

TABLE DES MATIÈRES

RÉSUMÉ	i
ABSTRACT.....	iii
REMERCIEMENTS.....	iv
LISTE DES TABLEAUX.....	ix
LISTE DES FIGURES	xi
LISTE DES ACRONYMES ET ABRÉVIATIONS.....	xii
INTRODUCTION	1
1. PROBLÉMATIQUE	3
2. ÉTAT DE LA QUESTION ET CADRE CONCEPTUEL.....	8
2.1 Typologie de l'évaluation.....	8
2.2 L'évaluation diagnostique	9
2.3 Le raisonnement clinique	10
2.4 Le raisonnement clinique infirmier (RCI).....	11
2.5 L'évaluation du raisonnement clinique	15
2.6 Définition, application et qualités psychométriques du TCS.....	18
2.7 De l'évaluation formative au test diagnostique.....	22
2.8 Questions de recherche.....	23
3. MÉTHODOLOGIE	25
3.1 Type de recherche	25
3.2 Description de l'étude 1	26
3.2.1 Objectif de l'étude 1.....	26
3.2.2 Participants sollicités dans l'étude 1	26

3.2.3	Instrument de cueillette de données de l'étude 1	27
3.2.4	Stratégies d'analyse des données de l'étude 1	29
3.2.5	Résultats de l'étude 1	30
3.3	Méthodologie de l'étude 2.....	31
3.3.1	Objectif de l'étude 2.....	31
3.3.2	Participants.....	31
3.3.3	Données.....	31
3.3.4	Méthode d'analyse	31
3.4	Éthique	32
4.	PRÉSENTATION DES RÉSULTATS	34
4.1.	Procédure de sélection et identification des catégories et des stratégies de pensée pertinentes à cette étude	34
4.2.	Mise en évidence des liens entre les 21 catégories/stratégies et les 13 questions pour l'ensemble des participantes.....	43
4.3	Relations entre les 21catégories/stratégies et les 13 questions pour chaque groupe de participantes	48
4.4	Accord entre les groupes	63
4.5	Accord entre les questions 2 à 2 sur l'ensemble des participants	65
4.6	Accord entre les questions groupe par groupe	68
4.7	Synthèse des résultats.....	72
5.	DISCUSSION.....	76
5.1	Première question de recherche : Dans quelle mesure chaque question d'un TCS visant à évaluer le RCI peut-elle être reliée à des catégories et des stratégies de pensée spécifiques?	76
5.1.1	Possibilité de construire des Matrices Q.....	76
5.1.2	Variations et disparités entre les questions et entre les groupes.....	77

5.1.2.1 Pour les catégories de pensée	77
5.1.2.2 Pour les stratégies de pensée	78
5.2 Deuxième question de recherche : Comment peut-on décrire le potentiel diagnostique d'un TCS pour évaluer le RCI?	79
5.3 Mise en évidence des différentes catégories et stratégies de pensée utilisées selon leur utilisation, pour chaque groupe	79
5.4 RCI des expertes versus RCI des étudiantes	80
5.5 TCS et matrice Q.....	81
5.7 Potentiel diagnostique des questions pour chacun des groupes	81
5.8 Adaptation des questions du TCS	82
6. LIMITES ET PISTES DE RECHERCHE.....	84
CONCLUSION.....	86
RÉFÉRENCES	87
ANNEXES.....	96
Annexe A : Questions de l'étude 2	96
Tableau A : Vignettes et questions du TCS utilisées.....	96
Annexe B : MCA de l'étude 1	99
Tableau B : Le modèle cognitif d'apprentissage (MCA) du raisonnement clinique infirmier (RCI) de Goudreau, Boyer, & Dumont (2012)	99
Annexe C : Calculs des accords entre les questions pour les catégories de pensées	100
Tableau C1 : Accord entre les questions pour les catégories, groupe 1 ^{ère} année	100
Tableau C2 : Accord entre les questions pour les catégories, groupe 2 ^e année	101
Tableau C3 : Accord entre les questions pour les catégories, groupe 3 ^e année	102
Annexe D : Calculs des accords entre les questions pour les stratégies de pensées.....	104
Tableau D1 : Accord entre les questions pour les stratégies, groupe 1 ^{ère} année	104

Tableau D2 : Accord entre les questions pour les stratégies, groupe 2e année.....	105
Tableau D3 : Accord entre les questions pour les stratégies, groupe 3e année.....	106
Tableau D4 : Accord entre les questions pour les stratégies, groupe expertes	107

LISTE DES TABLEAUX

Tableau 1: <i>13 stratégies de pensée utilisée durant le processus de RCI (Fonteyn, 1998)</i>	12
Tableau 2: <i>Les 2 études</i>	26
Tableau 3 : <i>Pourcentages de cas ayant utilisé la catégorie de pensée - Approfondir ou préciser une évaluation</i>	35
Tableau 4 : <i>Pourcentages de cas ayant utilisé la catégorie de pensée - Clarifier politique, procédure</i>	36
Tableau 5: <i>Les 21 catégories sélectionnées et les stratégies de pensées correspondantes</i>	37
Tableau 6 : <i>Les 40 catégories de pensée éliminées et les stratégies de pensées correspondantes</i>	38
Tableau 7 : <i>Pourcentages de cas le plus faible et le plus élevé pour chaque catégorie de pensée</i>	42
Tableau 8 : <i>Catégories de pensée, questions et groupes</i>	45
Tableau 9 : <i>Stratégies de pensée, questions et groupes</i>	47
Tableau 10 : <i>Catégories de pensée, questions, groupe 1^{ère} année</i>	51
Tableau 11 : <i>Stratégies de pensée, questions, groupe 1^{ère} année</i>	53
Tableau 12 : <i>Catégories de pensée, questions, groupe 2^e année</i>	54
Tableau 13 : <i>Stratégies de pensée, questions, groupe 2^e année</i>	56
Tableau 14 : <i>Catégories de pensée, questions, groupe 3^e année</i>	57
Tableau 15 : <i>Stratégies de pensée, questions, groupe 3^e année</i>	59
Tableau 16 : <i>Catégories de pensée, questions, groupe expertes</i>	60
Tableau 17 : <i>Stratégies de pensée, questions, groupe expertes</i>	62
Tableau 18 : <i>Degrés d'accord entre les groupes pour les catégories</i>	64
Tableau 19 : <i>Degrés d'accord entre les groupes pour les stratégies</i>	64
Tableau 20 : <i>Accord entre les questions pour les catégories</i>	67

Tableau 21 : <i>Accord entre les questions pour les stratégies</i>	68
Tableau 22 : <i>La fréquence des valeurs d'accord entre deux items pour les 4 groupes et les catégories de pensée</i>	70
Tableau 23 : <i>La fréquence des nombres de catégories de pensée réellement partagées par couples d'items</i>	70
Tableau 24 : <i>La fréquence des valeurs d'accord entre deux items pour les 4 groupes et les stratégies de pensée</i>	71
Tableau 25 : <i>La fréquence des nombres de stratégies de pensée réellement partagées par couples d'items</i>	71
Tableau 26 : <i>Questions requérant le plus et le moins de catégories de pensée</i>	74
Tableau 27: <i>Questions requérant le plus et le moins de stratégies de pensée</i>	75

LISTE DES FIGURES

Figure 1 : <i>Exemple d'une vignette du TCS en sciences infirmières comprenant deux questions</i>	29
Figure 2 : <i>Le pourcentage de cas ayant utilisé la Catégorie - Faire des choix d'actions pour chaque question</i>	40

LISTE DES ACRONYMES ET ABRÉVIATIONS

CRP : Clinical reasoning problem

ÉCOS : Examen clinique objectif structuré

MCA : Modèle cognitif d'apprentissage

PMP : Patient management problem

QCM : Questionnaire à choix multiples

RCI : Raisonnement clinique infirmier

TCS : Test de concordance de script

INTRODUCTION

L'évaluation est un aspect incontournable du domaine de l'éducation. Huba et Freed (2000, cités dans Tardif, 2006, p. 6) la définissent comme étant « un processus de recueil et de discussion de données à partir de sources multiples et variées dans le but de développer une estimation juste de ce que les étudiants savent, de ce qu'ils comprennent, de ce qu'ils peuvent accomplir avec leurs connaissances compte tenu de leurs expériences de formation ». Selon ces auteurs, le processus d'évaluation est vraiment significatif quand les résultats servent à améliorer les apprentissages subséquents. Or, l'évaluation diagnostique sert justement à améliorer le cheminement d'apprentissage de l'étudiant grâce à la connaissance de son environnement pédagogique et de ses caractéristiques (style cognitif, style d'apprentissage, intérêt, motivation, etc.) qui peuvent avoir des influences individuelles positives ou négatives sur le développement de ses compétences (Legendre, 2005, p.640).

Le raisonnement clinique infirmier est une compétence essentielle de la profession infirmière et il serait utile de mettre en place des mesures évaluatives diagnostiques afin de découvrir où sont les failles et ainsi soutenir efficacement son développement pendant la formation initiale. Dans ce but, cette étude vise à explorer le potentiel diagnostique des vignettes et des questions d'un test de concordance de script (TCS) pour évaluer le raisonnement clinique infirmier (RCI) des étudiants au cours de leur formation.

Pour ce faire, nous avons analysé les réponses apportées à des questions issues d'un TCS par des étudiantes au baccalauréat en sciences infirmières à l'aide de la méthode *Think aloud* et à la lumière des processus de pensée de Fonteyn (1998). Tout d'abord, ce mémoire présente la problématique de la recherche. Les questions telles que la difficulté d'évaluer le RCI dans les programmes de formation, le manque d'outils d'évaluation pertinents et les faiblesses du RCI sont abordées, suivies d'une brève présentation du test de concordance de script (TCS) qu'à l'instar de plusieurs chercheurs (par exemple, Charlin et al, 2010), nous considérons comme un outil d'évaluation du raisonnement clinique très prometteur.

La deuxième partie présente l'état de la question et le cadre conceptuel de cette recherche. Dans cette section, les types d'évaluation sont explicités, dont l'évaluation diagnostique qui nous intéresse principalement. Par la suite, le concept de raisonnement clinique est défini, d'abord dans le domaine médical où le TCS a été développé, puis plus particulièrement en sciences infirmières. Aussi, les principaux instruments d'évaluation du raisonnement clinique sont exposés. Enfin, notre hypothèse et nos questions de recherche sont présentées. La troisième partie explique notre méthodologie de recherche qui exploite les données recueillies d'une étude précédente, afin d'évaluer le potentiel diagnostique de certaines questions du TCS. En quatrième lieu, ce mémoire présente les résultats de la recherche. Ces résultats portent sur les liens et les accords entre certaines questions du TCS et le RCI. Cinquièmement, la section discussion apporte une interprétation des résultats obtenus afin de répondre aux questions de recherche. Enfin, avant de conclure, les limites de cette étude sont rapportées.

1. PROBLÉMATIQUE

Depuis une vingtaine d'année, des changements majeurs de paradigme ont eu lieu dans le domaine de l'éducation. Ces changements accordent désormais une place prépondérante à l'orientation pédagogique axée sur le développement des compétences espérant ainsi améliorer le rendement et l'adaptation des étudiants à leur vie professionnelle future (Lasnier, 2000). Maintenant, tout comme dans les autres domaines de la santé, plusieurs programmes de formation en sciences infirmières ont donc pour but principal le développement des compétences professionnelles (Sibert et al., 2001; Marie et al., 2005).

Compétence cruciale de la profession infirmière, le raisonnement clinique infirmier devrait être développé et soutenu par des programmes de formation adaptés aux besoins et aux caractéristiques des étudiants (Goudreau et al., 2009). En effet, la stimulation des habiletés de raisonnement clinique a été identifiée comme le défi prioritaire de la formation des infirmières afin de s'assurer que les nouvelles infirmières soient prêtes à assumer leur rôle dans la pratique (Lindeman, 2000).

En effet, dans le contexte actuel de complexification des soins, le raisonnement clinique constitue une compétence fondamentale de l'exercice de la profession infirmière. Toutefois, des difficultés de raisonnement clinique infirmier ont été mises en évidence, notamment sur les plans de l'évaluation de l'état de santé (Profetto-McGrath, 2005) et de l'utilisation des résultats probants (Tourangeau, Cranley & Jeffs, 2006). Comme Fero et ses collègues (2009), nous pensons qu'une étude plus approfondie est donc nécessaire afin d'identifier les difficultés des étudiants dans le domaine du raisonnement clinique. Puisque l'évaluation a un impact important sur les stratégies d'apprentissage des étudiants (Sibert & al., 2001; Durak, Caliskan, Bor, & Van der Vleuten, 2007), il apparaît essentiel de développer des instruments d'évaluation valides et fidèles pour identifier les problèmes de RCI dans une approche diagnostique. De tels instruments pourraient et devraient donner des renseignements précis sur le niveau de développement des compétences des candidats sur le plan cognitif et sur leurs difficultés, afin de guider les décisions pédagogiques (Gierl, 2007).

Primordial à tout professionnel de la santé qui travaille dans des situations complexes, le raisonnement clinique exige des connaissances, mais également des attitudes et une pratique professionnelle réfléchie (Deschênes, 2006). Aussi, Fero et ses collègues (2008) relèvent l'importance de développer et d'évaluer des stratégies éducatives pour aider les infirmières à développer leur raisonnement clinique. Il en résulterait des soins plus sécuritaires pour les patients et un rehaussement de la pratique infirmière.

Malheureusement, l'évaluation du raisonnement clinique demeure un problème récurrent en formation des professionnels de la santé (Caire, Sol, Moreau, Isodori & Charlin, 2004; Charlin, Gagnon, Sibert, & Van der Vleuten, 2002). De manière générale en éducation, bien que les pratiques de l'enseignement aient été modifiées, des embûches persistent toujours en matière d'évaluation (Scallon, 2004). Les instruments d'évaluation traditionnels portent leurs limites éducatives (Louis, 1999) pour répondre au mandat du développement des compétences chez les étudiants et les outils de mesure du raisonnement clinique « classiques » présentent quant à eux des problèmes majeurs (Kalamideres, Caire, Dauger, Brassier, & Moreau, 2008). Notre projet de recherche vise donc à explorer des stratégies novatrices pour évaluer le raisonnement clinique de manière diagnostique.

Par ailleurs, les formules pédagogiques visant le développement de compétences nécessitent une contextualisation de l'apprentissage et de l'évaluation afin d'y impliquer activement l'étudiant (Lasnier, 2000; Scallon, 2004). Ainsi, les professeurs doivent utiliser des méthodes d'évaluation qui permettent aux étudiants d'être confrontés à des tâches significatives reflétant la vie professionnelle de manière authentique. Évaluer la capacité des étudiants à utiliser un raisonnement clinique approprié aux situations cliniques s'avère ainsi difficile et exige des mesures d'évaluation pertinentes (Charlin & St-Jean, 2002; Charlin, Gagnon, Sibert, & Van der Vleuten, 2002).

Depuis plusieurs dizaines d'années, les chercheurs et les éducateurs dans le domaine des sciences de la santé ont travaillé à élaborer des instruments d'évaluation qui permettraient de mesurer efficacement le raisonnement clinique (Charlin, Bordage, & Van der Vleuten,

2003). L'un de ces outils, le test de concordance de script (TCS), a attiré particulièrement notre attention car plusieurs études semblent supporter le TCS en termes de validité, fiabilité, faisabilité et applicabilité dans différentes disciplines et divers contextes (Carrière, Gagnon, Charlin, Downing, & Bordage, 2009). Néanmoins, pour l'instant, le TCS semble surtout faire l'objet de recherches, mais être encore relativement peu utilisé dans les programmes de formation en sciences de la santé.

Le test de concordance de script (TCS), issu de la théorie des scripts, est un outil novateur pour l'évaluation du raisonnement clinique (Meterissian, Zabolotny, Gagnon, & Charlin, 2007). Le TCS est un examen standardisé, construit à partir de situations professionnelles authentiques où plusieurs scénarios cliniques sont présentés à l'étudiant. Cet examen évalue la qualité du raisonnement clinique, car il permet de juger si les connaissances de l'étudiant sont élaborées et organisées plutôt qu'accumulées (Charlin, Gagnon, Sibert, & Van der Vleuten, 20002; Charlin, Bordage, & Van der Vleuten, 2003; Charlin, & Van der Vleuten, 2004).

La théorie des scripts, fut amenée par Feltovich et Barrow (1984) et utilisée ensuite par Schimdt, Norman et Boshuizen (1990), puis par Charlin, Tardif et Boshuizen (2000). Cette théorie postule que face à une situation clinique, les professionnels utilisent des scripts mentaux lors d'un processus de raisonnement clinique (Charlin, Boshuizen, Custers, & Feltovich, 2007). Les scripts sont essentiellement décrits comme des structures de liens associatifs de connaissances spécifiquement adaptées aux différentes tâches cliniques (Meterissian, Zabolotny, Gagnon, & Charlin, 2007). Les scripts sont faits de relations entre les maladies, leurs caractéristiques cliniques et leur traitement. Ils permettent l'utilisation efficace des connaissances pour le diagnostic et les options d'enquête et de traitement. Les scripts commencent à apparaître lorsque les étudiants sont confrontés à leur premier cas clinique et sont ensuite développés et affinés tout au long de leur formation, puis de leur carrière. Ils sont à la base du raisonnement clinique qui résulte de multiples interprétations de données en interaction les unes avec les autres et nécessitant le développement de multiples habiletés (Charlin & Van der Vleuten, 2004).

Amenés à développer un raisonnement clinique façonné par un idéal moral à atteindre, à savoir le *Human Caring*¹, les étudiants ont besoin d'outils d'apprentissage et d'évaluation en cours de formation (Deschênes, 2006). Le TCS développé par Deschênes vise l'évaluation du raisonnement clinique afin d'estimer les ressources que les étudiants en sciences infirmières mobilisent lors du traitement d'une situation d'apprentissage nécessitant un raisonnement empreint de *Human caring*. Or, nous pensons que ce TCS pourrait aussi servir d'outil d'évaluation diagnostique de cette compétence. À cet effet, nous sommes d'accord avec l'avis de Kalamarides, Caire, Dauger, Brassier et Moreau (2008, p.25) à ce sujet. Selon eux, cette mesure de connaissances qu'est le TCS comporte plusieurs avantages. Premièrement, il confirme les besoins d'apprentissages, fait percevoir la pertinence des acquisitions à effectuer et permet d'activer les connaissances antérieures des étudiants. De plus, il permet de détecter aisément les champs de connaissances pour lesquels les scripts des étudiants s'avèrent semblables à ceux des experts (et donc pour lesquels il n'est pas nécessaire de faire de formation) et quels sont les champs dans lesquels il existe au contraire un écart qui démontre la nécessité d'une formation.

L'enjeu de l'évaluation diagnostique consiste à déterminer les processus cognitifs mobilisés afin de prendre des décisions permettant de fournir à chaque sujet en difficulté une aide appropriée (Loye, 2005). La présente étude vise à explorer la possibilité d'utiliser des vignettes et questions d'un TCS comme outil d'évaluation diagnostique, afin de cerner les particularités des étudiants en difficulté et ainsi déterminer leurs forces et leurs difficultés.

¹ Pour le concept de *Human caring*, Deschênes (2006 p.10) nous explique la théorie de Watson : « S'inscrivant dans le paradigme de la transformation, le *Human caring* selon Watson est une conception de la discipline infirmière fondée sur des principes phénoménologiques, existentiels, scientifiques et humanistes (Kérouac et al., 2003). Reflétant la richesse de la pensée infirmière, la conception de Watson précise la nature et les liens existant entre les quatre éléments du métaparadigme infirmier : la personne, l'environnement, la santé et le soin ».

Puisque « les tests d'évaluation diagnostique s'intéressent à la démarche qui leur permet d'aboutir à la réponse, plutôt qu'à la réponse elle-même » (Loye, 2005, p.52), nous croyons que le TCS pourrait s'avérer un outil intéressant pour évaluer le raisonnement clinique des étudiants en sciences infirmières. Mais quel est le potentiel d'un TCS comme outil d'évaluation diagnostique? Nous tenterons de répondre à cette question dans les paragraphes qui suivent.

2. ÉTAT DE LA QUESTION ET CADRE CONCEPTUEL

Ce chapitre présente la typologie de l'évaluation, les visées de l'évaluation et les particularités de l'évaluation diagnostique en éducation. De même, il sert à définir le raisonnement clinique dans le domaine de la santé en général, puis, en sciences infirmières. Enfin, une comparaison des différents outils permettant l'évaluation du raisonnement est explicitée afin d'appuyer le choix d'utiliser le test de concordance de script (TCS) dans cette étude. Comme ces outils sont issus d'études empiriques, ce chapitre s'intéressera aussi à certaines d'entre-elles.

2.1 Typologie de l'évaluation

Bloom et ses collaborateurs (1981, cités dans Scallon, 1988), ont proposé une typologie de l'évaluation. Ces auteurs distinguent trois types d'évaluation : formative, sommative (ou certificative) et diagnostique. L'évaluation sommative (ou certificative, termes discutés dans Tardif, 2006) sert de bilan, et intervient après un ensemble d'activités d'apprentissage. Cette évaluation a souvent pour objectif de classer les étudiants, d'attribuer un certificat (Durand & Chouinard, 2006) ou de déterminer un niveau de développement (Tardif, 2006). Pour sa part, l'évaluation formative est effectuée pendant l'apprentissage. Elle a pour objectif d'informer l'étudiant et l'enseignant du niveau de maîtrise atteint et de définir où et en quoi l'étudiant éprouve des difficultés d'apprentissage, afin de pouvoir lui suggérer ou lui faire découvrir de nouvelles stratégies pour progresser (Durand & Chouinard, 2006). Elle vise à réguler (Durand & Chouinard, 2006) et à soutenir (Tardif, 2006) les apprentissages. D'après Legendre (2005), la pratique de l'évaluation formative permet d'identifier les étudiants qui ne profitent pas suffisamment des correctifs qui leurs sont présentés et pour qui l'évaluation diagnostique serait profitable. L'évaluation formative peut se présenter sous forme d'évaluation diagnostique car elle a une perspective de relation d'aide et de régulation des apprentissages. Notre étude s'intéresse à l'évaluation diagnostique, qui doit aider à découvrir les forces, les faiblesses et le niveau de préparation des étudiants. Elle peut se

faire avant une période importante d'enseignement et d'apprentissage, un cours ou un programme d'études ou pendant le déroulement de cette même séquence (Scallon, 1988). L'évaluation diagnostique permet à l'enseignant de repérer les acquis et les difficultés des étudiants. Il peut alors adapter efficacement son enseignement à leurs besoins (Dossat, 2002).

Selon Rey, Carette, Defrance et Kahn (2006), toute évaluation se doit d'être diagnostique et doit permettre aux enseignants et aux chercheurs de récolter des informations concernant les capacités de chaque étudiant à mobiliser les connaissances et les procédures qu'il a acquises. « Si les épreuves sont diagnostiques, c'est-à-dire qu'elles permettent de mettre à jour des informations pertinentes quant aux facilités ou aux difficultés des élèves confrontés à la résolution de tâches complexes, c'est l'utilisation qu'en feront les enseignants qui déterminera son caractère formatif ou certificatif » (p.44).

Puisque notre étude s'intéresse particulièrement à l'évaluation diagnostique dans un contexte de formation, nous l'expliquerons plus exhaustivement dans les paragraphes suivants.

2.2 L'évaluation diagnostique

L'évaluation diagnostique est utilisée depuis longtemps dans les domaines de l'éducation et de la psychologie (Roussos, Templin, & Henson, 2007). En éducation, elle cherche à reconnaître les particularités des étudiants en examinant leurs réponses aux questions dans un test (Loye, 2008). Son enjeu vise à apprécier les caractéristiques individuelles du sujet (Legendre, 2005) en lien avec ses savoirs ou ses savoirs faire (Loye, 2008). L'évaluation diagnostique permet d'identifier les difficultés individuelles (Scallon, 1988; Legendre, 2005), d'en déceler les causes et de déterminer des mesures correctives appropriées. Elle favorise donc la prise de décision pour procurer de l'aide à chaque étudiant en difficulté (Scallon, 1988). Enfin, dans le cadre de notre étude, nous pensons que l'évaluation diagnostique devrait permettre une meilleure compréhension de la

manière de penser, de raisonner et de résoudre des problèmes des étudiantes en sciences infirmières qui doivent raisonner dans un contexte clinique.

La prochaine partie expliquera donc le concept de raisonnement clinique, tout d'abord dans le domaine de la santé en général, puis, particulièrement en sciences infirmières.

2.3 Le raisonnement clinique

On retrouve plusieurs similarités entre les études relatives aux stratégies cognitives impliquées dans la prise de décision clinique des infirmières et les études faites à ce sujet dans le domaine médical. Ces dernières étant davantage exhaustives et élaborées depuis la dernière décennie, elles nous permettent une meilleure conceptualisation du raisonnement clinique (Deschênes, 2006).

Dans le domaine de la santé, le raisonnement clinique est considéré comme l'ensemble des pensées et des processus de décision associés à la pratique clinique. Il permet aux praticiens de choisir la meilleure action dans un contexte particulier (Audétat & Laurin, 2010). Le raisonnement clinique consiste à acquérir, organiser et interpréter des informations; proposer et tester des hypothèses; et évaluer des diagnostics alternatifs pour, finalement, poursuivre un traitement (Elstein & Schwartz, 2002 cités dans Elizondo-Omana, 2010). Cette activité intellectuelle synthétise l'information obtenue à partir de la situation clinique et l'utilise pour prendre des décisions de diagnostic et de prise en charge du patient en intégrant les connaissances et expériences antérieures (Charlin, Bordage, & Van der Vleuten, 2003).

Le raisonnement clinique est donc un processus complexe multidimensionnel qui utilise la cognition, la métacognition et les connaissances spécifiques à chaque discipline pour recueillir et analyser l'information du patient, évaluer son importance et penser à des actions alternatives (Higgs, Jones, Loftus, & Christensen, 2008; Simmons, 2010; Dunphy, 2010).

2.4 Le raisonnement clinique infirmier (RCI)

Depuis les 25 dernières années, plusieurs études ont été réalisées dans le but de documenter le raisonnement clinique en sciences infirmières (Funkesson, Anbäcken, & Ek, 2007). Le raisonnement clinique infirmier est souvent associé aux phénomènes de *jugement clinique*, de *pensée critique*, de *prise de décision clinique*, de *résolution de problème*, de *raisonnement diagnostique* (Funkesson, Anbäcken, & Ek, 2007; Goudreau et al., 2009; Simmons, 2010). Afin de circonscrire la nature même de ce processus, nous utiliserons le terme *raisonnement clinique* pour spécifier cette compétence fondamentale de l'infirmière.

Certaines recherches ont décrit la nature complexe du raisonnement clinique chez les infirmières par les processus cognitifs que celles-ci utilisent afin de comprendre et donner une signification aux données recueillies auprès d'un patient et ainsi, établir un plan de traitement qui réponde à ses besoins (Kautz, Kuiper, Pesut, Knight-Brown, & Daneker, 2005).

Par exemple, Tanner, Padrick, Westfall et Putzier (1987) ont indiqué que le processus de raisonnement clinique des infirmières et des étudiantes implique l'activation précoce d'hypothèses et que des informations sont systématiquement recueillies pour retenir ou éliminer ces hypothèses, ainsi le RCI se dit de nature hypothético-déductive (Deschênes, 2006). Ces chercheurs ont observé un recueil plus systématique des données et une plus grande exactitude diagnostique chez les infirmières plus expérimentées.

De leur côté, Fonteyn et Ritter (2008) définissent le raisonnement clinique comme « l'ensemble des stratégies et processus cognitifs que les infirmières utilisent pour comprendre la signification des données du patient, pour identifier et diagnostiquer les problèmes actuels ou potentiels du patient, pour prendre des décisions cliniques qui contribuent à la résolution du problème, et pour atteindre des résultats positifs pour le patient. » (p. 107). Goudreau et ses collègues (2009) l'expliquent comme étant « le processus cognitif complexe de réflexion et de décision associé à l'action clinique » (p.9).

Les études menées en sciences infirmières ont permis d'identifier différents éléments du raisonnement clinique : connaissances spécifiques à chaque discipline, expérience et stratégies de pensée (Funkesson, Anbäcken, & Ek, 2007; Simmons, 2010). Ces dernières constituent un élément précieux du raisonnement clinique et ont été identifiées par Fonteyn (1998). Les résultats de son étude rapportent que le raisonnement clinique des infirmières se compose de différentes stratégies de pensée. Fonteyn a effectué une étude descriptive basée sur les stratégies de pensée de 15 infirmières reconnues pour leur expertise dans leur domaine. La méthode *Think aloud* a été utilisée sur 45 dilemmes cliniques. Les infirmières devaient donc dire à voix haute toutes leurs pensées, sans temps d'arrêt pour réfléchir. Le verbatim a été transcrit et analysé qualitativement à l'aide du logiciel NUD-IST. Cette analyse a permis d'identifier 13 opérations cognitives utilisées par les infirmières dans leur pratique : 1) reconnaître un pattern, 2) ordonner des concepts par priorités, 3) chercher des informations manquantes, 4) faire des hypothèses, 5) prédire, 6) relier des informations, 7) énoncer une proposition, 8) énoncer une règle de pratique, 9) faire des choix, 10) juger de la valeur, 11) conclure, 12) fournir des explications et 13) autres stratégies de pensée. Chacune de ces stratégies se décline en catégories. Il y a 61 catégories au total. Le modèle de Fonteyn a été choisi pour notre étude parce qu'il utilise des stratégies de cognitives détaillées et se base sur des résultats de recherche empirique. Le tableau 1 présente le modèle de Fonteyn (1998) : les 13 stratégies de pensée, ses catégories et ses définitions.

Tableau 1: 13 stratégies de pensée utilisée durant le processus de RCI (Fonteyn, 1998)

Stratégies de pensée	Catégories de pensée	Définitions
1. Reconnaître un pattern	1. Cas type 2. Protocole et traitement standard 3. Représenter une situation familière 4. Représenter une situation non-familière	Identifier un schéma, des informations qui sont liées entre elles.
2. Ordonner des concepts par priorités	5. Focus sur plan d'action 6. Focus sur	Ordonner des concepts en ordre

	préoccupation(s) du patient	d'importance ou d'urgence.
3. Chercher des informations manquantes	7. Signes vitaux 8. Approfondir ou préciser une évaluation 9. Information logistique 10. Histoire du patient 11. Déterminer le plan de soin 12. Résultats des tests 13. Médication 14. Ressources 15. Procédures, politiques, droits	Chercher des informations, des données manquantes.
4. Faire des hypothèses	16. Déterminer la cause 17. Identifier maladie/problème de santé 18. Spéculer sur le statut du patient 19. Identifier les besoins du patient 20. Intervention ou traitement	Énoncer des explications possibles, potentielles pour expliquer un ensemble de faits.
5. Prédire	21. Interventions 22. Résultats des soins (ou interventions) 23. Découvertes, résultats d'examens 24. Situation 25. Événements 26. Réponses 27. Besoins du patient	Faire des prédictions.
6. Relier des informations	28. Conclusions d'évaluations/autres informations 29. Information sur le traitement et autre type d'information 30. Histoire du patient et problèmes/situation du patient 31. Résultats des tests et contexte 32. Données organisationnelles	Relier des informations pour mieux comprendre.

7. Énoncer une proposition	33. Établir ou écarter un problème 34. Choisir un plan d'action 35. Déterminer la cause 36. Juger intervention, traitement 37. Évaluer les données 38. Clarifier une politique, procédure	Énoncer une règle basée sur si... alors...
8. Énoncer une règle de pratique	39. Politique 40. Procédures 41. Précepte, règle informelle	Énoncer une évidence basée sur la pratique
9. Faire des choix	42. Interventions infirmières 43. Actions 44. Traitement (pharmacologique ou procédural) 45. Tests	Choisir parmi des alternatives possibles.
10. Juger de la valeur	46. Des conclusions d'évaluation 47. Traitements (ou interventions) 48. Résultats de tests	Juger de la valeur en termes d'utilité, de sens, d'importance.
11. Conclure	49. Condition du patient ou statut 50. Une situation	Décider, se former une opinion.
12. Fournir des explications	51. Soins thérapeutiques 52. Interventions infirmières 53. Sur des actions 54. Test 55. Préoccupations (inquiétudes) 56. Prédications	Justifier ses actions, croyances, commentaires.
13. Autres stratégies de pensées	57. Faire une pause réflexion 58. Qualifier, nuancer 59. Poser une question 60. Supposer 61. Généraliser, inférer	

Les résultats de Fonteyn ont joué un rôle essentiel dans la recherche en formation en sciences infirmières, et nous portent à nous intéresser davantage à l'évaluation du

raisonnement clinique, notamment aux instruments, dans la prochaine partie de ce mémoire.

2.5 L'évaluation du raisonnement clinique

Plusieurs instruments pour évaluer le raisonnement clinique ont été développés dans le domaine des sciences de la santé. Quelques-uns de ces instruments seront ici présentés soit le *patient management problem (PMP)*, la grille d'évaluation globale, le *questionnaire à choix multiples (QCM)*, l'*examen clinique objectif structuré (ÉCOS)*, le *Clinical Reasoning Problem (CRP)* et enfin, le *test de concordance de script (TCS)*. Nous y porterons un regard critique dans le but de mettre en évidence la pertinence du TCS.

Le *patient management problem (PMP)* débute avec un nombre variable d'informations sur le patient. L'étudiant est alors invité à recueillir d'autres données de façon séquentielle, linéaire ou ramifiée, puis à écrire ses réponses (sur papier ou par ordinateur). Après avoir fait la collecte de données, comme s'il agissait en situation réelle, l'étudiant peut choisir des investigations et/ou prendre des décisions de diagnostic et de gestion. Le groupe d'étudiants est comparé à un groupe d'experts et des scores sont attribués en fonction du caractère complet de la collecte de données, de son efficacité et de sa pertinence (Charlin, Bordager, & Van der Vleuten, 2003). Le problème avec cette méthode est que la performance sur un PMP donne une mince prédiction de la performance sur un autre PMP. Selon plusieurs études la corrélation entre les problèmes est faible (0,1-0,3 d'après Norman et al., 2005 cités dans Higgs, Jones, Loftus, & Christensen, 2008). Le PMP a été contesté pour ses limites psychométriques, mais aussi en raison de l'impossibilité de déterminer un cheminement constant d'expert auquel serait comparé celui des étudiants (Higgs, Jones, Loftus, & Christensen, 2008).

La *grille d'évaluation globale* énumère une liste de critères et de comportements à observer chez l'étudiant. Elle est complétée après quelques rencontres entre l'étudiant et le(s) observateur(s). Cette grille est destinée à refléter l'ensemble des observations effectuées au cours d'un stage. Elle présente une certaine validité par l'exhaustivité des

critères d'évaluation de la compétence clinique (Norman, 1993 cité dans Charlin, Bordage, & Van der Vleuten, 2003). Par contre, les critères pour évaluer le raisonnement clinique ne représentent généralement qu'une portion de la grille (Charlin, Bordage, & Van der Vleuten, 2003, p. 45). Comme l'ont fait plusieurs auteurs, nous critiquons la faible fidélité de cet outil, causée par plusieurs facteurs. Ces facteurs sont le jugement par un nombre restreint d'observateurs, les erreurs causées par de longues périodes d'observation, le jugement porté sur quelques épisodes de soins seulement, ainsi que la nature subjective des données récoltées avec cet outil d'évaluation (Charlin, Bordage, & Van der Vleuten, 2003).

Le *questionnaire à choix multiples (QCM)* est construit à partir de questions qui sont suivies de choix de réponses. La plupart des critiques sur cet instrument ont trait à l'effet induit sur les apprentissages, avec une incitation à l'apprentissage superficiel dit « par cœur » et une survalorisation des connaissances factuelles au détriment de la capacité de résolution de problèmes (Charlin, Bordage, & Van der Vleuten, 2003). De plus, lorsqu'une sélection limitée de réponses possibles est présentée, un étudiant peut reconnaître la bonne réponse plutôt que de la générer, améliorant ainsi son rendement à l'examen (Charlin, Bordage, & Van der Vleuten, 2003; Park et al., 2010). Toutefois, le QCM est un outil qui affiche un haut degré de fidélité par le fait qu'il s'agit d'un choix de réponses et il permet d'évaluer rapidement un grand nombre d'étudiants sur un large spectre de connaissances. Il permet aussi l'attribution d'un score objectif (Charlin, Bordage, & Van der Vleuten, 2003).

L'*examen clinique objectif structuré (ÉCOS)* « évalue la démarche d'un étudiant par un observateur externe à partir de situations cliniques simulées (vrais patients, vrais acteurs) » (Charlin, Bordage, & Van der Vleuten, 2003, p.47). L'observateur utilise une grille d'évaluation prédéfinie comprenant les réponses attendues. Plusieurs comportements et habiletés distincts sont évalués lors d'un ÉCOS par la diversité des situations cliniques utilisées (Van der Vleuten & Swanson, 1990). Cet outil permet d'évaluer efficacement la démarche clinique par l'observation directe de comportements et d'aptitudes dans différentes situations cliniques simulées (Van der Vleuten &

Swanson, 1990). L'ÉCOS présente une bonne fidélité inter-juge, mais n'évalue que des comportements observables en lien avec l'évaluation du raisonnement clinique (Charlin, Gagnon, Sibert, & Van der Vleuten, 2002; Charlin, Bordage, & Van der Vleuten, 2003). En effet, bien que l'on perçoive une certaine habileté de l'étudiant dans sa collecte d'informations, il est difficile d'évaluer les subtilités du processus de raisonnement clinique qui se passe à l'interne (Charlin, Bordage, & Van der Vleuten, 2003).

Le *CRP (Clinical Reasoning Problem)* consiste en un scénario clinique incluant une présentation, une histoire et un examen physique présentés aux étudiants. À l'aide de ces informations, on demande aux sujets de désigner deux diagnostics et d'indiquer si les caractéristiques présentées supportent ou non les diagnostics choisis. Ensuite, on compare la réponse à celles d'un groupe d'experts (il n'y a pas nécessairement une seule bonne réponse). Selon Groves, Scott et Alexander (2002), le CRP est une méthode fiable pour évaluer le raisonnement clinique à différents niveaux de compétence. Le CRP est un instrument qui ressemble au test de concordance de script (TCS) (Higgs, Jones, Loftus & Christensen, 2008) car il vise l'évaluation du processus de raisonnement clinique et non un résultat. Cependant, il est différent car il peut évaluer deux autres aspects du raisonnement clinique à savoir l'identification des informations pertinentes et la génération d'hypothèses (Groves, Scott, & Alexander 2002).

Enfin, le *test de concordance de script (TCS)* est standardisé et construit à partir de situations professionnelles authentiques où plusieurs courts scénarios cliniques (ou vignettes) sont présentés à l'étudiant (Charlin, Gagnon, Sibert, & Van der Vleuten, 2002). Chaque vignette pose plusieurs hypothèses diagnostiques et vise à déterminer l'effet de plus amples informations sur ces hypothèses (Groves, Scott, & Alexander 2002). Ce test écrit évalue la qualité du raisonnement clinique, car il permet de juger si les connaissances de l'étudiant sont élaborées et organisées plutôt qu'accumulées (Charlin, Gagnon, Sibert, & Van der Vleuten, 2002; Charlin, Bordage, & Van der Vleuten, 2003; Charlin & Van der Vleuten, 2004). Contrairement au questionnaire à choix multiple, le TCS évalue les compétences des candidats et comment ils utilisent l'information pour

confirmer ou éliminer des hypothèses (Meteressian, Zabolotny, Gagnon, & Charlin, 2007).

Cette recension des instruments d'évaluation existants permet de clarifier leurs avantages et leurs limites. Bien que les études fassent la description d'instruments d'évaluation souvent intéressants, nous constatons un manque d'uniformité dans les démarches et le peu de développement en sciences infirmières. De même, nous décelons une différence de « sensibilité » et de spécificité entre les instruments utilisés pour évaluer le raisonnement clinique (Boychuk Duchscher, 2003). Enfin, comme le disent Charlin, Bordage et Van der Vleuten (2003), la plupart des instruments d'évaluation du raisonnement clinique, comme le QCM, l'ÉCOS ou la grille d'évaluation globale, visent à mesurer des comportements observables éloignés de la nature complexe de ce processus cognitif.

De fait, le TCS est le moins critiqué négativement, comporte plusieurs avantages énumérés ci-dessus et fait l'objet d'études prometteuses. Par exemple, des études, réalisées dans certaines spécialités médicales telles la gynécologie (Park et al., 2010), ou l'urologie (Sibert et al., 2001), ont démontré que le TCS permettrait de distinguer différents groupes de participants, en fonction de leurs niveaux d'expertise. De plus, le TCS vise à évaluer, outre la richesse des connaissances factuelles, le raisonnement et la compétence cliniques (Marie et al., 2005). Il apprécie la qualité de l'organisation des connaissances, de même que la pertinence de cette organisation pour agir dans des situations cliniques données qui appartiennent à la réalité de la pratique professionnelle. En conclusion, le choix du TCS semble donc des plus pertinents pour notre étude qui consiste à explorer le potentiel diagnostique du TCS pour évaluer le RCI.

2.6 Description, application et qualités psychométriques du TCS

Les scénarios (ou vignettes) présentés dans le TCS visent à évaluer le raisonnement clinique dans un contexte d'incertitude (Caire, Sol, Moreau, Isidori, & Charlin, 2004; Gagnon, Charlin, Lambert, Carrière, & Van der Vleuten, 2009; Ruiz et al., 2010;

Remaekers, Kremer, Pilot, Beukelen, & Keulen, 2010), illustrant une situation proche de la vie réelle et selon des circonstances professionnelles et authentiques (Charlin, Gagnon, Sibert, & Van der Vleuten, 2002; Charlin & Van der Vleuten, 2004). La situation décrite pose même un problème pour un expert puisqu'il manque un ou plusieurs éléments ou que certaines données sont difficilement interprétables, incomplètes ou ambiguës (Charlin, Gagnon, Sibert, & Van der Vleuten, 2002; Charlin & Van der Vleuten, 2004). Le TCS est conçu pour apprécier la qualité et la pertinence de l'organisation des connaissances de l'étudiant pour agir efficacement dans un contexte de pratique professionnelle (Deschênes, 2006).

Les éléments proposés dans le test ressemblent sensiblement à ceux qu'explorent les experts dans une situation clinique. De plus, les réponses proposées correspondent à des hypothèses ou à des options pertinentes et justifiées bien qu'elles puissent varier selon les experts du domaine. Avec l'aide d'un groupe d'experts, la conception du TCS est effectuée selon les objectifs d'apprentissage visés et selon une méthodologie bien précise (Charlin, Gagnon, Sibert, & Van der Vleuten, 2002; Deschênes, 2006, 2012).

Les praticiens d'expérience possèdent des réseaux de connaissances élaborées reliés à certaines tâches cliniques, appelés *scripts* (Meteressian, Zabolotny, Gagnon, & Charlin, 2007). Selon Kalamideres et ses collègues (2008, p.22) « les scripts sont des réseaux de connaissances qui permettent d'agir en situation clinique et qui, une fois activés, sont confrontés à la situation proposée et puis rejetés ou retenus ». Le TCS permet d'évaluer l'élaboration des scripts de l'étudiant en lien avec l'élaboration des scripts chez les experts, d'où le nom de test de concordance (Charlin, Gagnon, Sibert, & Van der Vleuten, 2002; Charlin & Van der Vleuten, 2004). Le principe de ce test consiste à présenter aux étudiants une série de situations cliniques authentiques, puis à leur demander de considérer l'effet d'une nouvelle information (positive, neutre, négative) sur une hypothèse crédible. Une échelle de type Likert à cinq niveaux² est proposée pour

² Légende : -2 : l'hypothèse est pratiquement éliminée; -1 : l'hypothèse devient moins probable; 0 : l'information n'a aucun effet sur l'hypothèse; 1 : l'hypothèse devient plus probable : 2 : il ne peut pratiquement s'agir que de cette hypothèse

répondre à chaque question. Un groupe d'experts du domaine sert de référence. Leurs réponses sont utilisées pour attribuer un score à chaque étudiant. Le crédit donné aux étudiants pour chaque réponse est fonction du nombre d'experts qui ont fourni la même réponse qu'eux (Sibert & al., 2001; Charlin, Bordage, & Van der Vleuten, 2003; Marie & al., 2005).

Dans plusieurs domaines de la médecine, diverses études visent à développer des TCS comme modalité d'évaluation formative. Citons, par exemple : la neurologie (Lubarsky, Chalk, Kazitani, Gagnon, & Charlin, 2009), la pédiatrie d'urgence (Carrière, Gagnon, Charlin, Downing, & Bordage, 2009), la radio-oncologie, la chirurgie, les sciences infirmières (Charlin et al., 2010), la gynécologie (Park et al., 2010), l'urologie (Sibert et al., 2001), la médecine interne (Marie et al., 2005), la chirurgie générale (Meteressian, Zabolotny, Gagnon, & Charlin, 2007), la gériatrie (Ruiz et al., 2010). Les caractéristiques étudiées sont : le moment de l'évaluation et les effets des différents formats, l'optimisation des méthodes de notation et la composition d'un panel de référence. Les résultats de ces études ont été comparés entre différents niveaux d'expérience clinique, de cultures et de contextes d'apprentissage (Remaekers, Kremer, Pilot, van Beukelen, & van Keulen, 2010, p.663).

Plusieurs recherches ont été conduites sur le TCS afin d'en évaluer les qualités psychométriques. Ces études ont abordé la validité, la fiabilité, la faisabilité et l'applicabilité dans différentes disciplines et divers contextes et semblent soutenir positivement le TCS (Carrière et al. 2009). En effet, le TCS a affiché une mesure de fiabilité élevée (Meteressian, Zabolotny, Gagnon, & Charlin, 2007), avec un coefficient alpha de Cronbach autour de 0,80, moyennant le recours à environ 80 items dans le test (Higgs, Jones, Loftus, & Christensen, 2008). La validité de construit de l'instrument a été démontrée par la progression linéaire des scores entre des groupes de niveaux d'expertise différents. La validité de construit signifie : « *l'adéquation entre un test et le cadre théorique dans lequel il est construit* » (Charlin, Gagnon, Sibert, & Van der Vleuten, 2002, p. 141). Le TCS a ainsi démontré sa capacité à détecter les candidats plus expérimentés cliniquement et à distinguer les différents groupes de participants en

fonction de leur niveau d'expertise; les scripts des experts étaient davantage élaborés, compilés et organisés que ceux des novices (Charlin & Van der Vleuten, 2004).

Dans le domaine des sciences infirmières, Deschênes (2006) a développé un TCS afin d'évaluer le raisonnement clinique empreint du *Human caring* dans des situations cliniques diverses de la pratique infirmière. Les vignettes visaient à évaluer le RCI dans des situations de soins courants. Deschênes a élaboré son TCS selon les quatre étapes suivantes : 1) la sélection des situations nécessitant un raisonnement empreint du *Human caring*, 2) la rédaction de scénarios cliniques et des items du test, 3) la validation par 2 juges-experts collaborateurs qui ont été sélectionnés selon les critères suivants : bonne connaissance de la théorie du *Human caring* et diplôme de deuxième ou troisième cycle en sciences infirmières, et enfin 4) la construction des grilles de réponses avec les réponses fournies par un panel de 15 experts (recrutés pour leur expérience et leur connaissance de la théorie du *Human caring*). Par la suite, le test a été complété par un échantillon de 30 étudiantes ayant accepté de participer à l'étude. Ces étudiantes étaient en première année de baccalauréat en sciences infirmières de l'Université de Montréal.

L'analyse des données comportait 2 volets : l'analyse psychométrique du TCS et la comparaison des scores des 2 groupes de répondants avec un test *t* et le coefficient alpha de Cronbach. Les résultats révèlent que cet indice de fiabilité est élevé (0,86) et que la corrélation entre les différentes dimensions du test varie entre 0,50 et 0,67 et est statistiquement significative. De plus, la différence de scores entre les experts du panel et les étudiantes s'est avérée statistiquement significative dans l'ensemble du test et dans chaque dimension d'évaluation du *Human caring que sont la personne, l'environnement, la santé et le soin*. Les étudiantes ont obtenu des scores plus faibles que les experts et leurs scores ont affiché une plus grande variabilité au sein du groupe.³ Nous pouvons conclure que le TCS de Deschênes (2006) a démontré des propriétés psychométriques adéquates et offre ainsi une méthode standardisée d'évaluation du RCI en sciences

³ Pour les experts, le score de chaque participant a été obtenu en utilisant, comme panel de référence pour la grille de cotation, la distribution des réponses de tous les autres experts à l'exclusion de ses propres réponses.

infirmières. Nous avons donc retenu ce test pour la présente étude afin d'en étudier le potentiel diagnostique.

2.7 De l'évaluation formative au test diagnostique

Étudier le potentiel diagnostique du TCS revient à s'interroger sur sa capacité à distinguer les particularités individuelles des étudiants et à déterminer leurs forces et leurs faiblesses en «s'intéress[ant] à la démarche qui [leur] permet d'aboutir à la réponse, plutôt qu'à la réponse elle-même » (Loye, 2005, p.52). Rappelons, que nous avons choisi le modèle des stratégies de pensée de Fonteyn pour définir les processus cognitifs sous-jacents au RCI.

Mais comment utiliser le TCS dans cette optique? Tout d'abord, l'objet d'évaluation, la forme du test et la façon dont les résultats seront interprétés devraient être déterminés lors de la conception d'un test diagnostique (Loye, 2005 p.52). Il faut donc chercher à identifier si des stratégies de pensée spécifiques sont sous-jacentes à chaque question du TCS et varient d'une question à une autre.

D'après nous, une excellente manière de formaliser les relations entre les processus cognitifs et les questions d'un test a été proposée par Tatsuoka (1983) sous la forme d'une matrice nommée Q. Cette dernière permet d'opérationnaliser les liens entre les processus cognitifs, par exemple les stratégies de pensées, et les questions (Loye, 2008, pp 3-4). Elle prend la forme d'un tableau dans lequel les questions constituent les colonnes et les processus cognitifs les lignes. Une valeur 1 à l'intersection d'une ligne et d'une colonne indique que le processus est requis par la question. Dans notre étude, nous tenterons donc d'élaborer une matrice Q à partir des vignettes et des questions du TCS de Deschênes (2006) qui ont été utilisées dans l'étude de Goudreau et al., (2009) en les croisant avec les stratégies (ou les catégories) de pensée de Fonteyn (1998) dans une procédure inductive telle que proposée par Loye (2008) :

« La matrice Q peut également être obtenue par induction, après la création du test. Les experts doivent identifier les attributs cognitifs et leurs liens avec les items avec la contrainte d'utiliser des items déjà existants. Ils peuvent utiliser une liste préalable

d'attributs cognitifs pour établir les liens ou élaborer eux-mêmes cette liste en plus d'établir les liens » (Loye, 2008, p.5).

La matrice Q, si elle est possible à mettre en évidence, fournit la structure diagnostique d'un test. Les réponses fournies aux questions du test par les étudiants pourraient alors être exploitées pour réaliser le diagnostic en identifiant les questions réussies ou non dans le test, voir ce que partagent ces questions et regarder les erreurs commises (Loye, 2008).

Ainsi, nous pensons que structurer un test de concordance de script (TCS) comme celui de Deschênes (2006) sous la forme d'une matrice Q, pourrait être la première étape pour diagnostiquer les forces et les faiblesses sous-jacentes au raisonnement clinique infirmier (RCI). Chercher à élaborer une telle matrice offre ainsi la possibilité d'étudier le potentiel diagnostique des vignettes et questions d'un TCS en analysant les liens avec les stratégies (et les catégories) de pensée de Fonteyn. Notre hypothèse est que les étudiants qui proposent des réponses correctes aux questions posées n'utilisent pas les mêmes stratégies que ceux dont les réponses sont erronées. Ainsi, si cette hypothèse s'avère exacte, il devrait être possible de formaliser les relations entre chaque question et les stratégies nécessaires à une réponse acceptable. Évidemment, l'utilisation du TCS dépend d'un panel d'experts pour juger de la qualité de la réponse proposée, cet aspect devrait donc être pris en compte pour différencier une réponse acceptable d'une réponse qui ne l'est pas.

2.8 Questions de recherche

Ainsi, de notre question générale de recherche qui était : « quel est le potentiel d'un TCS comme outil d'évaluation diagnostique? », nous en arrivons à nos questions spécifiques de recherche :

- 1- Dans quelle mesure chaque question d'un TCS visant à évaluer le RCI peut-elle être reliée à des catégories et à des stratégies et de pensée spécifiques ?

2- Comment peut-on décrire le potentiel diagnostique d'un TCS pour évaluer le RCI?

Afin de pouvoir répondre à ces questions de recherche, nous avons fait certains choix méthodologiques qui seront décrits dans la prochaine partie.

3. MÉTHODOLOGIE

La présente recherche fait suite à une étude menée par Goudreau et ses collègues en 2012. Dans ce chapitre, nous distinguons clairement les méthodologies des deux études qui utilisent les mêmes données. Dans les paragraphes qui suivent, nous identifions le type de recherche et ses visées, les participants sollicités et la méthode d'échantillonnage, l'instrument de la cueillette de données ainsi que ses procédures, les considérations éthiques, et les méthodes d'analyse.

3.1 Type de recherche

Une recherche exploratoire a pour but l'obtention d'une plus grande connaissance d'un phénomène et la clarification de concepts. Ce type d'étude consiste à explorer les domaines de recherche théoriquement peu développés. Puisque notre étude utilise les données qualitatives d'une étude antérieure et que nous n'interviendrons que très peu, nous en concluons que notre étude est de type *exploratoire* (Lamoureux, 2006).

Ce projet est une analyse secondaire des données de l'étude de Goudreau, Boyer et Dumont (2012), intitulée « *Un modèle cognitif d'apprentissage du raisonnement clinique infirmier* » subventionnée par la Fondation canadienne de la recherche sur les services de santé (FCRSS) dans lequel nous travaillons à titre de *coordonnatrice de recherche*. Le tableau 2 présente les titres et auteurs des deux études successives.

Tableau 2: Les 2 études

Ordre	Études
Étude 1	Un modèle cognitif d'apprentissage du raisonnement clinique infirmier Goudreau, Boyer, & Dumont (2012)
Étude 2	Étude exploratoire du potentiel diagnostique d'un test de concordance de scripts (TCS) pour évaluer le raisonnement clinique infirmier (RCI) Dumont (2013)

3.2 Description de l'étude 1

3.2.1 Objectif de l'étude 1

L'objectif de l'étude 1 était d'élaborer le modèle cognitif d'apprentissage (MCA) du raisonnement clinique infirmier (RCI), du début de l'apprentissage jusqu'à l'expertise, à savoir mettre en évidence les apprentissages critiques correspondant à des niveaux de compétence. Par conséquent, cette étude a précisé les processus en lien avec les participants.

3.2.2 Participants sollicités dans l'étude 1

Les participants sollicités lors de l'étude 1 sont 41 étudiants au baccalauréat en sciences infirmières de l'Université de Montréal dont 14 de 1^{ère} année, 14 de 2^e année, 13 de 3^e année, 11 infirmières nouvellement diplômées (ayant 12-18 mois d'expérience clinique) et 14 infirmières expertes (reconnues par leurs pairs comme exerçant un bon raisonnement clinique, ayant plus de 5 ans d'expérience auprès de la clientèle). La majorité de ces personnes sont des femmes.

Le recrutement des étudiants s'est fait tout d'abord par des invitations écrites à l'ensemble des cohortes d'étudiantes de l'université de Montréal, suivies du tour des

classes pour présenter la recherche. Quant au recrutement des infirmières nouvellement diplômées et des expertes, des lettres accompagnant le relevé de paie ont été envoyées dans les milieux partenaires (CHUM et CHU Ste-Justine).

L'échantillonnage est critérié, soit un groupe d'étudiants par année d'un programme de trois ans, avec recherche de variation quant à la formation antérieure au Québec et aux résultats scolaires; un groupe d'infirmiers en début de pratique afin de représenter l'étape « critique » d'intégration au milieu de travail et un groupe d'infirmiers experts afin de représenter un niveau élevé de compétence, avec recherche de variation quant à la formation initiale en soins infirmiers. Les données demeurent anonymes afin d'assurer aux participants et aux syndicats qu'aucun résultat ne pourra être utilisé dans l'évaluation du rendement. Les participants et les syndicats sont informés qu'afin d'assurer la confidentialité, les données seront conservées dans un lieu sûr. Ces données seront gardées pour une durée maximale de 7 ans après quoi elles seront détruites.

3.2.3 Instrument de cueillette de données de l'étude 1

Les données de l'étude 1 ont été recueillies à l'aide d'entrevues structurées. Les entrevues étaient basées sur 13 questions reliées à 5 vignettes (situations cliniques) provenant du TCS en sciences infirmières développé par Deschênes (2006). Ces questions et vignettes font l'objet de l'annexe A. Cependant, pour les besoins de l'étude 1, la question « *Quelles sont vos pensées?* » a été ajoutée à chaque vignette. En outre, notons que seules les vignettes et questions étaient utilisées et non la procédure complète de réponses aux questions du TCS développée par Deschênes et qui permet d'attribuer un score aux participants.

Lors de ces entrevues, pour chaque question nous demandions aux participantes de décrire leurs pensées à voix haute pour y répondre en utilisant la méthode du *Think aloud*. Selon Funkesson, Anbäcken, & Ek, (2007), cette méthode est reconnue pour son utilité à collecter des données permettant de mettre à jour les processus cognitifs sous-jacents au RCI. Décrite comme une technique de collecte de données verbales auprès d'une

personne qui accomplit une tâche complexe, cette méthode a permis de dégager les opérations cognitives d'infirmières dans différentes situations cliniques (Goudreau et al, 2009).

Les vignettes, et les questions s'y rapportant (voir annexe A), étaient présentées par écrit aux participantes et lues par l'interviewer. Toutes les entrevues furent enregistrées sur un support audio puis retranscrites au verbatim. La figure 1 présente la vignette 1 et les questions reliées utilisées lors de ces entrevues.

Vignette 4, question 2 (V4Q2)

Vous visitez à domicile M et Mme Lafrance. Mme Lafrance souffre d'un cancer des os en phase terminale, elle désire mourir chez elle entourée de sa famille. Vous devez lui administrer ses injections d'analgésique à des heures régulières. À votre arrivée, M Lafrance a le faciès tendu et les larmes aux yeux, il se dit dépassé par la situation et estime que sa femme est très souffrante.

Si vous pensez que :	Et qu'alors vous trouvez que :	
Si vous pensez à demander à M. Lafrance la meilleure façon de l'aider lui et sa femme pour le moment	M. Lafrance vous dit qu'il trouve que sa femme reçoit trop d'injections d'analgésiques et qu'il a peur que le médicament accélère son décès	Comment cette information influence-t-elle l'idée initiale?

Vignette 4, question 3 (V4Q3)

Vous visitez à domicile M et Mme Lafrance. Mme Lafrance souffre d'un cancer des os en phase terminale, elle désire mourir chez elle entourée de sa famille. Vous devez lui administrer ses injections d'analgésique à des heures régulières. À votre arrivée, M Lafrance a le faciès tendu et les larmes aux yeux, il se dit dépassé par la situation et estime que sa femme est très souffrante.

Si vous pensez que :	Et qu'alors vous trouvez que :	
Si vous pensez à faire verbaliser M. Lafrance sur sa propre perception de la mort	M. Lafrance vous dit que seul Dieu peut décider de l'heure de la mort d'une personne	Comment cette information influence-t-elle l'idée initiale?

Figure 1: Exemple d'une vignette du TCS en sciences infirmières comprenant deux questions

3.2.4 Stratégies d'analyse des données de l'étude 1

Le verbatim relatif à chaque vignette est appelé un « protocole ». La méthode d'analyse de protocole a été utilisée ; elle comporte trois étapes : (1) l'analyse de phrases visant à

identifier les concepts utilisés par le participant pendant qu'il raisonne; (2) l'analyse de sens visant à décrire les liens que le participant fait entre les concepts ; (3) l'analyse de scripts visant à produire une description globale des processus de raisonnement des participants. Cette analyse est basée sur le modèle de Fonteyn (1998) où 13 stratégies de pensée dominantes sont divisées en 61 catégories (voir tableau 1). Chaque catégorie de pensée est définie et devient un code. Un cahier de codification a été construit à l'aide du logiciel QDA Miner (version 4). Chaque entrevue est un cas, et chaque cas est identifié par une variable groupe qui précise si la personne est une étudiante de 1^e, 2^e, 3^e année, une infirmière nouvellement diplômée ou une experte.

Cette analyse des données est *déductive* parce que les 13 stratégies de pensée et leurs 61 catégories rapportées par Fonteyn ont servi de grille de codage. « La méthode déductive réfère à la vérification d'une théorie » (Fortin, Taggart, Kérouak, & Normand, 1988), dans ce cas, celle de Fonteyn. Les analyses portaient sur la description du raisonnement clinique des infirmières en formation, en début de carrière et expérimentées.

Les analyses ont été effectuées par une même personne et deux des chercheurs ont analysés environ 15% du matériel afin de valider les résultats. Par la suite, les données ont été étudiées groupe par groupe, qualitativement et quantitativement. Une description pour chaque groupe en est ressortie. Ensuite, une comparaison entre les groupes a été effectuée.

3.2.5 Résultats de l'étude 1

Cette étude a permis d'élaborer le MCA du RCI. Un MCA est un type de modélisation reposant sur des données scientifiques valides qui, à partir d'apprentissages critiques, circonscrit les étapes du développement d'une compétence (Tardif, 2006). Le premier apprentissage critique mis en évidence par les résultats pour la compétence de RCI consiste en l'appropriation des aspects scientifiques du rôle de l'infirmière chez les étudiantes infirmières. Il appert aussi que le passage des études au travail entraîne une mise en veilleuse du développement de cette compétence au profit de l'apprentissage de

routines propres aux environnements de travail. Cela soulève l'importance du soutien au développement continu de cette compétence infirmières tant en milieu clinique qu'en milieu académique.

3.3 Méthodologie de l'étude 2

3.3.1 Objectif de l'étude 2

L'objectif de ce projet de recherche est d'étudier le potentiel diagnostique du test de concordance de script (TCS) pour évaluer le raisonnement clinique infirmier (RCI).

3.3.2 Participants

Les participants sont les mêmes que ceux de l'étude 1, à l'exception des infirmières nouvellement diplômées. Étant donné que notre objectif est d'étudier le potentiel du TCS comme outil d'évaluation diagnostique à utiliser durant la formation initiale, nous avons choisi de ne pas exploiter leurs données.

3.3.3 Données

Les données sont constituées par les réponses aux 13 questions reliées à 5 vignettes du TCS de Deschênes (2006) fournies par les étudiants de 1^{ère}, 2^e, 3^e années et par les expertes. Notons que nous avons choisi de ne pas utiliser et recoder les réponses à la question « *Quelles sont vos pensées?* » qui a été ajoutée à chaque vignette car elles ne font pas directement partie du TCS, et sont donc jugées non pertinentes pour les visées de notre recherche.

3.3.4 Méthode d'analyse

La présente étude exploite les données des étudiantes et des infirmières expertes recueillies et déjà codées dans l'étude 1 puisque nous nous intéressons aux stratégies de pensée de Fonteyn (1998) sous-jacentes à chaque question. Un second codage, dans le fichier QDA Miner, a été ajouté au codage déjà présent, afin de repérer dans le corpus de données les informations pertinentes à l'étude 2 et relatives à chacune des questions du TCS. Ce codage superposé a été fait dans le but de pouvoir identifier le protocole relatif à

chacune des 13 questions (V1Q2, V1Q3, V1Q4, etc.). Il ne nécessitait donc pas de contre-codage puisque c'est un codage *objectif*, dans le sens qu'il n'y a pas d'ambiguïté à choisir le texte qui correspond à chaque question (protocole).

Dans un premier temps, les catégories/stratégies pertinentes de pensée à cette étude sont identifiées parmi l'ensemble de celles proposées par Fonteyn. En second lieu, les analyses visent à identifier les catégories/stratégies de pensée utilisées pour chacune des 13 questions du TCS par chaque groupe d'étudiantes et par les expertes. Cette identification aboutit à fabriquer des tableaux prenant la forme de matrices Q offrant une vision synthétique des relations et d'en faire l'analyse afin de répondre à nos questions de recherche.

Notre analyse diffère totalement de la première étude, puisqu'il s'agit de préciser les processus en lien avec chacune des questions du test, au lieu de les lier à des individus, afin d'explorer le potentiel diagnostique du TCS. Le groupe des infirmières expertes sert de référence dans le même esprit que l'approche qui permet de fabriquer les scores dans le TCS (voir la technique d'attribution des scores à la page 25).

3.4 Éthique

La 1^{ère} étude a été acceptée par les comités d'éthique des 2 milieux cliniques participants ainsi que par celui de l'Université de Montréal. Avant chaque entrevue, un formulaire de consentement devait être expliqué par l'agente de recherche (nous-mêmes) et signé par la participante. Les règles d'éthique étaient la confidentialité des informations, la participation volontaire et la possibilité de se retirer de la recherche à tout moment.

Pour la 2^e étude, les arguments suivants ont été présentés au comité d'éthique afin d'être entérinés. Étant donné que nous n'avons pas besoin des renseignements identificatoires, la confidentialité est complètement conservée. Les participants ont consenti à une étude sur le développement du raisonnement clinique et notre étude demeure complètement dans cette thématique. Nous avons collecté nous-mêmes les données et donc personne n'a eu accès aux renseignements identificatoires au moment de reprendre la base de données.

Pour toutes ces raisons, les participantes n'ont pas eu à signer de nouveau pour leur consentement. L'objectif de l'étude, se situant dans la même thématique, a été présenté au comité d'éthique de l'Université de Montréal de manière très explicite. Le certificat d'éthique a été obtenu.

4. PRÉSENTATION DES RÉSULTATS

Le but de cette étude était d'explorer le potentiel d'un TCS comme outil d'évaluation diagnostique. Pour atteindre ce but, la présentation des résultats doit permettre de répondre à ces deux questions de recherche :

- 1- Dans quelle mesure chaque question d'un TCS visant à évaluer le RCI peut-elle être reliée à des catégories et des stratégies de pensée spécifiques ?
- 2- Comment peut-on décrire le potentiel diagnostique d'un TCS pour évaluer le RCI?

4.1. Procédure de sélection et identification des catégories et des stratégies de pensée pertinentes à cette étude

Notons d'abord qu'étant donné la difficulté de faire la différence entre une réponse acceptable et une réponse qui ne l'est pas, l'ensemble des résultats portera indifféremment sur toutes les réponses des répondants.

Dans un premier temps, nous avons dû composer avec le fichier tel qu'organisé dans la première étude et dans lequel les cas étaient constitués par les participants. Nous avons ainsi groupé les codes représentant les catégories de pensée avec ceux représentant les questions (voir annexe A). Cette procédure a permis la création de nouveaux codes combinés (par exemple : choisir un plan d'action & V1Q2 (vignette 1, question 2), choisir un plan d'action & V1Q3 (vignette 1, question 3), et ainsi de suite, pour chaque catégorie de pensée et chaque question) afin de départager les questions les unes des autres et identifier les catégories de pensée qui leur étaient reliées. Au total, nous avons finalement obtenu 727 codes correspondant aux combinaisons existantes. Cette procédure (combiner les codes) a été effectuée avec la version 4 du logiciel QDA Miner.

À l'aide de QDA Miner, nous avons ensuite produit et analysé le tableau de fréquence d'utilisation de ces codes combinés incluant le nombre de cas et le pourcentage de cas concernés. Étant donné que notre objectif de recherche était d'étudier le potentiel

diagnostique du TCS et d'identifier les catégories de pensée reliées à chaque question, nous avons fait le choix de nous intéresser à celles qui variaient entre les questions et d'éliminer celles qui étaient sous-utilisées ou sur-utilisées. Notons que nous avons au départ travaillé avec l'unité la plus fine, à savoir les catégories de pensées.

Ainsi, nous avons choisi d'éliminer les catégories de pensée peu utilisées, c'est-à-dire pour lesquelles tous les codes combinés correspondaient à moins de 25% des cas, et celles très fréquemment utilisées pour lesquelles tous les codes combinés correspondaient à plus de 50% des cas). Les deux exemples suivants illustrent notre processus de sélection des catégories de pensée. Pour la catégorie *Approfondir ou préciser une évaluation*, les pourcentages de cas variaient entre 24% et 69% et cette catégorie a donc été retenue (voir tableau 3). Pour la catégorie de pensée *Clarifier politique, procédure*, les pourcentages de cas variaient de 0 à 15% et cette catégorie a donc été éliminée (voir tableau 4).

Tableau 3 : Pourcentages de cas ayant utilisé la catégorie de pensée - *Approfondir ou préciser une évaluation*

Questions	Pourcentages de cas
V1Q2	38%
V1Q3	33%
V1Q4	38%
V2Q2	69%
V2Q3	58%
V2Q4	58%
V3Q2	31%
V3Q3	24%
V4Q2	40%
V4Q3	31%
V5Q3	36%
V5Q4	56%
V5Q5	44%

Tableau 4 : *Pourcentages de cas ayant utilisé la catégorie de pensée - Clarifier politique, procédure*

Questions	Pourcentages de cas
V1Q2	2%
V1Q3	4%
V1Q4	
V2Q2	
V2Q3	
V2Q4	
V3Q2	
V3Q3	
V4Q2	
V4Q3	
V5Q3	2%
V5Q4	15%
V5Q5	4%

En procédant ainsi, nous avons retenu 21 catégories de pensée (voir tableau 5) et en avons rejeté 40 en raison d'une faible ou trop forte utilisation (voir tableau 6). De plus, 2 des 13 stratégies de pensée de Fonteyn (voir tableau 1) et tableau 6 en caractères gras) ont disparu, soit *Prédire* et *Reconnaître un pattern*.

Tableau 5: Les 21 catégories sélectionnées et les stratégies de pensées correspondantes

Stratégies de pensée	Catégories de pensée
1. Autres stratégies de pensée	– faire une pause réflexion
2. Chercher des informations manquantes	– sur les ressources – sur l’histoire du patient – sur la médication – pour approfondir ou préciser une évaluation
3. Conclure	– sur une situation – sur la condition du patient ou statut
4. Énoncer une proposition	– choisir un plan d’action
5. Énoncer une règle de pratique	– politique
6. Faire des hypothèses	– spéculer sur le statut du patient – identifier besoins du patient – déterminer la cause
7. Faire des choix	– d’actions
8. Fournir des explications	– sur des actions – interventions infirmières
9. Juger de la valeur	– des traitements (ou interventions) – des conclusions d’évaluation
10. Ordonner des concepts par priorités	– focus sur plan d’action
11. Relier des informations	– sur les conclusions d’évaluation/autres informations – histoire du patient et problèmes/ situation du patient – information sur le traitement et autre type d’information

Tableau 6 : Les 40 catégories de pensée éliminées et les stratégies de pensées correspondantes

Stratégies de pensée	Catégories de pensée
Autres stratégies de pensée	1. généraliser, inférer
	2. poser une question
	3. qualifier, nuancer
	4. supposer
Chercher des informations manquantes	5. pour déterminer le plan de soins
	6. information logistique
	7. procédures, politiques, droits
	8. sur les résultats des tests
	9. sur les signes vitaux
Énoncer une proposition	10. clarifier une politique, procédure
	11. établir ou écarter un problème
	12. évaluer les données
	13. juger intervention, traitement
	14. déterminer cause
Énoncer une règle de pratique	15. précepte, règle informelle
	16. procédures
Faire des hypothèses	17. identifier maladie/problème de santé
	18. intervention ou traitement
Faires des choix	19. de tests
	20. de traitement (pharmacologique ou procédural)
	21. d'interventions infirmières*
Fournir des explications	22. sur les prédictions
	23. sur les préoccupations (inquiétudes)
	24. sur les soins thérapeutiques
	25. test
Juger de la valeur	26. des résultats de tests
Ordonner des concepts par priorités	27. focus sur préoccupations du patient
Prédire**	28. découvertes, résultats d'examens
	29. des événements
	30. des interventions
	31. des réponses
	32. résultats des soins (ou interventions)
	33. une situation
	34. les besoins du patient
Reconnaître un pattern**	35. cas type
	36. protocole et traitement standard
	37. représenter une situation familière
	38. représenter une situation non-familière
Relier des informations	39. données organisationnelles
	40. sur les résultats des tests et contexte

Notes : * correspond aux catégories éliminées car fortement utilisées.

**Les stratégies de pensée inscrites en caractère gras ont été complètement éliminées.

Par la suite, nous avons construit un graphique pour chacune de ces 21 catégories de pensée à l'aide de QDA Miner. La figure 2 illustre la catégorie *Faire des choix d'action* et présente le pourcentage de cas relativement à chaque groupe (1^{ère}, 2^e, 3^e année, expertes) qui a utilisé cette catégorie de pensée pour chacune des questions identifiée par une couleur différente.

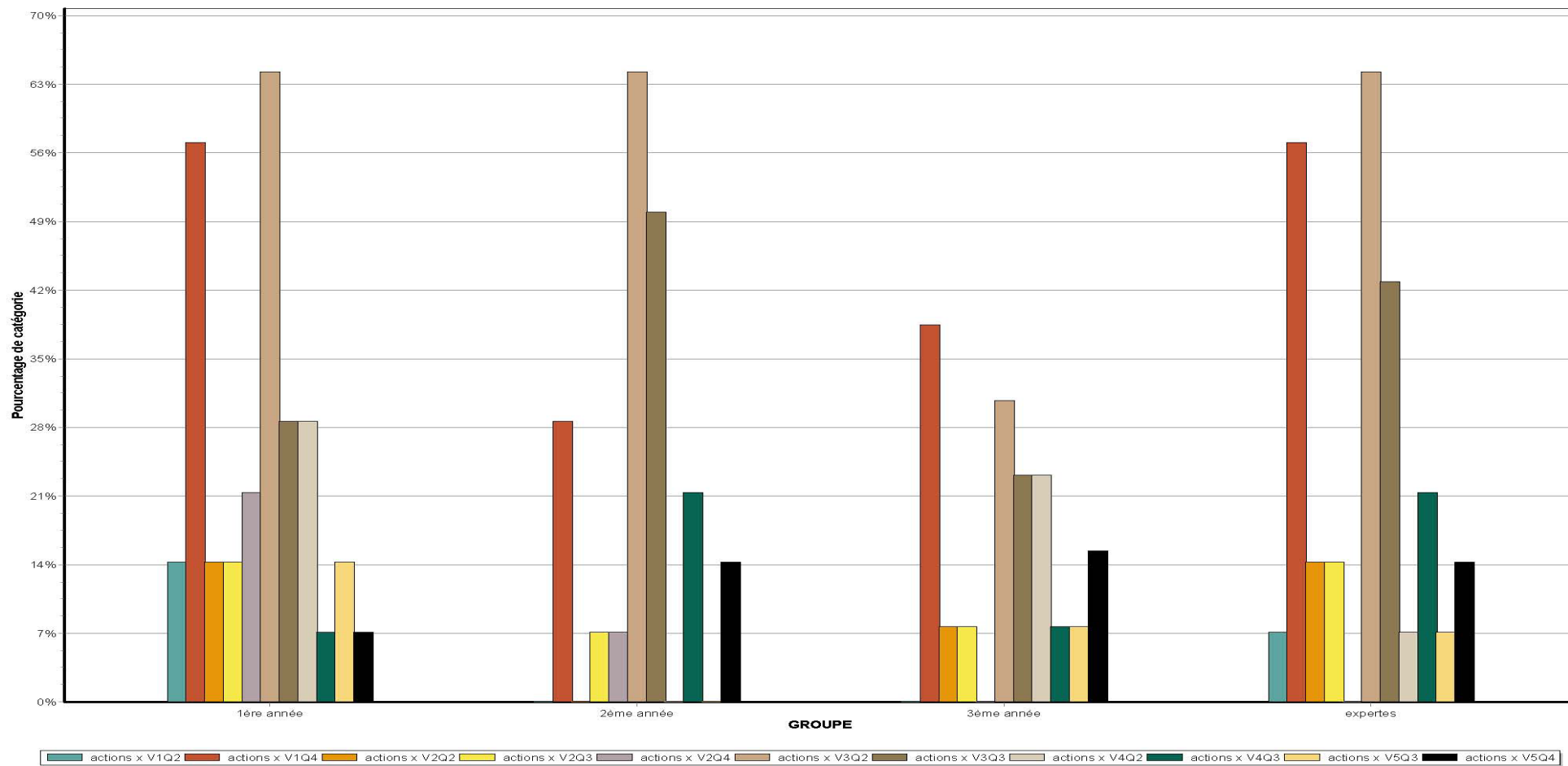


Figure 2 : Le pourcentage de cas ayant utilisé la Catégorie - Faire des choix d'actions pour chaque question

En tout 21 graphiques ont été produits et chacun d'eux a permis de constater les différences entre les pourcentages de cas de chaque question. Le tableau 7 regroupe le plus faible et le plus élevé pourcentage d'utilisation pour chaque catégorie de pensée. Il met en évidence les grandes variations d'utilisation de chaque catégorie de pensée avec le pourcentage le plus élevé allant de 36% à 93%, et le moins élevé variant entre 0 et 7%. Étant donné que nous souhaitons identifier les catégories de pensée utilisées pour chaque question, nous avons décidé d'établir un lien entre une question et une catégorie lorsqu'elle est utilisée par plus de 50% de l'utilisation maximum; la valeur seuil fait l'objet de la dernière colonne du tableau 7. Par exemple, pour la catégorie *faire des choix d'action* le pourcentage le plus élevé est 64%, nous avons donc seulement sélectionné les questions pour lesquelles le pourcentage est supérieur à 32%. Ainsi, dans la figure 2, les questions en rouge (V1Q4), en beige (V3Q2) et en brun (V3Q3) ont été retenues.

Tableau 7 : Pourcentages de cas le plus faible et le plus élevé pour chaque catégorie de pensée

Catégories de pensée	% le plus faible	% le plus élevé	Valeur seuil choisie
1- Faire des choix d'actions	0	64	32
2- Chercher des informations manquantes pour approfondir ou préciser une évaluation	7	92	46
3- Chercher des informations manquantes sur l'histoire du patient	0	46	23
4- Chercher des informations manquantes sur la médication	0	50	25
5- Chercher des informations manquantes sur les ressources	0	43	21,5
6- Juger de la valeur des conclusions d'évaluation	0	46	23
7- Juger de la valeur des traitements (ou interventions)	0	71	35,5
8- Relier des informations histoire du patient et problèmes/situation du patient	0	71	35,5
9- Relier des informations : information sur le traitement et autres type d'information	0	57	28,5
10- Relier des informations conclusions d'évaluations/ autres informations	7	71	35,5
11- Énoncer une règle de pratique : politique	0	71	35,5
12- Faire des hypothèses : identifier besoins du patient	0	50	25
13- Faire des hypothèses : spéculer sur le statut du patient	0	64	32
14- Faire des hypothèses : déterminer la cause	0	93	46,5
15- Énoncer une proposition : choisir un plan d'action	0	71	35,5
16- Fournir des explications sur des actions	0	43	21,5
17- Fournir des explications : Interventions infirmières	7	54	27
18- Ordonner des concepts par priorités : focus sur plan d'action	0	43	21,5
19- Conclure sur la condition du patient ou statut	0	43	21,5
20- Conclure sur une situation	0	36	18
21- Autres stratégies de pensée : faire une pause réflexion	7	86	43

4.2. Mise en évidence des liens entre les 21 catégories/stratégies et les 13 questions pour l'ensemble des participantes

Nous avons procédé ainsi pour les 21 catégories de pensée et produit le tableau 8 qui contient toutes les catégories sélectionnées et toutes les questions, ainsi que le tableau 9, qui contient toutes les stratégies de pensée et toutes les questions. Ces deux tableaux prennent la forme de matrices Q et permettent de visualiser les catégories de pensée (tableau 8) et les stratégies (tableau 9) spécifiques à chaque question pour chaque groupe. Ainsi, dans la case correspondant à une question et une catégorie (ou une stratégie), nous avons inscrit quel groupe (1^{ère} année=1, 2^e année=2, 3^e année=3, expertes=E) a utilisé chaque question selon notre mode de sélection. Ensuite, nous avons surligné en jaune les catégories/questions utilisées (n=124) et avons calculé le pourcentage d'utilisation de chaque question et de chaque catégorie (tableau 8) puis de chaque question et de chaque stratégie (tableau 9).

Par exemple, nous pouvons voir dans le tableau 8 que la catégorie de pensée *pause réflexion* est employée dans 9 des 13 questions soit 69%, ou encore que la question V1Q4 nécessite 13 des 21 catégories de pensée, soit 62%. Nous constatons que les questions vont chercher les catégories de pensée sélectionnées d'une manière assez partagée, dans une proportion se situant entre 29% à 67%. De plus, le tableau 8 permet de constater que la catégorie *Relier des informations sur les conclusions d'évaluations/autres informations* est utilisée dans presque toutes les questions (92%) et que la catégorie *Conclure sur une situation* est aussi fortement utilisée à 85%, de même que *Fournir des explications sur les interventions infirmières* (77%). Par contre, la catégorie *Énoncer une règle de pratique sur la politique* n'est utilisée que pour une question (8%). Du côté des questions, il apparaît que la V3Q2 incite plusieurs catégories de pensée (67%), de même que la V1Q4 (62%). Cependant, la V5Q3 n'implique que 29% des catégories de pensée. Enfin, nous voyons que pour certaines stratégies de pensée plusieurs catégories sont utilisées, telle *Chercher des informations manquantes* où l'on retrouve 4 catégories. Cependant, pour d'autres stratégies, comme *Faire des choix*, une seule catégorie est présente (la catégorie *Actions*).

Au tableau 9, nous avons une vision plus large, moins spécifique, sur les stratégies de pensée plutôt que sur les catégories. Ce tableau permet de constater que quatre stratégies de pensée sont utilisées pour répondre à toutes les questions. Ces stratégies sont *Chercher des informations manquantes*, *Conclure*, *Faire des hypothèses* et *Relier des informations*. Par contre, la stratégie *Énoncer règle de pratique* n'est employée que pour répondre à la question V5Q4. D'un autre côté, nous voyons que les questions requièrent plusieurs stratégies de pensée à la fois, soit entre 45% et 82%.

Tableau 8 : Catégories de pensée, questions et groupes

Questions/Catégories		V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des questions
Stratégies	Catégories														
Faire des choix	Actions			1 3 E				1 2 E	2 E						23%
Autres stratégies de pensée	Pause réflexion		1 2	1	1	2 3	2 3	1 2		1	1 2		1		69%
Chercher des infos manquantes	Histoire du pt		3	1 3		1		1 2 3 E				3 E	3		46%
	Approfondir ou préciser évaluation				1 2 3	1 2 3	1 3			1			1 2 3	1 2	46%
	Médication	1 3 E					1 2 3		1 2 3	1 E					31%
	Ressources			1				1 2			E				23%
Conclure	Condition du pt ou statut			1 2 3 E	3 E	3		3 E	3	E	1	1 3	1		69%
	Une situation	3	3 E	2 3		2 3	2	2	3	E	1 2 3 E		2 3 E	2	85%
Énoncer règle de pratique	Politique												2 E		8%
Énoncer une proposition	Choisir plan d'action	1 2				3		3	1				1	1 2 E	46%
Faire des hypothèses	Déterminer la cause	1 2 3 E	2 3 E		1 2 3 E	1 2 3	1 2 E							1 2 E	46%
	Spéculer sur statut pt		E	2 E	2	1 2 E		2 3 E			2 E	1 2	3		62%
	Identifier besoins pt	3						2 E	2	E			1		38%
Fournir des	Interventions inf.	1 3	3	3	3	3	3	1 3	3		1			E	77%

explications	Sur des actions			1 E				1 2 3 E	1 2 3 E	3					23%
Juger de la valeur	Des conclusions d'évaluation			1 3	3						2				23%
	Traitements (ou interventions)	E							3	1 2 E					23%
Ordonner des concepts par priorité	Focus plan d'action	3		3		3		1 3 E	1 3			2 3	3	1 2	62%
Relier des infos	Conclusions éval./autres infos	1 2 3	1 3	2	3	2 3	2	2		2	1 2	3	2 3	1	92%
	Histoire pt et probl./ situation du pt		1 2	2 3				2 3 E				2 E			31%
	Info traitement et autre type d'info						1 2 3		1	E					23%
% des catégories		43%	38%	62%	38%	52%	38%	67%	52%	48%	38%	29%	52%	33%	

Tableau 9 : Stratégies de pensée, questions et groupes

Questions/Stratégies	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des questions
1. Faire des choix			1 3 E				1 2 E	2 E						23%
2. Autres stratégies de pensée		1 2	1	1	2 3	2 3	1 2		1	1 2		1		69%
3. Chercher des infos manquantes	1 3 E	3	1 3	1 2 3	1 2 3	1 2 3	1 2 3 E	1 2 3	1 E	E	3 E	1 2 3	1 2	100%
4. Conclure	3	3 E	1 2 3 E	3 E	2 3	2	2 3 E	3	E	1 2 3 E	1 3	1 2 3 E	2	100%
5. Énoncer règle de pratique												2 E		8%
6. Énoncer une proposition	1 2				3		3	1				1	1 2 E	46%
7. Faire des hypothèses	1 2 3 E	2 3 E	2 E	1 2 3 E	1 2 3 E	1 2 E	2 3 E	2	E	2 E	1 2	1 3	1 2 E	100%
8. Fournir des explications	1 3	3	1 3 E	3	3	3	1 2 3 E	1 2 3 E	3	1			E	85%
9. Juger de la valeur	E		1 3	3				3	1 2 E	2				46%
10. Ordonner des concepts par priorité	3		3		3		1 3 E	1 3			2 3	3	1 2	62%
11. Relier des infos	1 2 3	1 2 3	2 3	3	2 3	1 2 3	2 3 E	1	2 E	1 2	2 3 E	2 3	1	100%
% des stratégies	73%	55%	82%	64%	73%	55%	82%	82%	64%	55%	45%	73%	64%	

4.3 Relations entre les 21 catégories/stratégies et les 13 questions pour chaque groupe de participantes

Un regard sur les tableaux 8 et 9 permet de voir la variabilité qui existe entre les groupes d'étudiantes et d'expertes. Nous avons donc produit les tableaux 10 à 17 dans le but de mettre en évidence les catégories et les stratégies de pensée utilisées pour chaque question, par chaque groupe, et ensuite d'étudier les différences et les similitudes qui existent entre les groupes. Pour construire ces tableaux, nous avons repris les catégories/questions, utilisées par chaque groupe séparément, et nous avons calculé le pourcentage d'utilisation de chaque catégorie et de chaque question. Par la suite, nous avons fait de même pour chaque stratégie et chaque question.

Les tableaux 10 et 11 mettent en évidence le nombre de catégories de pensée et de stratégies utilisées par le groupe d'étudiantes de 1^{ère} année. Par exemple, au tableau 10, nous voyons que la catégorie de pensée *Pause réflexion* est la plus fréquente étant requise par 7 des 13 questions (soit 54%). Quant aux questions qui vont chercher le plus de catégories de pensée chez ce groupe, ce sont la V1Q4 et la V3Q2 qui nécessitent 7 des 21 catégories de pensée, donc 33%.

Au tableau 11, nous voyons que la stratégie *Chercher des informations manquantes* est la plus utilisée par les étudiantes de 1^{ère} année, soit à 77%. Aussi, c'est la question V1Q4 qui stimule le plus de stratégies de pensée, à 55%. Par ailleurs, 5 questions requièrent 45% des stratégies de pensée. La question V5Q3 est la question qui nécessite le moins de stratégies de pensée chez ce groupe, soit 10%.

Au tableau 12, nous voyons que la catégorie de pensée *Relier des informations sur les conclusions d'évaluation et autres informations* est utilisée le plus fréquemment à 62% par le groupe de 2^e année tandis que *Fournir des explications sur les interventions infirmières* n'est pas du tout utilisée par ce groupe. En ce qui a trait aux questions, la V3Q2 est la plus productive pour ce groupe (à 48%) tandis que les autres questions ont un faible pourcentage, variant de 10 à 29%.

Au tableau 13, nous remarquons que les étudiantes de 2^e année font plus d'hypothèses que les autres groupes, à 85%. Elles utilisent toutes les stratégies de pensée pour répondre aux questions. Cependant, *Énoncer une règle de pratique* n'est utilisée qu'à 8%. La question V3Q2 va chercher 64% des stratégies de pensée, tandis que 4 autres questions stimulent 45% des stratégies de pensée.

Au tableau 14, nous voyons que la catégorie de pensée la plus utilisée par les étudiantes de 3^e année est *Fournir des explications sur les interventions infirmières* (62%). Par ailleurs, deux de ces catégories de pensée n'ont pas été utilisées pour ces questions, soit *Énoncer une règle de pratique sur la politique* et *Chercher des informations manquantes sur les ressources*. La question faisant appel au plus grand nombre de catégories de pensée est la V2Q3 à 43%. Il est à noter que la question V5Q5 n'est allée chercher aucune des catégories de pensée sélectionnées chez ce groupe, donc elle fait appel aux catégories de pensée que nous avons éliminées pour être trop ou peu utilisées (voir tableau 6).

Au tableau 15, la stratégie de pensée *Énoncer une règle de pratique* n'est pas utilisée par les étudiantes de 3^{ème} année, alors que les stratégies *Chercher des informations manquantes* et *Conclure* sont utilisées à 77%. La question V2Q3 va chercher 73% des stratégies de pensée alors que pour les questions V4Q2 et V4Q3, ce n'est que 9%. Les pourcentages des stratégies varient entre 0 et 77% et les pourcentages des questions entre 0 et 73%. Le tableau 16 illustre que 4 des catégories de pensée ne sont pas utilisées par les expertes soit : 1) *Chercher des informations manquantes pour approfondir ou préciser une évaluation*, 2) *Relier des informations sur les conclusions d'évaluation et autres informations*, 3) *Juger de la valeur des conclusions d'évaluation* et 4) *Pause Réflexion*. Cependant, il est intéressant de remarquer que les expertes utilisent peu de catégories de pensée mais font plus d'hypothèses, *afin de déterminer la cause d'un ensemble de faits* (38%) ou pour *spéculer sur le statut du patient* (38%). Par ailleurs, même si les questions sont toutes utilisées, la question suscitant le plus de catégories de pensée de la part de ce groupe est la V4Q2 (29%). Nous voyons dans ce tableau que les

variations sont plus faibles pour le groupe d'expertes, que ce soit pour les catégories (0 à 38%) ou les questions (0 à 29%).

Au tableau 17, les expertes utilisent la stratégie *Conclure* à 77% alors qu'elles n'utilisent pas la stratégie de pensée *Autres stratégies de pensée*. La question V3Q2 requiert 64% des stratégies de pensée et les questions V2Q2 et V2Q3, seulement 9%. Les pourcentages pour les stratégies et pour les questions varient de façon importante.

Tableau 10 : Catégories de pensée, questions, groupe 1^{ère} année

Questions/Catégories		V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des catégories
Stratégies	Sous-catégories														
Faire des choix	Actions			1				1							15%
Autres stratégies de pensée	Pause réflexion		1	1	1			1		1	1		1		54%
Chercher des infos manquantes	Histoire du pt			1		1		1							23%
	Approfondir ou préciser évaluation				1	1	1			1			1	1	46%
	Médication	1					1		1	1					31%
	Ressources			1				1							15%
Conclure	Condition du pt ou statut			1							1	1	1		31%
	Une situation										1				8%
Énoncer règle de pratique	Politique														0
Énoncer une proposition	Choisir plan d'action	1							1				1	1	31%
Faire des hypothèses	Déterminer la cause	1			1	1	1							1	38%
	Identifier besoins pt												1		8%
	Spéculer sur statut pt					1						1			15%

Fournir des explications	Interventions inf.	1						1			1				23%
	Sur des actions			1				1	1						23%
Juger de la valeur	Des conclusions d'évaluation			1											8%
	Traitements (ou interventions)									1					8%
Ordonner des concepts par priorités	Focus plan d'action							1	1					1	23%
Relier des infos	Conclusions éval./autres infos	1	1								1			1	31%
	Histoire pt et probl./ situation du pt		1												8%
	Info traite. + autre type info						1		1						15%
	% total	24%	19%	33%	14%	19%	19%	33%	24%	19%	24%	10%	24%	24%	

Tableau 11 : Stratégies de pensée, questions, groupe 1^{ère} année

Questions/Stratégies	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des stratégies
1. Faire des choix			1											15%
2. Autres stratégies de pensée		1	1	1			1		1	1		1		54%
3. Chercher des infos manquantes	1		1	1	1	1	1	1	1			1	1	77%
4. Conclure			1							1	1	1		31%
5. Énoncer règle de pratique														0
6. Énoncer une proposition	1							1				1	1	31%
7. Faire des hypothèses	1			1	1	1					1	1	1	54%
8. Fournir des explications	1		1				1	1		1				38%
9. Juger de la valeur			1						1					15%
10. Ordonner des concepts par priorité							1	1					1	23%
11. Relier des infos	1	1					1	1		1			1	46%
% total	45%	18%	55%	27%	18%	18%	45%	45%	27%	36%	18%	45%	45%	

Tableau 12 : Catégories de pensée, questions, groupe 2^e année

Questions/Catégories		V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des catégories
Faire des choix	Actions							2	2						15%
Autres stratégies de pensée	Pause réflexion		2			2	2	2			2				38%
Chercher des infos manquantes	Histoire du pt							2							8%
	Approfondir ou préciser évaluation				2	2							2	2	31%
	Médication						2		2						15%
	Ressources							2							8%
Conclure	Condition du pt ou statut			2											8%
	Une situation			2		2	2	2			2		2	2	54%
Énoncer règle de pratique	Politique												2		8%
Énoncer une proposition	Choisir plan d'action	2												2	15%
Faire des hypothèses	Déterminer la cause	2	2		2	2	2							2	46%
	Identifier besoins pt							2	2						15%
	Spéculer sur statut pt			2	2	2		2			2	2			46%
Fournir des	Interventions inf.														0

explications	Sur des actions							2	2						15%
Juger de la valeur	Des conclusions d'évaluation										2				8%
	Traitements (ou interventions)									2					8%
Ordonner des concepts par priorités	Focus plan d'action											2		2	15%
Relier des infos	Conclusions éval./autres infos	2		2		2	2	2		2	2		2		62%
	Histoire pt et probl./ situation du pt		2	2				2				2			31%
	Info traite. + autre type info						2								8%
	% total	14%	14%	24%	14%	29%	29%	48%	19%	10%	24%	14%	19%	24%	

Tableau 13 : Stratégies de pensée, questions, groupe 2^e année

Questions/Stratégies	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des Stratégies
1. Faire des choix							2	2						15%
2. Autres stratégies de pensée		2			2	2	2			2				38%
3. Chercher des infos manquantes				2	2	2	2	2				2	2	54%
4. Conclure			2		2	2	2			2		2	2	54%
5. Énoncer règle de pratique												2		8%
6. Énoncer une proposition	2												2	15%
7. Faire des hypothèses	2	2	2	2	2	2	2	2		2	2		2	85%
8. Fournir des explications							2	2						15%
9. Juger de la valeur									2	2				15%
10. Ordonner des concepts par priorités											2		2	15%
11. Relier des infos	2	2	2		2	2	2		2	2	2	2		77%
% total	27%	27%	27%	18%	45%	45%	64%	36%	18%	45%	27%	36%	45%	

Tableau 14 : Catégories de pensée, questions, groupe 3^e année

Questions/Catégories		V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des catégories
Faire des choix	Actions			3											8%
Autres stratégies de pensée	Pause réflexion					3	3								15%
Chercher des infos manquantes	Histoire du pt		3	3				3				3	3		38%
	Approfondir ou préciser évaluation				3	3	3						3		31%
	Médication	3					3		3						23%
	Ressources														0
Conclure	Condition du pt ou statut			3	3	3		3	3			3			46%
	Une situation	3	3	3		3			3		3		3		54%
Énoncer règle de pratique	Politique														0
Énoncer une proposition	Choisir plan d'action					3		3							15%
Faire des hypothèses	Déterminer la cause	3	3		3	3									31%
	Identifier besoins pt	3													8%
	Spéculer sur statut pt							3					3		15%
Fournir des	Interventions inf.	3	3	3	3	3	3	3	3						62%

explications	Sur des actions							3	3	3					23%
Juger de la valeur	Des conclusions d'évaluation			3	3										15%
	Traitements (ou interventions)								3						8%
Ordonner des concepts par priorités	Focus plan d'action	3		3		3		3	3			3	3		54%
Relier des infos	Conclusions éval./autres infos	3	3		3	3						3	3		46%
	Histoire pt et probl./ situation du pt			3				3							15%
	Info traite. + autre type info						3								8%
	% total	33%	24%	38%	29%	43%	24%	38%	33%	5%	5%	19%	29%	0	

Tableau 15 : Stratégies de pensée, questions, groupe 3^e année

Questions/Stratégies	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des stratégies
1. Faire des choix			3											8%
2. Autres stratégies de pensée					3	3								15%
3. Chercher des infos manquantes	3	3	3	3	3	3	3	3			3	3		77%
4. Conclure	3	3	3	3	3		3	3		3	3	3		77%
5. Énoncer règle de pratique														0
6. Énoncer une proposition					3		3							15%
7. Faire des hypothèses	3	3		3	3		3					3		46%
8. Fournir des explications	3	3	3	3	3	3	3	3	3					69%
9. Juger de la valeur			3	3				3						23%
10. Ordonner des concepts par priorité	3		3		3		3	3			3	3		54%
11. Relier des infos	3	3	3	3	3	3	3				3	3		69%
% total	54%	45%	64%	55%	73%	36%	64%	45%	9%	9%	36%	45%	0	

Tableau 16 : Catégories de pensée, questions, groupe expertes

Questions/Stratégies		V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des catégories
Faire des choix	Actions			E				E	E						23%
Autres stratégies de pensée	Pause réflexion														0
Chercher des infos manquantes	Histoire du pt							E				E			15%
	Approfondir ou préciser évaluation														0
	Médication	E								E					8%
	Ressources										E				8%
Conclure	Condition du pt ou statut			E	E			E		E					31%
	Une situation		E							E	E		E		31%
Énoncer règle de pratique	Politique												E		8%
Énoncer une proposition	Choisir plan d'action													E	8%
Faire des hypothèses	Déterminer la cause	E	E		E		E							E	38%
	Identifier besoins pt							E		E					15%
	Spéculer sur statut pt		E	E		E		E			E				38%
Fournir des	Interventions inf.													E	8%

explications	Sur des actions			E				E	E						23%
Juger de la valeur	Des conclusions d'évaluation														0
	Traitements (ou interventions)	E								E					15%
Ordonner des concepts par priorités	Focus plan d'action							E							8%
Relier des infos	Conclusions éval./autres infos														0
	Histoire pt et probl./ situation du pt							E				E			15%
	Info traite. + autre type info									E					8%
	% total	14%	14%	19%	10%	5%	5%	39%	10%	29%	14%	10%	10%	14%	

Tableau 17 : Stratégies de pensée, questions, groupe expertes

Questions/Stratégies	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5	% des stratégies
1. Faire des choix			E				E	E						23%
2. Autres stratégies de pensée														0
3. Chercher des infos manquantes	E						E		E	E	E			38%
4. Conclure		E	E	E			E		E	E		E		54%
5. Énoncer règle de pratique												E		8%
6. Énoncer une proposition													E	8%
7. Faire des hypothèses	E	E	E	E	E	E	E		E	E			E	77%
8. Fournir des explications			E				E	E					E	31%
9. Juger de la valeur	E								E					15%
10. Ordonner des concepts par priorités							E							8%
11. Relier des infos							E		E		E			23%
% total	27%	18%	36%	18%	9%	9%	64%	18%	45%	27%	18%	18%	27%	

En résumé, ces tableaux mettent en évidence les catégories et les stratégies de pensée utilisées pour chaque question, par chaque groupe et montrent d'importantes différences entre les groupes, ce qui nous a amené à quantifier l'accord entre eux, présenté dans la section suivante.

4.4 Accord entre les groupes

Nous avons calculé les pourcentages d'accord entre les groupes pour les catégories de pensée (tableau 18) et pour les stratégies de pensée (tableau 19). Le but de cette opération était de quantifier les différences entre les groupes. En nous basant sur le tableau 8, nous avons compté le nombre de fois où les groupes se retrouvaient ensemble dans la même case, et nous avons divisé par le nombre total de catégories/questions utilisées ($n=124$). Par exemple, il y a 29 cases sur 124 pour lesquelles on retrouve à la fois les groupes 1 et 2, et 7 cases sur 124 où tous les groupes sont présents. Ces calculs ont permis de créer le tableau 18 qui confirme des disparités importantes entre les groupes quant aux catégories utilisées. Par exemple, lorsqu'on considère deux groupes, le plus haut pourcentage d'accord se retrouve entre le groupe 1 et le groupe 2, ne s'élevant cependant qu'à 23%. Ce pourcentage passe à 13% entre le groupe 1 et les expertes. Il est à noter que les pourcentages sont plus bas lorsque l'on inclut les expertes dans la comparaison.

En nous basant sur le tableau 9, nous avons ensuite procédé de la même manière, mais en divisant par le nombre de stratégies/questions utilisées ($n=96$) pour produire le tableau 19. De manière assez logique, nous y voyons que le pourcentage d'accord entre les groupes est toujours plus élevé lorsque l'on considère les stratégies plutôt que les catégories de pensées, soit par exemple à 31% au lieu de 23 % pour les groupes 1 et 2. Étant donné le faible pourcentage d'accord entre les 4 groupes, soit 6% au tableau 18 et 9% au tableau 19, nous pouvons conclure que, de manière générale, les quatre groupes n'utilisent pas les mêmes catégories et stratégies de pensée.

Tableau 18 : Degrés d'accord entre les groupes pour les catégories

Groupes	Degrés d'accord
1 et 2	$29/124=0,23$
3 et 1	$23/124=0,19$
3 et 2	$27/124=0,22$
E et 1	$16/124=0,13$
E et 2	$22/124=0,18$
E et 3	$18/124=0,15$
1, 2, 3	$15/124=0,12$
1, 2, E	$10/124=0,08$
1, 3, E	$10/124=0,08$
2, 3, E	$11/124=0,09$
1, 2, 3, E	$7/124=0,06$

Tableau 19 : Degrés d'accord entre les groupes pour les stratégies

Groupes	Degrés d'accord
1 et 2	$30/96=0,31$
3 et 1	$27/96=0,28$
3 et 2	$29/96=0,30$
E et 1	$19/96=0,20$
E et 2	$24/96=0,25$
E et 3	$21/96=0,22$
1, 2, 3	$17/96=0,18$
1, 2, E	$14/96=0,15$
1, 3, E	$13/96=0,14$
2, 3, E	$13/96=0,14$
1, 2, 3, E	$9/96=0,09$

Suite à ces résultats, nous avons cherché si certaines questions ont des comportements semblables, autrement dit, si elles font appel à des combinaisons de catégories ou de stratégies comparables, ce qui permettrait éventuellement de créer des regroupements de questions faisant appel à des processus de réponses semblables.

4.5 Accord entre les questions 2 à 2 sur l'ensemble des participants

Finalement, les tableaux 20 et 21 ont été conçus dans le but d'identifier les ressemblances éventuelles des combinaisons de catégories de pensée et de stratégies de pensée, entre deux questions, en calculant les degrés d'accord entre les questions considérées 2 par 2. Pour produire le tableau 20, nous avons compté dans le tableau 8, le nombre de fois où les deux questions nécessitaient, ou non, les mêmes catégories de pensée, et divisé le résultat par 21 correspondant au nombre total de catégories de pensées possibles. Nous avons procédé de la même manière pour concevoir le tableau 21, mais à partir du tableau 9, et en divisant par 11, qui est le nombre total de stratégie de pensée. Nous avons proposé ces résultats sous forme de rapports et ajouté, entre parenthèses, le nombre de fois où les deux questions nécessitaient l'utilisation des mêmes catégories de pensée (tableau 20) ou des mêmes stratégies de pensée (tableau 21). Par exemple, dans le tableau 20, les questions V1Q4 et V3Q2 fonctionnent de manière identique pour 18 catégories de pensée sur 21, 6 sont des absences communes et 12 sont des présences communes. Nous avons fait ces distinctions afin d'assurer une interprétation correcte des valeurs d'accord. En tout, nous avons considéré les 78 comparaisons possibles.

Dans le tableau 20, les valeurs des accords varient de 6/21 à 18/21 suggérant qu'il y a toujours des ressemblances entre les questions prises 2 à 2. Nous pouvons en outre remarquer que certaines questions ont un pourcentage d'accord élevé, telles que les questions V1Q4 et V3Q2, à 86% (18/21). De plus, nous constatons que la moitié des questions ont un pourcentage d'accord de plus de 60% (13/21) quand on comptabilise autant les absences communes que les présences communes.

Lorsqu'on s'intéresse à l'accord calculé sur les catégories de pensées réellement partagées (nombres en parenthèses), les valeurs varient entre 2 et 12 sur 21. Par exemple V1Q4-V3Q2 ont 57% de présences de catégories communes (12/21). Ainsi, les nombres entre parenthèses montrent que les accords qui semblent être élevés, sont davantage constitués des absences communes de catégories/stratégies de pensée que des présences communes. Finalement, il est à noter que les questions partagent toujours au moins deux catégories de pensée et au plus 9.

Au niveau des stratégies de pensée (tableau 21), seulement 7 combinaisons de questions ont un pourcentage d'accord de 55% (6/11), alors que toutes les autres combinaisons ont un accord de 64% (7/11) et plus. Aussi, 4 combinaisons de questions ont un accord de 100% (11/11), soit V1Q3-V2Q4, V2Q2-V4Q2, V2Q2-V4Q3 et V4Q2-V4Q3.

Tableau 20 : Accord entre les questions pour les catégories

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	12/21 (4)												
V1q4	7/21 (4)	14/21 (7)											
V2q2	10/21 (3)	15/21 (5)	12/21 (6)										
V2q3	13/21 (6)	16/21 (7)	12/21 (8)	16/21 (7)									
V2q4	14/21 (5)	15/21 (5)	8/21 (4)	15/21 (5)	14/21 (6)								
V3q2	10/21 (6)	13/21 (7)	18/21 (12)	9/21 (5)	14/21 (9)	6/21 (4)							
V3q3	15/21 (7)	6/21 (2)	9/21 (6)	6/21 (2)	9/21 (5)	10/21 (4)	12/21 (8)						
V4q2	12/21 (5)	8/21 (3)	7/21 (5)	11/21 (4)	10/21 (5)	14/21 (6)	9/21 (6)	14/21 (7)					
V4q3	10/21 (3)	15/21 (5)	16/21 (9)	17/21 (6)	14/21 (6)	13/21 (4)	13/21 (7)	8/21 (3)	11/21 (4)				
V5q3	10/21 (2)	15/21 (4)	14/21 (6)	13/21 (3)	14/21 (5)	9/21 (1)	13/21 (6)	8/21 (2)	9/21 (2)	13/21 (3)			
V5q4	11/21 (4)	11/21 (5)	11/21 (7)	11/21 (5)	17/21 (9)	9/21 (4)	14/21 (9)	9/21 (5)	12/21 (6)	12/21 (5)	14/21 (5)		
V5q5	17/21 (6)	14/21 (4)	9/21 (5)	14/21 (4)	17/21 (7)	16/21 (5)	10/21 (5)	11/21 (4)	10/21 (3)	12/21 (3)	11/21 (2)	13/21 (5)	

Tableau 21 : Accord entre les questions pour les stratégies

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	7/11 (5)												
V1q4	8/11 (7)	8/11 (6)											
V2q2	7/11 (6)	9/11 (6)	9/11 (7)										
V2q3	9/11 (7)	9/11 (6)	8/11 (7)	8/11 (6)									
V2q4	7/11 (5)	11/11 (6)	8/11 (6)	9/11 (6)	9/11 (6)								
V3q2	8/11 (7)	8/11 (6)	9/11 (8)	7/11 (6)	10/11 (8)	8/11 (6)							
V3q3	10/11 (8)	6/11 (5)	9/11 (8)	7/11 (6)	8/11 (7)	6/11 (5)	9/11 (8)						
V4q2	9/11 (6)	10/11 (6)	9/11 (7)	11/11 (7)	8/11 (6)	10/11 (6)	7/11 (6)	7/11 (6)					
V4q3	8/11 (6)	10/11 (6)	9/11 (7)	11/11 (7)	8/11 (6)	10/11 (6)	7/11 (6)	7/11 (6)	11/11 (7)				
V5q3	8/11 (5)	8/11 (4)	7/11 (5)	7/11 (4)	8/11 (5)	8/11 (4)	7/11 (5)	7/11 (5)	7/11 (4)	7/11 (4)			
V5q4	7/11 (6)	7/11 (5)	6/11 (6)	6/11 (5)	9/11 (7)	7/11 (5)	8/11 (7)	6/11 (6)	6/11 (5)	6/11 (5)	8/11 (5)		
V5q5	10/11 (6)	8/11 (5)	7/11 (6)	7/11 (5)	10/11 (7)	8/11 (5)	9/11 (7)	9/11 (7)	7/11 (5)	7/11 (5)	9/11 (5)	8/11 (6)	

4.6 Accord entre les questions groupe par groupe

Étant donné les différences importantes observées jusqu'à présent entre les quatre groupes, nous avons ensuite calculé l'accord entre les questions pour les catégories, groupe par groupe (voir annexe C) et fait pareillement pour les stratégies de pensée (voir annexe D). En procédant de la même façon que pour les tableaux 20 et 21, nous avons construit les tableaux des annexes C et D à partir des tableaux 10, 12, 14, 16 pour les catégories, et les tableaux 11, 13, 15, 17 pour les stratégies. Nous avons compté le nombre de fois où deux questions nécessitaient, ou non, les mêmes catégories de pensée, et divisé le résultat par 21 correspondant au nombre total de catégories de pensées possibles (soit 21 possibilités). Ensuite, nous avons recompté les catégories de pensées

réellement partagées (nombres entre parenthèses). Nous avons procédé pareillement pour les stratégies de pensée, mais divisé le résultat par 11 plutôt que 21.

Nous avons synthétisé les informations contenues dans les annexes C et D, et produit les tableaux de fréquences des accords incluant absences et présences communes (tableaux 22 et 24) et ne comptabilisant que les présences communes (tableaux 23 et 25). Nous mettons ainsi en évidence les variations des valeurs d'accord entre les couples de questions selon le groupe.

De manière générale, l'accord augmente peu à peu au fil des années de formation puis de pratique lorsqu'on considère les catégories. Par exemple, on peut constater (tableau 22) que seules les étudiantes de 3^e année et les expertes ont des valeurs d'accord entre deux questions qui atteignent 19/21 (90%) ou 20/21 (95%). De plus, une grande majorité de couples de questions ont un accord supérieur à 12/21 (57%). En effet, il y en a 69 sur 78 en 1^e année, 71/78 en 2^e année, 73/78 en 3^e année et 74/78 pour les expertes. Il est toutefois impossible de mettre en évidence des couples de questions ayant un comportement semblable pour les quatre groupes de participantes.

Lorsqu'on s'intéresse au nombre de catégories de pensée réellement partagées dans chaque couple de questions selon le groupe, il est clair que l'augmentation d'accord observée au fil de l'évolution de l'expertise est essentiellement constituée de la non utilisation des catégories de pensées. En effet, le tableau 23 met en évidence le fait que les expertes utilisent très rarement les mêmes catégories de pensées dans deux questions. Enfin, nous remarquons qu'aucun groupe n'a de fréquence d'utilisation réellement partagée supérieure à 5/21 (24%), ce qui est faible et qui signifie que les questions nécessitent des combinaisons de catégories de pensée différentes quel que soit le groupe.

Du côté des stratégies de pensée (tableaux 24 et 25), l'accord a aussi tendance à augmenter au fil des années de formation et de pratique. La plupart des couples de questions ont un accord supérieur à 6/11 (55%). Nous pouvons en compter 59 sur 78 en 1^e année, 69/78 en 2^e année, 66/78 en 3^e année et 70/78 pour les expertes.

Quant aux stratégies de pensée vraiment partagées, les étudiantes de 3^e année ont des accords plus élevés que les autres groupes, soit jusqu'à 7/11 (64%). De plus, nous constatons une augmentation de l'accord vraiment partagé au fil des années de formation. Cependant, les expertes se différencient à ce niveau, utilisant peu les mêmes stratégies de pensée pour 2 questions.

Tableau 22 : La fréquence des valeurs d'accord entre deux items pour les 4 groupes et les catégories de pensée

Accord	1 ^e année	2 ^e année	3 ^e année	Expertes
8/21		1		
9/21	2			
10/21	2	2	3	2
11/21	5	4	2	2
12/21	8	6	11	4
13/21	13	6	14	4
14/21	13	13	17	7
15/21	12	14	15	8
16/21	11	17	9	15
17/21	9	12		14
18/21	3	3	4	13
19/21			1	8
20/21			2	1

Tableau 23 : La fréquence des nombres de catégories de pensée réellement partagées par couples d'items

Accord	1 ^e année	2 ^e année	3 ^e année	Expertes
0/21	20	17	25	44
1/21	34	30	13	27
2/21	22	18	14	6
3/21	1	8	14	
4/21		5	9	1
5/21	1		3	

Tableau 24 : La fréquence des valeurs d'accord entre deux items pour les 4 groupes et les stratégies de pensée

Accord	1e année	2 ^e année	3 ^e année	Expertes
2/11	1			
3/11			1	
4/11	5	2	4	2
5/11	13	7	7	6
6/11	18	18	15	12
7/11	14	17	20	14
8/11	16	16	15	22
9/11	8	17	8	14
10/11	2		8	6
11/11	1	1		2

Tableau 25 : La fréquence des nombres de stratégies de pensée réellement partagées par couples d'items

Accord	1e année	2 ^e année	3 ^e année	Expertes
0/11	8	3	16	20
1/11	30	24	17	34
2/11	26	30	3	19
3/11	10	14	10	3
4/11	4	4	14	2
5/11		3	15	
6/11			2	
7/11			1	

Les résultats ont été organisés de manière à expliciter la démarche utilisée pour identifier les catégories et les stratégies de pensées en lien avec les 13 questions étudiées. Par la suite, nous avons exploré les relations entre les questions et les catégories/stratégies de pensées pour l'ensemble des participantes, puis en considérant chaque groupe séparément. Nous nous sommes également intéressés à l'accord entre les groupes de participants, puis à la ressemblance entre les questions prises deux à deux. Les paragraphes qui suivent visent à faire une synthèse des résultats.

4.7 Synthèse des résultats

La synthèse des résultats permet de mettre en évidence les éléments suivants :

1. Les 2/3 des catégories de pensée de Fonteyn (1998) sont peu utilisées ou utilisées de manière quasi-systématique par les participantes pour les 13 questions de notre étude (voir tableau 6). Nous avons choisi de les éliminer puisque notre question de recherche est d'étudier le potentiel diagnostique du TCS et d'identifier les catégories et stratégies de pensée qui varient entre les questions. Nous avons ainsi retenu 21 catégories de pensée en lien avec 11 stratégies, sélectionnées en raison de leur importante variation d'utilisation entre les questions (voir tableau 7).
2. Pour tous les groupes confondus, les questions sollicitent entre 29% et 67% des catégories de pensées. Cependant, certaines catégories sont utilisées dans presque toutes les questions (92%) alors que d'autres ne sont utilisées que dans très peu de questions (8%) (voir tableau 8).
3. De manière générale, les stratégies de pensée sont utilisées pour plus de questions que les catégories puisque l'on regroupe plusieurs catégories en une seule stratégie. Par exemple, nous pouvons voir que quatre stratégies de pensée sont systématiquement utilisées pour répondre à toutes les questions. Aussi, les questions requièrent plus de stratégies de pensée que de catégories de pensée selon les pourcentages liés aux questions dans les tableaux 8 et 9.
4. Nous pouvons constater des disparités importantes entre les groupes de participants, c'est-à-dire que les quatre groupes n'utilisent pas les mêmes catégories de pensée pour répondre à chacune des questions (tableau 18) et les utilisent différemment (voir tableaux 10, 12, 14, 16). Il en est de même pour les stratégies de pensée (voir tableaux 11, 13, 15, 17, 19).
5. Nous avons cherché à identifier des questions qui pourraient être regroupées en fonction de leur recours aux mêmes combinaisons de catégories ou de stratégies de pensée. Nous avons constaté que les accords entre deux questions qui semblaient élevés au premier abord (tableaux 20, 21, 22 et 24) étaient en fait

essentiellement constitués des catégories ou stratégies non requises par les questions (tableaux 20, 21, 23 et 25). Nous avons quand même constaté que les pourcentages d'accord étaient plus élevés pour les stratégies que pour les catégories de pensées. Il n'a toutefois pas été possible de créer des groupes de questions de cette manière-là.

6. En général, l'accord entre deux questions pour les 4 groupes a tendance à augmenter au fil des années de formation puis de pratique (voir tableaux 22 et 23). Cependant, l'augmentation d'accord observée au fil du développement de l'expertise est essentiellement constitué de la non- utilisation des catégories/stratégies de pensées. Effectivement, les tableaux 23 et 25 montrent que les expertes utilisent très rarement les mêmes catégories/stratégies de pensée dans 2 questions.
7. L'accord pour les catégories et les stratégies vraiment partagées par deux questions pour chaque groupe est très faible. Il augmente au fil des années d'étude mais diminue ensuite chez les expertes. Nous en déduisons que chaque question est différente car elle requiert une nouvelle combinaison de catégories/stratégies de pensée pour y répondre, et que chaque groupe se distingue dans son raisonnement.
8. Dans l'intention d'utiliser une autre approche pour tenter de classer certaines questions, nous avons produit les tableaux 26 et 27. Pour ce faire, nous avons identifié dans les tableaux 8 à 17, les questions avec le pourcentage total d'utilisation de catégories /stratégies le plus élevé, et celles avec le plus faible, pour l'ensemble des participantes puis pour chaque groupe. Dans les cas où deux questions ou plus avaient un pourcentage identique, nous les avons inscrites ensemble dans les tableaux 26 et 27. Par exemple, pour le groupe de 1^{ère} année, deux questions requièrent le plus de catégories de pensée, soit les questions V1Q4 et V3Q2 à 33%. Les couleurs permettent de repérer les questions identiques. Ainsi, par exemple, la question V2Q3 est celle qui nécessite le moins de stratégies de pensée autant pour le groupe de 1^{ère} année que le groupe d'expertes. La

question V3Q2 requiert davantage de catégories et de stratégie pour tous les groupes à l'exception du groupe des étudiantes de 3^e année. Toutefois, les questions qui nécessitent le moins de catégories et stratégies de pensée diffèrent sensiblement d'un groupe à l'autre. Notons également que la question V2Q3 apparaît à la fois dans les deux colonnes des deux tableaux impliquant beaucoup de variations d'un groupe à un autre. Enfin, la question V5Q5 n'est reliée à aucune des catégories et stratégies que nous avons retenues dans notre étude pour le groupe de 3^e année. Les variations observées dans les tableaux 26 et 27 montrent que cette autre manière de procéder ne permet pas non plus de regrouper certaines questions.

Tableau 26 : Questions requérant le plus et le moins de catégories de pensée

Groupes	Questions requérant le plus de catégories de pensée	%	Questions requérant le moins de catégories de pensée	%
1-2-3-E	V3Q2	67	V5Q3	29
1 ^{ère} année	V1Q4 V3Q2	33	V5Q3	10
2 ^e année	V3Q2	48	V4Q2	10
3 ^e année	V2Q3	43	V5Q5	0
Expertes	V3Q2	39	V2Q3, V2Q4	5

*Note : Une même couleur aide à repérer une même question dans le tableau.

Tableau 27: Questions requérant le plus et le moins de stratégies de pensée

Groupes	Questions requérant le plus de stratégies de pensée	%	Questions requérant le moins de stratégies de pensée	%
1-2-3-E	V1Q4, V3Q2, V3Q3	82	V5Q3	45
1 ^{ère} année	V1Q4	55	V1Q3, V2Q3, V2Q4, V5Q3	18
2 ^e année	V3Q2	64	V2Q2, V4Q2	18
3 ^e année	V2Q3	73	V5Q5	0
Expertes	V3Q2	64	V2Q3 V2Q4	9

*Note : Une même couleur aide à repérer une même question dans le tableau.

Dans le prochain chapitre, nous tenterons de situer les observations qui se dégagent de nos résultats afin de répondre aux questions de recherche.

5. DISCUSSION

5.1 Première question de recherche : Dans quelle mesure chaque question d'un TCS visant à évaluer le RCI peut-elle être reliée à des catégories et des stratégies de pensée spécifiques?

Notre première question de recherche est reliée à la formalisation des relations entre les processus cognitifs et les questions d'un test proposée en 1983 par Tatsuoka sous la forme d'une matrice Q. Ainsi nous nous sommes intéressés à la possibilité de relier chacune des 13 questions du TCS utilisées dans notre étude à des catégories et à des stratégies de pensées spécifiques parmi celles proposées par Fonteyn (1998) afin de pouvoir évaluer le raisonnement clinique infirmier (RCI) des étudiants au cours de leur formation .

5.1.1 Possibilité de construire des Matrices Q

Dans un premier temps, nous avons dû élaborer une méthode pour identifier les catégories et les stratégies de pensées pertinentes pour les 13 questions de notre étude. Cette méthode nous a amenés à éliminer 40 catégories et deux stratégies de pensée sur la base de leur sur-utilisation ou sous utilisation.

Pour les 21 catégories et 11 stratégies de pensée retenues, nous avons ensuite proposé une méthode pour identifier les liens entre les catégories de pensée et les questions, puis entre les stratégies de pensée et les questions sur la base de l'analyse des verbatim des participantes. Cette méthode a permis d'identifier les catégories/stratégies de pensée qui sont utilisées pour répondre à chacune des questions et de produire dix matrices Q selon que nous avons considéré l'ensemble des participantes ou les quatre groupes séparément, les catégories ou les stratégies de pensée.

5.1.2 Variations et disparités entre les questions et entre les groupes

À partir de l'analyse de ces matrices, nous constatons que le nombre de catégories ou de stratégies de pensée varie beaucoup entre les questions, mais également d'un groupe à un autre. Ainsi, nous nous sommes demandé si le regard critique des expertes lors de la validation des vignettes et questions du TCS dans l'étude de Deschênes (2006) pouvait aider à documenter ces variations. Malheureusement, les commentaires des expertes que nous avons pu obtenir portaient plutôt sur la forme que sur le contenu de la question. Nous suggérons donc comme une piste de recherche intéressante d'avoir recours à une évaluation du contenu des questions du TCS par des expertes si nous voulons élaborer une matrice Q pour les questions d'un TCS. Ceci est dans la lignée de l'approche proposée par plusieurs auteurs (par exemple, Loye, 2011).

5.1.2.1 Pour les catégories de pensée

En ce qui a trait aux catégories de pensée, les matrices Q présentent d'importantes variations en termes de quantité utilisées pour chaque question. Nous en déduisons que certaines questions requièrent un petit nombre de catégories de pensées alors que d'autres doivent en combiner plusieurs. Par exemple, pour tous les groupes, la question V3Q2 nécessite plusieurs catégories/stratégies de pensée tandis que pour la question V5Q3, très peu sont employées. Notons ici que le fait d'avoir éliminé au départ de nos analyses les stratégies fortement utilisées ne change pas grand-chose car une seule a été éliminée à cause d'une trop grande utilisation (voir tableau 6).

De plus, les matrices Q nous permettent de voir qu'au-delà des variations quantitatives, les quatre groupes n'utilisent pas les mêmes catégories de pensée pour répondre à chacune des questions et les combinent de manière différente dans les questions. Les calculs des accords entre les groupes sont faibles, notamment lorsqu'on considère les expertes qui raisonnent de toute évidence différemment des étudiantes.

5.1.2.2 Pour les stratégies de pensée

Nos conclusions relativement aux stratégies de pensée sont semblables aux catégories. Toutefois, les stratégies sont davantage utilisées pour répondre à une question que les catégories, ce qui est normal puisqu'une seule stratégie de pensée de Fonteyn (1998) comprend quelques catégories (voir tableaux 1 et 5). Pour cette raison également, les stratégies sont utilisées pour un plus grand nombre de questions que les catégories de pensée.

Comme pour les catégories de pensée, les matrices Q incluant les stratégies de pensée illustrent des variations dans la quantité requise par chaque question. Par ailleurs, nous pouvons constater des disparités importantes entre les groupes de participants, mais moindre que du côté des catégories de pensée. Tout de même, en général les quatre groupes n'utilisent pas les mêmes stratégies de pensée pour répondre à chacune des questions et les utilisent différemment, particulièrement pour les expertes qui se distinguent des autres groupes. Enfin, nous n'avons pas réussi à établir des regroupements de questions qui auraient un fonctionnement comparable. En effet, nous avons d'abord tenté de les regrouper en fonction de leur recours aux mêmes combinaisons de catégories ou de stratégies de pensée. Ensuite, nous avons essayé de regrouper les questions selon le pourcentage total d'utilisation de catégories /stratégies le plus élevé, et celles avec le plus faible, pour l'ensemble des participantes puis pour chaque groupe.

En conclusion, nous avons montré qu'il est possible de relier les 13 questions du TCS utilisées dans notre étude à des catégories et stratégies de pensée spécifiques. Toutefois, il est clair qu'une matrice unique ne pourrait pas être utilisée pour les trois années de formation étant donné le manque de stabilité observé entre les 5 matrices impliquant les catégories de pensée ou les 5 matrices impliquant les stratégies de pensée. Ainsi, nous avons proposé une méthode pour élaborer de telles matrices, mais également montré que les variations importantes qui existent au fur et à mesure de l'apprentissage des étudiantes dans leur formation nécessitent que la matrice soit adaptée à leur niveau. En outre, ces 13 questions ne peuvent pas être classées selon des fonctionnements type en lien avec certaines combinaisons de catégories ou de stratégies de pensée.

5.2 Deuxième question de recherche : Comment peut-on décrire le potentiel diagnostique d'un TCS pour évaluer le RCI?

Notre deuxième question de recherche vise à décrire le potentiel diagnostique d'un TCS pour évaluer le RCI. Par potentiel, nous entendons la possibilité d'utiliser un TCS afin de mettre en évidence les catégories ou stratégies de pensée efficaces ou déficientes chez chaque étudiante en référence à un groupe d'experts, dans une démarche d'évaluation diagnostique.

Le fait d'être parvenue à mettre en place une méthode pour élaborer des matrices Q est le point de départ à la possibilité d'utiliser un TCS dans une perspective d'évaluation diagnostique. En outre, nos matrices Q mettent en évidence les catégories ou stratégies de pensée intéressantes à diagnostiquer aux diverses phases du développement cognitif d'apprentissage des étudiantes en sciences infirmières.

5.3 Mise en évidence des différentes catégories et stratégies de pensée utilisées selon leur utilisation, pour chaque groupe

Nous avons noté que les accords entre les groupes de 1^e et 2^e années étaient les plus élevés. Ces deux groupes ont nécessairement des ressemblances puisque ces étudiantes n'ont pas d'expérience, n'ont pas encore réalisé (ou très peu) de stages en milieux cliniques, et ont peu de connaissances. Ainsi, les étudiantes de 1^e et de 2^e années ont besoin de plus de pauses réflexion (catégorie la plus utilisée), cherchent de l'information (stratégie la plus utilisée *Chercher des informations manquantes*) et tentent de relier des informations (la catégorie de pensée *Relier des informations sur les conclusions d'évaluation et autres informations* est utilisée le plus fréquemment par les 2^e années). Par ailleurs, le fait qu'en 1^e année de formation les étudiantes ne sont pas au courant des politiques d'un milieu clinique explique probablement pourquoi elles n'utilisent pas la catégorie *Énoncer une règle de pratique sur la politique*. De plus, le groupe de 2^e année n'utilise pas du tout la catégorie de pensée *Fournir des explications sur les interventions*

infirmières, ce qui s'expliquerait sans doute par leur manque de connaissances et d'expérience de soins auprès des patients. Ces deux groupes correspondent de manière logique au stade 1 du modèle cognitif d'apprentissage de Goudreau, Boyer, & Dumont (2012) (voir tableau B en annexe) qui a été élaboré à partir des mêmes données (étude 1).

Nos résultats révèlent ensuite que le groupe de 3^e année quant à lui fait davantage de conclusions (stratégie de pensée *Conclure*) et que ces étudiantes sont en mesure de *fournir des explications* et *relier des informations*. Ces résultats coïncident avec ceux de Simmons (2003), car ces apprenties infirmières ont maintenant deux ans d'expérience mais ne sont pas considérées comme des expertes. D'un autre côté, certaines catégories/stratégies de pensée ne sont plus utilisées en 3^e année de formation. Ainsi la catégorie *Chercher des informations manquantes sur les ressources* et la stratégie *Énoncer une règle de pratique* disparaissent sans doute pour laisser la place à l'intuition et aux scripts. En effet, comme le rapportent Humbert, Besinger et Miech (2011), plus les étudiantes prennent de l'expérience et acquièrent des connaissances, plus elles développent des scripts. Elles développent aussi la capacité d'organiser les scripts et d'en traiter plusieurs en même temps. Le RCI commence donc à être plus évolué chez les étudiantes de 3^e année.

5.4 RCI des expertes versus RCI des étudiantes

Du côté des expertes, rejoignant les résultats de plusieurs auteurs (Deschênes, Charlin, Gagnon & Goudreau, 2012; Lucchiari & Pravettoni, 2012; Norman, 2005; Nouh & al., 2012), notre étude montre clairement que ce groupe ne peut servir de référence dans la perspective d'identifier les catégories ou stratégies de pensées qui posent problème aux étudiantes car leur raisonnement est très différent des étudiantes en formation. Les différences se situent notamment sur le plan de la fréquence des valeurs d'accord pour les catégories et les stratégies de pensée, ainsi que pour la fréquence des valeurs d'accord pour les questions. Ces accords sont plus faibles chez les expertes que chez les étudiantes. Cependant, certains spécialistes de l'évaluation diagnostique tels que Gierl (2007) et Loye (2005; 2008; 2011), rapportent que dans un but diagnostique, un « état de

connaissance », tel que l'expertise, est opérationnalisé comme un modèle de réponses idéales qui caractérise la performance à un test. En d'autres termes, en évaluation diagnostique, il est important d'avoir un modèle d'expert afin de pouvoir inférer les processus cognitifs maîtrisés ou non par un étudiant. Nous avons donc besoin d'experts pour réaliser une matrice Q de référence.

5.5 TCS et matrice Q

En effet, au départ, nous pensions que de structurer un test de concordance de script (TCS) comme celui de Deschênes (2006) sous la forme d'une matrice Q, pourrait être la première étape pour diagnostiquer les forces et les faiblesses sous-jacentes au raisonnement clinique infirmier (RCI). Chercher à élaborer une telle matrice offrait ainsi la possibilité d'étudier le potentiel diagnostique des vignettes et questions d'un TCS en analysant les liens avec les catégories/stratégies de pensée de Fonteyn. Notre hypothèse était que les étudiants qui proposent des réponses correctes aux questions posées n'utilisent pas les mêmes stratégies que ceux dont les réponses sont erronées. Étant donné le cadre exploratoire de cette étude, nous n'avons pas pu valider ou invalider cette hypothèse en faisant la différence entre les bonnes réponses et les moins bonnes réponses. Cependant, si cette hypothèse s'avère exacte, il serait possible de formaliser les relations entre chaque question et les stratégies nécessaires à une réponse acceptable. En revanche, l'utilisation du TCS dépend d'un panel d'experts pour juger de la qualité de la réponse proposée, mais puisque nos experts sont si différents des étudiantes, nous ne pouvons nous y référer pour différencier une réponse acceptable d'une réponse qui ne l'est pas.

5.7 Potentiel diagnostique des questions pour chacun des groupes

Notre étude met toutefois en évidence un potentiel diagnostique niveau par niveau (1^{ère}, 2^e, 3^e année, expertes) des questions du TCS de Deschênes pour évaluer le RCI. Notre argumentation porte sur le fait que nous avons pu produire dix matrices Q et ainsi relier chacune des 13 questions du TCS utilisées dans notre étude à des catégories et à des

stratégies de pensée spécifiques pour chacun des groupes. De plus, nous avons montré des variations importantes au fil de l'apprentissage des étudiantes dans leur formation. Ce qui nous porte à croire que les stratégies et catégories de pensée utilisées sont différentes pour chaque groupe puisque le raisonnement clinique se développe avec les apprentissages et l'expérience de l'infirmière. En effet, d'après Fero et al. (2009), la plupart des études montrent qu'il y a une progression dans la capacité des étudiants en soins infirmiers à développer une pensée critique par le biais de leur formation. De plus, selon Smith (2008), la transition des étudiants se fait à travers une série de stades de développement du raisonnement clinique, progressant du stade de novice à celui d'expert. Ces différentes étapes nécessitent différents progrès éducatifs et différentes méthodes d'évaluation des compétences. Cette découverte a donné lieu à des méthodes d'évaluation qui surpassent les méthodes traditionnelles. À cet effet, nous croyons que le TCS est un bon outil novateur d'évaluation du raisonnement clinique infirmier car il est sensible au changement, mais pour faire de l'évaluation diagnostique, il doit être adapté à chaque stade de développement de cette compétence qu'est le RCI.

5.8 Adaptation des questions du TCS

Pour adapter les questions du TCS à chaque stade de développement du RCI afin de les utiliser pour faire une évaluation diagnostique, il faudra, entre autre, définir qui est « expert ». Par exemple, les expertes pourraient être celles qui ont de bons résultats scolaires. Cette procédure permettrait d'identifier les expertes pour chaque groupe et de s'en servir comme référence.

Cette adaptation des questions du TCS devrait en outre tenir compte des 40 catégories de pensée que nous avons dû éliminer. Quelle est la cause de cette sous-utilisation ou sur-utilisation? Est-ce dû au petit nombre de questions utilisées dans l'étude 1? À ce sujet, Gagnon et al., 2009, rapportent que pour assurer la fiabilité d'un TCS, il faut un nombre suffisant de questions. Est-ce parce que les questions ne sont pas appropriées? En effet, nous remarquons que le TCS de Deschênes (2006) a été validé avec des étudiantes de 1^{ère} année du baccalauréat en sciences infirmières, alors que Fonteyn a interviewé des

expertes. Notre étude porte donc sur des questions adaptées à des étudiantes de première année de formation avec des catégories/stratégies de pensées attribuées à des expertes. Nos résultats montrent que les catégories/stratégies de pensées utilisées par les expertes sont très différentes des étudiantes. Alors peut-être que les questions devraient être adaptées aux étudiantes de chaque année de formation et selon le modèle cognitif d'apprentissage. Devrait-on se baser sur les scénarios cliniques utilisés dans l'étude de Fonteyn (1998) plutôt que ceux de Deschênes (2006) pour adapter le TCS? Ou alors, est-ce que les questions du TCS de Deschênes qui ne sont pas dans notre étude, peuvent stimuler les stratégies et catégorie de pensée sous-utilisées? Nous pensons que toutes ces questions sont des pistes de recherche pour des études ultérieures.

6. LIMITES ET PISTES DE RECHERCHE

Une première limite est en lien avec le nombre restreint de vignettes et de questions utilisées. En effet, il aurait été pertinent pour notre étude d'avoir des données pour toutes les vignettes et questions du TCS original, ce qui n'était pas possible étant donné que nous avons réalisé une analyse secondaire des données récoltées dans le cadre d'une autre étude (étude 1). Ainsi une piste de recherche serait de compléter les données et les analyses en utilisant les vignettes et questions du TCS de Deschênes (2006; Deschênes, Charlin, Gagnon & Goudreau 2011), qui ne sont pas dans l'étude 1. Elles stimulent sans doute d'autres stratégies et catégories de pensée de Fonteyn (1998), qui ne sont pas ou peu utilisées par nos participantes (tableau 7). Ainsi, nous aurions une vision plus globale du potentiel diagnostique des questions du TCS de Deschênes.

Une deuxième limite vient du fait que les participantes ont été sélectionnées dans une seule université, ce qui biaise la représentativité de la population étudiante en sciences infirmières. Il serait intéressant dans une étude ultérieure de comparer des étudiants de plusieurs universités ainsi que d'inclure les étudiantes en techniques infirmières au collégial. Ainsi, il est important de mentionner que nos résultats en lien avec les matrices élaborées ne sont pas généralisables. La méthode que nous avons développée pourrait toutefois être réutilisée dans divers contextes.

Par ailleurs, certains de nos choix méthodologiques sont discutables. Par exemple, il était arbitraire de choisir d'éliminer les catégories de pensée peu utilisées, c'est-à-dire pour lesquelles tous les codes combinés correspondaient à moins de 25% des cas, et celles très fréquemment utilisées pour lesquelles tous les codes combinés correspondaient à plus de 50% des cas).

Finalement, notre hypothèse de départ était que les étudiantes qui proposent des réponses correctes aux questions posées n'utilisent pas les mêmes stratégies que celles dont les réponses sont erronées. Étant donné la complexité du processus à mettre en place pour différencier les bonnes réponses des moins bonnes (par exemple le recrutement d'un panel d'experts) et que les conditions nécessaires débordaient les limites de cette étude,

nous avons choisi de ne pas tenir compte de la qualité de la réponse fournie par les participantes. Notre démarche repose donc sur l'ensemble des réponses aux questions sans nuances quant à leur qualité. Il serait néanmoins très pertinent de s'attarder à ces nuances dans une prochaine étude. Le fait de différencier une réponse acceptable d'une réponse qui ne l'est pas permettrait de mieux étudier le potentiel diagnostic des questions du TCS.

CONCLUSION

Les résultats de cette recherche mettent en évidence les catégories et stratégies de pensées utilisées pour répondre à certaines questions du TCS selon les groupes. Aussi, nos résultats ont permis de découvrir des disparités importantes entre les groupes, notamment que le RCI du groupe d'expertes est si différent des étudiantes, qu'il ne peut servir de référence.

De plus, cette étude démontre que le TCS a un potentiel diagnostique mais niveau par niveau (1^{ère}, 2^e, 3^e année et expertes). Or, étant donné le manque de procédures systématiques pour gérer les difficultés de raisonnement clinique des étudiantes (Audétat & al., 2012), le TCS pourrait s'avérer utile pour supporter le programme d'enseignement en sciences infirmières. Ainsi, il aiderait à identifier les difficultés de raisonnement clinique des étudiantes éprouvant certains problèmes, et ainsi permettre d'offrir une aide adaptée pour stimuler le développement des stratégies de pensée nécessaires à leur stade cognitif d'apprentissage.

Selon le courant constructiviste en éducation, l'évaluation est une bonne stratégie pour enseigner et favoriser l'apprentissage. Elle peut être utilisée afin de déterminer comment les étudiants traitent l'information et intègrent les connaissances (Nendez & al., cité dans Deschênes, Charlin, Gagnon, & Goudreau, 2012). Afin de faire progresser la pratique, il est nécessaire de développer et d'évaluer des stratégies visant à aider les infirmières à développer les compétences essentielles à l'exercice de leur pratique. Il semble donc pertinent que les instruments de mesure en éducation évoluent afin de permettre l'évaluation du RCI. Malheureusement, le TCS est employé dans le domaine de la recherche, mais encore peu en éducation. Or, malgré que cette étude ne soit qu'exploratoire, nous croyons qu'elle peut tout de même contribuer à l'avancée des connaissances dans ce domaine.

RÉFÉRENCES

- Audétat, M.C., & Laurin, S. (2010). Supervision of clinical reasoning: Methods and a tool to support and promote clinical reasoning. *Can Fam Physician*, 56, e127-e129.
- Boychuk Duchscher, J.E. (2003). Critical thinking: perceptions of newly graduated female baccalaureate nurses. *Journal of Nursing Education*, 42 (1), 14-27.
- Caire, F., Sol, J-C., Moreau, J.J., Isidori, P., & Charlin, B. (2004). Auto-évaluation des internes en neurochirurgie par tests de concordance de script (TCS) : le processus d'élaboration des tests. *Neurochirurgie*, 50 (1), 66-72.
- Carrière, B., Gagnon, R., Charlin, R., Downing, S., & Bordage, G. (2009). Assessing clinical reasoning in pediatric emergency medicine: validity evidence for a script concordance test. *Annals of Emergency Medicine*, 53 (5), 647-652.
- Charlin, B., Bordage, G., & Van der Vleuten, C. (2003). L'évaluation du raisonnement clinique. *Pédagogie Médicale*, 4 (1), 42-52.
- Charlin, B., Boshuizen, H.P., Custers, E.J., & Feltovich, P.J. (2007). Script and clinical reasoning. *Med Educ*, 41 (12), 1178-1184.
- Charlin, B., Brailowski, R.L., & Van der Vleuten, C.P.M. (2000). The script concordance test: a tool to assess the reflective physician. Teaching and learning. *Med Educ*, 12, 189-195.
- Charlin, B., Desaulniers, M., Gagnon, R., Blouin, D., & Van der Vleuten, C. (2002). 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., Lubarsky, S., Meteressian, S., Chalk, C., Goudreau, J., & Van der Vleuten, C. (2010). Assesment in the context of uncertainty using the script concordance test: more meaning for scores. *Teaching and Learning in Medicine*, 22 (3) 180-186.

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

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

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 (2), 182-190.

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

Deschênes, MF. (2006). *Élaboration et évaluation des propriétés psychométriques d'un instrument d'évaluation du raisonnement clinique empreint de Human caring*. Mémoire de maitrise, Faculté des sciences infirmières, Université de Montréal.

Deschênes, MF, Charlin, B., Gagnon, R., & Goudreau, J. (2011). Use of a Script Concordance Test to Assess Development of Clinical Reasoning in Nursing Student. *Journal of Nursing Education*, 50 (7): 381-387.

Dossat, L. (2002). *L'évaluation des élèves : Université d'été*. Compte-rendu de Dominique Seguetchian, Collège Larabière.

Durand, M-J., & Chouinard, R. (2006). *L'évaluation des apprentissages, de la planification de la démarche à la communication des résultats*, Montréal, éd. HMH.

Dunphy, B., Cantwell, R., Bourke, S., Fleming, M., Smith, B., Joseph, K.S., & Dunphy, S.L. (2010). Cognitive elements in clinical decision making: Toward a cognitive model for medical education and understanding clinical reasoning. *Adv in Health Sci Educ*, 15, 229-250.

Durak, H.I. , Caliskan, S.A., Bor, S., & Van Der Vleuten, C. (2007). Use of case-based exams as an instructional teaching tool to teach clinical reasoning, *Medical Teacher*, 29, e170-e174.

Elizondo-Omana, RE, Morales-Gomez, JA, Morquecho-Espinoza, O., Hinojosa-Amaya, JM, Villarreal-Silva, EE, Garcia-Rodriguez, M., Guzman-Lopez, S. (2010). Teaching skills to promote clinical reasoning in early basic science courses. *Anatomical Science Education*, September-october, 3, 267-271.

Elstein, AS, & Swchartz A. (2002). Clinical problem solving and diagnostic decision making: Selective review of the cognitive literature. *BMJ*, 324, 729-732.

Faculté des sciences infirmières de l'Université de Montréal (2003). *Révision du programme de Baccalauréat en sciences infirmières*. Montréal : Université de Montréal.

Feltovich, P. J., & Barrow H. S. (1984). Issues of generality in medical problem solving. Dans H. G. Schimdt & M. L. De Volder (Eds), *Tutorials in Problem-Based Learning : A new Direction in Teaching the Health Professions*. Allen, Hollande : Van Gorcum.

Fero, LJ, Witsberger, CM, Wesmiller, SW, Zullo, TG, & Hoffman, LA. (2009). Critical thinking ability of new graduate and experiences nurces. *Journal of advanced Nursing*, 65 (1), 139-148.

Fonteyn, M. (1998). *Thinking strategies for nursing practice*. Philadelphie, Lippincott.

Fonteyn, M., & Ritter, B. (2008). Clinical reasoning in nursing. In J. Higgs and M. Jones (Eds), *Clinical reasoning in the health professions*. (3rd ed, pp. 235-244). Oxford, England, Butterworth Heinemann.

Fortin, M.F., Taggart, M.E., K  rouac, S., & Normand, S. (1988). *Introduction    la recherche : auto-apprentissage assist   par ordinateur*. D  carie, Montr  al.

Funkesson, K.H., Anb  cken, EL., & Ek, A.C. (2007). Nurses' reasoning process during care planning taking pressure ulcer prevention as an exemple. A think aloud study. *International Journal of Nursing Studies*, 44, 1109-1119.

Gagnon, R., Charlin, B., Lambert, C., Carri  re, B., & Van der Vleuten, C. (2009). Script concordance testing: more cases or more questions? *Adv in Health Scien Educ*, 14, 367-375.

Gierl, MJ. (2007). Making diagnostic inferences about cognitive attributes using the rule-space model and attribute hierarchy method. *Journal of Education Measurement*, 44 (4), 325-340.

Goransson, KE, Ehnfors, M., Fonteyn, ME, Ehrenberg, A. (2008). Thinking strategies used by Registered Nurses during emergency department triage. *Journal of Advanced Nursing*. 61 (2), 163-172.

Goudreau, J., Pepin, J., Larue, C., Legault, A., Dubois, S., Boyer, L., Lang, A., Tardif, J., Fleury, D., Desc  teaux, R. (ao  t 2009). *Approche int  gr  e de la formation initiale et continue des infirmi  res pour am  liorer la qualit   du raisonnement clinique et du leadership cliniques infirmiers*. Recension des   crits. Recherche financ  e par la Fondation canadienne de la recherche sur les services sociaux, Concours Recherche,   change et impact pour le syst  me de sant  . Non publi  .

Goudreau, J., Boyer, L., & Dumont, K. (2012). Un mod  le cognitif d'apprentissage du raisonnement clinique infirmier. Communication orale, SIDIIEF, Gen  ve, Suisse, mai 2012.

Goulet, F., Jacques, A., Gagnon, R., Charlin, B., & Shabah, A. (2010). Poorly performing physicians : Does the script concordance test detect bad clinical reasoning? *Journal of Continuing Education in The Health Professions*, 30 (3), 161-166.

Groves, M., Scott, I., & Alexander, H. (2002). Assessing clinical reasoning: a method to monitor its development in a PBL curriculum. *Medical Teacher*, 24 (5), 507-515.

Higgs, J., Jones, M., Loftus, S., & Christensen, N. (2008). *Clinical reasoning in the health professions* (3rd ed.), (pp.413-421), éd. Elsevier.

Humbert, A. J., Besinger, B., & Miech, E.J. (2011). Assessing clinical reasoning skills in scenarios of uncertainty : convergent validity for a script concordance test in an emergency medicine clerkship and residency. *Academic emergency medicine*, 18, 627-634.

Kalarimades, M., Caire, F., Dauger, F., Brassier, G., & Moreau, J.-J. (2008). (Creation of self-assessment tools for on-line continuing medical education). (Création d'outils d'autoévaluation pour une formation médicale continue en ligne. Modélisation à propos d'une formation). *Neurochirurgie*, 54 (1), 21-27.

Kautz, D.D., Kuiper, R.A., Pesut, D.J., Knight-Brown, P., & Daneker, D. (2005). Promoting clinical reasoning in undergraduate nursing students: Application and evaluation of the Outcome Present State Test (OPT) Model of clinical reasoning. *International Journal of Nursing Education Scholarship*, 2 (1), 4-19.

Lamoureux, A. (2006). *Recherche et méthodologie en sciences humaines*. 2^e édition, Beauchemin Chenelière Éducation, Montréal.

Lasnier, F. (2000). *Réussir la formation par compétences*. Montréal : Guérin.

Legendre, R. (2005). *Dictionnaire actuel de l'éducation* (3^e édition). Montréal: Guérin.

Lindeman, C.A. (2000). The future of nursing education. *Journal of Nursing Education*, 39 (1), 5-12.

Louis, R. (1999). *L'évaluation des apprentissages en classe*. Laval : Éditions Études Vivantes.

Loye, N. (2005). Quelques nouveaux modèles de mesure. *Mesure et Évaluation*, 28 (3), 51-68.

Loye, N. (2008). *Conditions d'élaboration de la matrice Q des modèles cognitifs et impact sur sa validité et sa fidélité*. Thèse de doctorat non-publiée, Université d'Ottawa, Ottawa.

Loye, N. (2010). 2010, odyssée des modèles de classification diagnostique (MCD). *Mesure et évaluation en éducation*, 33 (3), 75-98.

Loye, N., Caron, F., Pineault, J., Tessier-Baillargeon, M., Burney-Vincent, C., et Gagnon, M. (2011). La validité du diagnostic issu d'un mariage entre didactique et mesure sur un test existant. Dans G. Raïche, K. Paquette-Côté et D. Magis (Dir.), *Des mécanismes pour assurer la validité de l'interprétation de la mesure en éducation*, volume 2 (pp. 11-30). Ste-Foy, Québec: Presses de l'Université du Québec.

Lubarsky, S., Chalk, C., Kazitani, D., Gagnon, R., & Charlin, B. (2009). The script concordance test: a new tool assessing clinical judgement in neurology. *Canadian Journal of Neurological Science*, 36, 326-331.

Lucchiari, C., & Pravettoni, G. (2011). Cognitive balanced model : a conceptual scheme of diagnostic decision making. *Journal of Evaluation in Clinical Practice*, 18, 82-88.

Marie, I., Sibert, L., Roussel, F., Hellot, M-F., Lechevallier, J., & Weber, J. (2005). Le test de concordance de scripts : un nouvel outil d'évaluation du raisonnement et de la compétence clinique en médecine interne? *La revue de médecine interne*, 26, 501-507.

Meterissian, S., Zabolotny, B., Gagnon, R., & Charlin, B. (2007). Is the script concordance a valid instrument for assessment of intraoperative decision-making skills? *American journal of surgery*, 193(2), 248-251.

Miles, M.B., & Huberman, A.M. (2005). *Analyse des données qualitatives*. 2^e édition, De boeck, Bruxelles.

Newble, D., Norman, G., & Van der Vleuten, C. (2000). Assessing clinical reasoning. Dans J. Higgs & M. Jones (Eds), *Clinical reasoning in the health professions* (2nd ed.), 156-165. Oxford : Butterworth Heinemann.

Nouh, T., Boustros, M., Reid, S., Pace, D., Walker, R., MacLean, A., Hameed, M., Charlin, B., & Meteressian, S.H. (2012). The script concordance a measure of clinical reasoning: a national validation study. *The American Journal of surgery*, 203, 530-534.

Norman, G. (2005). Research in clinical reasoning: past history and current trends. *Medical Education*, 39, 418-427.

Park, AJ, Barber, MD, Bent, AE, Dooley, YT, Dancz, C., Sutkin, G., & Jelovsek, J.E. (2010). Assesment of intraoperative judgment during gynecologic surgery using the Script Concordance Test. *American Journal of Obstetric & Gynecology*, 203 (240), e1-e6.

Perrenoud, P. (2000). L'école saisie par les compétences, dans C. Bosman., F. M. Gérard., & X. Roegiers (Éds), *Quel avenir pour les compétences?* (pp. 21-41). Bruxelles : Éditions De Boeck Université.

Profetto-McGrath, J. (2005). Critical thinking and evidence-based practice. *J Prof Nurs*, 21(6), 364-371.

Remaekers, S., Kremer, W., Pilot, A., van Beukelen, P., & van Keulen, H. (2010). *Assesment & Evaluation in Higher Education*, 35 (6), 661-673.

Rey, B., Carette, V., Defrance, A., & Kahn, S. (2006). *Les compétences à l'école. Apprentissage et évaluation*. Bruxelles : De Boeck.

- Roussos, L.A., Templin, J.L., & Henson, R.A. (2007). Skills diagnosis using IRT-based latent class models. *Journal of Educational Measurement*, 44 (4), 293-311.
- Ruiz, JG., Tunuguntla, R., Charlin, B., Ouslander, JG., Symes, SN., Gagnon, R., Phanco, F., Roos, BA. (2010). The script concordance test as a measure of clinical reasoning skills in geriatric urinary incontinence. *JAGS*, 58, 2178-2184.
- Scallon, G. (1988). *L'évaluation formative des apprentissages : la réflexion*. Les Presses de l'université Laval.
- Scallon, G. (2000). *L'évaluation formative*. St-Laurent : Éditions du Renouveau Pédagogique.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Saint-Laurent : Éditions du Renouveau Pédagogique.
- Schmidt, H.G., Norman, G.R., & Boshuizen, H.P.A. (1990). A cognitive perspective on medical expertise. Theory and implications. *Academic Medicine*, 65 (10), 611-621.
- Sibert, L., Charlin, B., Gagnon, R., Corcos, J., Khalaf, A., & Grise, P. (2001). Evaluation du raisonnement clinique en urologie : l'apport du Test de Concordance de Script. *Progrès en urologie*, 11, 1213-1219.
- Simmons, B. (2010). Clinical reasoning: concept analysis. *Journal of Advanced Nursing*, 66 (5), 1151-1158.
- Smith, C.S. (2008). A developmental approach to evaluating competence in clinical reasoning. *JVME*, 35 (3), 375-381.
- Tanner, C.A., Padrick, K.P., Westfall, U.E., & Putzier, D.J. (1987). Diagnostic reasoning strategies of nurses and nursing students. *Nursing Research*, 36 (6), 358-363.
- Tardif, J. (2006). *L'évaluation des compétences : documenter le parcours de développement*. Montréal : éd. De la Chenelière, coll. Chenelière Éducation.

Tatsuoka, K. K. (1983). Rule-space: an approach for dealing with misconceptions based on item response theory. *Journal of educational measurement*, 20, 345-354.

Tourangeau, AE, Cranley, LA, & Jeffs, L. (2006). Impact of nursing on hospital patient mortality: a focused review and related policy implications. *Qual Saf Health Care*. Feb, 15 (1), 4-8.

Van der Vleuten, C.P.M., & Swanson, D.B. (1990). Assessment of clinical skills with standardized patients : The state of the art. *Teaching and Learning in Medicine*, 2, 58-76.

Xin, T., Xu, Z., & Tatsuoka, K. (2004). Linkage between teacher quality, student achievement, and cognitive skills: a rule-space model. *Studies in Educational Evaluation*, 30, 205-223.

ANNEXES

Annexe A : Questions de l'étude 2

Tableau A : Vignettes et questions du TCS utilisées

Vignettes	Questions
<p>V1- M Forget, a eu une chirurgie afin de lui poser une prothèse à la hanche droite il y a 2jours. Ce matin, il refuse toujours de se lever disant qu'il en est incapable. Vous remarquez que M Forget refuse l'aide de sa conjointe que est présente au chevet. Il refuse également une médication analgésique avant la mobilisation.</p>	<p>Q2-Si vous pensez que M. Forget se méfie des effets secondaires des opiacés et qu'alors, vous remarquez que M. Forget a refusé toute médication analgésique depuis plus de 24 heures. Comment cette information influence-t-elle l'idée initiale ?</p>
	<p>Q3-Si vous pensez que M. Forget a peur de mobiliser sa hanche opérée et qu'alors vous remarquez que M. Forget souffre d'arthrose et de douleurs chroniques. Comment cette information influence-t-elle l'idée initiale ?</p>
	<p>Q4-Si vous pensez que M. Forget craint d'être une charge pour sa famille à son retour à son domicile et qu'alors vous remarquez que la conjointe de M Forget a été diagnostiquée d'un cancer du sein il y a 6 mois. Comment cette information influence-t-elle l'idée initiale ?</p>
<p>V2-Vous travaillez aux soins à domicile du CLSC Rosemont et vous visitez M Pouliot qui est diabétique de Type II depuis 10 ans. Il vient de commencer à prendre de l'insuline depuis 2 semaines. À votre visite matinale, vous observez qu'il présente une glycémie à 22, outre la présence d'une polydipsie, M Pouliot ne présente pas de signe d'hyperglycémie. Sa femme est présente lors de cette visite, c'est elle qui vient tout juste d'administrer l'insuline selon l'échelle prescrite par le médecin.</p>	<p>Q2-Si vous pensiez à explorer avec M. Pouliot et sa femme la signification de cette expérience de santé et du contrôle de son diabète et qu'alors vous trouvez que M Pouliot a un langage agressif dès que l'on fait mention du contrôle de son diabète. Comment cette information influence-t-elle l'idée initiale ?</p>
	<p>Q3-Si vous pensiez à valider les connaissances du couple en lien avec les signes et symptômes à surveiller et qu'alors, vous trouvez que M. Pouliot n'écoute pas les recommandations et</p>

	laisse sa femme prendre les informations nécessaires. Comment cette information influence-t-elle l'idée initiale ?
	Q4-Si vous pensiez à demander à la femme de M. Pouliot si cette situation se présente fréquemment et comment il gère habituellement l'administration de l'insuline et qu'alors, vous trouvez que M Pouliot dit s'être administré une dose plus élevée d'insuline hier soir. Comment cette information influence-t-elle l'idée initiale ?
V3-Nicolas, 2 ans ½ est admis pour un bronchospasme. Vous surveillez constamment l'état respiratoire de Nicolas qui présente un tirage sous costal et xiphoïdien, une saturation à 92% à l'air ambiant et une respiration à 42/minutes. Inquiète, la mère de Nicolas est à son chevet.	Q2-Si vous pensiez à demander à la mère de Nicolas ce qu'elle fait habituellement afin de contrôler les crises d'asthmes de son fils et qu'alors, vous trouvez que la mère de Nicolas se sent dépassée par les événements, monoparentale, elle cherche du soutien. Comment cette information influence-t-elle l'idée initiale ?
	Q3-Si vous pensiez à informer la mère de Nicolas des mesures de traitements d'une crise d'asthme vous la renseignez entre autres sur les médicaments administrés et qu'alors, vous trouvez que la mère de Nicolas craint un ralentissement de croissance de son enfant effet secondaire des médicaments glucocorticoïdes utilisés lors des crises d'asthme. Comment cette information influence-t-elle l'idée initiale ?

<p>V4-Vous visitez à domicile M et Mme Lafrance. Mme Lafrance souffre d'un cancer des os en phase terminale, elle désire mourir chez elle entourée de sa famille. Vous devez lui administrer ses injections d'analgésique à des heures régulières. À votre arrivée, M Lafrance a le faciès tendu et les larmes aux yeux, il se dit dépassé par la situation et estime que sa femme est très souffrante.</p>	<p>Q2-Si vous pensez à demander à M. Lafrance la meilleure façon de l'aider lui et sa femme pour le moment et qu'alors M. Lafrance vous dit qu'il trouve que sa femme reçoit trop d'injections d'analgésiques et qu'il a peur que le médicament accélère son décès. Comment cette information influence-t-elle l'idée initiale ?</p>
	<p>Q3-Si vous penser à faire verbaliser M. Lafrance sur sa propre perception de la mort et qu'alors, M. Lafrance vous dit que seul Dieu peut décider de l'heure de la mort d'une personne. Comment cette information influence-t-elle l'idée initiale ?</p>
<p>V5-Vous travaillez à l'unité des soins coronariens. M Marquis, 54 ans, se plaint d'une douleur accablante et soudaine à la région épigastrique à 10/10. Anxieux il devient de plus en plus dyspnéique. Vous initiez des mesures diagnostiques tel que prescrit, vous faites entre autres un électrocardiogramme et un dosage des enzymes cardiaques.</p>	<p>Q2-Si vous pensez à soutenir M. Marquis en lui disant que nous pouvons l'aider face à sa douleur soudaine et qu'alors vous notez que M Marquis estime que cette fois-ci c'est trop dur pour lui, il ne tiendra pas le coup. Comment cette information influence-t-elle l'idée initiale ?</p>
	<p>Q3-Si vous pensez à contacter la famille de M. Marquis et qu'alors, vous notez que M. Marquis désire qu'on le laisse seul, il ne veut pas que sa famille le voie dans cet état. Comment cette information influence-t-elle l'idée initiale ?</p>
	<p>Q4-Si vous pensez à expliquer vos interventions diagnostiques au fur et à mesure et qu'alors, vous notez que M. Marquis présente une augmentation marquée de son anxiété par les explications reçues. Comment cette information influence-t-elle l'idée initiale ?</p>

Annexe B : MCA de l'étude 1

Tableau B : Le modèle cognitif d'apprentissage (MCA) du raisonnement clinique infirmier (RCI) de Goudreau, Boyer, & Dumont (2012)

OBSTACLES AU DÉVELOPPEMENT DU RCI : LA PRIORITÉ À L'INTERVENTION PAR RAPPORT À LA RÉFLEXION; LA JUSTIFICATION DES INTERVENTIONS

1 ^{ère} année J'AI BESOIN D'INFORMATIONS POUR INTERVENIR	2e année JE DOIS JUSTIFIER SCIENTIFIQUEMENT MES INTERVENTIONS	3e année J'AJUSTE MES INTERVENTIONS SELON LES SITUATIONS	IND J'AJUSTE MES INTERVENTIONS SELON LES PRATIQUES DU MILIEU	Experte J'AJUSTE MES INTERVENTIONS SELON MON DOMAINE
Aider en consultant des ressources OBSTACLE : PERCEPTION ET EXPÉRIENCE ANTÉRIEURES « TASK- ORIENTED » DU RÔLE DE L'INFIRMIÈRE APPROPRIATION (INTÉGRATION- ACCEPTATION) DE L'ASPECT SCIENTIFIQUE DE LA PRATIQUE INFIRMIÈRE	Aider en utilisant des résultats probants APPROPRIATION (MAÎTRISE) D'UNE DÉMARCHE D'UTILISATION DE LA LITTÉRATURE SCIENTIFIQUE INFIRMIÈRE	Aider en visant des buts réalistes pour le patient/famille ARTICULATION DES OPÉRATIONS : LIER DES INFORMATIONS DE DIVERSES NATURES/FAIRE DES HYPOTHÈSES	Aider en utilisant les protocoles, procédures et routines OBSTACLE : INTÉGRATION DES ROUTINES EN GARDANT UN ESPRIT CRITIQUE INTÉGRATION DANS LE MILIEU APPROPRIATION DU RÔLE TEL QUE PRATiqué	Aider en utilisant son expérience OBSTACLE : DISPERSION DE L'EXPÉRIENCE MAÎTRISE D'UN DOMAINE DES SOINS INFIRMIERS

Annexe C : Calculs des accords entre les questions pour les catégories de pensées

Tableau C1 : Accord entre les questions pour les catégories, groupe 1^{ère} année

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	15/21 (1)												
V1q4	9/21 (0)	13/21 (1)											
V2q2	15/21 (1)	17/21 (1)	13/21 (1)										
V2q3	14/21 (1)	14/21 (0)	12/21 (1)	18/21 (2)									
V2q4	16/21 (2)	14/21 (0)	10/21 (0)	18/21 (2)	17/21 (2)								
V3q2	11/21 (1)	13/21 (1)	17/21 (5)	13/21 (1)	12/21 (1)	10/21 (0)							
V3q3	15/21 (2)	13/21 (0)	11/21 (1)	13/21 (0)	12/21 (0)	16/21 (2)	13/21 (2)						
V4q2	14/21 (1)	16/21 (1)	12/21 (1)	18/21 (2)	15/21 (1)	17/21 (2)	12/21 (1)	14/21 (1)					
V4q3	15/21 (2)	17/21 (2)	13/21 (2)	16/21 (1)	12/21 (0)	12/21 (0)	13/21 (2)	11/21 (0)	14/21 (1)				
V5q3	14/21 (0)	16/21 (0)	14/21 (1)	16/21 (0)	17/21 (1)	15/21 (0)	12/21 (0)	14/21 (0)	15/21 (0)	16/21 (1)			
V5q4	13/21 (1)	15/21 (1)	13/21 (2)	17/21 (2)	14/21 (1)	14/21 (1)	11/21 (1)	13/21 (1)	16/21 (2)	15/21 (2)	16/21 (1)		
V5q5	17/21 (3)	15/21 (1)	9/21 (0)	17/21 (2)	16/21 (2)	16/21 (2)	11/21 (1)	15/21 (2)	14/21 (1)	13/21 (1)	14/21 (0)	15/21 (2)	

Tableau C2 : Accord entre les questions pour les catégories, groupe 2^e année

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	17/21 (2)												
V1q4	15/21 (1)	15/21 (1)											
V2q2	17/21 (1)	17/21 (1)	15/21 (1)										
V2q3	16/21 (2)	16/21 (2)	16/21 (3)	18/21 (3)									
V2q4	16/21 (1)	16/21 (2)	14/21 (2)	14/21 (1)	17/21 (4)								
V3q2	10/21 (1)	12/21 (2)	14/21 (4)	10/21 (1)	13/21 (4)	11/21 (3)							
V3q3	14/21 (0)	14/21 (0)	12/21 (0)	14/21 (0)	11/21 (0)	13/21 (1)	13/21 (3)						
V4q2	18/21 (1)	16/21 (0)	16/21 (2)	16/21 (0)	15/21 (1)	15/21 (1)	11/21 (1)	15/21 (0)					
V4q3	15/21 (1)	15/21 (1)	17/21 (3)	15/21 (1)	18/21 (4)	16/21 (3)	14/21 (4)	12/21 (0)	16/21 (1)				
V5q3	15/21 (0)	17/21 (1)	17/21 (2)	17/21 (1)	14/21 (1)	12/21 (0)	12/21 (2)	14/21 (0)	16/21 (0)	15/21 (1)			
V5q4	16/21 (1)	14/21 (0)	16/21 (2)	16/21 (1)	17/21 (3)	15/21 (2)	11/21 (2)	13/21 (0)	17/21 (1)	16/21 (2)	14/21 (0)		
V5q5	17/21 (2)	15/21 (1)	13/21 (1)	17/21 (2)	16/21 (3)	14/21 (2)	8/21 (1)	12/21 (2)	14/21 (0)	13/21 (1)	15/21 (1)	16/21 (2)	

Tableau C3 : Accord entre les questions pour les catégories, groupe 3^e année

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	17/21 (4)												
V1q4	12/21 (3)	14/21 (3)											
V2q2	14/21 (3)	16/21 (3)	13/21 (3)										
V2q3	15/21 (5)	15/21 (4)	12/21 (4)	16/21 (5)									
V2q4	13/21 (2)	13/21 (1)	10/21 (1)	14/21 (2)	13/21 (3)								
V3q2	10/21 (2)	12/21 (2)	15/21 (5)	11/21 (2)	12/21 (4)	10/21 (1)							
V3q3	15/21 (4)	13/21 (2)	14/21 (4)	12/21 (2)	13/21 (4)	13/21 (2)	14/21 (4)						
V4q2	13/21 (0)	15/21 (0)	12/21 (0)	14/21 (0)	11/21 (1)	15/21 (0)	14/21 (1)	15/21 (1)					
V4q3	15/21 (1)	17/21 (1)	14/21 (1)	14/21 (0)	13/21 (1)	15/21 (0)	12/21 (0)	15/21 (1)	19/21 (0)				
V5q3	14/21 (2)	16/21 (2)	15/21 (3)	15/21 (2)	14/21 (3)	12/21 (0)	15/21 (3)	14/21 (2)	16/21 (0)	16/21 (0)			
V5q4	14/21 (3)	16/21 (3)	13/21 (3)	13/21 (2)	14/21 (4)	12/21 (1)	13/21 (3)	12/21 (2)	14/21 (0)	16/21 (1)	17/21 (3)		
V5q5	14/21 (0)	16/21 (0)	13/21 (0)	15/21 (0)	12/21 (0)	16/21 (0)	13/21 (0)	14/21 (0)	20/21 (0)	20/21 (0)	17/21 (0)	15/21 (0)	

Tableau C4 : Accord entre les questions pour les catégories, groupe d'expertes

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	17/21 (1)												
V1q4	14/21 (0)	16/21 (1)											
V2q2	18/21 (1)	18/21 (1)	17/21 (1)										
V2q3	17/21 (0)	19/21 (1)	18/21 (1)	18/21 (0)									
V2q4	19/21 (1)	19/21 (1)	16/21 (0)	20/21 (1)	19/21 (0)								
V3q2	10/21 (0)	12/21 (1)	17/21 (4)	13/21 (1)	14/21 (1)	12/21 (0)							
V3q3	16/21 (0)	16/21 (0)	19/21 (2)	17/21 (0)	18/21 (0)	18/21 (0)	15/21 (2)						
V4q2	16/21 (2)	14/21 (1)	13/21 (1)	15/21 (1)	14/21 (0)	14/21 (0)	11/21 (2)	13/21 (0)					
V4q3	15/21 (0)	19/21 (2)	16/21 (1)	16/21 (0)	19/21 (1)	17/21 (0)	12/21 (1)	16/21 (0)	14/21 (1)				
V5q3	16/21 (0)	16/21 (0)	15/21 (0)	17/21 (0)	18/21 (0)	18/21 (0)	15/21 (2)	17/21 (0)	13/21 (0)	16/21 (0)			
V5q4	16/21 (0)	18/21 (1)	15/21 (0)	17/21 (0)	18/21 (0)	18/21 (0)	11/21 (0)	17/21 (0)	15/21 (1)	18/21 (1)	17/21 (0)		
V5q5	17/21 (1)	17/21 (1)	14/21 (0)	18/21 (1)	17/21 (0)	19/21 (1)	10/21 (0)	16/21 (0)	12/21 (0)	15/21 (0)	16/21 (0)	16/21 (0)	

Annexe D : Calculs des accords entre les questions pour les stratégies de pensées

Tableau D1 : Accord entre les questions pour les stratégies, groupe 1^{ère} année

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	6/11 (1)												
V1q4	4/11 (2)	5/11 (1)											
V2q2	7/11 (2)	8/11 (1)	6/11 (2)										
V2q3	8/11 (2)	7/11 (0)	5/11 (1)	10/11 (2)									
V2q4	8/11 (2)	7/11 (0)	5/11 (1)	10/11 (2)	11/11 (2)								
V3q2	7/11 (3)	8/11 (2)	6/11 (3)	7/11 (2)	6/11 (1)	6/11 (1)							
V3q3	9/11 (4)	6/11 (1)	4/11 (2)	5/11 (1)	6/11 (1)	6/11 (1)	9/11 (4)						
V4q2	5/11 (1)	8/11 (1)	8/11 (3)	9/11 (2)	8/11 (1)	8/11 (1)	7/11 (2)	5/11 (1)					
V4q3	5/11 (2)	9/11 (2)	7/11 (3)	6/11 (1)	5/11 (0)	5/11 (0)	8/11 (3)	6/11 (2)	6/11 (1)				
V5q3	6/11 (1)	7/11 (0)	5/11 (1)	8/11 (1)	9/11 (1)	9/11 (1)	4/11 (0)	4/11 (0)	6/11 (0)	7/11 (1)			
V5q4	7/11 (3)	6/11 (1)	6/11 (3)	9/11 (3)	8/11 (2)	8/11 (2)	5/11 (2)	5/11 (2)	7/11 (2)	6/11 (2)	8/11 (2)		
V5q5	8/11 (4)	6/11 (1)	2/11 (1)	7/11 (2)	8/11 (2)	8/11 (2)	7/11 (3)	9/11 (4)	5/11 (1)	4/11 (1)	6/11 (1)	7/11 (3)	

Tableau D2 : Accord entre les questions pour les stratégies, groupe 2e année

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	9/11 (2)												
V1q4	9/11 (2)	9/11 (2)											
V2q2	8/11 (1)	8/11 (1)	8/11 (1)										
V2q3	7/11 (2)	9/11 (3)	9/11 (3)	8/11 (2)									
V2q4	7/11 (2)	9/11 (3)	9/11 (3)	8/11 (2)	11/11 (5)								
V3q2	6/11 (2)	7/11 (3)	7/11 (3)	6/11 (2)	9/11 (5)	9/11 (5)							
V3q3	6/11 (1)	6/11 (1)	6/11 (1)	9/11 (2)	6/11 (2)	6/11 (2)	8/11 (4)						
V4q2	8/11 (1)	8/11 (1)	8/11 (1)	7/11 (0)	6/11 (1)	6/11 (2)	5/11 (1)	5/11 (0)					
V4q3	7/11 (2)	9/11 (3)	9/11 (3)	6/11 (1)	9/11 (4)	9/11 (4)	7/11 (4)	4/11 (1)	8/11 (2)				
V5q3	9/11 (2)	9/11 (2)	9/11 (2)	8/11 (1)	7/11 (2)	7/11 (2)	5/11 (2)	6/11 (1)	8/11 (1)	7/11 (2)			
V5q4	6/11 (1)	6/11 (1)	8/11 (2)	7/11 (1)	8/11 (3)	8/11 (3)	6/11 (3)	5/11 (1)	7/11 (1)	6/11 (1)	6/11 (1)		
V5q5	7/11 (2)	5/11 (1)	7/11 (2)	8/11 (2)	7/11 (3)	7/11 (3)	5/11 (3)	6/11 (2)	4/11 (0)	5/11 (2)	7/11 (2)	6/11 (2)	

Tableau D3 : Accord entre les questions pour les stratégies, groupe 3e année

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	10/11 (5)												
V1q4	8/11 (5)	7/11 (4)											
V2q2	9/11 (5)	10/11 (5)	8/11 (5)										
V2q3	9/11 (6)	8/11 (5)	6/11 (5)	7/11 (5)									
V2q4	7/11 (3)	8/11 (3)	6/11 (3)	7/11 (3)	7/11 (4)								
V3q2	10/11 (6)	9/11 (5)	7/11 (5)	8/11 (5)	10/11 (7)	6/11 (3)							
V3q3	8/11 (4)	7/11 (3)	9/11 (5)	8/11 (4)	6/11 (4)	6/11 (2)	7/11 (4)						
V4q2	6/11 (1)	7/11 (1)	5/11 (1)	6/11 (1)	4/11 (1)	8/11 (1)	5/11 (1)	7/11 (1)					
V4q3	6/11 (1)	7/11 (1)	5/11 (1)	6/11 (1)	4/11 (1)	6/11 (0)	5/11 (1)	7/11 (1)	9/11 (0)				
V5q3	9/11 (4)	8/11 (3)	8/11 (4)	7/11 (3)	7/11 (4)	7/11 (2)	8/11 (4)	8/11 (3)	6/11 (0)	8/11 (1)			
V5q4	10/11 (5)	9/11 (4)	7/11 (4)	8/11 (4)	8/11 (5)	6/11 (2)	9/11 (5)	7/11 (3)	5/11 (0)	7/11 (1)	10/11 (4)		
V5q5	5/11 (0)	6/11 (0)	4/11 (0)	5/11 (0)	3/11 (0)	7/11 (0)	4/11 (0)	6/11 (0)	10/11 (0)	10/11 (0)	7/11 (0)	6/11 (0)	

Tableau D4 : Accord entre les questions pour les stratégies, groupe expertes

Questions	V1q2	V1q3	V1q4	V2q2	V2q3	V2q4	V3q2	V3q3	V4q2	V4q3	V5q3	V5q4	V5q5
V1q2													
V1q3	8/11 (1)												
V1q4	6/11 (1)	9/11 (2)											
V2q2	8/11 (1)	11/11 (2)	9/11 (2)										
V2q3	9/11 (1)	10/11 (1)	8/11 (1)	10/11 (1)									
V2q4	9/11 (1)	10/11 (1)	8/11 (1)	10/11 (1)	11/11 (1)								
V3q2	5/11 (2)	6/11 (2)	8/11 (4)	6/11 (2)	5/11 (1)	5/11 (1)							
V3q3	6/11 (0)	7/11 (0)	9/11 (2)	7/11 (0)	8/11 (0)	8/11 (0)	6/11 (2)						
V4q2	9/11 (3)	8/11 (2)	6/11 (2)	8/11 (2)	7/11 (1)	7/11 (1)	7/11 (4)	4/11 (0)					
V4q3	9/11 (2)	10/11 (2)	8/11 (2)	10/11 (2)	9/11 (1)	9/11 (1)	7/11 (3)	6/11 (0)	9/11 (3)				
V5q3	8/11 (1)	7/11 (0)	5/11 (0)	7/11 (0)	8/11 (0)	8/11 (0)	6/11 (2)	7/11 (0)	8/11 (2)	8/11 (1)			
V5q4	6/11 (0)	9/11 (1)	7/11 (1)	9/11 (1)	8/11 (0)	8/11 (0)	4/11 (1)	7/11 (0)	6/11 (1)	8/11 (1)	7/11 (0)		
V5q5	7/11 (1)	8/11 (1)	8/11 (2)	8/11 (1)	9/11 (1)	9/11 (1)	5/11 (2)	8/11 (1)	5/11 (1)	7/11 (1)	6/11 (0)	6/11 (0)	