance entre les connaissances antérieures et le cas concret qui est présenté comme opportunité d'apprentissage (Taylor & Hamdy, 2013). Lorsque nous sommes confrontés à quelque chose de nouveau, nous comparons ce que nous voyons avec ce que nous connaissons déjà et réfléchissons à la différence entre les deux. L'apprentissage prend du sens lorsque nous établissons des connexions entre les informations nouvelles et les connaissances antérieures (Regan-Smith et al. 1994). Suite à ces connexions, les informations sont triées et rangées de façon à leur redonner du sens et, enfin, les connaissances sont restructurées dans de nouveaux réseaux plus complets (Norman et al. 2006), ainsi se construisent petit à petit des réseaux de connaissances de plus en plus riches, organisés et hiérarchisés.

La validation du processus de raisonnement appliqué au cas clinique :

La conscience que le clinicien peut avoir de ses acquis et ses limites ainsi que sur le caractère probabiliste de son raisonnement est essentielle pour le contrôle qu'il doit avoir sur son processus de raisonnement (Mamede, Schmidt & Rikers, 2007). La qualité du

raisonnement dépend de la façon dont le clinicien gère la coordination et l'intégration des connaissances et du processus de raisonnement en utilisant la métacognition (Mamede, Schmidt, & Rikers, 2007). Lorsque le clinicien n'obtient pas le résultat escompté par rapport à une option prise, il doit pouvoir revenir sur son processus de raisonnement pour en reconsidérer les choix et tenter de comprendre où sont les failles de son processus de raisonnement. Les études notamment de Mamede et al. (2007, 2012) témoignent de l'intérêt d'une analyse réflexive du clinicien sur son propre processus de raisonnement. Ces études ont été menées principalement autour de la réflexion sur les erreurs de raisonnement. Les erreurs de raisonnement peuvent être liées à un problème à chacune des étapes du processus et seront abordées ultérieurement.

Dans nos travaux, l'exploitation des degrés de certitude, témoins implicites d'une validation (articles 2 et 3) et les effets d'une consigne explicite de validation individuelle du raisonnement (article 4) concourent à montrer que les capacités métacognitives de validation sont liées à la capacité de raisonnement et à la réussite.

3.1. Les limites du modèle

Notre modèle s'inscrit dans les théories cognitives du traitement de l'information. Ces théories attribuent une grande valeur aux processus mentaux des cliniciens. Cette vision du raisonnement clinique implique que l'apprentissage du raisonnement clinique soit fortement lié à l'acquisition d'un réseau de connaissances hautement organisé et que la notion d'expérience y joue un rôle capital (Durning, Artino, Schuwirth, & Van Der Vleuten, 2013). Il ne tient donc pas compte des éléments spécifiques liés à la situation et aux interactions avec le patient et les autres membres de la famille ou de l'équipe médicale (Durning & Artino, 2011). Il ne tient pas compte non plus des émotions qui ont une influence énorme dans la façon d'appréhender un cas, dans les éléments qui seront identifiés et pris en compte dans le cadre de prise en charge d'une situation particulière (McConnell & Eva, 2012).

Bien que nous soyons conscients des limites de ce modèle et particulièrement pour l'appréhension du raisonnement clinique des cliniciens plus expérimentés, il nous semble apporter des repères appropriés et efficaces pour penser l'enseignement, l'évaluation et la remédiation du raisonnement en début de cursus médical.

« No single theory of learning (or reasoning) can fully and accurately represent the underlying mechanisms of complex cognitive processes. »
(Durning et al. 2013).

4. LES ERREURS DE RAISONNEMENT

Les erreurs de raisonnement ont principalement été étudiées à travers les erreurs de diagnostic. La plupart des difficultés sont de nature cognitive (Audetat et al. 2012). Les études sur les erreurs diagnostiques mettent en évidence que les mécanismes par lesquels les erreurs diagnostiques apparaissent peuvent notamment être liés au processus de raisonnement ainsi qu'aux connaissances et leur utilisation. Les erreurs diagnostiques peuvent être dues à des connaissances inadéquates, à une collecte des données imprécise ou à une analyse incorrecte des données ; l'analyse des données correspond à l'interprétation des informations cliniques pertinentes (Friedman, Connell, Olthoff, Sinacore, & Bordage, 1998).

Ces éléments sont en lien avec le modèle proposé ci-dessus. Un travail de formation au raisonnement biomédical et puis clinique basé sur ce modèle pourrait dès lors consister à analyser le raisonnement pour le cautionner quand il est correct mais aussi identifier les erreurs de raisonnement et y remédier. En effet, ces erreurs sont souvent identifiées tard dans le développement professionnel du futur clinicien et leur remédiation est trop peu envisagée lors de l'apprentissage.

Audetat, Laurin et Sanche (2011a) ont proposé une taxonomie des difficultés de raisonnement clinique dont un certain nombre sont également liées aux connaissances et processus de raisonnement. D'autres difficultés liées aux attitudes, à la communication et aux facteurs relationnels sont mises en évidence mais ne seront pas abordées dans le cadre de cette discussion. En effet, de la même manière que nous avons opté pour ne pas intégrer dans notre approche précoce du raisonnement les éléments relationnels et émotionnels tels qu'envisagés par Charlin (2012) au terme de la formation, nous ne discuterons pas les erreurs qui impliquent ces mêmes aspects. A travers différentes publications, Audetat, Laurin et Sanche (2011b) ont mis en évidence 5 erreurs potentielles de raisonnement en lien avec les étapes du raisonnement diagnostique pour lesquelles des pistes de remédiation ont été envisagées pour les jeunes cliniciens. Les erreurs identifiées et remédiations proposées sont résumées ci-dessous. Nous soulignerons

ensuite les aspects convergents et complémentaires qui ressortent de nos travaux.

- Difficultés au niveau du recueil initial des données et génération d'hypothèses (Audetat et al. 2011b) : Il est proposé que ces erreurs soient liées à un déficit en connaissances, une difficulté à se faire une représentation du problème, une analyse séquentielle du problème, ou la méconnaissance du processus hypothético-déductif.

Les stratégies de correction proposées pour ces difficultés sont : faire expliciter le processus de raisonnement, verbaliser sa représentation du problème, développer la génération précoce d'hypothèses, développer l'identification d'indices discriminants.

- Difficultés au niveau du traitement et raffinement des hypothèses – fermeture prématurée (Sanche, Audetat, & Laurin, 2012a) : Le clinicien se centre très rapidement et exclusivement sur une seule hypothèse. Outre des explications liées à la personnalité du médecin ou aux circonstances cliniques, il est proposé d'expliquer ces difficultés par un déficit en connaissances ou une absence d'intégration de l'ensemble des éléments pertinents du cas.

Les stratégies de corrections proposées sont : amener l'étudiant à réfléchir aux éléments impliqués dans son raisonnement, stimuler la génération de différentes hypothèses, encourager l'étudiant à justifier ses hypothèses et encourager l'organisation des connaissances.

- Difficultés au niveau du traitement et raffinement des hypothèses – priorisation (Laurin, Audetat, & Sanche 2012a) : Des éléments contextuels ou liés à la personnalité ou l'expérience de l'étudiant sont mentionnés pour expliquer ces difficultés. A nouveau l'organisation des connaissances et leur potentiel à identifier les éléments clés sont proposés comme causes à ces difficultés.

Les stratégies de remédiation proposées sont : faire expliciter les éléments impliqués dans le processus de raisonnement et faire réfléchir aux éléments pris en compte en priorité et amener à en considérer

d'autres et enfin encourager à développer ses connaissances en contrastant les éléments clés qui permettent de distinguer les pathologies entre elles.

- Difficultés à l'étape de l'élaboration d'un portrait global de la situation clinique (Laurin, Audetat, & Sanche, 2012b) : Les éléments d'explication proposés concernent ici davantage l'inexpérience de l'étudiant, sa personnalité, la capacité à gérer le contexte psychosocial, les émotions, la communication, etc.

Les stratégies de remédiation proposées sont : au-delà des techniques de communication et de vision du patient, il est proposé de faire réfléchir sur les liens entre les différents éléments de la situation clinique, de faire construire un schéma ou une carte conceptuelle de la situation clinique et d'en discuter avec l'apprenant.

- Difficultés à l'étape de l'élaboration du plan d'intervention (Sanche, Audetat, & Laurin, 2012b) : Les éléments d'explication proposés sont un déficit en connaissances, une difficulté à intégrer les données et les nouvelles informations, et enfin les influences du contexte, des circonstances cliniques ou des difficultés liées à la perception du patient.

Les stratégies de remédiation proposées sont : faire expliciter le raisonnement, inciter l'étudiant à se positionner, encourager les lectures comparées entre les différentes options d'investigations.

Nos travaux ont démontré l'importance de la base de connaissances pour les performances en raisonnement et notamment en début de cursus (article 2 et 3). Nous avons également mis en évidence l'importance de distinguer, parmi ses connaissances, celles qui sont maîtrisées de celles qui ne le sont pas (article 2, 3 et 4). Ces 2 éléments interviennent dans les erreurs liées aux déficits en connaissances. L'importance de la base de connaissances et de sa validation démontrées dans nos différents travaux sont identifiées comme essentielles dans ces différentes étapes du processus du raisonnement diagnostique. Par ailleurs, notre étude sur le transfert ainsi que le modèle présenté insiste sur l'importance de la représentation du problème dans le processus de raisonnement. Nous

avons également abordé dans notre modèle l'analyse séquentielle et peu hiérarchisée des novices dans leur processus de raisonnement par rapport aux experts. L'absence d'intégration des différents éléments du cas renvoie aussi à la représentation du problème et à l'importance de faire schématiser les liens que l'étudiant peut faire entre les différents éléments constituant le cas. Tous ces éléments ont été identifiés comme sources d'erreur dans les étapes du raisonnement diagnostique.

Ces différentes publications concernent le développement et les erreurs du raisonnement diagnostique pour les étudiants en fin de formation clinique. Néanmoins, remarquons que la plupart des difficultés énoncées et leur remédiation peuvent tout à fait s'appliquer au développement du raisonnement biomédical dès les premières années du cursus. Ces erreurs et remédiations impliquent pour la plupart les éléments présentés dans notre modèle. Le Tableau 2 ci-dessous répertorie les différentes erreurs de raisonnement et les remédiations proposées par l'équipe Audetat, Laurin et Sanche en lien avec les différentes étapes de notre modèle. Ce Tableau permet de mettre en évidence que les remédiations proposées peuvent être d'application pour déjà travailler à ces difficultés dès le début de la formation au raisonnement biomédical.

Tableau 2. Le processus de raisonnement, les erreurs et les pistes de remédiation

Les étapes du processus de raisonnement en lien avec l'utilisation des connaissances	Les mécanismes d'erreurs de raisonnement identifiés dans ces 6 publications	Les pistes de solutions potentielles proposées dans ces 6 publications
1. Confrontation au cas clinique	Inexpérience dans la gestion de la complexité du contexte authentique	Augmenter et diversifier la confrontation à des situations cliniques
2. Activation des connaissances en MLT	Déficit en connaissances Absence d'intégration	Encourager le développement et l'organisation des connaissances Faire réfléchir sur les liens entre les différents éléments de la situation clinique
3. Sélection des informations	Difficulté à identifier les éléments clés du cas	Développer l'identification d'indices pertinents Faire réfléchir aux indices considérés en priorité
4. Distinction éléments connus vs douteux ou ignorés, à clarifier et Abstraction du connu	Connaissances inadéquates	Faire construire un schéma ou une carte conceptuelle de la situation clinique et en discuter avec l'apprenant
5. Représentation du problème	Difficulté de se faire une représentation du problème	Verbaliser sa représentation du problème
6. Identification du (des) script(s)/réseau(x)	Analyse séquentielle du problème	Développer la génération précoce d'hypothèse
7. Sélection du script	Méconnaissance du processus hypothético-déductif	Stimuler la génération de différentes hypothèses et leur justification
Validation du processus de raisonnement		Faire expliciter le processus de raisonnement Encourager la lecture comparée

Un processus d'identification des connaissances maîtrisées versus celles qui ne le sont pas a rarement été abordé dans les études sur l'origine des erreurs de raisonnement. Or un déficit en connaissances, une difficulté à identifier les éléments clés et à se représenter le problème sont directement liés à la maîtrise des

connaissances ainsi que leur organisation et accès en mémoire. Développer cette capacité métacognitive chez l'étudiant est essentiel pour la validation du processus de raisonnement et pour lui permettre d'identifier la plupart des failles de ce processus. C'est un apport original de nos travaux.

Les travaux d'Audetat, Laurin et Sanche, mettent quant à eux, plus particulièrement en évidence l'importance du contexte émotionnel et les caractéristiques personnelles qui interviennent dans la qualité du processus de raisonnement.

5. EVOLUTION DU PBL A LA FACULTÉ DE MÉDECINE DE L'ULG ET PISTES D'ADAPTATION DE LA FORMATION

5.1. *Le PBL à la Faculté de Médecine de l'ULg depuis 2011*

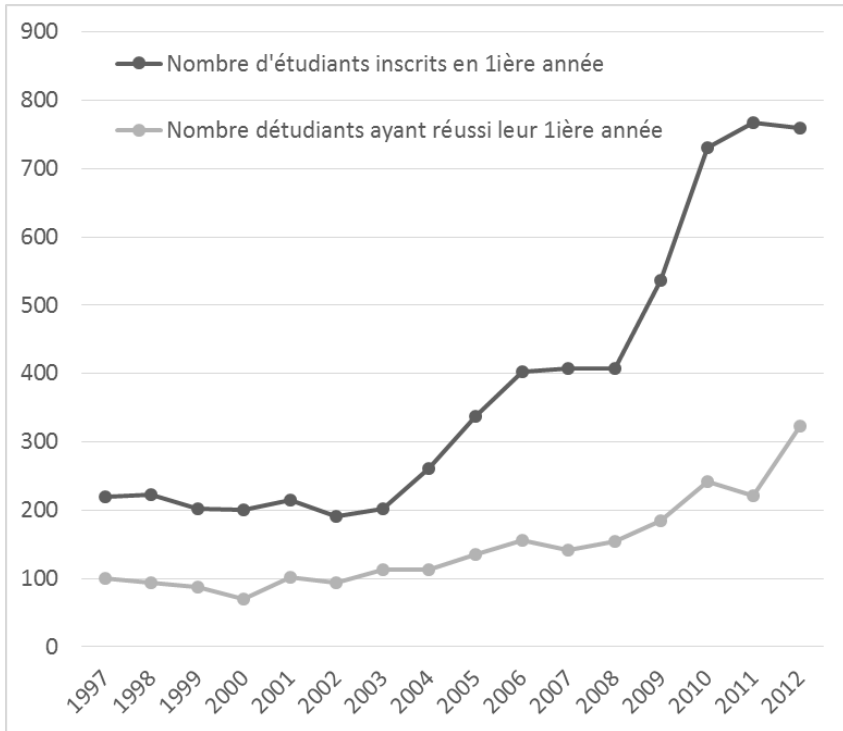


Figure 3. Nombre d'étudiants inscrits et nombre de réussites en première année à la Faculté de Médecine de l'ULg de 1997 à 2012

Comme expliqué dans l'introduction, dans les années 2000 durant lesquelles le nombre d'étudiants était limité par les concours d'admission et le contingentement (Figure 3), une opportunité était offerte pour mettre en place un apprentissage actif en petits groupes tel que les séminaires PBL. Cette modalité de formation a ainsi pris une grande place dans le cursus des futurs médecins à l'Université de Liège. Or, ces dernières années, la Faculté de médecine a vu son nombre d'étudiants inscrits augmenter de façon très importante (Figure 3) sans

une croissance parallèle des ressources matérielles ni des moyens humains disponibles. La Faculté de médecine s'est alors trouvée contrainte d'adapter ses modalités pédagogiques en diminuant drastiquement la place du PBL dans les premières années du cursus. Notamment en 2^{ème} et 3^{ème} années, les séminaires d'Apprentissage Par Problèmes ont été considérablement restreints pour laisser plus de place aux cours magistraux nécessitant moins de ressources. Actuellement et depuis 3 années, trois vignettes cliniques sont proposées aux étudiants lorsqu'ils abordent la structure et le fonctionnement d'un nouveau système anatomique. Les étudiants sont répartis en 3 groupes, chaque groupe travaillant plus particulièrement une de ces vignettes. Ces groupes sont divisés en sous-groupes de 8 à 10 étudiants se réunissant une première fois sans tuteur ou encadrant afin de lire, résumer, identifier et proposer des hypothèses explicatives aux phénomènes présentés dans la vignette. Une proposition de schéma physiopathologique est envoyée aux enseignants responsables à la fin de cette première séance. Les étudiants disposent ensuite de 2 demi-journées de travail afin de compléter et corriger ce schéma avant une séance retour en grand amphithéâtre avec l'ensemble des étudiants (les sous-groupes d'étudiants ayant travaillé sur la même vignette ainsi que les groupes d'étudiants ayant travaillé sur les autres vignettes). Ces séances retour sont animées par les enseignants responsables qui interrogent les sous-groupes ayant travaillé sur la vignette en question et reconstruisent ensemble un schéma physiopathologique de référence, approfondissant et éventuellement réexpliquant les différents concepts et liens sous-jacents. Jusqu'à présent, la cohérence formation-évaluation a pu être préservée : les étudiants passent au terme de chaque quadrimestre un examen basé sur l'analyse d'une vignette multidisciplinaire et une discussion avec un jury d'enseignants. Ceux-ci qui étaient initialement au nombre de 5 (baccalauréat) ne sont plus que 3 pour les raisons de contrainte de nombre évoquées plus haut. Par ailleurs, les études médicales passant de 7 à 6 années, d'autres modifications concernant la structure des études ont été indispensables.

Malgré ces changements structurels, le Collège d'enseignants de 2^{ème} et 3^{ème} années souhaite conserver les lignes directrices qu'il s'était fixées. C'est-à-dire :

- organiser l'enseignement sur un mode multidisciplinaire, par modules basés sur les systèmes anatomiques plutôt que sur les disciplines, définis par les enseignants réunis en collège ;
- privilégier les méthodes pédagogiques qui visent à stimuler une participation active de l'étudiant à son apprentissage ;
- organiser l'évaluation des connaissances et du raisonnement (examens) sur un mode multidisciplinaire.

Les mesures réalisées dans le cadre de notre travail ont essentiellement été récoltées avant 2011 et n'ont donc pas été influencées par ces nouvelles modalités. Néanmoins, le modèle proposé et les conclusions sur les bénéfices du PBL pour le développement du raisonnement biomédical et puis clinique devraient pouvoir servir de base à la réflexion sur des modalités pédagogiques adaptées aux nouvelles conditions structurelles.

5.2. Propositions d'adaptations pour les modalités de formation PBL à la Faculté de Médecine

Les deux nouveaux défis majeurs de la Faculté de médecine nous semblent être d'une part, optimiser les séances PBL qui ont été fortement réduites afin d'en conserver les principaux bénéfices et proposer des modalités d'évaluation du raisonnement en adéquation avec cette nouvelle modalité PBL et le nombre d'étudiants croissant.

Dans les choix qui seront posés, il serait souhaitable de prendre en compte nos travaux de recherche qui ont mis en évidence l'importance d'inciter davantage l'étudiant à réfléchir à ses propres acquis et ses limites, notamment dans le cadre des séances PBL. Ceci permet de renforcer la validation du processus de raisonnement et favorise l'apprentissage. Bien que les principes de base du PBL invitent à cet exercice, un travail plus formalisé de délimitation du champ du connu et de l'ignoré au cours de ces séances devrait être mis en place de

façon systématique afin de renforcer le potentiel d'apprentissage du raisonnement des séances PBL. En outre, la démarche essentielle du positionnement critique vis à vis de son propre raisonnement sera ainsi entraînée et plus facilement poursuivie par l'étudiant, une fois devenu professionnel sur le terrain.

Optimiser les séances PBL en 2^{ème} et 3^{ème} année

En ce qui concerne la nouvelle modalité PBL, nous pouvons retenir de ces travaux que l'essentiel est de faire réfléchir les étudiants sur leur base de connaissances antérieures : quelles sont celles qui sont maîtrisées et celles qui ne le sont pas encore mais qui sont nécessaires pour avoir une représentation globale du problème et des questions posées par le cas présenté. Parmi les autres éléments à mettre en avant pour l'apprentissage du raisonnement, nous avons discuté l'importance de travailler l'identification des éléments clés d'une situation et de les hiérarchiser ; faire expliciter et justifier le raisonnement de l'étudiant, les liens qu'il établit entre les différents éléments de la situation ; faire schématiser le raisonnement par chacun et inciter les étudiants à comparer leur représentation aux autres schématisations d'une même situation ; et enfin diversifier la confrontation à des cas cliniques.

Dans la modalité PBL actuelle de début de cursus, ces exercices sont suggérés, notamment en demandant aux étudiants de remplir un tableau qui reprend les éléments du cas, les hypothèses d'explication et leur justification. Il leur est aussi demandé d'indiquer, sur un document en fin de séance aller, les points à clarifier durant la phase de travail personnel. Cependant ce document n'est à aucun moment visible par l'enseignant et l'exercice collectif ne permet pas nécessairement à chaque étudiant de se situer par rapport à lui-même. Un seul schéma physiopathologique est demandé pour l'ensemble du groupe et ces schémas sont retravaillés et confrontés à ceux produits par les autres groupes, lors d'une séance retour en amphithéâtre avec l'ensemble des étudiants de la cohorte. Cette modalité répond potentiellement aux recommandations ci-dessus et reste praticable compte tenu des conditions et ressources actuelles. Cette modalité risque toutefois de réduire le travail individuel de réflexion en raison d'une « prise de pouvoir » par le groupe qui impose un fonctionnement dans lequel le

rythme propre à chacun est difficilement respecté. Qui plus est, l'étudiant généralement préoccupé par la réussite plutôt que par l'apprentissage, laissera faire pour autant que le groupe produise le schéma coopératif qui est faussement perçu comme un gage de réussite (ce qu'il faut savoir pour l'examen, en fin de compte).

Sur base des conclusions de nos travaux, différentes améliorations pourraient être proposées et évaluées à travers des recherches ultérieures.

En ce qui concerne le travail collectif versus individuel en APP :

- Augmenter les moments de réflexion individuelle au cours de la séance aller ;
- Demander à chaque étudiant de compléter son propre tableau en identifiant clairement ce qui lui paraît maîtrisé et ce qui lui semble encore flou ;
- Demander à chaque groupe d'envoyer le tableau reprenant les éléments clés, les hypothèses et leurs justifications ainsi que les points à clarifier en complément du schéma physiopathologique ;
- Demander à chaque étudiant de soumettre un document reprenant le travail de correction et progression effectué lors de la période de travail personnel. Certains pourraient être exploités pour la séance retour ;
- Inciter les étudiants à expliciter leur raisonnement et le comparer à celui des autres lors de la séance retour en renforçant les interactions, notamment grâce à des outils tels que les boitiers de vote par exemple ;
- Introduire, dès le premier module PBL, des exercices en petits groupes où chaque étudiant analyse une vignette, construit un schéma et le justifie/valide devant ses pairs qui lui renvoient des questions ;
- Lors de la journée de formation au PBL, mettre l'accent sur l'importance de la production et la validation individuelles.

Au-delà de ces séances collectives, d'autres situations cliniques pourraient être proposées comme exercices individuels. Ces exercices individuels pourraient être proposés en ligne avec une progression pas à

pas prévue telle que c'est le cas lors des séances collectives. Un correctif pourrait être soumis en fin d'exercice avec comme consigne de comparer sa propre production au correctif.

Les modalités d'évaluation du raisonnement

Etant donné le nombre d'étudiants croissant, l'évaluation du raisonnement biomédical par une discussion multidisciplinaire d'une vignette clinique semble être compromise. D'autres modalités d'évaluation exigeant moins de ressources humaines et matérielles sont souhaitables.

Bien que nos travaux de recherche n'aient pas directement travaillé les modalités d'évaluation du raisonnement, nous pouvons proposer les recommandations suivantes :

- Contextualiser l'évaluation en proposant des situations cliniques adaptées au niveau de la formation ;
- Proposer différentes modalités d'évaluation ; parmi celles-ci, nos travaux indiquent qu'un TCS pourrait être intégré à l'examen écrit multidisciplinaire ;
- Varier les situations cliniques proposées ;
- Même si le format d'évaluation choisi exige de proposer une solution, tenter d'obtenir des informations sur le cheminement vers cette solution (en demandant une justification par exemple).

CONCLUSION ET PERSPECTIVES

Résumé

Le développement du raisonnement clinique constitue un des objectifs de base de la formation médicale. La littérature sur le sujet est centrée sur le développement du raisonnement diagnostique, finalité de la formation clinique et souligne combien il est important de développer cette compétence au cours du cursus médical. Nous avons exploré, à travers notre travail, le développement du raisonnement biomédical, dès le début de la formation préclinique et positionné les raisonnements biomédical et clinique sur un même continuum de développement.

La Faculté de médecine de l'Université de Liège a fait le choix, au début des années 2000, de proposer dès la deuxième année de son cursus un enseignement de type PBL dont un des objectifs principaux était le développement précoce du raisonnement clinique. Dans ce cadre, nous avons analysé le processus de raisonnement dans différents contextes et les facteurs qui impactent son développement. Ce développement du raisonnement en lien avec la base de connaissances ainsi que l'utilisation et l'évaluation de cette base de connaissances ont été explorés lors de nos recherches dont les résultats sont rapportés à travers 4 articles.

Nos travaux démontrent qu'une capacité de raisonnement est manifeste dès la 2^{ème} année du cursus et se développe avec l'expérience. Ils indiquent que les étudiants sont capables de transférer les connaissances acquises dans le cadre des séminaires PBL pour les appliquer dans un contexte inédit. Par son côté inédit, le contexte paraît stimuler le raisonnement davantage que le contexte original d'apprentissage. Ces résultats suggèrent que les capacités de raisonnement peuvent être évaluées très tôt dans le cursus médical et que le transfert du raisonnement dans différents contextes est opérationnel.

En ce qui concerne la progression des performances en raisonnement et en connaissances, ces 2 capacités, bien qu'elles soient

fortement liées, n'évoluent pas parallèlement au cours du cursus. Le gain en capacité de raisonnement est plus important et est observé même lorsqu'il porte sur une matière pour laquelle la réactivation des connaissances subit une réduction bien connue avec le temps. Nous avons largement discuté l'importance d'une base de connaissances structurée et hiérarchisée pour favoriser les performances en raisonnement.

Un des apports majeurs de nos travaux est la mise en évidence d'une relation étroite entre les capacités de raisonnement et la capacité à auto-évaluer ses propres connaissances. L'évolution de cette capacité avec l'expérience, chez un même étudiant, a pu être démontrée ainsi que son lien avec la réussite. Ces résultats démontrent l'importance de travailler cette capacité dans le cadre des séminaires PBL afin de promouvoir le développement et la validation du processus de raisonnement. Il est essentiel de pouvoir distinguer ce que l'on connaît de ce que l'on ne connaît pas ou insuffisamment afin d'orienter et de structurer l'acquisition de nouveaux apprentissages. C'est donc le processus métacognitif qui est souligné par nos observations. Il est encore davantage mis en valeur par l'impact d'un exercice de validation d'un schéma explicatif qui renforce la démarche individuelle de l'étudiant.

Notre réflexion a abouti à la proposition d'un modèle permettant de mettre en perspective les différents facteurs impliqués dans le processus d'apprentissage du raisonnement en lien avec la base de connaissances et son utilisation. Bien que conscients des limites de ce modèle, nous pensons qu'il peut apporter des repères utiles pour l'analyse des activités pédagogiques proposées dès le début de la formation médicale. Il pourrait également permettre d'identifier à quel niveau de développement se situe un étudiant afin de lui proposer des exercices adaptés à son parcours. Ainsi, un même contenu d'une vignette clinique peut être exploité pour développer et évaluer le raisonnement biomédical chez un étudiant de niveau pré-clinique et le raisonnement clinique chez un étudiant en master. C'est donc la tâche demandée qui importe avant tout.

Enfin, nous espérons que notre travail pourra susciter et guider les réflexions pédagogiques des enseignants de la Faculté de médecine de l'Université de Liège confrontés à de nouveaux défis avec l'augmentation du nombre d'étudiants et le passage des études médicales de 7 à 6 années. Des pistes concrètes ont été proposées dans la discussion.

Perspectives et Transfert

Le raisonnement biomédical et clinique est un processus multidimensionnel complexe mais indispensable à l'exercice de la médecine. Les recherches concernant son enseignement et son évaluation sont essentielles. Elles permettent d'avancer vers une meilleure compréhension de cette compétence et de proposer des options d'amélioration des dispositifs pour la développer. Notre travail aura permis une meilleure compréhension des facteurs impliqués dans le raisonnement biomédical et son développement vers la compétence clinique. Des recherches complémentaires sur les facteurs impliqués dans la validation du processus de raisonnement nous paraissent essentielles, à la fois pour l'enseignement du raisonnement en début de formation préclinique, mais également pour le développement professionnel de chaque médecin.

Parmi les pistes envisagées, le DTI (Diagnostic Thinking Inventory) pourrait être exploité. Il semble être un outil de choix pour l'auto-évaluation de l'organisation de ses propres connaissances ainsi que pour l'auto-évaluation de son processus de raisonnement. En vue de confirmer l'intérêt de cet outil pour la formation, les résultats au DTI pourraient être comparés avec les mesures déjà prises dans le cadre de nos travaux. Les résultats de ces mesures devront être actualisés étant donné le changement de contexte que connaît la formation médicale à l'ULg.

L'ensemble de notre travail s'inscrit dans le paradigme du raisonnement vu comme la résolution d'un problème. Ce paradigme est centré sur les recherches en psychologie cognitive et vise à comprendre

les processus et stratégies utilisés par les professionnels lorsqu'ils résolvent une situation problème, soit une situation clinique dans le cas du clinicien. Ce paradigme postule que le raisonnement est un processus qui peut être appris. Cette vision du raisonnement nous permet d'élargir la portée de nos travaux à toute formation impliquant la résolution de problème. Le raisonnement se conçoit ici comme le processus de pensée qui implique l'identification d'éléments clés d'une situation et la mise en évidence des liens qui les unissent, à la lumière des connaissances de base permettant la compréhension et la représentation globale d'un problème. Dès lors, nous pouvons appliquer cette acception du raisonnement à toute forme de contexte indépendamment de la formation médicale. Nous pouvons ainsi envisager la transférabilité de nos pratiques relatives à l'apprentissage contextualisé dès les premières années d'autres cursus que la médecine. D'autres filières de formation pourraient en effet bénéficier du transfert précoce des connaissances dans différents contextes ainsi que d'un apprentissage précoce de la résolution de problème.

L'élément central à promouvoir à travers les différentes disciplines concernées est la validation des connaissances et, par là, la validation des apprentissages et du processus de résolution de problèmes qui sont trop peu prises en compte dans la plupart des dispositifs de formation. Une telle démarche prend une signification particulière au moment où la masse du savoir explose dans beaucoup de disciplines et impose, dès lors, des questions inédites en matière de gestion des limites et de la certitude, dans la base de connaissances de chacun.

Conclusions générales

Traditionnellement, l'acquisition des performances professionnelles liées à l'exercice du métier sont investies en fin de formation et les premières années sont, quant à elles, dédiées à l'acquisition des connaissances et capacités de base. Nous démontrons que l'acquisition d'une performance aussi complexe que le raisonnement biomédical et clinique, qui est à la base de l'exercice de la profession médicale, peut déjà se travailler très tôt dans le cursus.

Néanmoins, la question subsiste de savoir dans quelle mesure un apprentissage précoce de ces performances complexes permet de former des professionnels plus ou plus rapidement compétents sur le terrain.

L'importance du processus d'auto-validation de ses connaissances et de son raisonnement est largement soulignée par nos travaux. Jusqu'à présent, l'enseignement privilégie plutôt une démarche cognitive de l'étudiant face à son apprentissage ; celui-ci est davantage préoccupé par l'acquisition de savoir et de savoir-faire afin de réussir ses examens et passer à l'étape de formation suivante. Ce souci de la réussite pourrait constituer un court-circuit de la démarche métacognitive face à l'apprentissage et au processus de réflexion en termes de parcours professionnel. Ce biais est aggravé par la pression liée à la réussite dont le niveau conditionne l'accès à la discipline spécialisée visée par l'étudiant. Nous démontrons que l'apprentissage de performances professionnelles et la démarche métacognitive sont des facteurs également associés à la réussite.

Nos travaux s'inscrivent dans un contexte spécifique de formation médicale qui implique partiellement le PBL au cours d'un cursus de 7 années à l'Université de Liège avec une pression de la réussite qui ne quitte pas l'étudiant tout au long du parcours de formation. La littérature nous apprend que nos questionnements, difficultés, résultats et préoccupations sont partagés par d'autres dans des contextes de formation parfois fort différents. Notamment, on retrouve la question du développement des connaissances et du raisonnement dans des Universités fondées sur le PBL comme à Maastricht ou encore l'importance de la base de connaissances et de la validation du processus de raisonnement dans les recherches canadiennes où la formation médicale s'étale sur 4 ou 5 années après une formation scientifique de base et une sélection préalable.

Le cadre de la Faculté de médecine de l'Université de Liège dans lequel nos travaux ont pu prendre place a été déterminant. Le contexte de questionnement des enseignants initié par la réforme de 2000 s'est

avéré être un moteur de progression des pratiques de formation mais aussi de la recherche. Si un élément est à cultiver dans cette Faculté comme dans d'autres, il s'agit du questionnement permanent des enseignants, travaillant de manière collégiale plutôt qu'individuelle, exposant leurs idées et leurs modalités d'enseignant aux regards critiques des collègues. L'essentiel n'est pas le choix d'une méthode plutôt qu'une autre mais le questionnement qui sous-tend ce choix. Il en va donc de la formation comme de la pratique de la profession : la qualité de l'action se mesure à la profondeur du questionnement qui la justifie.

RÉFÉRENCES

Allen, S.W., Norman, G.R., & Brooks, L.R. (1992). Experimental studies of learning dermatological diagnosis : the impact of examples. *Teaching and Learning in Medicine*, 4, 35-44.

Anderson, K.J. (2006). *Factors affecting the development of undergraduate medical students' clinical reasoning ability* (Doctoral Dissertation). Faculty of Health Science, University of Adelaide, Australia.

Audetat, M.C., Dory, V., Nendaz, M., Vanpee, D., Pestiaux, D., Junod Perron, N., & Charlin, B. (2012). What is so difficult about managing clinical reasoning difficulties? *Medical Education*, 46, 216-227.

Audetat, M.C., Laurin, S., & Sanche, G. (2011a). Aborder le raisonnement clinique du point de vue pédagogique I. Un cadre conceptuel pour identifier les problèmes de raisonnement clinique. *Pédagogie Médicale*, 12(4), 223-229.

Audetat, M.C., Laurin, S., & Sanche, G. (2011b). Aborder le raisonnement clinique du point de vue pédagogique II. Les difficultés de raisonnement clinique à l'étape du recueil initial des données et de la génération d'hypothèses. *Pédagogie Médicale*, 12(4), 231-236.

Ark, T.K., Brooks, L.R., & Eva, K.W. (2007). The benefits of flexibility: the pedagogical value of instructions to adopt multifaceted diagnostic reasoning strategies. *Medical Education*, 41(3), 281-287.

Arocha, J.F., Patel, V.L., & Patel, Y.C. (1993). Hypothesis generation and the coordination of theory and evidence in novice diagnostic reasoning. *Medical Decision Making*, 13(3), 198-211.

Askew, K., Manthey, D., & Mahler, S. (2012). Clinical reasoning: are we testing what we are teaching? *Medical Education*, 46, 540-542.

Avrahami, J., Kareev, Y., Bogot, Y., Caspi, R., Dunaevsky, S., & Lerner, S. (1997). Learning by examples: Implications for the process of category

acquisition. *Quarterly Journal of Experimental Psychology A*, 50 A(3), 586-606.

Barrows, H. (1984). A specific problem-based, self-directed learning method designed to teach medical problem-solving skills, and enhance knowledge retention and recall. In H.G. Schmidt & M. de Volder (Eds.), *Tutorials in Problem-based Learning* (pp.16-32). Assen: van Gorcum.

Barrows, H.S., & Bennett, K. (1972). The diagnostic (problem solving) skill of the neurologist. Experimental studies and their implications for neurological training. *Archives of neurology*, 26(3), 273-277.

Barrows, H.S., Norman, G.R., Neufeld, V.R., & Feightner, J.W. (1982). The clinical reasoning of randomly selected physicians in general medical practice. *Clinical Investigative Medicine*, 5, 49-55.

Barrows, H., & Tamblyn, R.M. (1980). *Problem-Based Learning: An approach to Medical Education*. New-York: Springer Publishing Co.

Bell, K., Boshuizen, H., Scherpbier, A., & Dornan, T. (2009). When only the real thing will do: junior medical students' learning from real patients. *Medical Education*, 43, 1036-1043.

Belleza, F., & Bower, G. (1981). The representational and processing characteristics of scripts. *Bulletin of the Psychonomic Society*, 18, 1-4.

Bernstein, D., Penner, L.A., Clarke-Stewart, A., & Roy, E.J. (2007). *Psychology* (8th ed.). Boston: Cengage Learning. pp262-263.

Beulens, J., Struyf, E., & van Damme, B. (2006). Diagnostic ability in relation to clinical seminars and extended-matching questions examinations. *Medical Education*, 40, 1173-1179.

Blake, R.L., Hosokawa, M.C., & Riley, S.L. (2000). Student performances on Step 1 and Step 2 of the United States Medical Licensing Examination following implementation of a problem-based learning curriculum. *Academic medicine*, 75(1), 66-70.

Blumberg, P., & Michael, J. (1992). The development of self-directed learning behaviors in a partially teacher-centred, problem-based learning curriculum. *Teaching and learning in Medicine*, 4, 3-8.

Boniver, J. (2004). Les études de médecine à l'Université de Liège: le renouveau pédagogique de la Faculté de Médecine. *Revue Médicale de Liège*, 59(12), 717-30.

Bordage, G., & Zacks, R. (1984). The structure of medical knowledge in the memories of medical students and general practitioners: categories and prototypes. *Medical Education*, 18, 406-416.

Bordage, G., Grant, L. & Marsden, P. (1990) Quantitative assessment of diagnostic ability. *Medical Education*, 24, 413-425.

Bordage, G., & Lemieux, M. (1991). Semantic structures and diagnostic thinking of experts and novices. *Academic Medicine*, 66(9 Suppl), S70-72.

Boshuizen, H.P.A., & Schmidt, H.G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, 16, 153-184.

Boshuizen, H.P., & Schmidt, H.G. (2008). The development of clinical reasoning expertise. In J. Higgs, M.A. Jones, S. Loftus, & N. Christensen (Eds.), *Clinical reasoning in the Health professions* (pp.113-121). Amsterdam: Elsevier.

Boshuizen, H., Schmidt, H.G., Custers, E., & van der Wiel, M. (1995). Knowledge development and restructuring in the domain of medicine: the role of theory and practice. *Learning & instruction*, 5, 269-289.

Brooks, L.R., Norman, G.R., & Allen, S.W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology General*, 120, 278-287.

Carveni, J.P., Nguyen-Xuan, A., Hoc, J.M., & Politzer, G. (1990). Raisonement formel et raisonnement en situation. In J.F. Richard, C. Bonnet, & R. Ghiglione (Eds.), *Traité de Psychologie Cognitive 2: le*

traitement de l'information symbolique (pp.105-165). Paris: Dunod, Bordas.

Case, S.M., Swanson, D.B., & Stillman, P.S. (1988). Evaluating diagnostic pattern recognition: The psychometric characteristics of a new item format. *Research in Medical Education*, 27, 3-8.

Chamberland, M. (1998). Les séances d'apprentissage au raisonnement clinique (ARC). Un exemple d'activité pédagogique contextualisée adaptée aux stages cliniques en médecine. *Annales de Médecine Interne*, 149, 479-84.

Charlin, B., Bordage, G., & van der Vleuten, C. (2003). L'évaluation du raisonnement clinique. *Pédagogie Médicale*, 4, 42-52.

Charlin, B., Boshuizen, H.P.A., Custers, E.J., & Feltovich, P.J. (2007) Scripts and clinical reasoning. *Medical Education*, 41, 1178–1184.

Charlin, B., Desaulniers, M., Gagnon, R., Blouin, D., & van der Vleuten, C. (2002b). Comparison of an aggregate scoring method with a consensus scoring method in a measure of clinical reasoning capacity. *Teaching and Learning in Medicine*, 14(3), 150-156.

Charlin, B., Gagnon, R., Sibert, L., & Van der Vleuten, C. (2002a). Le test de concordance de script : un instrument d'évaluation du raisonnement clinique. *Pédagogie Médicale*, 3, 135-144.

Charlin, B., Lubarsky, S., Millette, B., Crevier, F., Audétat, M.C., Charbonneau, A., Caire Fon, N., Hoff, L, & Bourdy, C. (2012). Clinical reasoning processes: unravelling complexity through graphical representation. *Medical Education*, 46, 454-463.

Charlin, B., Tardif, J., & Boshuizen, H.P.A. (2000). Scripts and Medical Diagnostic Knowledge: Theory and Applications for Clinical Reasoning Instruction and Research. *Academic Medicine*, 75,182-190.

Charlin, B., Roy, L., Brailovsky, C., Goulet, F., & Van der Vleuten, C. (2000). The script concordance test, a tool to assess the reflective clinician. *Teaching and Learning in Medicine, 12*, 189-195.

Charlin, B., & Van der Vleuten, C. (2004). Standardized assessment in context of uncertainty: The script concordance approach. *Evaluation & the Health Professions, 27*, 304-319.

Chi, M., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Chi, M., Glaser, R., & Farr, M. (1988). *The nature of expertise*. Hillsdale, J.J.: Erlbaum.

Coderre, S., Mandin, H., Harasym, P.H., & Fick, G. H. (2003). Diagnostic reasoning strategies and diagnostic success. *Medical Education, 37*, 695-703.

Coles, C.R. (1985). Differences between conventional and problem-based curricula in their student's approach to studying. *Medical Education, 19*(4), 308-309.

Côté, D.J., Bellavance, C., Chamberland, M., & Graillon, A. (2005). Un programme pour aider les étudiants en médecine à développer leurs stratégies d'apprentissage. *Pédagogie Médicale, 5*, 95-102.

Custers, E.J. (2010). Long-term retention of basic science knowledge: a review study. *Advance in Health Science Education, 15*, 109-128.

Custers, E.J., Boshuizen, H.P., & Schmidt, H.G. (1996b). The influence of medical expertise, case typicality, and illness script component on case processing and disease probability estimates. *Memory and Cognition, 24*(3), 384-399.

Custers, E.J.F.M., Regehr, G., & Norman, G.R. (1996a). Mental representations of medical diagnostic knowledge: A review. *Academic Medicine, 71*(10 suppl), S55-61.

de Volder, M., Schmidt, H., Moust, J., & de Grave, W. (1986). Problem-based learning and intrinsic motivation. In J. van der Berchen, T. Bergen, & E. de Bruyn (Eds.), *Achievement and Task Motivation*.

Lisse, The Netherlands: Swets and Zeitlinger and Swets, North America.

Diemers, A.D., Dolmans, D.H., Verwijnen, M.G., Heineman, E., Scherpbier, A.J. (2008). Students' opinions about the effects of pre-clinical patient contacts on their learning. *Advance in Health Science Education: Theory and Practice*, 13, 633–647.

Diemers, A.D., van de Wiel, M.W., Scherpbier, A.J., Heineman, E., & Dolmans, D.H. (2011). Pre-clinical patient contacts and the application of biomedical and clinical knowledge. *Medical Education*, 45, 280-288.

Dixon R.A., & Brown, R.A. (2012). Transfer of learning: connecting concepts during problem solving. *Journal of Technology Education*, 24, 2-17.

Dornan, T., & Bundy, C. (2004). What can experience add to early medical education? Consensus survey. *BMJ*, 329, 834–837.

Dory, V., Gagnon, R., Vanpee, D., Charlin, B. (2012). How to construct and implement script concordance tests : insights from a systematic review. *Medical Education*, 46(6), 552-563.

Durning, S.J., & Artino, A.R. (2011). Situativity theory: A perspective on how participants and the environment can interact. *Medical Teacher*, 33, 188-199.

Durning, S.J., Artino, A.R., Schuwirth, L., & Van Der Vleuten, C. (2013). Clarifying assumptions to enhance our understanding and assessment of clinical reasoning. *Academic Medicine*, 88(4), 442-448.

Eisenstaedt, R., Barry, W., & Glanz, K. (1990). *Problem-based Learning: cognitive retention and cohort traits of randomly selected participants*

and decliners. Paper presented at the Research in Medical Education, Washington, DC.

Elstein, A., Shulman, L., & Sprafka, S. (1978). *Medical problem solving: an analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.

Elstein, A., Shulman L., & Sprafka S. (1990). Medical Problem Solving, a Ten-Year Retrospective. *Evaluation & the Health Professions*, 13, 5-36.

Eva, K.W. (2005). What every teacher needs to know about clinical reasoning. *Medical Education*, 39, 98-106

Eva, K.W., Neville, A.J., & Norman, G.R. (1998). Exploring the etiology of content specificity: factors influencing analogic transfer and problem solving. *Academic Medicine*, 73(10 Suppl), S1-5.

Eylon, B., & Reif, F. (1984). Effects of knowledge organization on task performance. *Cognitive Instruction*, 1, 5-44.

Fayol, M., & Monteil, J-M. (1988). The notion of script: from general to developmental and social psychology. *European Bulletin of Cognitive Psychology*, 8, 335-61.

Fazey D. (1993). Self-assessment as a generic skill for enterprising students: the learning process. *Assessment and Evaluation in Higher Education*, 18, 235-250.

Feletti, G.I. (1980). Reliability and validity on modified essay question. *Medical Education*, 55, 933-941.

Feltovich, P.J., & Barrows, H.S. (1984). Issues of Generality in Medical Problem Solving. In H.G. Schmidt, & M. De Volder (Eds.), *Tutorials in Problem-Based Learning: A new Direction in Teaching the Health Professions* (pp.128-142). Assen(Holland): Van Gorcum.

Frenay, M. (2004). Du transfert des apprentissages au transfert des connaissances. In A. Presseau, & M. Frenay (Eds.), *Le transfert des*

apprentissages, comprendre pour mieux intervenir (pp. 7-48). Quebec: Les presses de l'Université de Laval.

Gick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1-38.

Gick, M.L., & Holyoak, K.J. (1987). The cognitive basis of knowledge transfer. In S.M. Cormier, & J.D. Hagman (Eds.), *Transfer of Learning: Contemporary Research and Applications* (pp. 9-47). New-York: Academic Press.

Giet, D., Massart, V., Stir, A., Freyens, A., Firket, P., & Boniver, J. (2005). Approche de la complexité contextuelle et de la pluridisciplinarité de l'action médicale : mise en place de séances d'apprentissage à la résolution de problèmes complexes (ARPC) en fin de 2e cycle d'études médicales. *Pédagogie Médicale*, 6, 88-97.

Gilhooly, K.J. (1990). Cognitive psychology and medical diagnosis. *Applied Cognitive Psychology*, 4, 261-272.

Glaser, R. (1991). The maturing of the relationship between the science of learning and cognition and educational practice. *Learning and instruction*, 1, 3-8.

Grant, J., & Marsden, P. (1987). The Structure of Memorized Knowledge in Students and Clinicians: An Explanation for Diagnostic Expertise. *Medical Education*, 21, 92-98.

Groves, M. (2008). *Clinical reasoning and diagnostic expertise*. Saarbrücken : VDM Verlag Dr. Müller.

Groves, M. (2012). Understanding clinical reasoning: the next step in working out how it really works. *Medical Education*, 46, 444-446.

Harden, R.M., & Gleeson, F.A. (1979). Assessment of medical competence using an Objective Structured Clinical Examination. *Medical Education*, 13, 39-54.

Hassebrock, F., & Prietula, M. (1992). A protocol-based coding scheme for the analysis of medical reasoning. *Journal of Man-machine studies*, 37, 613-652.

Hatala, R.M., Brooks, L.R., & Norman, G.R. (2003). Practice makes perfect: the critical role of mixed practice in the acquisition of ECG interpretation skills. *Advance in Health Science Education: Theory and Practice*, 8, 17-26.

Hatala, R.A., Norman, G.R., Brooks, L.R. (1996). The effect of clinical history on physicians' ECG interpretation skills. *Academic Medicine*, 71(10 Suppl), S68-70.

Hatala, R., & Norman, G.R. (2002). Adapting the Key Features Examination for a clinical clerkship. *Medical Education*, 36, 160-165.

Higgs, J., Jones, M.A., Loftus, S., & Christensen, N. (2008). *Clinical reasoning in the Health Professions*. Oxford: Butterworth-Heinemann.

Hmelo, C.E. (1998). Cognitive consequences of problem-based learning for the early development of medical expertise. *Teaching and learning in medicine*, 10(2), 92-100.

Hmelo-Silver, C.E. (2004). Problem-based learning: what and how do students learn? *Educational Psychology Review*, 16, 235-266.

Hoff, L., Bestawros, A., Kassis, J., & Charlin, B. (2010). Le test de concordance de script comme outil d'enseignement et d'apprentissage : un projet-pilote pour les étudiants de première année de médecine. *Pédagogie Médicale*, 11(1), 51-56.

Holmboe, E.S., & Hawkins, R.E. (1998). Methods for evaluating clinical competence of residents in internal medicine: a review. *Annals of Internal Medicine*, 129, 42-48.

Humbert, A.J., Johnson, M.T., Miech, E., Friedberg, F., Grackin, J.A., & Seidman, P.A. (2011). Assessment of clinical reasoning: A Script

Concordance test designed for pre-clinical medical students. *Medical Teacher*, 33, 472-477.

Jean, P., DesMarchais, J.E., & Delorme, P. (1993). Apprendre à enseigner les sciences de la santé. Guide de formation pratique. Faculté de médecine des universités de Montréal et de Sherbrooke, 4e édition.

Jolly, B., & Grant, J. (1997). The Good Assessment Guide. A practical Guide to Assessment and Appraisal for Higher Specialist Training. Joint Center for Education in Medicine. London: UK.

Joseph, G.M., & Patel, V.L. (1990). Domain knowledge and hypothesis generation in diagnostic reasoning. *Medical Decision Making*, 10(1), 31-46.

Kassirer, J.P. (1983). Teaching clinical medicine by iterative hypothesis testing. Let's preach what we practice. *New England Journal of Medicine*, 309, 921-923.

Kassirer, J., & Kopelman, R. (1991). *Learning clinical reasoning*. Baltimore: Williams & Wilkins.

Klayman, J., & Brown, K. (1993). De bias the environment instead of the judge: An alternative approach to reducing error in diagnostic (and other) judgment. *Cognition*, 49, 97-122.

Koh, G., Khoo, H.E., Wong, M.L., Koh, D. (2008). The effects of problem-based learning during medical school on physician competency: a systematic review. *Canadian Medical Association Journal*, 178(1), 34-41.

Kulatunga-Moruzi, C., Brooks, L.R., Norman, G.R. (2001). Coordination of analytic and similarity-based processing strategies and expertise in dermatological diagnosis. *Teaching and Learning Medicine*, 13, 110-6.

Langendyk, V. (2006). Not knowing that they do not know: self-assessment accuracy of third-year medical students. *Medical Education*, 40, 173-179.

La Rochelle, J.S, Durning, S.J., Pangaro, L.N., Artino, A.R., Van der Vleuten, C., & Schuwirth, L. (2012). Impact of Increased Authenticity in Instructional Format on Preclerkship Students' Performance: A Two-Year, Prospective, Randomized Study. *Academic Medicine*, 87(10), 1341-1347.

Laurin, S., Audetat, M.C., Sanche, G. (2012a). Aborder le raisonnement clinique du point de vue pédagogique IV. Les difficultés de raisonnement clinique à l'étape du raffinement et du traitement des hypothèses : les difficultés de priorisation. *Pédagogie Médicale*, 13(2) : 109-114.

Laurin, S., Audetat, M.C., Sanche, G. (2012b). Aborder le raisonnement clinique du point de vue pédagogique V. Les difficultés de raisonnement clinique à l'étape de l'élaboration d'un portrait global de la situation clinique. *Pédagogie Médicale*, 13(3), 203-208.

LeBlanc, V., Norman, G., & Brooks, L. (2001). Effect of a diagnostic suggestion on diagnostic accuracy and clinical feature identification. *Academic Medicine*, 76(Suppl 10), S18-S20.

LeBlanc, V.R., Dore, K., Norman, G.R., & Brooks, L.R. (2004). Limiting the playing field: does restricting the number of possible diagnoses reduce errors due to diagnosis-specific feature identification? *Medical Education*, 38, 17-24.

Lemieux, M., & Bordage, G. (1992) Propositional vs Structural Semantic Analyses of Medical Diagnostic Thinking. *Cognitive Science*, 16, 185-204.

Lesgold, A., Rubinson, H., Feltovitch, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures. In M. Chi, R. Glaser, & M. Farr, (Eds.), *The Nature of Expertise* (pp. 311-342). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Mamede, S, Schmidt, H.G., & Rikers, R. (2007). Diagnostic errors and reflective practice in medicine. *Journal of Evaluation in Clinical Practice*, 13, 138-145.

Mamede, S, van Gog, T., Moura, A.S., de Faria, R., Peixoto, J.M., Rikers, R., & Schmidt, H.G. (2012). Reflection as a strategy to foster medical students' acquisition of diagnostic competence. *Medical Education*, 46, 464-472.

Mandin, H., Jones, A., Woloschuk, W., & Harasym, P. (1997). Helping students learn to think like experts when solving clinical problems. *Academic Medicine*, 72, 173-179.

Marchand, C., & Gagnayre, R. (2005). Utilisation de la carte conceptuelle auprès de formateurs en santé pour l'apprentissage de concepts pédagogiques. *Pédagogie Médicale*, 5, 13-23.

Martenson, D., Eriksson, H., & Ingelman-Sundberg, M. (1985). Medical chemistry: evaluation of active and problem-oriented teaching methods. *Medical education*, 19(1), 34-42.

McConnell, M.M., & Eva, K.W. (2012). The role of emotion in the learning and transfer of clinical skills and knowledge. *Academic Medicine*, 87(10), 1316-1322.

Mc Guire, C.H., & Babbott, D. (1967). Simulation techniques in the measurement of problem solving skills. *Journal of Educational Measurement*, 4, 1-10.

Nandi, P.L., Chan, J.N.F., Chan, C.P.K., Chan, P., Chan, L.P.K. (2000). Undergraduate medical education: comparison of problem-based learning and conventional teaching. *Hong Kong Medical Journal*, 6, 301-306.

Needham, D.R., & Begg, I.M. (1991). Problem-oriented training promotes spontaneous analogical transfer: memory-oriented training promotes memory for training. *Memory and Cognition*, 19(6), 543-557.

Nendaz, M.R. & Bordage, G. (2002). Promoting diagnostic problem representation. *Medical Education*, 36, 760-766.

Nendaz, M., Charlin, B., Leblanc, V., Bordage, G. (2005). Le raisonnement clinique : données issues de la recherche et implications pour l'enseignement. *Pédagogie Médicale*, 6, 235-254.

Nelson, K. (1986). Event knowledge and cognitive development. In K. Nelson (Ed.), *Event Knowledge: Structure and Function in Development*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Newble, D.I., & Clark, R.M. (1986). The approaches to learning of students in a traditional and in an innovative problem-based medical school. *Medical Education*, 20(4), 267-273.

Newell, A., & Simon, H. (1972). *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall.

Norcini, J.J., Diserens, D., Day, S.C., Cebul, R.D., Schwartz, J.S., Beck, L.H. et al. (1990) The scoring and reproductibility of an essay test of clinical judgement. *Academic Medicine*, 65(Supp 9), S41-S42.

Norman, G.R. (1993). Theoretical and psychometric considerations. In Report on the evaluation system for specialist certification (pp 73-80). Task force of the evaluation committee. The Royal College of Physicians and Surgeons of Canada. Ottawa.

Norman, G.R. (2005). From theory to application and back again: Implications of research on medical expertise for psychological theory. *Canadian Journal of Experimental Psychology*, 59, 35-40.

Norman G. (2009). Teaching basic science to optimize transfer. *Medical teacher*, 31, 807-811.

Norman, G.R., Brooks, L.R., Regehr, G., Marriott, M., & Shali, V. (1996). Impact of feature interpretation on medical student diagnostic performance. *Academic Medicine*, 71(1 Suppl), S108-109.

Norman, G., Eva, K.W., Brooks, L., Hamstra, S. (2006). Expertise in medicine and surgery. In K.A. Ericsson, N. Charness, P.J. Feltovich, R.R.

Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp.339-353). Cambridge: Cambridge University Press.

Norman, G.R., & Feightner, J.W. (1981). A comparison of behavior on simulated patients and patient management problems. *Medical Education, 55*, 529-537.

Norman, G.R., Rosenthal, D., Brooks, L.R., Allen, S.W., & Muzzin, L.J. (1989). The development of expertise in dermatology. *Archives of dermatology, 125*(8), 1063-8.

Norman, G.R., & Schmidt, H.G. (1992). The psychological basis of problem-based learning: a review of the evidence. *Academic Medicine, 67*(9), 557-565.

Nouns, Z., Schaubert, S., Witt, C., Kingreen, H., & Schüttpeitz-Brauns, K. (2012). Development of knowledge in basic sciences: a comparison of two medical curricula. *Medical Education, 46*, 1206-1214.

O'Neill, P.A. (2000). The role of basic sciences in a problem based learning clinical curriculum. *Medical Education, 34*(8), 608–613.

Page, G., & Bordage, G. (1995). The Medical Council of Canada's Key Features Project: a more valid written examination of clinical decisions skills. *Academic Medicine, 70*, 104-110.

Papa, F.J., Stone, R.C., & Aldrich, D.G. (1996). Further evidence of the relationship between case typicality and diagnostic performance: implications for medical education. *Academic Medicine, 71*(1 Suppl), S10-12.

Patel, V.L., Glaser, R., & Arocha, J.F. (2000). Cognition and expertise: acquisition of medical competence. *Clinical and Investigative Medicine, 23*(4), 256–260.

Patel, V.L., Groen, G.J., & Arocha, J.F. (1990). Medical expertise as a function of task difficulty. *Memory and cognition, 18*(4), 394-406.

Patel, V.L., Groen, G.J., & Norman, G.R. (1991). Effects of conventional and problem-based medical curricula on problem solving. *Academic medicine*, 66(7), 380-389.

Patel, V., & Kaufman, D. (2000). Clinical reasoning and biomedical knowledge: implications for teaching. In J.J. Higgs (Ed), *Clinical reasoning in the health professions* (pp.33-44). Oxford: Butterworth-Heinemann.

Patel, V.L., & Groen, G.J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive science*, 10, 91-116.

Prince, K.J., van de Wiel, M., Scherpbier, A.J., van der Vleuten C., Boshuizen, H.P. (2000). A qualitative analysis of the transition from theory to practice in undergraduate training in a PBL medical school. *Advance in Health Science Education: Theory and Practice*, 5, 105–116.

Prince, K.J.A.H., van Mameren, H.V., Hylkema, N., Drukker, J., Scherpbier, A.J.J.A., & van der Vleuten, C.P.M. (2003). Does problem based learning lead to deficiencies in basic science knowledge? An empirical case on anatomy. *Medical Education*, 37, 15–21.

Regan-Smith, M., Obenshain, S., Woodward, C., Richards, B., Zeits, H., & Small, P.J. (1994). Rote learning in medical school. *Journal of the American Medical Association*, 272, 1380–1381.

Regehr, G., & Norman, G.R. (1996). Issues in cognitive psychology: implications for professional education. *Academic Medicine*, 71, 988-1001.

Rumelhart, D.E. (1981). Schemata: The building blocks of cognition. In J.T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp.3-26). Newark: International Reading Association.

Sanche, G., Audetat, M.C., Laurin, S. (2012a). Aborder le raisonnement clinique du point de vue pédagogique III. Les difficultés de raisonnement clinique à l'étape du traitement et du raffinement des hypothèses : la fermeture prématurée. *Pédagogie Médicale*, 13(2), 103-108.

Sanche, G., Audetat, M.C., Laurin, S. (2012b). Aborder le raisonnement clinique du point de vue pédagogique VI. Les difficultés de raisonnement clinique à l'étape de l'élaboration du plan d'intervention. *Pédagogie Médicale*, 13(3), 209-214.

Schank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.

Schmidt, H.G. (1983). Problem-based learning: rationale and description. *Medical Education*, 17(1), 11-16.

Schmidt, H.G. (1993). Foundations of problem-based learning: some explanatory notes. *Medical Education*, 27(5), 422-432.

Schmidt, H.G., & Boshuizen, H.P. (1993a). On the origin of intermediate effects in clinical case recall. *Memory and cognition*, 21(3), 338-351.

Schmidt, H.G., & Boshuizen, H.P.A. (1993b). On acquiring expertise in medicine. *Educational Psychology Review* 5, 1–17.

Schmidt, H.G., Boshuizen, H., & Norman, G. (1992). Reflections on the nature of expertise in medicine. In E. Keravnou (Ed.), *Deep models for medical knowledge engineering* (pp.231-248). Amsterdam: Elsevier, North-Holland.

Schmidt, H.G., de Grave, W., de Volder, M., Moust, J., & Patel, V. (1989). Explanatory models in the processing of science text: the role of prior knowledge activation through small-group discussion. *Journal of Educational Psychology*, 81, 481-491.

Schmidt, H.G., Norman G.R., & Boshuizen H.P. (1990) A cognitive perspective on medical expertise: theory and implication. *Academic Medicine*, 65, 611-621.

Schmidt, H.G., & Rikers, R. (2007). How expertise develops in medicine: knowledge encapsulation and illness script formation. *Medical Education*, 41, 1133–1139.

Schön, D.A. (1983). *The reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.

Seabrook, M. (2004). Clinical students' initial reports of the educational climate in a single medical school. *Medical Education*, 38, 659–669.

Shin, J., Haynes, R., & Johnston, M. (1993). Effect of Problem-Based, self-directed undergraduate education on life-long learning. *Canadian Medical Association Journal*, 148, 969-976.

Smith, S.M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin and Review*, 8(2), 203-220.

Spiro, R.J., Vispoel, W.P., Schmitz, J.P., Samarapungavan, A., & Boerger, A.E. (1987). Knowledge acquisition for application: cognitive flexibility and transfer in complex domains. In B.K. Britton, & S.M. Glynn (Eds.), *Executive control processes in reading* (pp.177-199). Hillsdale, N.J.: Erlbaum.

Streiner, D.L. (1985). Global rating scales. In V.R. Neufeld, & G.R. Norman (Eds.), *Assessing Clinical Competence* (pp.119-141). New-York: Springer.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257–285.

Tardif J. (1992). *Pour un enseignement stratégique: l'apport de la psychologie cognitive*. Montréal: Les Editions Logiques.

Taylor, D.C., Hamdy, H. (2013). Adult learning theories: Implications for learning and teaching in medical education: AMEE Guide No. 83. *Medical Teacher*, 35, e1561-e1572

Thomas, R.E. (1997). Problem-based Learning: measurable outcomes. *Medical Education*, 31(5), 320-329.

Tulving, E., & Thomsons, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), 352-373

- Van der Vleuten, C., & Newble, D.I. (1995). How can we test clinical reasoning? *Lancet*, *345*, 1032-1034.
- Van der Wiel, M., Boshuizen, H., & Schmidt, H. (2000). Knowledge restructuring in expertise development: evidence from pathophysiological representations of clinical cases by students and physicians. *European Journal of Cognitive Psychology*, *12*(3), 323-355.
- Vernon, D.T., & Blake, R.L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine*, *68*, 550-563.
- Wass, V., & Jolly, B. (2001). Does observation add to the validity of the long case? *Medical Education*, *35*, 729-734.
- Wigton, R.S., Connor, J.L., & Centor, R.M. (1998). Transportability of a decision rule for the diagnosis of streptococcal pharyngitis. *Archives of Internal Medicine*, *146*, 81-83.
- Zimmerman, B. (2002). Becoming a self-regulated learner: an overview. *Theory into Practice*, *41*, 64-70.