

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

DÉVELOPPER UN PROCESSUS POUR GÉNÉRER UN JUGEMENT CRÉDIBLE
EN ÉVALUATION DE PROGRAMME

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN ÉDUCATION

PAR

VÉRONIQUE LEMIEUX

MAI 2014

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Avant de présenter le contenu de mon mémoire, je souhaite remercier particulièrement certaines personnes qui, sans elles, l'aboutissement de ce dernier n'aurait pas été possible.

Je tiens à remercier vivement ma directrice, Marthe Hurteau, professeure à la faculté d'éducation de l'Université du Québec à Montréal, pour sa patience et ses encouragements soutenus. Tout au long de cette recherche, ses suggestions et commentaires, toujours pertinents ont rendu le dépôt de ce mémoire possible. Merci de m'avoir intégrée dans un projet de recherche tant intéressant que pertinent et de m'avoir transmis votre passion pour la recherche. Encore une fois, merci pour tout.

J'aimerais également remercier Monsieur Sylvain Houle et Monsieur Pascal Ndinga, qui ont accepté, à titre de correcteurs, d'apporter des commentaires constructifs à ce mémoire. Je les remercie pour l'intérêt qu'ils ont manifesté à l'évaluation de mon mémoire.

Je tiens aussi particulièrement à remercier sincèrement toutes les personnes qui ont participé à ce projet de recherche, entre autres Marie-Pierre Marchand et tous les intervenants qui ont participé à cette étude. Merci d'avoir partagé avec nous votre expertise et merci de nous avoir accordé votre temps précieux.

Enfin, merci à ma famille pour leur soutien, sans vous, jamais ce projet n'aurait été possible. Merci à Dominic, Danielle, Rolland et Olivier.

TABLE DES MATIÈRES

REMERCIEMENTS	ii
LISTE DES FIGURES.....	vi
RÉSUMÉ.....	vii
INTRODUCTION	1
PROBLÉMATIQUE	4
1.1 L'évaluation de programme.....	4
1.2 Son évolution.....	5
1.3 L'état actuel de la situation.....	7
1.3.1 Les remises en question concernant l'évaluation de programme.	7
1.3.2 Les pistes d'explication suggérées	8
1.4 Les dernières avancées concernant le raisonnement évaluatif.....	10
1.5 Les dernières avancées sur le jugement crédible.....	12
1.5.1 La définition du jugement crédible	12
1.5.2 Les conditions essentielles en vue de générer un jugement crédible.....	13
1.5.3 Le problème : Comment développer une argumentation pour appuyer le jugement crédible.	13
1.6 La question de recherche	14
1.7 La pertinence de l'étude.....	15
1.7.1 La pertinence scientifique	15
1.7.2 La pertinence sociale.....	16
CHAPITRE II.....	18
CADRE THÉORIQUE	18
2.1 Introduction	18
2.2 La notion de jugement.....	19
2.2.1 Le courant philosophique.....	19
Le jugement comme une aptitude qui se définit comme une sagesse.....	20
2.2.2. Le second courant : une compétence.....	23
2.2.3 La compétence à juger.....	24
2.3 Les processus permettant de générer un jugement : Les modélisations	25
2.3.1. La logique de l'évaluation (Scriven, 1980).....	26
2.3.2 La logique de travail de Fournier (1995)	27
2.3.3 La contribution de Stake (2004): la pensée interprétative.....	27
2.3.4 La modélisation du processus spécifique à l'évaluation de programme (Hurteau, Lachapelle et Houle, 2006).....	28
2.3.5 Le raisonnement évaluatif (Arens, 2006).....	30
2.4 Les processus permettant de générer un jugement crédible.....	30
2.4.1 La notion de crédibilité : sa définition et son développement.....	31

2.4.2 La notion de crédibilité appliquée à l'évaluation de programme. Comment s'applique-t-elle au jugement?.....	34
2.5 Synthèse et objectifs de recherche.....	36
CHAPITRE III	39
DÉMARCHE MÉTHODOLOGIQUE	39
3.1 Le type de recherche	39
3.2 Les choix méthodologiques	41
3.2.1 L'échantillon.....	41
3.2.2 La méthode de collecte de données.....	42
3.2.3 La modalité retenue : l'entrevue semi-dirigée exploratoire	43
3.3 L'analyse et l'interprétation des données exploratoire.....	46
3.4 Les considérations éthiques	50
3.5 Les forces et les limites de l'étude	51
3.5.1 Les forces.....	51
3.5.2 Les limites	52
CHAPITRE IV.....	53
RÉSULTATS	53
4.1 Rappel des objectifs de recherche	53
4.2 L'analyse des données.....	54
4.2.1 Le codage des données.....	54
4.2.2 La synthèse des données.....	58
4.3 Les résultats de recherche visant la description d'un processus de production du jugement crédible.....	59
4.4 Atteinte des objectifs	67
4.3.1 Synthèse des résultats.....	67
CHAPITRE V.....	69
DISCUSSION	69
5.1 Les considérations méthodologiques.....	69
En premier lieu, il est important de mentionner que les résultats de la présente étude ont été réalisés à partir d'un questionnaire qui a été validé par Marchand (2011). En effet, pour s'assurer de son efficacité de répondre à la question de recherche, Marchand (2011, p. 75) affirme que :	69
5.2 Les résultats révélateurs.....	70
5.3 La modélisation de Hurteau et Houle.....	73
5.2 Les questions qui restent en suspens	74
CONCLUSION	75
APPENDICE A	76
FORMULAIRE DE CONSENTEMENT.....	77
APPENDICE B	80

QUESTIONNAIRE D'ENTREVUE.....	81
BIBLIOGRAPHIE.....	85

LISTE DES FIGURES

Figure		Page
2.1	Schéma de la compétence professionnelle.....	25
2.2	La modélisation du processus spécifique à l'évaluation de programme et les fondements théoriques de ses composantes.....	30
2.3	La démarche sous-jacente au raisonnement évaluatif : l'élaboration du jugement et de son argumentation	31
2.4	Schéma de la crédibilité	33
2.5	Processus de «crédibilisation» d'une découverte.....	34
2.6	Schéma de la crédibilité en évaluation de programme	36
2.7	Processus de «crédibilisation» du jugement.....	37
3.1	Démarches méthodologiques de la théorie ancrée.....	49
5.1	Processus d'ajustement des stratégies.....	71
5.2	Modélisation de production et de crédibilisation du jugement en évaluation de programme.....	74

RÉSUMÉ

L'évaluation de programme est de plus en plus utilisée dans notre société afin d'estimer la pertinence des programmes, projets et politiques. Paradoxalement, ce domaine fait face à de nombreuses et constantes critiques en ce qui concerne sa fonction première, celle de générer un jugement suffisamment crédible pour être accepté des clients, et ainsi collaborer au processus de prise de décision afin de guider les grandes orientations. Sur ce, Toulemonde (2005) signale que très peu de clients ayant fait appel à des évaluateurs utilisent les résultats finaux contenus dans le rapport d'évaluation (jugements et recommandations) puisque ces derniers ne correspondent pas, ou presque, aux attentes du milieu, et ne réussissent que peu à participer à la prise de décision. D'ailleurs, pour tenter de résoudre ce phénomène, plusieurs théoriciens ont tenté, sans succès, d'élaborer des modélisations du processus sous-jacent le jugement. En effet, tous ont proposé une conceptualisation linéaire en considérant le jugement comme un produit final, ce qui a rendu ces modélisations inopérantes. Par conséquent, quelques chercheurs ont entrepris de délimiter les conditions requises afin d'engendrer un jugement crédible, ce qui a fait remonter en surface un manque concernant la production de l'argumentation et de la synthèse des informations (Scriven, 1990; Bossiroy, 2009). Ainsi, l'exploration du corpus théorique concernant le jugement et la crédibilité permet d'innover en percevant le jugement comme un processus entier et comme une compétence propre à l'être humain. De même, la crédibilité de ce dernier serait assujettie à la mise en place de conditions idéales tout en demeurant propre aux personnes ayant généré ce jugement.

Plusieurs professionnels, de divers domaines d'intervention socialement reconnus, sont amenés quotidiennement à porter des jugements lors de leur pratique. À cet effet, la présente recherche, de type exploratoire, vise la cueillette d'informations permettant de poser un jugement fondé et crédible. Puisque l'objet de cette recherche porte sur le jugement, l'information sera recueillie auprès de professionnels, qui par leur reconnaissance sociale, tendent à confirmer qu'il est possible de poser un jugement professionnel crédible. Ainsi, il sera possible d'approfondir les conditions nécessaires afin de générer un jugement crédible et de reconnaître les éléments transférables à l'évaluation de programme. L'analyse des données obtenues dans le cadre de cette recherche présente des résultats à la fois convaincants et prometteurs dans le cadre du processus de formulation du jugement crédible.

Mots clés :

Compétence, crédibilité, évaluation de programme, jugement, raisonnement.

INTRODUCTION

Depuis les dernières décennies, l'évaluation de programme a vu augmenter sa popularité tant au niveau de la pratique qu'au niveau théorique. L'engouement pour la discipline de l'évaluation a donné naissance à une multitude de revues spécialisées en évaluation ainsi qu'à des publications, ce qui a permis d'enrichir et de solidifier la discipline ainsi que son corpus théorique. Ainsi, l'évaluation de programme s'est imposée dans les milieux éducatifs, gouvernementaux et politiques. Paradoxalement, même si l'évaluation de programme est de plus en plus utilisée de nos jours afin d'estimer la pertinence de nombreux programmes, celle-ci fait face à de nombreuses critiques en ce qui concerne sa fonction première, celle de générer un jugement crédible (Dubois & Marceau, 2005). Tout récemment, Hurteau, Valois et Boissiroy (2011) se sont intéressés à cette problématique en proposant une définition préliminaire du jugement crédible et des conditions requises pour l'obtenir. Les résultats de leur étude ont permis de révéler une nouvelle piste de questionnement en proposant une condition encore peu explorée à ce jour: l'élaboration de l'argumentation.

La présente étude s'inscrit dans le cadre du troisième volet du projet de recherche de Hurteau et Houle, amorcé en 2005, visant l'étude des pratiques évaluatives et le développement global du domaine de l'évaluation de programme. Elle poursuit la visée de Hurteau et Bossiroy (2009) en tentant de mieux comprendre les processus sous-jacents du jugement professionnel crédible. Pour ce faire, la présente recherche favorise l'approche exploratoire en effectuant des entrevues auprès de professionnels de différents milieux et qui sont appelés à porter un jugement dans le cadre de leur

profession. Ainsi, le présent mémoire présente l'ensemble de la démarche scientifique réalisée.

À cet effet, le premier chapitre établit la problématique de cette recherche en s'assurant ainsi de la pertinence de cette dernière. L'évolution de la discipline est présentée afin d'assurer une meilleure compréhension de la situation actuelle. Les remises en question récurrentes et les causes de ces dernières seront explorées pour ensuite faire place à la recherche de Hurteau et Bossiroy (2009), qui a servi de tremplin à cette recherche. Ensuite, l'identification de la question de recherche sera annoncée pour conclure en exposant les pertinences scientifiques et sociales de cette recherche.

Le deuxième chapitre établit les bases théoriques et les concepts du jugement et de la crédibilité utiles à la compréhension de ce mémoire. Conséquemment, une définition de l'évaluation de programme et du jugement est d'abord présentée pour faire place au processus sous-jacent le jugement et au concept de crédibilité, particulièrement lors de la pratique de l'évaluation de programme. Les objectifs spécifiques et généraux visés seront ensuite identifiés.

Le troisième chapitre expose la méthodologie utilisée afin de répondre aux objectifs préalablement établis, à savoir, l'entrevue exploratoire semi-dirigée. Finalement, les types de recherche, d'entrevue et d'approche retenus pour réaliser cette recherche seront décrits pour en arriver aux considérations éthiques et aux forces et limites de la recherche.

Le quatrième chapitre décrit et met en lumière les résultats obtenus à la suite de la passation du questionnaire aux sujets retenus et soumet des conclusions partielles afin

d'expliquer les choix de l'analyse aux lecteurs et d'expliquer l'atteinte des objectifs préalablement établis.

Enfin, une discussion des résultats de recherche sera présentée dans le dernier et cinquième chapitre, ainsi que certaines pistes de réflexion pour de futures recherches.

CHAPITRE I

PROBLÉMATIQUE

Ce premier chapitre est consacré à l'établissement de la problématique sur laquelle se penche cette recherche. Pour ce faire, une définition ainsi qu'un survol de l'évolution du domaine de l'évaluation de programme sont présentés. L'état actuel de la discipline y est décrit, le malaise auquel ce domaine est confronté y est exposé et des pistes d'explications y sont proposées. Ensuite, le chapitre présente l'étude de Hurteau et Bossiroy (2009), étude servant de racines à cette recherche. Pour terminer, la question de recherche est présentée ainsi que sa pertinence sociale et scientifique.

1.1 L'évaluation de programme

Afin de mieux comprendre les enjeux auxquels est confrontée l'évaluation de programme aujourd'hui, il importe de bien la définir et de la situer dans le temps, en établissant son historique et ses assises théoriques.

Les théories se développant de plus en plus dans le domaine, il existe aujourd'hui un large éventail de définitions en évaluation de programme et elle peut sembler difficile à définir étant donné la diversité des programmes auxquels elle se réfère. Il apparaît donc pertinent, selon cette perspective, de camper la définition retenue pour cette recherche. Si plusieurs auteurs ont habilement défini le concept de l'évaluation,

la présente recherche s'inspirera de Champagne, Contandriopoulos, Brouselle, Hartz et Denis (2009, p.38) qui l'ont décrit ainsi :

L'évaluation consiste fondamentalement à porter un jugement de valeur sur une intervention en mettant en œuvre un dispositif capable de fournir des informations scientifiquement valides et socialement légitimes sur cette intervention ou sur n'importe laquelle de ses composantes, l'objectif étant de faire en sorte que les différents acteurs concernés, dont les champs de jugement sont parfois différents, soient en mesure de prendre position sur l'intervention pour qu'ils puissent construire, individuellement ou collectivement, un jugement susceptible de se traduire en actions.

Nous retenons, pour la présente recherche, cette définition parce qu'elle inclut deux dimensions qui nous intéressent plus particulièrement, à savoir, le jugement et l'implication des parties prenantes.

1.2 Son évolution

Depuis l'Antiquité, beaucoup d'efforts ont été déployés afin de mieux saisir, décrire, comprendre et améliorer la condition humaine. Bien que les termes *évaluation de programme* et *recherche en évaluation* soient relativement récents, il en est autrement des activités qu'elles sous-tendent. Il est possible de retracer les activités en évaluation dès le tout début de la science moderne. En effet, il y a plus de trois siècles, Thomas Hobbes et ses contemporains ont tenté, par des mesures numériques, d'évaluer les conditions sociales et d'identifier les causes de mortalité, de morbidité et de désorganisation sociale (Rossi, Lipsey et Freeman, 2004).

Même si ses origines sont lointaines, ce n'est toutefois qu'au début des années 60 que l'évaluation de programme est apparue comme une réelle discipline, à part entière. C'est dans le cadre de l'élaboration de politiques pour contrer la pauvreté de la fin de la Deuxième Guerre mondiale que le gouvernement américain a investi en masse pour l'évaluation de l'efficacité et du rendement des organisations dépendantes des fonds publics (Madaus et Stufflebeam, 2004). À ce moment, l'évaluation de programme et le domaine de la recherche scientifique en sciences sociales sont si près qu'ils sont facilement confondus.

L'engouement pour la discipline de l'évaluation a donc donné naissance, dans les années 1970, à une multitude de revues spécialisées en évaluation ainsi que des publications, ce qui a permis d'enrichir et de solidifier la discipline ainsi que son corpus théorique (Rossi, Lipsey et Freeman, 2004). La professionnalisation de l'évaluation de programme entraîne donc l'élaboration de normes (principes et standards établis par le *Joint Committee on Standards for Educational Evaluation* et les lignes directives en matière d'éthique, élaborées par la Société canadienne d'évaluation et autres sociétés), la création d'associations à un niveau international, plusieurs publications scientifiques, et l'apparition de programmes d'études en évaluation au sein des universités. C'est à ce moment que l'évaluation de programme se libère peu à peu de la recherche en sciences sociales en s'imposant particulièrement en gestion des programmes sociaux. Ainsi, par ce changement de paradigme, de scientifique à constructiviste, l'évaluation de programme a su se distinguer en visant la production d'un jugement concernant un programme. En effet, nous faisons présentement face à une période de remise en question. Les modèles présentement établis accordent une attention particulière aux dialogues entre les parties prenantes et les personnes impliquées par l'évaluation de programme. Alors que l'évaluation de programme consistait au départ à mesurer, sa fonction première est aujourd'hui d'émettre un jugement sur la qualité d'un programme ou d'une

composante de ce dernier, en s'appuyant sur une argumentation solide et valide afin que les parties prenantes l'acceptent. Il est possible d'affirmer que l'évaluation ne semble pas avoir réussi encore à relever ce défi si l'on prend en considération les réactions qu'elle soulève encore aujourd'hui.

1.3 L'état actuel de la situation

Même si l'évaluation de programme est de plus en plus utilisée de nos jours afin d'estimer principalement la pertinence et l'efficacité de nombreux programmes, celle-ci fait toutefois face à de nombreuses critiques en ce qui concerne sa fonction première, c'est-à-dire celle de générer un jugement crédible (Dubois & Marceau, 2005). Sur ce, certains chercheurs s'intéressant à ce problème proposent des pistes d'explications à ces diverses remises en question.

1.3.1 Les remises en question concernant l'évaluation de programme.

Les résultats de l'étude de McDavid (2001) portant sur des évaluations de programme de Colombie-Britannique concluent que ces dernières ne parviennent toujours pas à émettre un jugement convenablement fondé pour provoquer une prise de décision. Dans un même ordre d'idée, une étude publiée en 2004 par le Secrétariat du Conseil du Trésor du Canada indique que pour 115 rapports d'évaluation étudiés, 32% possèdent un raisonnement inadéquat pour appuyer le jugement rendu et que les jugements ne sont pas soutenus par une information pertinente dans 50 % des cas. De plus, les travaux de Toulemonde (2005) dévoilent que très peu de clients ayant fait appel à des évaluateurs utilisent les résultats finaux contenus dans le rapport d'évaluation (jugements et recommandations) puisque ces derniers ne concordent pas,

ou presque pas, aux attentes du milieu, et ne réussissent que peu à participer à la prise de décision. Ainsi,

La recherche de Hurteau, Houle et Mongiat (2007) met en évidence que, sur les 40 rapports d'évaluation étudiés, seulement 50 % génèrent un jugement, et que dans 90 % des cas, le jugement ne s'avère pas pertinent puisqu'il ne répond pas à la question qui avait alors amorcé le processus d'évaluation. Leurs travaux ont également mis en lumière que, dans plusieurs des cas, les évaluateurs ne vont pas plus loin dans leur démarche que la simple description du programme. Finalement, il va de soi que si l'évaluation de programme peut utiliser l'enquête pour exposer la nature de l'évaluation (Fournier et Smith, 1993), elle doit aller plus loin que la simple description pour exposer un jugement fiable (Houle, 2006).

1.3.2 Les pistes d'explication suggérées

Au fil du temps, plusieurs auteurs ont tenté de trouver des pistes de solutions aux remises en question récurrentes dont souffre l'évaluation de programme. En effet, au début des années 1970, Guba (1972), dans son célèbre article *The failure of Educational Evaluation*, a été le premier à mentionner le malaise au sein de l'évaluation de programme et à le documenter en se questionnant sur les difficultés qu'ont les évaluateurs à développer les critères dans le but de générer un jugement, ce qui devrait logiquement se trouver dans une évaluation de programme. Dans le même ouvrage, il reproche également aux évaluateurs leur manque d'expertise, ce qui expliquerait leur difficulté à cerner le caractère spécifique de la démarche d'évaluation de programme.

Dans le même ordre d'idée, Scriven (1995) prétend que l'évaluation de programme est confrontée à un problème structurel en ce qui a trait à la logique qui accompagne le jugement posé et les recommandations émises et que les praticiens se réfèrent rarement aux fondements théoriques, ce qui affecterait la validité du jugement conclu dans le rapport d'évaluation. Fournier et Smith (1993) abordent la question en affirmant que la justification de la conclusion de la démarche évaluative est souvent délaissée en comparaison de la collecte de données, et le manque d'argumentation de cette dernière la rendrait moins crédible et légitime aux yeux des clients de l'évaluation.

Dans leur étude, Virtanen et Uusikyla (2002) avancent que les jugements posés lors de l'évaluation de programme, tel qu'elle est pratiquée présentement, ne prendraient pas en considération la complexité des programmes, mais se contenteraient plutôt de diffuser des certitudes de façon machinale sur des problèmes complexes de la société d'aujourd'hui. Par ailleurs, les travaux de Dubois et Marceau (2005) ainsi que ceux de Arens (2006) abordent le manque de cohérence entre la pratique et la théorie dans le domaine de l'évaluation de programme. Ainsi, les praticiens ne se renseigneraient pas au sujet des nouvelles avancées et découvertes en recherche évaluative alors que les chercheurs n'utiliseraient pas ce qui a été produit sur le « terrain ». De leur côté, Brousselle, Champagne et Contrandopoulos (2006) ont tenté d'illustrer l'écart qui existe entre les théories de l'évaluation de programme et la pratique de l'évaluation en aboutissant au constat qu'il y aurait un besoin de développer des connaissances intégrées entre la théorie et la pratique. Enfin, Donaldson et Lipsey (2006) situent une part du laxisme qui existe présentement dans la pratique de l'évaluation de programme dans le fait que les évaluateurs manqueraient de formation ou que cette dernière ne serait pas appropriée.

En résumé, les auteurs semblent tous s'entendre sur le fait qu'il existe présentement un obstacle majeur dans la pratique de l'évaluation, celui visant à émettre un jugement et une argumentation crédible afin qu'ils soient acceptables aux yeux des parties prenantes. C'est sur cet obstacle que quelques chercheurs ont orienté leurs recherches sur le plan du raisonnement évaluatif et qu'ils en ont proposé des pistes d'explication intéressantes et envisageables.

1.4 Les dernières avancées concernant le raisonnement évaluatif

La section suivante présente l'état général de la question sur les recherches concernant le raisonnement évaluatif. Encore aujourd'hui, peu de recherches scientifiques portent sur cet aspect et celles qui s'y sont intéressées sont pour la plupart encore relativement récentes.

Arens (2006), dans la lignée des auteurs cités, s'est intéressée au raisonnement sous-jacent de la démarche évaluative, et plus particulièrement à la façon dont les praticiens de l'évaluation produisent et justifient les jugements posés. L'auteure note que les évaluateurs documentent peu la démarche méthodologique à laquelle ils ont recours pour générer un jugement et que celle-ci s'inscrit rarement au sein d'un parcours logique. Ces manquements pourraient avoir un impact sur la crédibilité et la validité du jugement ainsi que sur les recommandations formulées par l'évaluateur dans son rapport final. De plus, le développement théorique de l'évaluation de programme étant relativement récent, ses enjeux sont, encore aujourd'hui, étroitement liés à ceux des différents domaines qui ont contribué à son évolution. Arens (2006) reconnaît qu'elle ne s'est pas attardée aux raisonnements utilisés par les évaluateurs, ce qui aurait requis de les interviewer. Le raisonnement évaluatif étant un processus

intellectuel difficilement perceptible, il aurait possiblement été plus facile à cerner en dialoguant avec les évaluateurs qu'en épluchant les rapports d'évaluation.

L'étude de Stake et Schwandt (2006) entrevoit le problème de la crédibilité du jugement émis différemment en affirmant que le jugement évaluatif nécessite de s'appuyer sur des informations relevant autant de la légitimité scientifique que sur des perceptions, ce qui les amène à introduire les notions de qualité mesurée (quality as measured) et de qualité appréciée (quality as experienced). La première fait référence aux données qualitatives et quantitatives qui servent à nourrir la question d'évaluation qui avait alors démarré l'évaluation, et la deuxième aux préoccupations, attitudes et opinion parties prenantes qui émergent en cours de route lors du processus. Les auteurs en viennent à la conclusion que les deux types de qualité devraient toujours être présents dans une démarche évaluative afin que le jugement évaluatif soit crédible.

Dans leur article intitulé *How Legitimate and Justified are Judgments in Program Evaluation?*, Hurteau, Houle et Mongiat (2009), se questionnent sur la production d'un jugement évaluatif. Ils en viennent à la conclusion que seulement 50% des rapports généraient un jugement et que si ces jugements semblaient légitimes, ils étaient rarement justifiés. Leur réflexion les amène à revenir à la définition de base de l'évaluation pour inclure la légitimité sociale – crédibilité - qui inclut par le fait même la perception des parties prenantes et les a conduit à effectuer davantage de recherches pour développer une théorie qui pourrait décrire et expliquer le processus requis pour générer un jugement crédible.

1.5 Les dernières avancées sur le jugement crédible

Le présent mémoire s'inscrit dans un projet de recherche amorcé en 2008 par Hurteau et Houle s'intéressant à développer un corpus théorique sur les pratiques évaluatives dans le domaine de l'évaluation de programme. Hurteau et Bossiroy (2009) ont tenté de définir le sens à donner au terme jugement crédible et à cerner les conditions requises pour le générer. Les auteures ont effectué des entrevues auprès d'évaluateurs chevronnés et de clients de l'évaluation. Les conclusions de cette étude nous ont permis d'analyser les conditions nécessaires à la crédibilité du jugement émis, c'est-à-dire la recevabilité et de l'acceptation du dit jugement,

1.5.1 La définition du jugement crédible

Selon les personnes interviewées, un jugement est considéré comme crédible s'il est satisfaisant et raisonnable aux yeux des personnes concernées. Plus encore, il se doit de posséder les caractéristiques suivantes pour être considéré comme tel. Le jugement se doit d'être : 1) argumenté, il appuie son raisonnement sur des arguments solides et nuancés; 2) utile, c'est-à-dire qu'il permet aux parties prenantes de prendre les décisions nécessaires quant aux orientations futures du programme; 3) explicite, afin qu'il soit clair et dénué d'ambiguïté pour les personnes recevant le rapport. Les caractéristiques dénotées ci-haut permettent de justifier la crédibilité du jugement en s'appuyant sur la satisfaction des parties prenantes et d'assurer sa recevabilité.

1.5.2 Les conditions essentielles en vue de générer un jugement crédible

Les répondants identifient quatre conditions essentielles contribuant à générer un jugement crédible : la rigueur méthodologique, la qualité de l'information recueillie par la démarche évaluative, la représentation des différents points de vue et la solidité de l'argumentation.

Cette dernière condition requiert une harmonisation des informations recueillies en vue de produire un raisonnement rigoureux qui supporte le jugement généré. Cependant, même si tous les répondants s'accordent sur l'importance de cette opération, ils soutiennent également qu'ils sont abandonnés à eux-mêmes puisqu'il n'existe aucun cadre afin de baliser la démarche pour générer un jugement crédible et que cette dernière est laissée à l'initiative des évaluateurs, ce qui ne crée pas toujours de bons résultats.

Enfin, il ressort de cette étude que toutes les personnes interrogées s'entendent sur l'importance d'argumenter le jugement par une démarche solide, mais tous s'accordent également sur le fait qu'on ne la comprend pas et qu'elle semble relever de l'initiative de chaque évaluateur.

1.5.3 Le problème : Comment développer une argumentation pour appuyer le jugement crédible.

Même si de nombreux efforts ont été déployés au cours des trente dernières années afin de développer des modélisations en vue de générer un jugement rigoureux (Scriven; 1980, 1990, 1995, Hurteau; 1991, Fournier; 1995, Stake; 2004, Arens; 2006, Hurteau, Lachapelle et Houle; 2006), aucune ne s'est avérée efficace (Stake et

coll., 1997; Arens, 2006; Hurteau, Houle et Mongiat, 2009). À cet égard, ce constat est pour Scriven (1990) le talon d'Achille de la pratique évaluative.

En ce qui concerne l'étude exploratoire de Hurteau et Bossiroy (2009), bien qu'elle ait innové par des données probantes en identifiant les éléments essentiels à la crédibilité du jugement émis et en offrant une piste de solution intéressante en suggérant que le jugement ne soit plus considéré comme un produit, mais aussi comme une compétence, il reste encore à découvrir le processus sous-jacent au jugement afin de générer une argumentation solide et crédible. D'autres parts, le processus argumentatif de l'évaluateur dans le but de générer un jugement est une perspective qui n'a jamais été étudiée jusqu'à présent. C'est donc dans cette perspective que les processus associés au développement du jugement crédible seront abordés dans la présente étude.

1.6 La question de recherche

Comme il a été possible de le constater dans les lignes précédentes, même si la pratique actuelle de l'évaluation de programme est de plus en plus utilisée dans notre société, cette dernière fait face à de nombreuses critiques en ce qui concerne sa fonction première, celle de générer un jugement fondé. Ainsi, il demeure essentiel de se pencher sur la problématique du jugement crédible au sein de la démarche évaluative.

L'ensemble des études portant sur la production d'un jugement s'entend sur le fait que la production d'un jugement n'est pas uniquement le fruit d'une démarche

logique, mais il est plus complexe, et intègre d'autres éléments. De plus, comme le souligne Arens, l'étude des rapports est insuffisante. Il faut interroger les gens.

La recherche de Hurteau et Bossiroy (2009) s'inscrit dans cette lignée. Elle a commencé à identifier des éléments et il en ressort que la mise en commun de l'information pour générer un jugement argumenté s'avère essentielle, mais n'est pas encore vraiment comprise. Elle semble être laissée à l'initiative des évaluateurs. Il est donc nécessaire d'approfondir cet aspect et c'est dans cette perspective que la présente recherche exploratoire, s'inscrivant dans le troisième volet d'un programme de recherche amorcé en 2005 par Hurteau et Houle portant sur l'étude des pratiques évaluatives, visera à répondre à la question suivante :

Quels processus sous-tendent la formulation d'un jugement crédible dans le cadre de l'évaluation de programme?

1.7 La pertinence de l'étude

Tenant compte de l'importance pour une recherche de contribuer à l'avancement des connaissances dans un milieu donné ainsi que de participer à l'amélioration de la société, la pertinence scientifique et la pertinence sociale concluront ce chapitre de la problématique pour faire place au cadre théorique.

1.7.1 La pertinence scientifique

Même si l'évaluation de programme a acquis ses lettres de noblesse grâce à de nombreuses contributions spécifiques dans le domaine, encore aujourd'hui, peu d'études se sont intéressées à la praxie en évaluation de programme (Hurteau,

Lachapelle et Houle, 2006) et à son développement théorique (Alkin, 2003). En effet, les écrits en évaluation de programme s'intéressent aux croyances, témoignages et anecdotes plutôt qu'aux données empiriques et le peu d'études théoriques existantes s'intéresse à l'instrumentation ou à la méthodologie (Arens, 2006; Smith, 1993; King, 2003). Ainsi, cette recherche permettra de contribuer au développement théorique d'un aspect encore peu exploité en évaluation de programme, la crédibilité du jugement émis et ainsi solidifier la discipline. Sur ce, Datta (2006) et Stern (2006) affirment qu'il serait important de diriger les nouvelles recherches en évaluation de programme vers la problématique impliquant la crédibilité du jugement émis.

De plus, même si la présente recherche s'intéresse davantage à l'avancement de la discipline de l'évaluation de programme, elle pourra tout de même contribuer à bonifier les théories actuelles dans différents domaines où les praticiens sont appelés à émettre un jugement professionnel.

1.7.2 La pertinence sociale

Au cours des dernières années, les gouvernements ont de plus en plus fait appel à l'évaluation de programme afin d'éclairer et de guider les orientations de divers programmes, projets et politiques sociales. De ce fait, il est donc primordial que l'évaluation de programme puisse s'appuyer sur un cadre assez rigoureux pour assurer un jugement crédible et fondé. Plus précisément, les résultats de l'évaluation de programme se doivent d'être utilisables et fiables afin de répondre à la question qui avait alors démarré l'évaluation et éviter de fermer un programme dont les résultats étaient bénéfiques aux participants ou encore de garder un programme n'ayant aucun effet.

Le jugement crédible en évaluation étant un aspect primordial, cette recherche permettra de soutenir théoriquement le domaine et ainsi participer au développement des connaissances permettant aux évaluateurs de faire des évaluations de plus grande qualité. Sur ce, elle permettra également de règlementer adéquatement la pratique lors de l'accréditation des membres par *l'association canadienne d'évaluation* et ainsi rassurer les parties prenantes et les clients de l'évaluation.

De plus, étant d'actualité depuis plusieurs années, la réforme des programmes scolaires de l'école québécoise a incité la population à s'intéresser aux résultats des évaluations de programmes. Il est donc capital que les jugements émis soient crédibles afin de justifier les changements. Dans le même ordre d'idée, Houle (2006, p.2) conclut :

[...] l'application de l'évaluation de programme au secteur de l'éducation ne pourra que bénéficier des éventuelles percées en évaluation de programme; cette perspective est d'autant plus importante que ce secteur y a de plus en plus recours, tant pour encadrer ses activités que pour appuyer les prises de décisions.

CHAPITRE II

CADRE THÉORIQUE

Le deuxième chapitre est destiné à dresser un portrait des principaux concepts en jeu dans cette recherche ainsi que des théories et modèles auxquels ils font référence. Plus spécifiquement, le concept du jugement sera défini en présentant diverses définitions, modèles et théories concernant le jugement comme produit et le jugement comme compétence. Puis, différents concepts touchant le jugement crédible et son argumentation seront développés afin d'apporter des pistes de réponse à la question de recherche formulée plus tôt. Enfin, une synthèse des recherches, permettant la formulation des objectifs de recherche, sera apportée.

2.1 Introduction

Fitzpatrick (2009, p.153) confirme le statut du jugement dans l'évaluation de programme en énonçant que : « [...] *Evaluation is fundamentally about judging the value of something* ». Dans le chapitre précédent, il a été conclu que l'ensemble des études portant sur la production d'un jugement en évaluation s'entend sur le fait que la production d'un jugement n'est pas uniquement le fruit d'une démarche logique, mais plus complexe, qui intègre d'autres éléments. C'est donc sur ces éléments que le présent chapitre s'attardera afin de mieux cerner la problématique et d'élaborer des objectifs de recherche.

2.2 La notion de jugement

Différents auteurs, de différentes disciplines, se sont intéressés à la question du jugement et en ont proposé diverses définitions et théorisations. Bien que le concept du jugement soit assez commun, il importe de bien le situer et de le définir de façon adéquate afin d'en avoir une compréhension complète et singulière. Il est possible de retracer deux grands courants de pensée lorsque l'on s'intéresse au jugement, un premier philosophique et le second comme une compétence.

2.2.1 Le courant philosophique

La philosophie étudie le jugement en le considérant comme un acte ou comme une aptitude. Le premier désigne le jugement comme un acte d'entendement et le deuxième comme une sagesse.

En ce qui concerne le jugement comme acte d'entendement, Reboul (1992), dans le collectif *La formation du jugement, peut-on apprendre le jugement*, questionne le sens du mot jugement et vise à connaître de quoi on parle lorsqu'on en fait référence. Inspirée du philosophe Descartes, et intellectualiste, la première définition du jugement le considère comme une « décision mentale par laquelle nous arrêtons d'une façon réfléchie le contenu d'une assertion et la posons à titre de vérité ». C'est par le contenu de cette décision mentale qu'il sera possible d'établir des affirmations et c'est par l'assertion qu'il sera possible de décider du bien-fondé de cette décision mentale.

Le jugement comme une aptitude qui se définit comme une sagesse

Cette deuxième définition du jugement (Reboul, 1992) considère le jugement comme une aptitude, une compétence, comme le fait d'être capable de bien juger. C'est sur cette dernière que cette recherche s'appuiera lorsqu'elle fera référence au jugement. Le jugement, prenant la forme d'une sagesse pratique, réfère à la capacité d'adopter, après réflexion, un verdict ou une décision. Cette réflexion, ou délibération est essentielle afin de ne pas tomber dans le hasard ou l'arbitraire. Quintilien (Reboul, 1992), quant à lui, définissait le jugement comme une aptitude propre à l'avocat qui ne se référerait qu'à une question d'habileté. Enfin, Aristote (Reboul, 1992) voyait dans cette sagesse pratique (*phronèsis*) non pas les méthodes et solutions visant à la production d'un jugement, mais plutôt le résultat, la virtuosité et le bonheur, et c'est pour cette raison que la compétence à juger appartient à l'éthique et est une caractéristique de la vertu.

En bref, la seconde définition, « *qui fait du jugement l'aptitude à se faire une opinion là où manque la connaissance certaine et où la délibération ne donne que du probable* », attribuée au jugement, particulièrement en tant que finalité première de l'enseignement, cinq traits indispensables (Reboul, p.25):

1. Le jugement se situe dans le monde du contingent et de l'incertain, qui n'est pas pour autant livré au hasard et au chaos : le monde des relations humaines.
2. Le jugement dépend donc de l'objet, mais aussi du juge; il est une décision qui crée une solution qui n'était écrite nulle part et qui, une fois trouvée, transforme le cours des choses; un verdict judiciaire, une critique de livre, un jugement sur autrui ou sur soi-même sont toujours quelque peu une *self-fulfilling prediction*.
3. C'est pourquoi le jugement engage ou compromet celui qui le porte; une erreur judiciaire, une critique aveugle ont des conséquences pratiques.

4. Ces deux exemples montrent que la question du vrai n'est pas absente du jugement ainsi défini. Même s'il est une décision, elle doit être motivée, et sera ensuite, d'après l'événement, estimée juste ou fausse.

5. Enfin, du moment qu'il engage la responsabilité de celui qui le porte, le jugement n'est pas seulement d'ordre technique, il est d'ordre moral; un mauvais diagnostic, une mauvaise expertise sont non seulement des erreurs, mais des fautes.

Dans le même ordre d'idées, Anger (1992) réfléchit à la formation du jugement en déclarant : « Un jugement équilibré et réaliste affirme les choses comme elles sont, mais il atteste en même temps le caractère pondéré et raisonnable du sujet » tout en tenant compte que le jugement suppose de « peser les éléments de preuve ». Le jugement rendu doit donc s'appuyer sur des raisons et des évidences suffisantes lorsqu'il est posé, sans cela, il sera irréfléchi. Il est donc vrai de dire que le jugement provoque et divulgue la dimension rationnelle de l'être humain. De plus, l'auteur souligne qu'émettre un jugement, c'est aussi en assumer la responsabilité. Anger (1992) prétend également qu'il est primordial d'ancrer le jugement à un processus cognitif puisque ce dernier ne se manifeste pas comme une opération séparée de l'activité cognitive. Au contraire, le jugement ferait partie d'un mécanisme comportant plusieurs niveaux tels que les perceptions, les questionnements, les intuitions et les conceptualisations.

De plus, Anger (1992) a relevé quatre niveaux d'opération menant à l'élaboration du jugement : le niveau de l'expérience, de la compréhension, de la réflexion et de la décision. Le niveau de l'expérience réfère aux sens externes (voir, toucher, sentir, goûter) et aux sens internes (éprouver, ressentir, souffrir, se réjouir). Il est possible de se préoccuper de ces sens ou non. En y prêtant attention, ces derniers susciteront des questionnements qui feront investiguer la raison de leur présence. Ainsi, l'interrogation posée entraînera le passage au second niveau. Le deuxième niveau, la

compréhension, contient les opérations de l'intelligence (chercher, intuitionner, conceptualiser, nommer, formuler) qui permettent le déclenchement de questionnements permettant à l'intuition de comprendre les liens entre les informations émergentes. Complétant le processus de la compréhension, le troisième niveau renvoie à la réflexion et au jugement. Ce niveau consolide les informations retenues en s'assurant que ces dernières sont exactes, véridiques. C'est la quête de la vérité qui conduira au jugement. Somme toute, le dernier niveau s'approprie les étapes précédentes afin de poser l'action finale et en assumer la responsabilité. De ce processus cognitif, Anger (1992) signale l'interdépendance des niveaux cognitifs nécessaires afin de prendre une décision responsable.

D'autre part, Lafortune (2006) définit le jugement professionnel comme un mécanisme permettant de faire un choix en se fiant à son expertise professionnelle et personnelle (expérience et formation). Il va de soi que ce processus demande une certaine *rigueur, cohérence et transparence*. Si Anger (1992) considérait le jugement comme un produit, Lafortune (2006) le considère comme un produit entier et avance l'idée d'un jugement comme une compétence. Au demeurant, dans son essai, *Everyday Practice of Science, Where Intuition and Passion Meet Objectivity and Logic*, Grinnell (2009) expose les propos du philosophe Thomas Kuhn qui décrit les paradigmes en sciences comme un ensemble de croyances et de valeurs partagées par les membres de la communauté scientifique comme une façon acceptable et établie de résoudre les problèmes. Il met également l'emphase sur le jugement scientifique qui dépend de l'expérience personnelle et de la personnalité du chercheur.

2.2.2. Le second courant : une compétence

Comme il a été possible de le constater dans le point précédent, plusieurs auteurs considèrent le jugement comme une compétence, c'est-à-dire comme une habileté ou une aptitude à prendre des décisions justes et éclairées, mais imprévisibles (Reboul, 1992). Lafortune et Allal (2008, p.22) considèrent aussi le jugement comme une compétence en le définissant ainsi :

[...] un processus qui mène à une prise de décision, laquelle prend en compte différentes considérations issues de son expertise (expérience et formation) professionnelle. Ce processus exige rigueur, cohérence et transparence. En ce sens, il suppose des moyens en lien avec les visées ou intentions et le partage des résultats de la démarche dans une perspective de régulation.

En ce sens, étant orienté vers la prise de décision et qu'une action devrait en découler, le jugement peut ainsi être considéré comme une compétence.

À son tour, Schleifer (1992) l'entrevoit comme possédant des caractéristiques d'incertitude et d'imprévisibilité, mais pas de hasard. Dans le même ordre d'idée, Lipman (1992) indique que le jugement est une représentation de celui qui le pose puisqu'il reflète la vision du monde de ce dernier et son impression d'une réalité ou d'une situation. Sur ce, Schleifer (1992), soulève l'idée qu'il serait impossible de programmer une machine capable de jugement, puisque l'acte de juger relève de l'incertitude, qui elle, est propre à l'être humain.

Traditionnellement, le jugement en évaluation de programme a toujours été envisagé comme un produit d'une démarche logique qui s'appuie sur des évidences crédibles (Donaldson, Christie et Mark, 2009). En le dégageant de cette perspective

réductionniste et en le considérant maintenant comme une compétence, il fait maintenant appel à un ensemble de connaissances et de ressources chez l'évaluateur.

2.2.3 La compétence à juger

Tel que vu précédemment, plusieurs auteurs considèrent le jugement comme une compétence. Il importe donc de définir cette notion afin d'en saisir le sens et tous ses aspects. Boudreault (2002) a proposé une définition complète de la compétence qui s'applique tout à fait à l'évaluation de programme. En effet, pour cette auteure, la compétence professionnelle se manifeste par le regroupement du savoir, savoir-faire, savoir-être et du contexte signifiant identifiable dans les programmes ou dans la vie de tous les jours. Elle expose également son raisonnement par le schéma suivant :

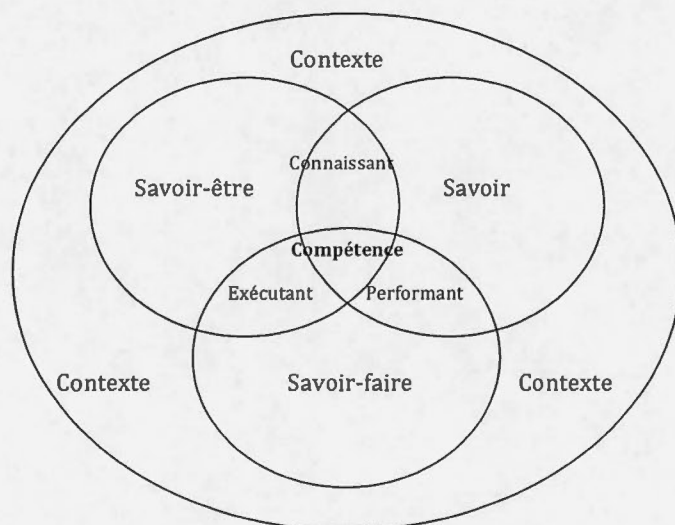


Figure 2.1 Schéma de la compétence professionnelle (tirer de Boudreault, 2002, p.10)

Par ce schéma, Boudreault (2002) confirme qu'il est maintenant difficile de percevoir le jugement comme un simple produit d'une démarche. Il importe de tenir compte de toutes les composantes pour en admirer sa complexité.

2.3 Les processus permettant de générer un jugement : Les modélisations

Le jugement en évaluation de programme s'appuie sur un ensemble d'opinions indépendantes et sur des évidences à propos d'un objet (Schwantd, 2002). La procédure sous-jacente au jugement implique un raisonnement et une argumentation qui se basent sur des conclusions évaluatives justifiables, précises, légitimes, acceptables et valides (House, 1993; Fournier et Smith, 1993; Arens, 2006). Ce sont ces derniers aspects qui permettent aux parties prenantes d'accepter le jugement et de ce fait, lui accorder une crédibilité.

Plusieurs auteurs ont tenté de théoriser le processus sous-jacent la production d'un jugement en le percevant comme le résultat final d'une démarche singulière mettant l'accent sur des étapes précises à suivre. C'est donc dans cet esprit de conceptualisation du jugement que plusieurs modèles ont été élaborés afin de fournir aux professionnels de l'évaluation un cadre rigoureux visant à baliser la production du jugement rendu. C'est dans cette optique que seront présentés ci-dessous certains modèles des plus influents.

2.3.1. La logique de l'évaluation (Scriven, 1980)

Scriven (1981), par sa conception de la logique de l'évaluation, offre une métathéorie qui place le jugement de valeur au cœur de la démarche d'évaluation. La logique de l'évaluation de Scriven (1980) a fermement participé à développer l'acte spécifique d'évaluer au moyen des 4 opérations ci-dessous :

1. Établir les critères (les éléments qui seront évalués);
2. Élaborer les standards (le point de coupure entre un jugement favorable et un jugement défavorable);
3. Mesurer la performance et la mettre en relation avec les standards;
4. Synthétiser et intégrer les données pour ensuite porter un jugement;

Les travaux de Scriven (1980) ont introduit les assises théoriques de l'évaluation de programme. Les éléments ci-dessus constituent l'acte d'évaluer lui-même lorsqu'ils sont pris logiquement. Par le fait même, il permet d'accorder une valeur par l'attribution d'une note, d'un rang, d'un indicateur de proportion ou d'un résultat à l'objet d'évaluation. En 1990, Scriven complètera son modèle de la pratique évaluative en dressant une double pyramide (voir fig. 1.3). La première pyramide expose le processus analytique qui sert à identifier les critères, les indicateurs et les données. La deuxième pyramide, renversée, permet d'arriver à un jugement à partir des inférences et synthèses des informations cumulées plus tôt.

2.3.2 La logique de travail de Fournier (1995)

Sous le nom de la *logique du travail*, Fournier ajoute de nouveaux éléments, plus pratiques et plus opérationnels, en contexte d'évaluation de programme. Contrairement à la logique générale ayant inspiré la logique de l'évaluation de Scriven (1980), ceux-ci seraient spécifiques à l'évaluation de programme, et seraient utilisés en alternance avec les éléments proposés par Scriven:

1. Déterminer les besoins des gestionnaires qui recevront l'évaluation;
2. Cerner avec précision l'objet à l'étude selon les besoins identifiés;
3. Formuler une question d'évaluation qui prend en considération tous ces éléments.

Or, par l'introduction de cette nouvelle logique, Fournier (1995) a contribué à développer le processus spécifique à l'évaluation de programme par l'ajout de ces nouveaux éléments.

2.3.3 La contribution de Stake (2004): la pensée interprétative

Tout comme les auteurs précédents, Stake (2004) a enrichi le processus spécifique à l'évaluation de programme en mettant de l'avant la théorie liée à la *pensée interprétative*. Cette théorie renferme deux types de pensées évaluatives utilisées conjointement, la *pensée critériée* et la *pensée interprétative*.

Selon Stake (2004), la *pensée critériée* représente l'objectivité de l'évaluation qui fait notamment référence à la rationalité, à la rigueur et au raisonnement logique. Elle comporte l'identification de critères et leur mise en relation avec des standards afin de porter un jugement et fait référence à l'ensemble des éléments apporté par Scriven (1980) dans *la logique de l'évaluation*. La *pensée critériée* permet également de mesurer les variables étudiées le plus précisément possible et ainsi d'envisager des conclusions crédibles. La *pensée interprétative* quant à elle représente l'opposé de la *pensée critériée* et fait ainsi référence à la subjectivité de l'évaluateur et des personnes impliquées, de près ou de loin, dans le programme à l'étude. La *pensée interprétative* fera surface lorsque l'évaluateur devra émettre un jugement sur la valeur de l'objet d'évaluation par des critères et des standards. L'évaluateur sera en mesure de justifier sa propre subjectivité s'il possède une grande connaissance du contexte de son programme et de son évaluation. Finalement, Stake (2004), par l'élaboration de sa théorie, comme plusieurs auteurs présentés dans ce chapitre, a grandement contribué au développement du processus spécifique à l'évaluation de programme et à l'avancée de la discipline en général.

2.3.4 La modélisation du processus spécifique à l'évaluation de programme (Hurteau, Lachapelle et Houle, 2006)

La participation au processus spécifique à l'évaluation de Hurteau et Houle (2006) est colossale puisqu'ils ont réuni l'ensemble des contributions précédentes dans le domaine (Scriven, 1980; Hurteau, 2001; Fournier, 1995; Stake, 2004; et Arens, 2006) selon une anasynthèse afin d'élaborer une modélisation du processus d'évaluation, soit l'acte spécifique d'évaluer. La modélisation proposée s'effectue à l'aide de six étapes distinctes : 1) description du programme 2) problématique de l'évaluation 3) opérationnalisation de l'évaluation 4) choix stratégique 5) déclaration évaluative et

synthèse 6) jugement. Cette modélisation a été validée dans le cadre de la première phase du programme de recherche de Hurteau et Houle, 2005, et a fait l'objet de publication scientifique.

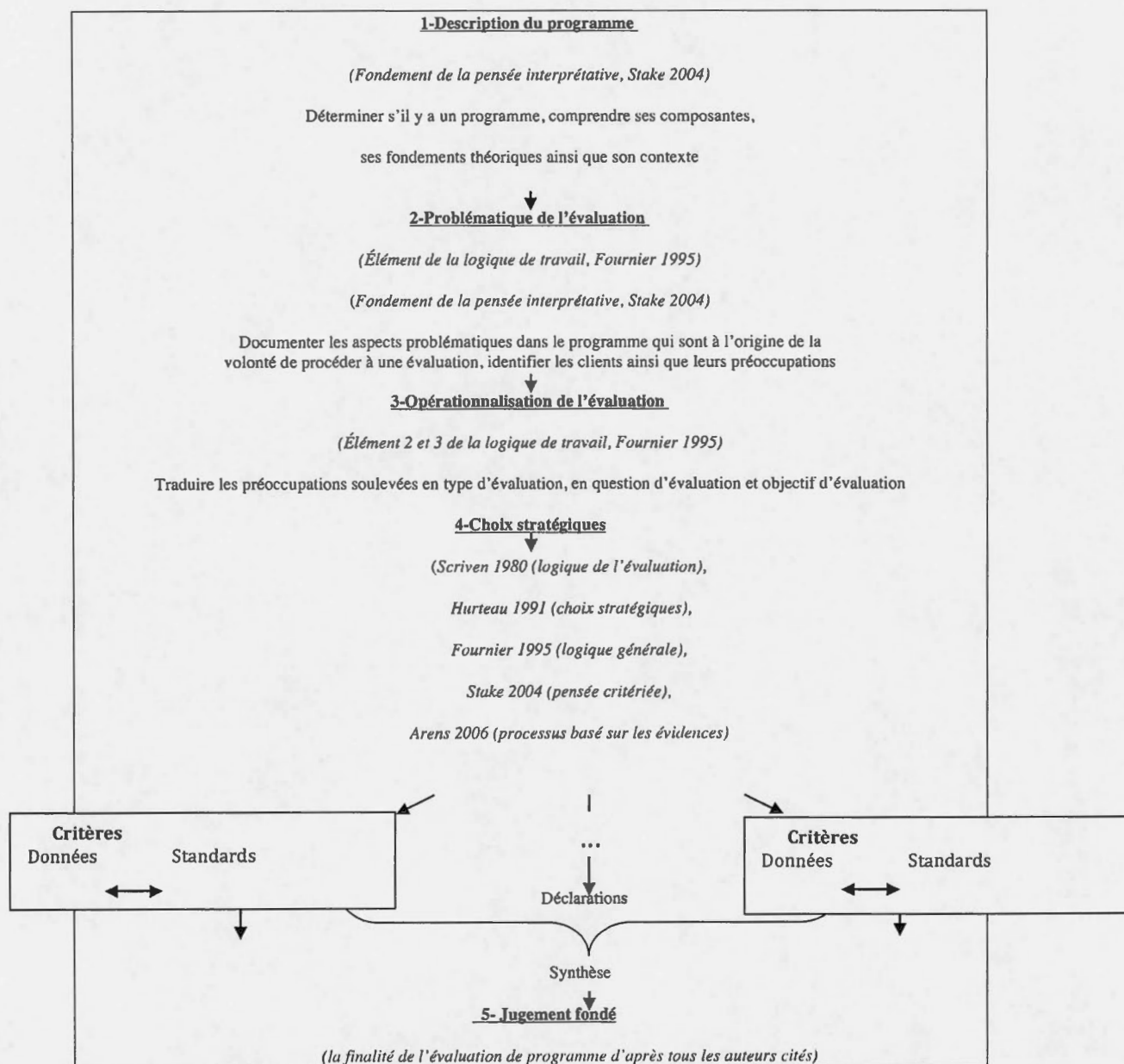


Figure 2.2 La modélisation du processus spécifique à l'évaluation de programme et les fondements théoriques de ses composantes (Hurteau et Houle, 2006)

2.3.5 Le raisonnement évaluatif (Arens, 2006)

Afin de produire une conclusion évaluative acceptable, l'évaluateur devra suivre un raisonnement évaluatif. C'est au processus sous-jacent ce raisonnement évaluatif que Arens (2006) s'est particulièrement intéressée puisque c'est durant ce raisonnement que l'évaluateur devra forger un jugement, jugement qui est au cœur même de la démarche évaluative. Arens (2006) avance que ce raisonnement doit se fonder sur des conclusions valides provoquées par un processus logique, soit la production d'un jugement et son argumentation (voir fig. 2.3).

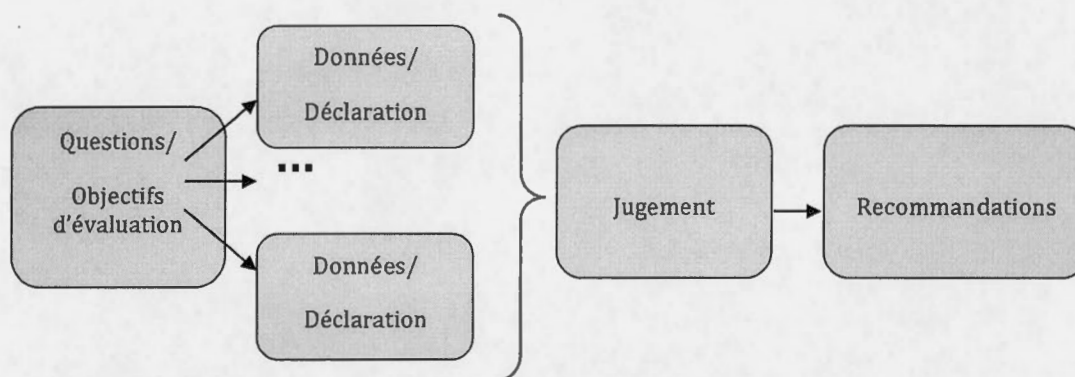


Figure 2.3 La démarche sous-jacente au raisonnement évaluatif : l'élaboration du jugement et de son argumentation. (Tirée d'Arens, 2006).

2.4 Les processus permettant de générer un jugement crédible

La prochaine section portera dans un premier temps sur la notion de la crédibilité et à son application à l'évaluation de programme. Cette section aidera à mieux comprendre cette notion qui ne s'inscrit pas dans la perspective d'un critère de subjectivité. À cet effet, la notion de crédibilité vue par Grinnell (2009) sera présentée

en détail pour ensuite faire place à la transformation de ce modèle vers l'évaluation de programme.

2.4.1 La notion de crédibilité : sa définition et son développement

Le terme crédibilité fait référence à une situation où l'on s'interroge sur la véracité d'une affirmation, d'un document ou d'une donnée auquel on a accès ou non.

Une information, une donnée, un « dire », un document, une statistique, une expertise, etc. sera jugé plus ou moins crédible selon l'idée que le récepteur se fait du niveau de véracité et/ou de vérifiabilité, de preuves de sa source et de son vecteur, et par extension, du niveau général de confiance qu'il place dans la personne, l'institution ou l'entité source ou vecteur de l'information (Guba, 2004).

La théorie de Grinnell offre un nouveau regard sur la rationalisation de la crédibilité dans le domaine de la science. En effet, dans son essai *Everyday Practice of Science, Where Intuition and Passion Meet Objectivity and Logic* (2009), Grinnell expose un modèle où il entrevoit la crédibilité comme un processus externe au chercheur et où la crédibilité apparaît à la suite d'une conversation interactive entre le monde à étudier (world to be studied) et la communauté scientifique (research community). C'est lors de cette conversation que le chercheur tentera de persuader ses pairs de la convenance et de la nouveauté de sa découverte, qui lui concéderont une crédibilité : « *The former conversation gives rise to the circle of discovery-learning new things. The latter gives rise to the circle of credibility-trying to convince others that new findings are correct. These conversations are dialogs that proceed in an interactive manner* » (Grinnell, 2009, p.5).

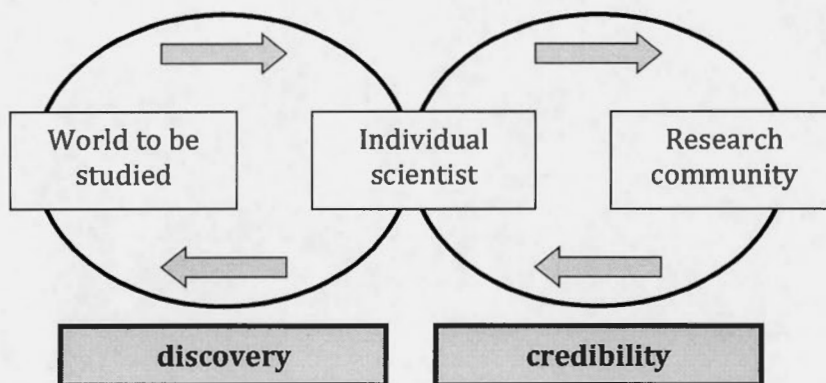


Figure 2.4 Schéma de la crédibilité (tiré de Grinnell, 2009, p.5)

C'est ainsi que ces deux dimensions, la découverte permettant de générer des résultats inédits, et la crédibilité permettant de tenter de convaincre ses pairs de la valeur de ses résultats, s'inscrivent au sein d'un dialogue constant. Ce nouveau modèle innove en introduisant une nouvelle dimension à la crédibilité qui n'avait jamais alors été explorée. Alors que traditionnellement les chercheurs associaient la crédibilité à la rigueur méthodologique, Grinnell repousse les limites en introduisant de nouveaux concepts qui assurent cette crédibilité.

2.4.1.2 Le processus de « crédibilisation » d'une découverte (Hurteau, Houle et Guillemette, 2012).

Hurteau et al. (2012) mentionnent l'importance et introduisent le processus de « crédibilisation d'une découverte » en proposant le schémas ci-dessous et en affirmant que : « le chercheur doit adopter une tout à fait contraire à celle qui consiste à se réfugier dans ses propres cadres de références et ses propres croyances, parce qu'un tel comportement risque d'entraîner un échec, sans compter que cela pose un problème de nature éthique ».

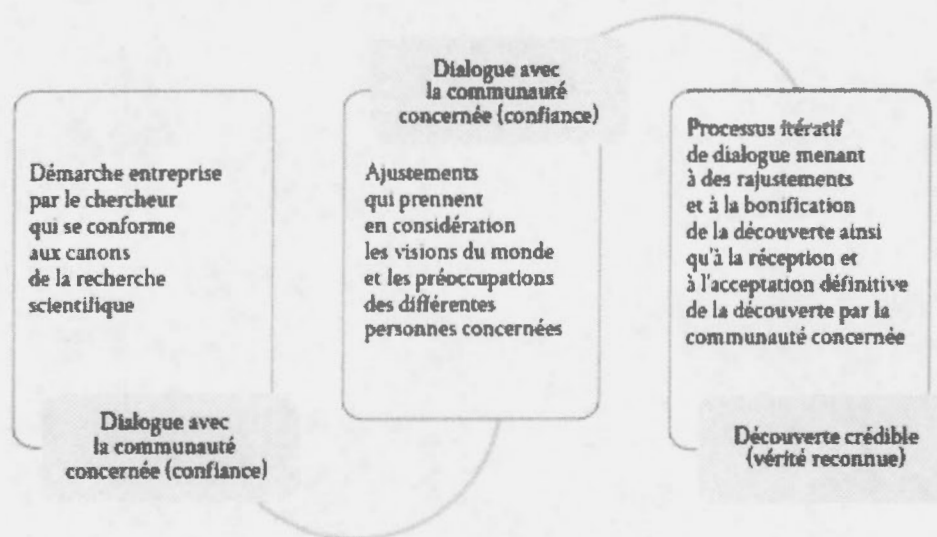


Figure 2.5 Schéma du processus de « crédibilisation » d'une découverte (tiré de Hurteau, Houle et Guillemette, 2012, p. 73)

2.4.2 La notion de crédibilité appliquée à l'évaluation de programme. Comment s'applique-t-elle au jugement?

Tout comme mentionné au premier chapitre, la notion de crédibilité est incluse dans la définition de Champagne, Contandriopoulos, Brouselle, Hartz et Denis (2009, p. 38) qui l'ont décrite ainsi :

L'évaluation consiste fondamentalement à porter un jugement de valeur sur une intervention en mettant en œuvre un dispositif capable de fournir des informations scientifiquement valides et socialement légitimes sur cette intervention ou sur n'importe laquelle de ses composantes, l'objectif étant de faire en sorte que les différents acteurs concernés, dont les champs de jugement sont parfois différents, soient en mesure de prendre position sur l'intervention pour qu'ils puissent construire, individuellement ou collectivement, un jugement susceptible de se traduire en actions.

Tout au long de son évolution au sein de la pratique évaluative, le jugement a toujours été envisagé comme le produit d'une démarche logique s'appuyant essentiellement sur des données probantes (Donaldson et coll. 2009). Hurteau (2009), inspirée de Green (dans Fitzpatrick, Christie & Mark, 2009), a transposé à l'évaluation de programme le schéma de Grinnell qui prend la forme ci-dessous:

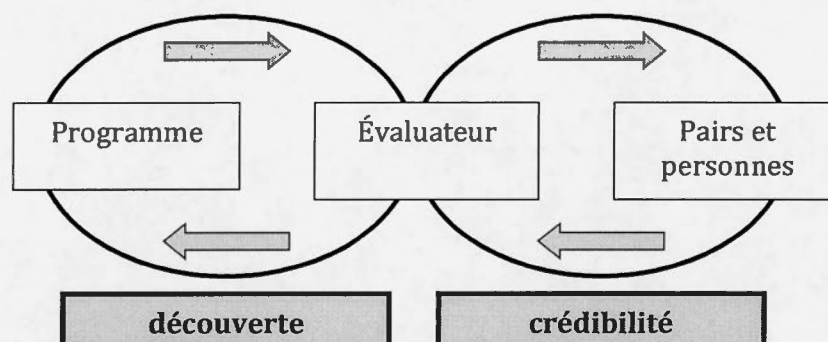


Figure 2.6 Schéma de la crédibilité en évaluation de programme (tiré de Hurteau, 2009)

Ainsi, l'évaluation de programme est perçue comme une démarche basée sur une conversation interactive continue avec les personnes concernées, portant sur les enjeux du programme, visant à décrire le programme et à porter un jugement (découverte), de façon à ce que les personnes concernées puissent avoir confiance en l'information ainsi qu'au jugement généré (crédibilité) (Hurteau, 2009).

Or, pour Hurteau et al. (2012), le jugement doit maintenant être envisagé « comme le fruit d'une négociation entre l'évaluateur et les différentes parties prenantes dans le but de dégager une vision intégrée du programme dont la facture se rapproche le plus possible de la réalité ». En effet, pour les auteurs, la modélisation présentée à la figure 2.4 offre une option intéressante puisqu'elle fournit une attention particulière au dialogue en énonçant que « celui-ci favorise le déploiement graduel de la découverte par un va-et-vient d'échanges entre les différents acteurs ». Dans le contexte de l'évaluation de programme, les auteurs ont présenté la modélisation qui suit :

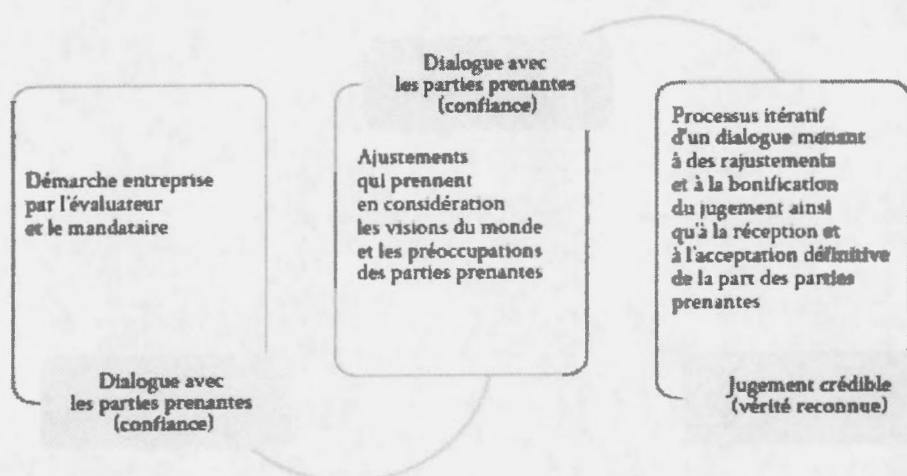


Figure 2.7 Schéma du processus de « crédibilisation » du jugement (tiré de Hurteau, Houle et Guillemette, 2012, p. 83)

2.5 Synthèse et objectifs de recherche

Les lignes précédentes ont permis de concevoir le jugement non plus comme un produit, mais comme une compétence, ce qui implique plusieurs éléments complexes. Malheureusement, ces différentes contributions s'avèrent peu ou pas opérationnelles (Stake, Davis, Cisneros, DePaul, Dunbar, Famer, Feltovich, Johnson, Williams, Zurita et Chaves, 1997; Hurteau, Houle et Mongiat, 2009) et les théoriciens ne sont pas en mesure d'offrir une alternative au constat de Boissiroy (2009) à l'effet qu'il n'y a pas d'argumentation pour soutenir leur jugement.

Comme il a été possible de le constater à la lecture du cadre théorique, l'évaluation de programme n'est présentement pas en mesure d'émettre un jugement suffisamment fondé pour le rendre acceptable aux yeux des clients et des utilisateurs de

l'évaluation. Il est donc fondamental d'aller recueillir l'information nous permettant de poser un jugement fondé et crédible. Puisque l'objet de cette recherche porte sur le jugement, l'information sera recueillie auprès de professionnels, qui par leur reconnaissance sociale, tend à confirmer qu'il est possible de poser un jugement professionnel crédible.

Un des problèmes majeurs que rencontre l'évaluation de programme est que les chercheurs ont toujours tenté de comprendre le jugement à partir de l'exercice de sa fonction. C'est dans cette perspective que nous avons décidé d'opter pour une approche originale afin de répondre à la question initiale Or, plusieurs professionnels, de différents domaines, pour ne nommer que les juges, les enquêteurs et les psychologues praticiens, arrivent à émettre un jugement crédible et potentiellement acceptable socialement par leur reconnaissance professionnelle.

À la question : *Quels processus sous-tendent la formulation d'un jugement crédible dans le cadre de l'évaluation de programme?* nous envisageons le jugement comme une compétence qui serait transversale à différents actes professionnels et transférable à la fonction évaluative. C'est dans cette perspective que le programme de recherche subventionné de Hurteau et Houle vise à :

1. Cerner des pratiques professionnelles qui conduisent à un jugement professionnel
2. Documenter le processus vécu par les personnes
3. Induire une éventuelle modélisation à partir de ces témoignages

Le projet s'inscrit dans une perspective exploratoire et utilise la technique de l'entrevue en vue de collecter des informations sur le jugement émis par des professionnels de différents domaines dans le cadre de leur pratique. Un questionnaire descriptif a été validé (Marchand, 2011) dans le but de s'assurer une information de qualité. C'est dans cette lancée que l'étude présentée, visant les objectifs énoncés ci-haut, vise l'objectif spécifique suivant :

Puisqu'aucune solution n'a encore été trouvée acceptable pour le moment, nous proposons d'apporter un regard nouveau sur le jugement produit lors de l'évaluation de programme en l'observant dans une variété de disciplines afin de déclencher une compréhension appuyée par des preuves.

En adoptant un nouveau regard sur le phénomène, les chercheurs ont supposé que le jugement n'est pas le résultat exclusif au processus formel d'évaluation. Stake (2011) a soutenu ce point de vue lors de sa présentation à la conférence annuelle de la Société canadienne d'évaluation en soulignant que le jugement est présent aussi bien dans la loi (évaluation de la légalité), la médecine (évaluation de la santé), l'éducation (évaluation de l'enseignement et de l'apprentissage), le journalisme (journalisme d'investigation), les entreprises (évaluation de la gestion et de profit), les organismes de réglementation (évaluation de la sécurité et le code de l'adhésion), l'intelligence (évaluer la sécurité et le code de l'adhésion) et le militaire (évaluation de la stratégie, de la tactique, de la menace et des risques).

CHAPITRE III

DÉMARCHE MÉTHODOLOGIQUE

Le troisième chapitre de ce mémoire révèle en détail la méthodologie utilisée afin de répondre aux objectifs exposés dans le cadre théorique. En premier lieu, le type de recherche effectuée sera présenté sommairement. Ensuite, les choix méthodologiques concernant l'échantillon et la méthode de collecte de données seront exposés puis suivront l'analyse et l'interprétation des données. En dernier lieu, les considérations éthiques ainsi que les forces et limites de l'étude concluront cette section.

3.1 Le type de recherche

Les objectifs spécifiques appuyant la présente recherche étant les suivants :

1. Cerner des pratiques professionnelles qui conduisent à un jugement professionnel
2. Documenter le processus vécu par les personnes
3. Induire une éventuelle modélisation à partir de ces témoignages

l'approche de cette recherche s'inscrit dans une perspective qualitative fondamentale de type exploratoire et s'appuie sur la théorie ancrée « *Grounded Theory* ». En effet, la présente recherche vise la production de nouvelles connaissances en ce qui concerne la nature du jugement crédible tout en se situant dans un contexte de découverte puisque cette dernière s'inscrit dans une démarche heuristique et inédite (Lessard-

Hébert, Goyette et Boutin, 1990). Stake (2010) définit l'essence de l'approche qualitative comme étant interprétative, basée sur des expériences, situationnelle et personnalisée. Il ajoute que chaque chercheur a sa propre façon de concevoir sa recherche, mais que tous travaillent très fort sur l'interprétation des données. Les chercheurs vont essayer de transmettre en termes expérientiels l'histoire de chacun tout en essayant de traiter chaque sujet comme étant unique, mais similaire à d'autres.

L'approche de cette recherche s'inscrit aussi dans un paradigme phénoménologique que Legendre (2005, p.1036) définit comme l' « étude qui a pour but d'identifier les structures fondamentales des phénomènes étudiés en mettant l'accent sur le sens conféré par les personnes à leurs expériences vécues », parce qu'elle est basée sur le vécu et l'expérience des gens au sein de l'exercice de leur fonction respective.

C'est en prenant ses racines dans l'étude exploratoire de Hurteau et Bossiroy (2009) qui ont offert une piste de solution intéressante en considérant le jugement non plus comme un produit, mais aussi comme une compétence que la présente recherche a pour objet d'étude les éléments constituant un jugement crédible ainsi que les conditions requises pour l'obtenir. Comme il a été possible de le constater à la lecture du cadre théorique, l'évaluation de programme n'est présentement pas en mesure d'émettre un jugement suffisamment fondé pour le rendre acceptable aux yeux des clients et des utilisateurs de l'évaluation. Il est donc fondamental d'aller recueillir l'information nous permettant de poser un jugement fondé et crédible. Puisque l'objet de cette recherche porte sur le jugement, l'information sera recueillie auprès de professionnels, qui par leur reconnaissance sociale, tend à confirmer qu'il est possible de poser un jugement professionnel crédible, tels les juges, médecins, psychologues cliniciens, enquêteurs, etc.

En s'inspirant de la théorie ancrée, on oublie temporairement le cadre théorique pour adopter un regard frais et nouveau. L'analyse n'est pas effectuée en fonction du cadre théorique, mais plutôt avec ce qui émerge des données. Par la suite, un examen de concordance avec le cadre théorique est fait pour vérifier si cela concorde ou pas. La théorie ancrée a été introduite en 1967 par Glaser et Strauss qui la définissait comme une méthodologie générale permettant d'engendrer de nouvelles théories en sciences humaines et sociales. Cette approche favorise l'innovation en s'appuyant sur une méthode d'enracinement de l'analyse des données produites sur le terrain. Ainsi, *« non seulement l'analyse prend comme point de départ les premiers épisodes de collecte de données, mais elle se poursuit dans un processus de validation qui consiste à revenir constamment, soit aux données déjà collectées, soit à de nouvelles données. »* (Guillemette, 2006).

3.2 Les choix méthodologiques

La présente section exposera en détail les choix méthodologiques effectués pour la constitution de l'échantillon, la modalité, ainsi que pour la méthode de collecte de données.

3.2.1 L'échantillon

Puisque l'échantillon retenu pour cette recherche doit posséder les caractéristiques suivantes : 1) l'importance du jugement au sein de sa pratique professionnelle; 2) la diversité des secteurs d'intervention (santé, éducation, justice, relation d'aide, comptabilité, gestion d'entreprises, etc.); 3) la diversité des contextes (situation

d'urgence, ou non, nature de l'impact du jugement); 4) la crédibilité professionnelle, la technique d'échantillonnage intentionnel est retenue. De ce fait, en retenant seulement les candidats possédant les caractéristiques voulues, la recherche s'assure de répondre aux critères de sélection quant au choix des participants et de posséder un corpus de données fructueux et riche en informations (Karsenti et Savoie-Zajc, 2004). L'échantillon intentionnel a également été utilisé pour atteindre une saturation théorique.

La recherche principale de Hurteau et al. (2011) Comptait 19 intervenants provenant de plusieurs disciplines, mais pour les besoins du présent exercice, un échantillon de six répondants a été retenu. Le choix des participants a été fait en fonction des différentes professions. Le nombre de 6 participants (n=6) constitue un échantillon suffisant pour un mémoire puisqu'il permet d'aller en profondeur et de lier les expériences afin de faire ressortir des données représentatives. Le choix des entrevues a été fait dans le but d'avoir un éventail le plus large possible de domaine d'intervention. Ainsi, un pompier (n=1), un médecin (n=1), un policier (n=1), un avocat (n=1), un juge (n=1) et un professeur (n=1) constituent l'échantillon de cette recherche.

3.2.2 La méthode de collecte de données

Afin de cerner et comprendre la nature du jugement crédible et les conditions requises pour l'obtenir, les entrevues individuelles semi-dirigées en personne ont été sélectionnées afin de recueillir les données. Cette méthode apparaît comme la plus pertinente puisqu'elle permet aux chercheurs de préserver l'interaction avec les

participants et de créer une situation d'échange contrôlée (Karsenti et Savoie-Zajc, 2004).

3.2.3 La modalité retenue : l'entrevue semi-dirigée exploratoire

Considérant le sujet à l'étude (le jugement), l'entrevue en personne constitue le meilleur choix afin de réunir des données sur l'expérience des participants. Savoie-Zajc (2003, p.295) définit l'entrevue comme « une interaction verbale entre des personnes qui s'engagent volontairement dans pareille relation afin de partager un savoir d'expertise, et ce, pour mieux dégager conjointement une compréhension d'un phénomène d'intérêt », à savoir ici, le jugement. Considérant cette méthode comme une possibilité de laisser émerger des informations pertinentes quant au jugement, cette dernière s'est avérée plus que pertinente. Pour Stake (2010) l'entrevue est utilisée afin d'obtenir des informations ou des interprétations uniques et pour trouver « la chose » que le chercheur n'a pas été capable d'observer par lui-même. De ce fait, l'entrevue semi-dirigée permet d'explorer des thèmes négligés et émergents au cours de l'entrevue afin d'obtenir les informations les plus pertinentes pour le chercheur.

Dans le même ordre d'idée, appliqué au sujet de la présente recherche, pour Stewart (1997), il y a deux façons principales de décrire le processus du jugement: l'analyse du protocole et l'analyse du jugement. L'analyse du protocole repose sur la description verbale et la description de la façon dont les gens raisonnent et elle se compose d'individus qui demandent à « penser tout haut » pendant qu'ils travaillent à travers une série de jugements (Kleinnuntz, 1968; Ericsson & Simon, 1993). Les descriptions sont par la suite analysées afin de développer un modèle qui consiste en un certain nombre de "si ... alors ...". Les relations trouvées peuvent être schématisées

par la suite sous la forme d'un organigramme. Les recherches de Fitzpatrick, Christie et Mark (2009) constituent une bonne illustration de cette approche. Dans leur étude, ils ont interrogé des théoriciens et leur ont demandé de "penser à haute voix" pour réfléchir à leur pratique.

D'autres chercheurs ont utilisé des méthodes d'analyse du jugement qui ne dépend pas de la capacité d'une personne à fournir des descriptions verbales précises du processus. La personne est tenue de ne faire que ce qu'il ou elle ferait naturellement. L'analyste développe un modèle pour décrire le processus d'inférence qui a produit le jugement (Hammond et al, 1975; Brehmer & Joyce, 1988; Cooksey, 1996). L'analyse du jugement offre un avantage majeur comme outil pour développer des modèles descriptifs, car il fournit une méthode de modèle qui ne repose pas sur la capacité du décideur de décrire son processus de réflexion. Le cas Bunche-Da Vinci par Alkin et Christie (2004) constitue une bonne illustration de quelques théoriciens bien connus qui ont été invités à préparer des évaluations. Pour Einhorn, Kleinmuntz et Kleinmuntz (1979), ces approches peuvent être combinées parce qu'elles se complètent bien, et parce que l'enquêteur peut tirer parti des perspectives offertes par les deux approches tout en évitant certains pièges. Cette approche combinée se rapporte à celle décrite par Strauss et Corbin (1998) qui décrivent l'évaluateur comme une capacité démontrée de prendre du recul et d'analyser la situation avec un point de vue critique, de reconnaître la partialité possible, de penser de manière abstraite, d'être flexible et ouvert à la critique, et d'être sensible au monde qui nous entoure et à l'action des parties prenantes.

La procédure choisie pour la présente recherche s'appuie sur la dernière stratégie qui combine des analyses de protocoles et de jugements. Les entrevues ont été

enregistrées après avoir obtenu un accord signé de chaque personne. Le chercheur a mené des entrevues en profondeur qui ont duré de 90 à 120 minutes dans un dialogue entre le chercheur et les participants, en utilisant des questions ouvertes. Tout d'abord, les participants et le chercheur se sont entendus sur une compréhension commune de ce qu'est le jugement et de ce qu'est la crédibilité. Ensuite, le chercheur a demandé aux participants de rapporter deux expériences professionnelles qu'ils ont considérées comme ayant généré une « bonne » décision, et une seconde ayant généré un « moins bon » jugement. Ils devaient justifier leurs choix et décrire les situations en insistant sur les processus qu'ils avaient alors utilisés. Par la suite, ils ont été invités à « penser tout haut » en comparant les deux situations (ce qui était semblable et ce qui était différent) et de spéculer sur ce qui pourrait expliquer les différents résultats. Durant l'échange, le chercheur a eu l'occasion de tester une hypothèse formulée par les personnes interrogées précédemment.

L'entrevue semi-structurée individuelle apparaît comme le meilleur choix pour cette recherche puisque cette technique permet des échanges sur des thèmes suggérés et invite le participant à s'engager sur les sujets amenés. De plus, elle permet de garder une « certaine constante dans les entrevues même si l'ordre des questions, des détails et de la dynamique particulière varient » (Karsenti et Savoie-Zajc, 2004, p.133). L'entrevue individuelle a été préférée par rapport à l'entrevue de groupe pour que les personnes interviewées soient libres d'exprimer leurs opinions et sentiments sans gêne ou peur d'être jugées. De plus, elle a permis d'investiguer les points de vue personnels des répondants en produisant ainsi plusieurs discours individuels : un par entrevue.

Un questionnaire a été développé à cet effet, ce qui constitue une considération importante puisqu'il témoigne de la qualité de l'information générée. Ce questionnaire a fait l'objet d'une validation rigoureuse (Marchand, 2011).

3.3 L'analyse et l'interprétation des données exploratoire

Après la transcription complète des entrevues en verbatim, chaque entrevue a été soigneusement analysée à partir des objectifs de recherche exposés dans le cadre théorique. La démarche *d'analyse inductive générale* proposée par Thomas (2006), procédure de réduction de données, s'est avérée une option pertinente afin d'analyser les données qualitatives recueillies. Blais et Martineau (2006, p.1) définissent les objectifs liés à l'utilisation de ce concept comme suit :

1. Condenser des données brutes dans un format résumé;
2. Établir des liens entre les objectifs de la recherche et les catégories découlant de l'analyse des données brutes;
3. Développer un cadre de référence ou un modèle à partir des nouvelles catégories émergentes.

En résumé, l'analyse inductive générale consiste à forger des résultats de recherche en dégagant les données à partir de l'examen du corpus et en les interprétant afin de les catégoriser.

Dans le même ordre d'idée, la présente recherche s'inspire de la *grounded-theory* que Glaser (2005) définit comme étant une méthode de recherche qui fonctionne presque à l'inverse de la recherche traditionnelle et qui peut, aux premiers abords, sembler

être en contradiction avec la méthode scientifique. En effet, plutôt que de commencer par la recherche et l'élaboration d'une hypothèse, la première étape consiste à faire la collecte de données. D'après les informations recueillies, les principaux points seront identifiés par une série de codes extraits du texte, que nous nommons *unités de sens*. Ces dernières seront par la suite regroupées en concepts similaires et prendront la forme de catégories. Ces catégories favoriseront, entre autres, la découverte et l'émergence de nouveaux concepts, de nouvelles théories.

Comme mentionné dans le précédent paragraphe, *l'unité de sens* (codification) consiste à attribuer un nom aux différents segments des transcriptions des données et à mettre en évidence les diverses idées contenues dans le texte.

La *catégorie* quant à elle se définit comme étant un deuxième niveau d'analyse qui permet d'identifier, à partir des *unités de sens*, des catégories ou des thèmes, qui émergent des données, et qui ont une vie externe aux évidences. Elles résultent d'une démarche intuitive et peuvent aussi se fonder sur l'atmosphère qui se dégage des données. Elles se doivent de respecter quelques conditions telles que : refléter la problématique, être exhaustives, mutuellement exclusives, sensibles et conceptuellement congruentes (Hurteau, 2009).

Pour Stake (2010), le codage est une caractéristique commune à toutes les microrecherches et analyses qualitatives et de synthèses. Le codage est le tri de tous les ensembles de données du sujet, des thèmes et des questions importantes à l'étude. Il peut être structuré par la question de recherche, par une carte conceptuelle, et par la plupart des données collectées. Il est possible de commencer le codage dès le

début des entretiens ou encore d'attendre à la toute fin. À mesure que la question de recherche prend un nouveau sens et que le travail sur le terrain apporte de nouvelles notions, les catégories prennent elles aussi une nouvelle signification. Le codage est donc en mouvance jusqu'à ce que les choix s'arrêtent sur les catégories finales qui seront contenues dans le rapport final.

Luckerhoff et Guillemette (2012) proposent la démarche méthodologique suivante afin d'expliquer les processus de la théorisation et de l'émergence des données dans une perspective de recherche appuyée sur la théorie ancrée.

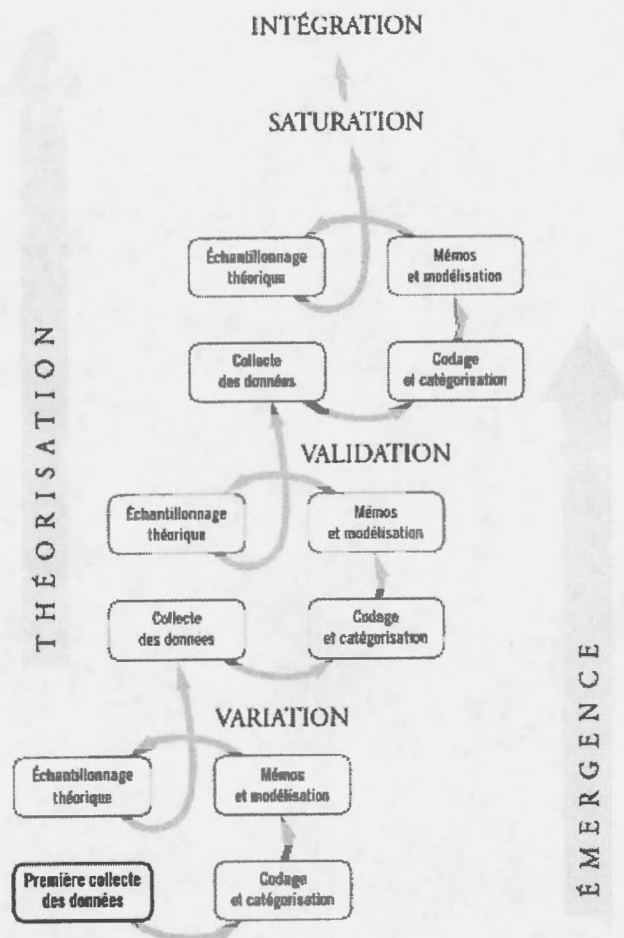


Figure 3.1 Démarches méthodologiques de la théorie ancrée (Tiré de Luckerhoff et Guillemette, 2012).

Le présent projet s'est inspiré de cette théorie. En effet, le processus a été itératif et les actions de «variation» et de «validation» se sont répétées jusqu'à la saturation. Le codage et la catégorisation a changé et évolué à chaque entrevue jusqu'à une intégration complète des catégories. Cette intégration a été possible par le traitement de l'information.

C'est à partir du logiciel QDA MINER que le traitement de l'information a été réalisé puisque ce dernier s'agence parfaitement au traitement de verbatims d'entrevues. De même, l'examen des concordances et divergences a favorisé l'émergence de ce que constitue un jugement crédible et les conditions requises pour l'obtenir.

3.4 Les considérations éthiques

Les répondant à cette recherche ont participé à l'étude de façon volontaire. Compte tenu de ce qui précède, les participants avaient la liberté de se rétracter ou de mettre fin à leur participation à tout moment, sans justification et sans pénalité. De plus, les participants ont formellement été informés, avant leur implication, du sujet de la recherche, de ses objectifs et de son déroulement. À cet effet, chaque participant a reçu une copie du questionnaire quelques jours avant l'entrevue, et a dû signer un formulaire de consentement (voir appendice A). L'anonymat et la confidentialité des participants et des données recueillies ont été respectés selon le désir des participants et des règlements institutionnels en vigueur lors du traitement et de l'analyse des données ainsi que lors de la diffusion des résultats de recherche. De plus, toute information permettant d'identifier les participants a été remplacée par un code lors de la transcription. Sur ce, seuls les membres de l'équipe de recherche ont été autorisés à traiter les données et les enregistrements utilisés ont demeuré sous clé jusqu'à leur destruction lors de la fin de l'étude. À noter également que la présente recherche s'est vue émettre un certificat d'approbation du *Comité Institutionnel d'Étique de la Recherche avec des êtres humains de l'UQAM* (CIÉR). Pour terminer, les participants ont été remerciés et informés des résultats et des conclusions de la recherche.

3.5 Les forces et les limites de l'étude

Les considérations méthodologiques effectuées pour cette étude assurent l'atteinte des objectifs présentés dans le cadre théorique. Sur ce, ci-dessous seront présentées les principales forces et limites de cette recherche au niveau méthodologique.

3.5.1 Les forces

Tout d'abord, la recherche présente une méthodologie intéressante puisque son échantillon multidisciplinaire, représentant des praticiens issus de diverses activités professionnelles, permet de s'intéresser au jugement professionnel global, le sortant ainsi de son enfermement lié à la discipline de l'évaluation de programme. En s'intéressant au jugement professionnel en général, il est possible d'aller plus loin dans les recherches et d'élargir le cadre théorique, qui demeure peu étoffé en évaluation de programme.

Ensuite, la méthode *d'analyse inductive générale* et la transcription intégrale des six entrevues en *verbatim* à l'aide du logiciel *QDA MINER*, assurent une rigueur à la démarche méthodologique proposée et aux résultats obtenus.

Enfin, cette recherche s'inscrit dans un plus grand projet, qui consiste à prendre les résultats obtenus lors de cette étude et d'établir, à l'aide d'entrevues de groupe avec des professionnels de l'évaluation de programme et des clients de l'évaluation, les éléments du jugement transférables à la discipline de l'évaluation de programme.

3.5.2 Les limites

En tenant compte que la recherche s'inscrit dans un courant qualitatif, il importe de soulever le fait que les résultats de cette recherche puissent être subjectifs. Les unités de sens et les catégories retenues sont le fruit de la personnalité et de la manière personnelle de réfléchir de l'étudiante chercheuse. Sur ce, Stake (2010, p.29) affirme : « Qualitative research is subjective. It's contribution toward an improved and disciplined sciences are slow and tendentious. New questions emerge more frequencely than new answers ». De plus, Silverman (2000) prétend que les risques éthiques associés à la recherche qualitative sont substantiels et que le coût est très élevé. De plus, il est probable que la position des interviewés change selon les situations puisqu'ils ne livrent que la synthèse de leur expérience. Le chercheur reconnaît que le produit de la recherche peut être interprété différemment puisque la réalité est partagée. Les expériences et les connaissances propres à chaque individu influencent leur interprétation.

CHAPITRE IV

RÉSULTATS

Ce quatrième chapitre est consacré aux résultats de la recherche. Six professionnels ayant à poser un jugement crédible lors de leur pratique ont été sélectionnés de façon intentionnelle afin de prendre part à des entrevues individuelles. Ce chapitre est composé de deux parties distinctes : la première reprend les différentes étapes auxquelles nous avons eu recours lors de l'analyse des données, tandis que la deuxième partie consiste à présenter les résultats de recherche en tant que tels, lesquels permettront de répondre à nos objectifs de recherche.

4.1 Rappel des objectifs de recherche

En premier lieu, il importe de rappeler la question de recherche qui a conduit à la présente étude : *Quels processus sous-tendent la formulation d'un jugement crédible dans le cadre de l'évaluation de programme?*

Cette question s'explique en l'objectif général suivant, soit :

- 1- Cerner et comprendre la façon de générer un jugement crédible, et ce, par l'analyse des processus issus de divers domaines d'intervention et pratiques professionnelles.

La présente recherche s'intéresse à l'exploration des différentes pratiques professionnelles conduisant à la formulation d'un jugement crédible. Pour ce faire, un

questionnaire a été élaboré afin de récolter les informations désirées. Ledit questionnaire a tout d'abord été validé par Hurteau et Marchand (2011) dans la première phase de ce projet de recherche. Une fois le questionnaire validé et prêt à l'emploi, la présente recherche vise les objectifs spécifiques suivants :

1. Cerner des pratiques professionnelles qui conduisent à un jugement professionnel
2. Documenter le processus vécu par les personnes
3. Induire une éventuelle modélisation à partir de ces témoignages

4.2 L'analyse des données

Avant de présenter les résultats de recherche, il est important d'indiquer la méthode suivie afin d'analyser les données amassées lors des six entrevues. Deux phases ont composé cette analyse, le codage des données, et la synthèse des données. En effet, selon les principes de l'analyse inductive de Thomas (2006), la valeur et la qualité de l'information ont été vérifiées pour l'ensemble des entrevues.

4.2.1 Le codage des données

Pour le présent mémoire, six entrevues sur dix-neuf ont été sélectionnées pour des fins d'analyse. Le choix des entrevues a été fait dans le but d'avoir un éventail le plus large possible de domaine d'intervention. Ainsi, un pompier, un médecin, un policier, un avocat, un juge et un professeur constituent l'échantillon de cette recherche.

Tel que mentionné dans la section 3.3, le présent mémoire s'inspire de la théorie ancrée. La première étape consiste à faire la collecte de données et, d'après les informations recueillies, les principaux points seront identifiés par une série de codes extraits du texte, que nous nommons *unités de sens*. Ces dernières seront par la suite regroupées en concepts similaires et prendront la forme de catégories. Ces catégories favoriseront, entre autres, la découverte et l'émergence de nouveaux concepts, de nouvelles théories.

En ce sens, les six entrevues ayant été réalisées et retranscrites en *verbatim*, une première lecture a été effectuée afin de recueillir des données brutes et de se familiariser avec les sujets à l'étude. Cette première étape a permis d'identifier les thèmes principaux. Par la suite, une deuxième lecture des entrevues a été faite afin de trouver les extraits qui nécessitaient un intérêt particulier. C'est de cette manière que les unités de sens ont été trouvées. L'élaboration des unités de sens est une étape subjective dans le processus puisque les données conservées pour le codage et le traitement des données dépendent uniquement du chercheur. À cet effet, Boutin (2008) affirme que le choix d'un passage en particulier découle de l'expérience du chercheur, de la connaissance qu'il a de son sujet ainsi que de sa sensibilité. Cette étape a permis de mettre en évidence les catégories importantes et de lui d'attribuer des unités pour chaque catégorie.

Il existe différents types d'encodage dans le cadre de la *grounded theory* : l'échantillonnage par choix raisonné, la collecte des données, l'échantillonnage théorique, le codage et la conceptualisation, la théorie de fond et enfin la saturation théorique.

Tout d'abord, la collecte des données s'appuie sur l'échantillonnage par choix raisonné qui permet au chercheur de choisir les participants en se basant sur son propre jugement par rapport à leur caractère typique ou atypique. Un échantillon choisi à dessein, aussi communément appelé un échantillon discrétionnaire, est celui qui est choisi en fonction de la connaissance de la population et de l'objet de l'étude. Les sujets sont choisis en raison de certaines caractéristiques pour procéder par la suite à la collecte de données qui s'appuie également sur un échantillonnage théorique.

Ce dernier est un élément central de la théorie ancrée préconisée par Glaser et Strauss (1967). L'échantillonnage théorique est lié à l'objectif de produire et de développer des idées théoriques, plutôt que d'être destiné à produire des résultats qui sont représentatifs d'une population ou à tester des hypothèses. Cette forme de prélèvement n'a pas lieu à un moment précis dans le processus d'enquête, mais est une caractéristique récurrente: à plusieurs reprises le chercheur doit se demander quels paramètres, quels événements, quelles populations, etc., il serait intéressant d'étudier afin de développer les aspects de la théorie émergente. De cette façon, l'échantillonnage théorique est guidé par et contribue à générer, la « sensibilité théorique » qui est nécessaire à la théorisation ancrée.

Le codage et la conceptualisation des données sont circulaires. Ils permettent de confronter les premiers résultats de l'analyse aux données suivantes. En effet, le chercheur peut vérifier au fur et à mesure les hypothèses de travail qui se dégagent des données lors des entretiens suivants

La théorie de fond peut être construite dans le processus d'identification des différences et des similitudes entre les cas contextualisés, entre et au sein des études de cas axées sur un thème similaire. Le contenu de la théorie de fond est essentiellement descriptif, et dans la plupart des cas, comporte un élément d'intentionnalité.

La saturation théorique est la phase d'analyse de données qualitatives dans laquelle le chercheur a poursuivi l'échantillonnage et l'analyse des données jusqu'à ce qu'aucune nouvelle donnée n'apparaisse et que tous les concepts de la théorie se soient bien développés. Un examen des concepts et des liens entre les concepts qui forment la théorie a été vérifié, et aucune donnée supplémentaire n'est nécessaire pour la suite de l'analyse.

À la suite de plusieurs lectures du contenu des entrevues, des catégories ont émergé pour ensuite y associer les unités de sens préalablement identifiées dans le codage de données. Il est important, à cet égard, de mentionner que certaines catégories ont été élaborées à partir du cadre théorique tandis que d'autres, ont été élaborées lors du codage de données. Ensuite, à l'aide du logiciel QDA MINER, il a été possible d'allouer une couleur pour chaque catégorie afin de détecter des unités de sens. Le logiciel a également permis de repérer, classer et organiser les données. Il est à noter que, tel le suggère Thomas (2006), une attention particulière a été apportée afin d'arriver à un maximum de huit catégories.

4.2.2 La synthèse des données

Afin d'exposer les résultats obtenus par l'analyse des données, une deuxième grande étape s'impose : l'examen des données et le traitement des données. Afin de les traiter de façon appropriée, un examen concis des données codées a dû être réalisé.

4.2.2.1 L'examen des données

Afin d'en arriver à un portrait global des données suite à l'obtention des catégories finales, un examen des unités de sens exposé dans le paragraphe précédent s'avère nécessaire. En effet, grâce au logiciel QDA MINER, il a été possible de s'assurer que les unités de sens étaient correctement reliées à sa catégorie d'appartenance. De plus, cette étape a permis de s'assurer qu'il n'y avait aucune erreur de codage au sein des catégories ou encore d'ordonner les unités de sens qui avaient pu être disséminées. L'examen des données a enfin permis d'examiner la congruence des informations trouvées en fonction des objectifs de recherche.

Étant donné l'échantillon limité ($n=6$) et les visées de cette recherche, le recours à des tableaux de fréquences s'est avéré inutile. En effet, étant donné que cette étude se situe dans une perspective qualitative d'exploration et qu'elle ne tend pas vers une généralisation des résultats, le but intrinsèque est plutôt de s'intéresser au vrai sens des données en évitant tout rapprochement quantitatif.

4.2.2.2 Le traitement des données

L'étape du traitement des données vise essentiellement à réaliser une réduction des données amassées préalablement afin d'en arriver à une forme plus compréhensible afin de s'orienter graduellement vers la présentation des résultats de recherche. Afin d'arriver à une analyse des résultats plus approfondie, pour chacune des catégories, les unités de sens sont structurées et condensées. C'est ainsi qu'il a été possible de regrouper les catégories où les unités de sens étaient trop similaires et d'établir des liens étroits entre les différentes entrevues.

Les étapes mentionnées précédemment ont ainsi permis d'en arriver aux résultats qui seront présentés à la section suivante. Des extraits des *verbatim* (citations) seront présentés au besoin afin d'appuyer les résultats obtenus afin de comprendre le sens exact des propos des sujets et ainsi faire valoir la teneur des données recueillies.

4.3 Les résultats de recherche visant la description d'un processus de production du jugement crédible.

Tel que mentionné précédemment, les six *verbatim* ont été analysés afin d'examiner la possibilité d'engager de nouvelles données pour parfaire notre compréhension du jugement crédible. En effet, à partir des six *verbatim*, six catégories ont été retenues afin d'expliquer le processus de formulation du jugement crédible, soit :

- 1- La cueillette et l'examen des données
- 2- La souplesse de l'intervenant

- 3- Le protocole
- 4- L'argumentation et le raisonnement
- 5- Les compétences et l'expérience de l'intervenant
- 6- La confirmation et la vérification

Dans les prochaines sections, les six catégories seront expliquées et justifiées en détail à l'aide de passages de verbatim spécialement choisis.

1) La cueillette et l'examen des données

Les professionnels ont mentionné l'importance de préciser d'abord le problème ou la situation avant de poser un jugement. Il est important de bien comprendre les enjeux et d'analyser les informations préalables avant de poser toutes actions. C'est la base qui va permettre de prendre la décision et d'effectuer une démarche pour poser un jugement crédible en bout de processus. « *Quand tu es à la caserne et que l'appel rentre, déjà on a quelques informations qui sont transmises par la station, déjà avec ces informations-là on commence à se préparer mentalement à quel genre de situation on va avoir à faire.* » « *On voit notre patiente pour une problématique et on se fait un algorithme pour arriver à régler la problématique.* » « *La qualité de l'information et des sources est aussi très importante.* » « *D'abord il faut préciser le problème, c'est la base de tout. Et assez souvent, quand le problème est bien présenté et clair, la plupart du temps, la décision s'en suit est très facilement.* »

D'après les intervenants interrogés, il est très important que la collecte d'information soit complète avant d'émettre un jugement et que cette dernière soit faite avec brio et sérieux. Il faut recueillir le plus d'information possible afin de rendre une information de qualité. Dépendant du milieu, cette dernière doit être réalisée objectivement selon les informations disponibles à cet instant. Si certaines informations sont manquantes, il faut aller les chercher, ce qui rend le processus itératif. Ce processus peut se prolonger tout au long de l'évènement jusqu'à ce que le jugement soit fait, ou alors être plus restreint dans le temps. *»On pense prendre la bonne décision sur le moment, mais on n'avait pas toutes les informations nécessaires à ce moment-là.» « Le temps nous permet d'avoir le plus de données possible. Plus on a d'information, plus c'est facile de faire un jugement et de prendre une bonne décision. Quand on fait des mauvais jugements, c'est souvent quand on manque de données ou que l'on ne peut tout simplement pas aller les chercher.» « L'important c'est d'avoir le plus de données, d'information, ce qui va permettre de porter un bon jugement et de finalement arriver à une bonne décision. »*

Les intervenants insistent également sur le fait de valider les informations reçues à cette étape du processus. En effet, il est primordial que le jugement rendu s'appuie sur des données valides afin de s'assurer de sa crédibilité. *« Pour que la décision soit claire, il faut que la personne devant toi ou qui pose le problème soit capable de la manifester, capable de dire pourquoi il y a un problème. Et le préciser dans des mots clairs. Et là je tombe dans l'autre aspect, qui est la décision que moi j'ai à prendre, c'est au niveau de l'analyse, l'analyse des faits. C'est quoi les faits, j'ai le problème, j'ai les faits. » « Recueillir les faits, pas les opinions, les faits. Tu as beau avoir la meilleure théorie juridique, si tu n'as pas les faits pour t'appuyer, ça ne marche pas, tu as besoin de la preuve. »*

2) La souplesse de l'intervenant

Selon les intervenants interviewés, il importe de faire preuve de souplesse, de remettre en question nos conceptions initiales avant de poser un jugement sur la situation. « *Pas de décision hâtive. Laisser murir* ». « *Mon jugement était influencé par ma zone de confort, s'était une situation facile. Mon jugement au départ était mauvais.* »

« *On a toujours des surprises sur les lieux, parce que la prise d'appel vient d'informations de citoyens qui sont soit en situation de panique ou qui n'ont pas très bien jugé la situation. [...] C'est une prise de décision action-réaction.* »

Si la situation le permet, il est préférable, selon certains, de laisser murir la décision avant de rendre son jugement. « *Avoir le temps d'analyser la situation et de se remettre en question.* »

Même si des protocoles sont instaurés, à quelques occasions, il faut écouter son instinct et y aller selon le "bon sens". Il est important de prendre une décision qui va correspondre aux valeurs de la personne. « *La loi me permet de prendre une décision, d'y aller selon mon bon sens là dedans.* »

3) Protocole

La majorité des intervenants ont mentionné à plusieurs reprises lors des entrevues l'importance de se baser sur des protocoles déjà établis ou des règles de fonctionnement. Cela leur permet de baliser leur jugement selon des normes préétablies et de rester dans le cadre légal de leur profession. Les protocoles établis assurent aussi à l'intervenant une crédibilité du jugement émis et encadrent leur démarche tout au long du processus. En effet, les participants ont mentionné :

- *« On travaille toujours dans un contexte, un schéma ou on a des pratiques bien établies pour chacune des situations. »*
- *« Dans notre travail, chacune des interventions à certaines balises déjà établies que l'on peut appelé les règles de l'art donc à partir du moment ou l'on reste dans ces balises là, tu te protèges comme tel. »*
- *« On a des critères de descente qui nous indique si cela a des chances de fonctionner ou pas. »*
- *« Il faut comprendre que la médecine est parfois très précise, mais d'autres fois, on se trompe. Ce qui fait qu'on doit comprendre les limites de nos outils. »*
- *« Quand tu es au niveau du droit fondé, pour moi c'est crédible. »*
- *« Tu as toujours le garde-fou du droit (déontologie). »*
- *« Il n'y a pas juste l'expérience, il y a un domaine théorique. »*
- *« Dans notre métier, la prise de décision va se faire beaucoup parce que l'on a des directives à respecter. On a des codes à respecter et on a des assises à respecter que ce soit le Code criminel, le Code pénal, le CSR, etc. Donc la prise de décision va beaucoup être influencée au départ. »*
- *« C'est ce que la loi nous dicte, ce que le code de déontologie nous dit de faire. Donc on n'a pas le choix de prendre certaines décisions. Des fois on n'a pas le choix. On doit répondre à des questions. »*

4) L'argumentation et le raisonnement

L'argumentation et le raisonnement à la suite de la production du jugement sont également très importants pour les personnes interviewées. En effet, il est important que les personnes puissent justifier pourquoi telle ou telle décision a été prise et pourquoi certaines informations ont pesées plus dans la balance et pourquoi d'autres ont tout simplement été éliminés. Les intervenants, quel que soit le domaine professionnel, doivent être en mesure d'expliquer son raisonnement et de l'argumenter afin de le rendre crédible aux yeux des autres. La démarche effectuée pour en arriver à la prise de décision se doit aussi d'être détaillée et logique :

- *« Il y a toujours une période d'analyse très courte dans le domaine de l'urgence. Tu arrives et tu as quelques secondes pour analyser les événements et pour poser les premiers gestes. Naturellement après cela il peut y avoir rectification parce que des fois une situation a l'air d'être différente de ce qu'elle est en réalité, ou alors elle se transforme. »*
- *« Il faut comprendre ce qui fait que ça ne va pas, et une fois qu'on va trouver ce qui ne va pas, il y a toujours des solutions médicales. »*
- *« Pour le deuxième cas, j'avais l'expérience du cas avant. À cause de l'autre cas, je me suis remis en question. »*
- *« Le jugement peut être tout à fait crédible et il va quand même y avoir des gens qui ne seront pas en accord avec cela. »*
- *« Il ne peut pas toujours y avoir consensus sur le jugement, il est crédible pour certains et pas pour d'autres. »*
- *« J'ai quand même validé des choses, j'ai quand même validé avec sa conjointe, elle m'a confirmé tout cela. »*

5) Les compétences et l'expérience de l'intervenant

Les compétences et l'expérience de l'intervenant sont un ensemble de savoir-être, de savoir-faire et de valeur. Les professionnels mentionnent à plusieurs reprises la subjectivité du jugement puisque les personnes ne réagissent pas toutes de la même façon à un événement. C'est le bagage spécifique de chaque personne qui influencera le jugement rendu, et c'est souvent le même bagage qui rendra son jugement crédible aux yeux des autres:

- *« Le but pour nous c'est de faire machinalement certaines manœuvres. »*
- *« C'est quand on sort de notre contexte, et qu'on prend encore des bonnes décisions à mon avis c'est parce que tu es capable de t'ajuster et de garder ton sang-froid. Pour moi s'était un bon test à ce moment-là. Prendre les décisions qui s'imposent. »*
- *« Quand tu gagnes de l'expérience, il faut que tu sois capable de prendre des bonnes décisions, quand tu n'es plus dans ton protocole justement. »*
- *« Tu es capable de prendre les bonnes décisions avec l'information que tu as. Tu es capable d'évaluer la situation. »*
- *« Quand tu es dans ta zone de confort, tu es moins sur tes gardes. Il faut toujours être alerte et éviter le plus possible les zones de confort, resté concentré. »*
- *« Il y a un feeling clinique qui vient avec l'expérience que l'on ne peut pas lire dans les livres. »*
- *« L'expérience peut t'aider à te faire d'une étape à l'autre, mais non, il faut des algorithmes, c'est certain. Il faut s'en faire un dans notre tête s'il n'en a pas. À l'urgence, il faut penser vite. »*
- *« Pourquoi moi dans telles circonstances je vais décider telle chose, c'est parce que j'ai fait tels métiers, telles lectures, tels voyages, etc. Mes intrants vont m'amener à prendre la bonne décision. »*
- *« C'est sûr que quand on prend une décision, c'est le jugement qui rentre en ligne de compte en premier lieu, le jugement est influencé par plein de facteurs, dont l'expérience de vie, les valeurs de chaque personne, l'expérience professionnelle, etc. »*
- *« Chaque personne a un sixième sens, si vous n'êtes pas à l'aise dans cette décision-là, dans ce jugement-là, c'est parce que ça ne va pas. Il y a quelque chose qui vous dit que ce n'est pas correct. À ce moment-là, il faut se faire rassurer, aller chercher d'autres informations, ou il faut valider avec quelqu'un. »*

6) La confirmation et la vérification

La confirmation avec un pair ou un supérieur a été fréquemment relevée par les professionnels comme une composante importante à la production du jugement final. Ils ont tous mentionné l'extrême importance de valider son jugement ou sa décision. En effet, cette étape permet de donner une crédibilité au jugement émis et également de se rassurer sur des doutes préalablement établis chez l'intervenant:

- « *Le fait d'expliquer aux patientes rend notre jugement plus acceptable.* »
- « *Réévaluer la situation.* »
- « *On a pas juste le regard de notre patiente, on a aussi le regard des infirmières, leurs feelings. Des fois on doit convaincre nos collaborateurs.* »
- « *Souvent, les meilleures décisions que j'ai prises ou qui ont donné les meilleurs résultats, s'étaient parfois des décisions qui n'ont jamais été acceptées et qui ne le seront jamais pour certain. Si on parle de l'honnêteté de la décision, de sa non-partialité, pour moi c'est plus cela un jugement crédible, c'est que je n'ai pas été payée pour prendre cette décision-là, il est plus éthique.* »
- « *La crédibilité vient de l'honnêteté et puis de la franchise et de la vérité.* »
- « *Pour moi ma décision est crédible si moi je me sens bien avec, peut importe qu'elle soit acceptable par les autres. La première crédibilité est par rapport à moi.* »
- « *La décision la meilleure avec ce que j'avais et avec ce que j'étais et selon les faits que j'avais au moment où je l'ai prise.* »
- « *Le jugement des pairs et l'information donnée par les pairs, c'est gros, c'est primordial.* »
- « *C'est de valider ce que j'ai comme décision. Je prends mes données et j'analyse la situation et après je me demande si je peux faire autrement. Je viens de valider ma décision. Quand je dis valider, ce n'est pas prendre la même décision deux fois de la même façon, c'est aller chercher d'autres hypothèses et parfois on va aller la valider avec un confrère, un supérieur, quand on a le temps.* »

4.4 Atteinte des objectifs

Les résultats présentés dans les pages précédentes permettent de mieux comprendre le processus argumentatif lié à la formulation d'un jugement crédible. Il est donc possible de conclure que l'objectif principal de la recherche qui visait à *cerner et comprendre la façon de générer un jugement crédible par l'analyse des processus issus de divers domaines d'intervention et pratiques professionnelles* est atteint. En effet, l'analyse présentée a permis de dégager des catégories valables à l'atteinte de cet objectif et transférables au contexte de l'évaluation de programme. Il est cependant important de noter que la présente recherche, ne reposant que sur l'analyse d'un échantillon de six participants (n=6), ne permet pas de confirmer que les résultats présentés sont aliénables à la pratique évaluative.

À cet égard, le présent mémoire suppose l'atteinte des objectifs mentionnés précédemment. Les pratiques professionnelles ont bien été choisies et l'échantillonnage raisonné bien utilisé afin de faire ressortir les éléments recherchés. Ensuite, le processus vécu par les personnes a été documenté et des catégories fidèles ont permis de dégager les principaux éléments permettant la formulation d'un jugement crédible.

4.3.1 Synthèse des résultats

En sommes, l'ensemble des résultats de cette étude démontre que le processus sous-jacent au jugement doit s'appuyer sur un ensemble de caractéristiques qui s'influencent l'une et l'autre selon le contexte. L'expérience et les compétences de

l'intervenant le guideront tout au long de l'élaboration du jugement. De plus, les informations disponibles et le raisonnement des dites données joueront un rôle primordial dans la formulation de la crédibilité du jugement final. En effet, ces dernières doivent être prises au sérieux et afficher une rigueur. De plus, il sera important de revenir sur certains points et d'aller chercher d'autres informations si le contexte l'exige. La souplesse de l'intervenant y jouera pour beaucoup. La formulation du jugement devra également, si possible, s'appuyer sur un protocole précis et vérifier et confirmer son jugement ou sa décision auprès des pairs pour valider sa crédibilité. Enfin, l'intégralité des composantes proposées dans l'étude agissent en interrelation les unes avec les autres et ne peuvent être prises et analysées seules. Dès lors, le processus d'élaboration d'un jugement crédible sera en tout temps itératif. Les résultats qui ont donné lieu à la modélisation du processus inhérent au jugement crédible par Hurteau et Houle seront présentés dans la section discussion.

CHAPITRE V

DISCUSSION

Le cinquième chapitre s'appuie sur les résultats du chapitre précédent pour engager une discussion quant aux points révélateurs de l'étude. Les éléments importants seront mis de l'avant et une discussion sera engagée à partir des liens suggérés dans le cadre théorique. Ce dernier chapitre se terminera par les questions qui restent encore en suspens et inclura par le fait même de nouvelles pistes de réflexion.

5.1 Les considérations méthodologiques

En premier lieu, il est important de mentionner que les résultats de la présente étude ont été réalisés à partir d'un questionnaire qui a été validé par Marchand (2011). En effet, pour s'assurer de son efficacité de répondre à la question de recherche, Marchand (2011, p. 75) affirme que :

les avantages liés à la validation présentée dans ce mémoire se manifesteront dans les suites à donner à cette étude, puisqu'elle s'inscrit dans une recherche qui utilisera ce questionnaire et la démarche qu'il favorise, en conduisant des entretiens auprès de professionnels issus de domaines distincts.

En effet, la validation du questionnaire a permis de s'assurer de la structure et de la qualité des informations générées. Sur ce, les entrevues effectuées pour le présent mémoire ont donné lieu à une information riche et pertinente en ce qui a trait au jugement crédible. La structure utilisée permettait aux participants de s'exprimer sur

une situation où il avait produit un « bon » jugement et une situation où il avait produit un « mauvais » jugement. C'est à partir de ces deux situations qu'ils ont pu réfléchir au processus qui les avait conduits à la production de ces jugements. Cette méthode de collecte de données, par l'entrevue semi-dirigée, a permis aux participants de mettre en lumière la qualité et la nature du jugement émis et de revenir sur les éléments clefs du processus. De plus, avant de procéder à l'entretien, tous les participants avaient reçu préalablement le questionnaire de recherche ce qui leur permettait de réfléchir aux situations qu'ils allaient proposer et de se questionner sur leurs actions. Un questionnement sur la conception des termes « jugement » et « crédibilité » était également effectué avant chaque début de rencontre, ce qui permettait aux chercheurs et aux participants d'être sur le même diapason. Ainsi, l'ensemble des choix méthodologiques de la présente étude a permis de favoriser l'émergence des concepts liés à la formulation du jugement crédible dans une diversité de professions et de s'appliquer, par la suite, à l'évaluation de programme.

5.2 Les résultats révélateurs

Les résultats présentés dans le chapitre précédent permettent d'améliorer le processus inhérent à la formulation du jugement crédible en affirmant que 6 catégories sont nécessaires à son aboutissement. S'appuyant sur les résultats de Bossiroy (2009) qui ne permettaient pas de cibler les composantes du volet argumentaire du processus de formulation crédible, il est possible maintenant de proposer une alternative, soit de s'appuyer sur : La cueillette et l'examen des données, la souplesse de l'intervenant, l'argumentation et le raisonnement, les compétences et l'expérience de l'intervenant, la confirmation et la vérification et le protocole.

La cueillette et l'examen des données nécessaires à la prise de décision sont une étape importante dans la production du jugement. Ainsi, la méthodologie utilisée influencera le résultat final. De plus, il est nécessaire de s'assurer que l'on a toutes les données disponibles et probantes pour mener à bien l'opération. Ce processus exige de nombreux va-et-vient tout au long de l'enquête. Transposée à l'évaluation de programme, la rigueur scientifique assure également une crédibilité à la démarche et au produit final. Tous comme Donaldson, Christie et Mark (2009), tous les professionnels rencontrés ont mentionné lors de leur entretien l'importance de s'appuyer sur des données valides.

Les participants ont mentionné à plusieurs reprises l'importance de s'ajuster tout au long du processus aux différentes situations imprévues. Il est important de prendre autant en compte ce qui avait été prévu et ce qui émerge au cours du processus. Le modèle de Mintzberg et Waters (1995) dans Hurteau, Houle et Guillemette (2012) tend à expliquer le processus d'ajustement des stratégies.



Figure 5.1 Processus d'ajustement des stratégies (Tiré de Hurteau, Houle et Guillemette 2012, p.105)

L'argumentation et le raisonnement soutenant la formulation d'un jugement crédible professionnel ont été à de nombreuses reprises mentionnées comme une caractéristique hautement importante par les interviewés. En effet, ils soutiennent que le jugement se doit d'être soutenu par une argumentation solide, cohérente, et appuyé par des preuves. C'est cette argumentation et ce raisonnement qui définiront dans plusieurs cas la recevabilité du jugement émis par les parties prenantes. Selon Hurteau, Houle et Guillemette (2012, p. 113) « *la force de l'argumentation d'un évaluateur réside dans sa capacité d'identifier, de traiter et de présenter logiquement l'information qui sous-tend son jugement et non d'expliquer comment elle a été recueillie* ».

Les compétences et l'expérience de l'intervenant jouent un rôle primordial dans la production du jugement. Comme tous les participants interviewés, Schwandt (2007) mentionne que les données et les évidences recueillies sont influencées par les croyances et les valeurs et demeurent particulières au contexte du déroulement de l'évaluation. De même, les compétences personnelles s'avèrent tout aussi importantes que les compétences professionnelles. Les propos de Grinnell (2009) tendent également dans le même sens en mentionnant l'influence de la personnalité du chercheur et son expérience personnelle dans sa pratique.

La confirmation et la vérification de son jugement par des pairs permettent également de s'assurer de son jugement et de refaire les vérifications si des lacunes se présentent. Or, cette étape s'avère essentielle lorsque la profession exercée ne s'appuie pas sur des assises légales. Elle permet de vérifier la qualité de l'information discutée dans le cadre théorique par Stake et Schwandt (2006), à savoir la qualité mesurée et la qualité intuitive.

Si pour plusieurs personnes interviewées lors de l'étude il était évident de s'appuyer sur des protocoles déjà établis par la discipline pour assurer de la qualité, des résultats et du cadre légal du jugement, il n'existe toujours pas à ce jour de « cadre » pour appuyer la démarche évaluative (Fitzpatrick, Sanders et Worthen, 2011). Même si certains auteurs ont tenté de proposer des modèles afin de guider les évaluateurs, ces modèles servent plutôt de guide et assurent une certaine crédibilité aux yeux des parties prenantes (Stake, 1981).

5.3 La modélisation de Hurteau et Houle

S'inscrivant dans le programme de recherche de Hurteau et Houle, les résultats de recherche présentée dans le chapitre précédent abondent dans le même sens. Même si les résultats diffèrent des catégories proposées par les chercheurs, il est possible d'établir des liens avec les résultats de recherche de Hurteau et Houle (2012) et de la modélisation qu'ils proposent quant aux processus de production et de crédibilisation du jugement en évaluation de programmes :

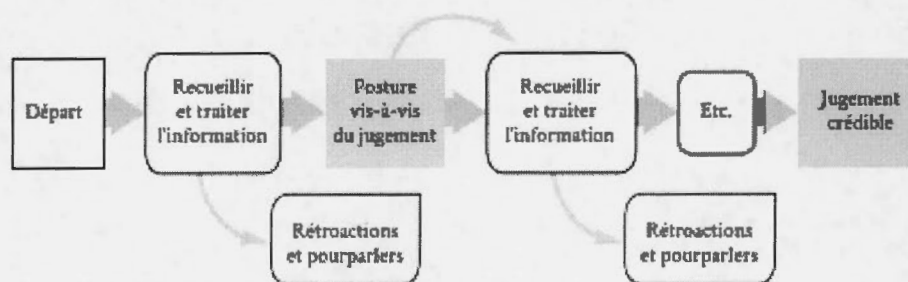


Figure 5.2 Modélisation de production et de crédibilisation du jugement en évaluation de programme (Tiré de Hurteau, Houle et Guillemette 2012, p.95)

En effet, dans leur ouvrage publié en 2012, Hurteau, Houle et Guillemette affirment que « *le jugement crédible commence à se former dès que le processus évaluatif est mis en branle* ». Ainsi, tel que mentionné dans les résultats de recherche, les premières informations disponibles sont continuellement améliorées des ressources supplémentaires tout au long du processus.

5.2 Les questions qui restent en suspens

Bien que la présente recherche est réussie à cerner et comprendre la façon de générer un jugement crédible, et ce, par l'analyse des processus issus de divers domaines d'intervention et pratiques professionnelles, il reste que le modèle présenté ne peut prétendre tout expliquer. Il reste dans le processus inhérent le jugement une dimension « humaine » qui ne peut s'enseigner. Ainsi, tel que le mentionnent Schleifer et Hurteau (2012, p. 118), « *Il existe une compétence qui va au-delà d'un raisonnement et qui se manifeste dans des actions qui se détachent des règles, sans pour autant les nier.* »

En résumé, les résultats de recherche présentés constituent une piste d'explication intéressante quant à la difficulté à laquelle fait face l'évaluation de programme, c'est-à-dire celle de générer un jugement crédible. Tout comme les autres domaines, l'évaluation de programme évolue et change avec le temps. Dans cette optique, il serait intéressant de s'intéresser dans une future recherche à la place que l'on accorde au « jugement » lors de l'évaluation de programme par des « Data Mining » (Spector, 2014) qui sont de plus en plus utilisés et populaires à l'heure actuelle.

CONCLUSION

L'évaluation de programme étant souvent remise en cause sur sa fonction première, soit celle d'émettre un jugement crédible et fondé, la présente recherche exploratoire s'est intéressée aux processus qui sous-tendent la formulation du jugement crédible dans sa pratique, en envisageant le jugement comme une compétence qui serait transférable à différents actes professionnels. Les résultats obtenus suite à la cueillette des données qui avait pour objectif de *cerner et comprendre la façon de générer un jugement crédible, et ce, par l'analyse des processus issus de divers domaines d'intervention et pratiques professionnelles* révèlent l'atteinte de cet objectif.

Après avoir cerné les pratiques professionnelles conduisant à un jugement professionnel, en s'appuyant de la théorie ancrée, le chapitre décrivant les résultats de la recherche a permis d'atteindre les objectifs spécifiques, c'est-à-dire de documenter le processus vécu par les personnes et d'induire une éventuelle modélisation de ces témoignages. En somme, la démarche analytique a permis de dégager des résultats significatifs qui permettent de porter un regard nouveau en établissant les mécanismes soutenant l'énonciation d'un jugement crédible qui sont également, sous toutes réserves, transférables à la pratique évaluative. La modélisation proposée par Hurteau et Houle appuie le processus qui a été ici présenté.

APPENDICE A
FORMULAIRE DE CONSENTEMENT

FORMULAIRE DE CONSENTEMENT

Projet de recherche sur la production d'un jugement crédible en évaluation de programme

Monsieur, Madame,

Je sollicite votre participation à un projet de recherche financé par le Conseil de recherches en sciences humaines du Canada. Ce projet vise à explorer les processus pour produire un jugement professionnel crédible en évaluation de programme. S'il est vrai que vous n'œuvrez pas dans ce domaine, il est aussi vrai que vous êtes régulièrement appelé, dans le cadre de vos activités professionnelles, à poser un diagnostic et émettre un jugement, et ce sont les processus mis en œuvre pour y parvenir qui m'intéressent plus spécifiquement. Votre participation favorisera ainsi grandement la compréhension des processus en jeu et, ultimement, l'amélioration de la pratique en évaluation de programme.

Avec votre autorisation, je vous contacterai pour convenir d'un moment pour effectuer une entrevue en personne qui devrait durer environ une heure. Préalablement, je vous aurai fait parvenir un questionnaire qui permettra d'orienter nos échanges. Je tiens à préciser que si vous réferez à des cas particuliers dans votre pratique, ils n'ont nullement besoin d'être identifiés de façon nominale aux fins de l'exercice puisque je m'intéresse au processus en général. La rencontre aura lieu dans un local de l'université ou dans un lieu choisi à votre convenance. Nos échanges feront l'objet d'un enregistrement audio.

Votre participation à cette recherche est volontaire. Cela signifie que même si vous consentez aujourd'hui, vous demeurez entièrement libre de ne pas participer ou de mettre fin à sa participation en tout temps sans justification ni pénalité.

Les enregistrements audio seront utilisés uniquement par l'équipe de recherche et conservés sous clé dans les locaux de recherche de la chercheuse principale à l'UQAM pour la durée du projet (2 ans). Toute information sur l'enregistrement audio permettant de vous identifier sera enlevée au moment de sa transcription.

Si vous avez besoin de renseignements additionnels concernant cette étude, avant de donner votre autorisation ou même après, n'hésitez pas à me contacter. Il nous fera plaisir de répondre à vos questions. Sachez aussi que notre étude a reçu l'approbation

du Comité institutionnel d'éthique de la recherche avec des êtres humains de l'UQAM (CIÉR). Tout commentaire ou toute inquiétude relativement aux questions d'éthique de la recherche peut être adressé au président du CIÉR, M. Joseph Josy Lévy. Il peut être joint au numéro (514) 987-3000 # poste 4483 ou au poste 7753.

Signature de la chercheure principale :

Numéro de téléphone :

Adresse courriel :

AUTORISATION PERSONNELLE

J'accepte de participer à la recherche :

Oui

Non

J'accepte de participer à une entrevue qui fera l'objet d'un enregistrement audio :

Oui

Non

Signature :

Date :

Coordonnées :

APPENDICE B

QUESTIONNAIRE D'ENTREVUE

QUESTIONNAIRE D'ENTREVUE

INTRODUCTION

Le présent projet de recherche vise à explorer les processus pour produire un jugement professionnel crédible en évaluation de programme. S'il est vrai que vous n'œuvrez pas dans ce domaine, il est aussi vrai que vous êtes régulièrement appelé, dans le cadre de vos activités professionnelles, à poser un diagnostic et émettre un jugement, et ce sont les processus mis en œuvre pour y parvenir qui m'intéressent plus spécifiquement.

Plus spécifiquement, nous aimerions comprendre ce qui se « passe dans votre tête ». Ainsi : à quelle information avez-vous recours? Comment l'obtenez-vous et la triezy-vous? Comment la traitez-vous et la présentez-vous? Etc.

Pour y parvenir, nous allons nous appuyer sur :

- Une situation où vous avez eu l'impression de porter un **bon** jugement professionnel (**positif, adéquat**), de prendre une **bonne** décision;
- Et une situation où vous avez eu l'impression de porter un **mauvais** jugement professionnel (**néгатif, inadéquat**), de prendre une **mauvaise** décision.

Vous pouvez décider de l'ordre dans lequel elles seront traitées dans l'entrevue.

De plus, il est important de comprendre que les questions permettent de baliser l'entrevue et que celle-ci va se dérouler sous la forme d'une conversation. Il n'est pas requis que vous ayez une réponse pour chacune des questions.

Avez-vous des questions avant que nous amorcions les questions?

QUESTIONS

1. Dans le cadre de votre profession, de quelle façon votre jugement est-il sollicité quotidiennement? Dans quelles circonstances devez-vous faire appel à votre jugement?
2. Établissez-vous une distinction entre porter un jugement professionnel et prendre une décision? Si oui, laquelle?
3. SI je définis le jugement crédible comme étant « acceptable, recevable, par toutes les personnes concernées ». Cela implique qu'il peut être « acceptable » sans pour autant qu'il leur convienne. Est-ce que cette définition vous convient? Est-ce que vous la complèteriez, la modifieriez? Par quoi?
4. Comme mentionnez, vous pouvez décider que nous allons inverser l'ordre des situations. Ainsi, sans nous fournir de l'information nominale, pourriez-vous me décrire brièvement la situation positive afin que je comprenne le contexte?
 - a. Qu'est-ce qui vous fait dire que c'était un bon jugement (positif, adéquat), une bonne décision?
 - b. Est-ce que vous avez l'impression qu'il était malgré tout « acceptable par toutes les personnes concernées » ou non? Partageaient-elles votre point de vue? Pouvez-vous justifier votre réponse?
 - c. Pouvez-vous me raconter comment vous êtes parvenu à ce jugement, à cette décision? plus spécifiquement :
 - Vous êtes-vous basé sur un cadre de référence régissant votre pratique professionnelle? Si oui, lequel et comment vous a-t-il été utile?
 - Vous êtes-vous référé à des situations antérieures similaires?
 - Sur quelle-s information-s vous vous êtes appuyé pour y parvenir? Quelles étaient vos sources d'information?
 - Comment avez-vous traité cette information pour produire le jugement, la décision?

- Est-ce que quelqu'un d'autre vous a aidé à porter ce jugement, à prendre la décision? Si oui, quelle a été la nature de sa contribution? Son importance?
5. Toujours sans nous fournir de l'information nominale, pourriez-vous me décrire brièvement la situation négative afin que je comprenne le contexte?
- d. Qu'est-ce qui vous fait dire que c'était un mauvais jugement (négatif, inadéquat), une mauvaise décision?
 - e. Est-ce que vous avez l'impression qu'il était malgré tout « acceptable par toutes les personnes concernées » ou non? Partageaient-elles votre point de vue? Pouvez-vous justifier votre réponse?
 - f. Pouvez-vous me raconter comment vous êtes parvenu à ce jugement, à cette décision? plus spécifiquement :
 - Vous êtes-vous basé sur un cadre de référence régissant votre pratique professionnelle? Si oui, lequel et comment vous a-t-il été utile?
 - Vous êtes-vous référé à des situations antérieures similaires?
 - Sur quelle-s information-s vous vous êtes appuyé pour y parvenir? Quelles étaient vos sources d'information?
 - Comment avez-vous traité cette information pour produire le jugement, la décision?
 - Est-ce que quelqu'un d'autre vous a aidé à porter ce jugement, à prendre la décision? Si oui, quelle a été la nature de sa contribution? Son importance?

Nous allons maintenant comparer les deux situations :

6. Est-ce que vous pouvez identifier ce qui a été différent dans les deux cas et qui pourrait expliquer la différence entre les situations? Plus spécifiquement, en ce qui concerne :
 - a. L'information (la nature? Les modalités pour la recueillir? Les sources d'information?)
 - b. Les modalités pour trier, traiter et diffuser l'information
 - c. Considérez-vous qu'il y a de l'information, des façons de faire, des... qu'il faudrait valoriser parce qu'elles semblent efficaces? Si oui, lesquelles?
 - d. Considérez-vous qu'il y a de l'information, des façons de faire, des... qu'il faudrait éviter parce qu'elles constituent des embûches? Si oui, lesquelles?
7. En vous basant sur ce que nous venons de parler, pensez-vous qu'il se dégage une procédure « idéale » pour porter un jugement crédible? Prendre une décision? Et qui peut même s'avérer efficace dans la vie de tous les jours?

Ceci met fin aux questions que j'avais à vous poser. Est-ce qu'il y a quelque chose que vous aimeriez ajouter parce qu'il n'a pas été abordé par les questions? Suggestion : Y a-t-il autre chose que vous aimeriez ajouter, un point ou un aspect qui n'a pas été abordé? Avez-vous des questions?

J'aimerais finalement connaître votre réaction par rapport au déroulement de l'entrevue.

Je vous remercie pour votre collaboration

BIBLIOGRAPHIE

- Angers, P. et C. Bouchard. (1992). «Le jugement, les valeurs et l'action». In *La formation du jugement*, sous la dir. de M. Schleifer, p. 135-157. Montréal: Les éditions Logiques.
- Alkin, M.C. (2003). *Evaluation Theory and Practice: Insights and New Directions*. *New Direction for Evaluation*, no 97, p. 81-90.
- Alkin, M.C. (2004). *Evaluation Roots. Tracing Theorists' Views and Influences*. Thousand Oaks (Calif): Sage Publications.
- Alkin, M.C. et C.A Christie (2004). «An evaluation theory tree», dans M.C. Alkin (dir.), *Evaluation Roots : Tracing Theorist' views and Influences*, Thousand Oaks, SAGE Publication, p. 12-65.
- Arens, S. (2006). L'étude du raisonnement dans les pratiques évaluatives. *Mesure et évaluation en éducation*. vol. 29, no 3, p. 45-56.
- Blais, M., et S. Martineau. 2006. «L'analyse inductive générale: description d'une démarche visant à donner un sens à des données brutes». *Recherche qualitative*, vol.26. no 2, p.1-18.
- Bossiroy, A. 2009. «Mieux comprendre la pratique évaluative : les éléments qui contribuent à la production d'un jugement fondé et argumenté». Mémoire de maîtrise, Montréal, Université du Québec À Montréal.
- Boudreault, H. (2002). Vie pédagogique :
http://www.viepedagogique.gouv.qc.ca/numeros/124/vp124_9-11.pdf
- Boutin, G. 2008. *L'entretien de recherche qualitatif*. Québec: Presses de l'Université du Québec.

- Brief, J.C (1992). Le jugement : une vision périphérale. Dans, M. Schleifer, La formation du jugement. (pp.37-52). Les éditions Logiques : Montréal.
- Brousselle, A., Champagne, F., & Contandriopoulos, A.P. (2006.) Vers une réconciliation des théories et de la pratique de l'évaluation, perspective d'avenir. *Mesure et évaluation en éducation*, no spécial sur l'évaluation de programme, 29(3), 57-73.
- Champagne, F., Contandriopoulos, A-P., Brousselle, A., Hartz, Z., Denis, J-L. (2009). L'évaluation dans le domaine de la santé, concepts et méthodes. Sous la direction de Brousselle, A. Champagne, F., Contandriopoulos, A-P. et Hartz, Z. (pp. 35-56), Montréal : Les Presses de l'Université de Montréal.
- Chen, H.-T. (2005). *Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness*. Thousand Oaks (Calif.): Sage Publications.
- Datta, L-E. (2006). The practice of evaluation challenges and new directions. Dans . Shaw, I.F , Greene J.C., & Mark, M.M. (Eds), *The Sage Handbook of Evaluation* (pp. 419-438). Thousand Oaks: Sage Publications Inc.
- Donalson, S.I., Christie, C.A., & Mark, M.M. (Eds), (2009). *What Counts as Credible Evidence in Applied Research and Evaluation practice?* Los Angeles: Sage Publication.
- Donaldson, S. L, et M. W. Lipsey. 2006. «Roles for Theory in Contemporary Evaluation Practice : DeveJoping Practical Knowledge». In *The Sage Handbook of Evaluation*, sous la dir. de I. F. Shaw, I. C. Greene et M. M. Mark, p.56-75. Thousand Oaks (Calif.) : Sage Publications.
- Dubois, N. et Marceau, R. (2005). Un état des lieux théoriques de l'évaluation: une discipline à la remorque d'une révolution qui n'en finit pas. *La revue canadienne en évaluation de programme*, vol. 20, no 1, p. 1-36.
- Einhorn, HJ., Kleinmuntz, DN. et Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review*. APA.

- Fitzpatrick, J., Christie, C., & Mark, M.M. (2009). *Evaluation in Action*. Los Angeles : Sage Publications
- Fitzpatrick, J., Sanders, J.R., & Worthen, B.R. (2011). *Program evaluation. Alternative Approaches and Practical Guidelines*, 4th Ed, Boston, Pearson.
- Fournier, D. M. (1995). Establishing Evaluative Conclusions : A Distinction Between General and Working Logic. Dans Fournier, D.M., (Éd.), *Reasoning in Evaluation: Inferential Links and Leaps* (pp. 15-32). *New Direction for Evaluation*, 68. San Francisco: Jossey-Bass.
- Fournier, D., et Smith, N. L. (1993). Clarifying the Merits of Argument in Evaluation Practice. *Evaluation and Program Planning*, vol. 16, no 4, p. 315-323.
- Glaser, B. G. (2005). *The grounded theory perspective 3: Theoretical coding*. Mill Valley, CA: Sociology Press.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Pub. Co.
- Guba, E.G. (1972). The Failure of Educational Evaluation. Dans *Evaluating Action Programs: Readings in Social Action and Education*, sous la dir. de Weiss C. H., p. 251-266. Boston: Allyn et Bacon.
- Guillemette, F. 2006. «L'approche de la *Grounded Theory*; pour innover ?». *Recherches qualitatives*, vol. 26, no 1, p. 32-50.
- Guillemette, F., & Luckerhoff, J. (2012). *Méthodologie de la théorisation enracinée: Fondements, procédures et usages*. Québec [Que.: Presses de l'Université du Québec.

- Grinnell, F. (2009) *Everyday Practice of Science : Where Intuition and Passion Meet Objectivity and Logic*. New York : Oxford University Press.
- Houle, S. (2006). La relation entre la théorie et la pratique en évaluation de programme : dialogue et monologue. *Mesure et évaluation en éducation*, vol. 29, no 3, p. 1-3.
- House, E. R. 1993. *Professional Evaluation: Social Impact and Political Consequences*. Thousand Oaks (Calif.) : Sage Publications.
- Hurteau, M. (1991). Strategic Choices in Programs Evaluation. Dans A.E. Love. (Éd.), *Sourcebook in Evaluation*, (pp.106-125). Ottawa: Les publications de la Société canadienne en évaluation de programmes.
- Hurteau, M. (2009). Le jugement généré par l'évaluation de programme : une valeur sûre ? (présentation)
- Hurteau, M. (2009). CAR3400- Note du cours 4 [Présentation PowerPoint]. Récupéré le 26 octobre 2009 de <http://www.moodle.uqam.ca/>
- Hurteau, M., Houle, S. (2005a.) Évaluation de programme et recherche évaluative: des activités distinctes. *Mesure et évaluation en éducation*. vol. 28, no 3, p. 83-95.
- Hurteau, M., Houle, S. (2006). L'évaluation de programme: réflexion concernant la théorie de l'évaluation. *Actes du 19^e colloque international de l'ADMEE-Europe* (Université du Luxembourg, septembre 2006).
- Hurteau, M., Houle, S. et Guillemette, F. (2012). *L'évaluation de programme axée sur le jugement crédible*. Québec: Presses de l'Université du Québec.
- Hurteau M., Lachapelle G. & Houle S. (2006). Comprendre les pratiques évaluatives afin de les améliorer: la modélisation du processus spécifique à l'évaluation de programme. *Revue mesure et évaluation en éducation*, 29 (3), 27-

- Hurteau, M., Houle, S., & Mongiat, S. (2009). Do program evaluation practitioners generate legitimate and justified judgments? *Evaluation*. Vol 15(3): p. 307–319. Sage Publication
- Karsenti, T., Savoie-Zajc, L. (2004). *La recherche en éducation: étapes et approches*. Sherbrooke: Éditions du CRP.
- King, J. A. (2003). The Challenge of Studying Evaluation Theory. *New Directions for Evaluation*, no 97, p. 57-67.
- Lafortune, L., Allal, L. (2007). Le jugement professionnel de l'évaluation :Pratique enseignante au Québec et à Genève. PUQ : Québec
- Le Bortef., G (2002). Développer la compétence des professionnels. Paris : Éditions d'Organisation 2002.
- Legendre, R. (2005). *Dictionnaire actuel de l'éducation*. Montréal, Guérin.
- Lessard-Hébert, M., G. Goyette et G. Boutin. 1990. *Recherche qualitative : fondements et pratiques*. Coll. «Éducation». Montréal: Agence d'Arc.
- Lipman, M. 1992. «L'éducation au jugement». In *La formation du jugement*, sous la dir. de M. Schleifer, p. 99-123. Montréal: Les éditions Logiques.
- Madaus, G.F. et Stufflebeam, D.L. (2004). Program Evaluation: A Historical Overview . In *Evaluation Models : Viewpoints on Educational and Human Services Evaluation* (2e éd.), sous la dir. de Stufflebeam, D.L. Madaus, G.F. et Kellagan T., p. 3-19. Boston: Kluwer Academic Publishers.
- Marchand, M-P. 2011. «Élaboration et validation d'un questionnaire descriptive du processus inhérent au développement professionnel». Mémoire de maîtrise, Montréal, Université du Québec À Montréal.

- Mathison, S. M. (2005). *Encyclopedia of Evaluation*. Thousand Oaks (Calif.): Sage Publications.
- McDavid, J.C. (2001). Program Evaluation In British Columbia in a Time of Transition: 1995-2000. *The Canadian Journal of Program Evaluation*, vol.16, p. 3-28.
- Mintzberg, H. & Waters, J.A., (1995). Of strategy, deliberate and emergent. *Strategic management Journal*, 6, 257-272.
- Mongiat, S. (2007). Analyse des pratiques en évaluation de programme : les éléments qui contribuent à formuler un jugement fondé. *Mémoire*. Montréal (Québec, Canada), Université du Québec à Montréal, Maîtrise en éducation.
- Reboul, O. (1992). *Qu'est ce que le jugement?* Dans M. Schleifer, La formation du jugement, (Pp. 17-36). Les éditions logiques : Montréal.
- Rossi, P. H., Lipsey, M.W. et Freeman, H.E. (2004). *Evaluation : A Systematic Approach*, Thousand Oaks (Calif.): Sage Publications, 470 p.
- Schleifer, M. (1992). Le jugement et le raisonnement. Dans, M. Schleifer, La formation du jugement. (Pp. 53-60). Les éditions Logiques : Montréal
- Schwandt, T.A. (2002). *Evaluation Practice Reconsidered*. New York : Peter Lang.
- Schwandt, T.A. (2007). *Educating for Intelligent Belief in Evaluation: Actes du congrès annuel de l'Association Américaine d'Évaluation* (Baltimore).
- Scriven, M. (1980). *The Logic of Evaluation*. Inverness (Calif.): Edgepress.125 p.
- (1991). *Evaluation Thesaurus* (4^e éd.). Newbury Park (Calif.) : Sage Publications, 408 p.

----- (1993).« Hard-won Lessons in Program Evaluation. *New Directions for Evaluation*, no 58, p.1-103.

----- (1995). The Logic of Evaluation and Evaluation Practice. *New Direction for Evaluation*, no 68, p.49-70.

----- (2001). Evaluation: Future Tense. *American Journal of Evaluation*, vol. 22, no 3, p. 301-307.

----- (2006). Value-Driven Evaluation : Actes du congrès annuel de la Société Américaine d'Évaluation (Portland, novembre 2006). Acte du congrès électronique : <http://lg.evaluationcanada.ca>.

Secrétariat du Conseil du Trésor du Canada. (2004). *Examen de la qualité des évaluations dans les ministères et les organismes*, Document interne, Gouvernement du Canada (http://www.tbs-sct.gc.ca/eval/pubs/rev-exam_f.asp).

Silverman, D. (2000). *Doing qualitative research : a practical handbook*. SAGE.

Spector, J. M. (2014). *Handbook of research on educational communications and technology*. New York, N.Y: Springer.

Stake, R.E. (2004). *Standards-based & Responsive Evaluation*, Thousand Oaks: Sage Publications.

Stake, R. E., C. Migotsky, R. Davis, E. I. Cisneros, G. DePaul, C. Dunbar, R. Famer, I. Feltovich, E. Johnson, B. Williams, M. Zurita, et I. Chaves. 1997. «The Evolving Syntheses of Program Value». *Evaluation Practice*, vol. 18, no 2, p. 89-103.

Stake, R. E., et Schwandt, A. (2006). *On Discerning Quality in Evaluation*. In *The Sage Handbook of Evaluation*, sous la dir. de I. F. Shaw, J. C. Greene, et M. M. Mark, p. 404-418. Thousand Oaks: Sage publications.

- Stake, R.E. (2010). *Qualitative Research. Studying how things work*. The Gilford Press.
- Stern, E. (2006). *Contextual challenges for evaluation practice*. Dans I.F. Shaw, J.C. Greene & M.M. Mark (éds), *The Sage Handbook of Evaluation* (pp. 292-314). Thousand Oaks: Sage Publications Inc.
- Strauss, A. et Corbin, J. (1998). *Basics of qualitative research : Techniques and Procedures for Developing Grounded Theory*. SAGE publications.
- Thomas, D. R. 2006. «A General Inductive Approach for Analyzing Qualitative Evaluation Data». *American Journal of Evaluation*, vol. 27, p. 237-246.
- Toulemonde, J. (2005). *Appropriation des résultats de l'évaluation: Leçons de la pratique en Région Limousin: Acte du colloque de la Société Française de l'Évaluation* (mai 2005).
- Virtanen, P., Uusikylä, P. (2002). Exploring the Missing Links Between Causes and Effects. Conférence annuelle de la Société européenne d'évaluation, Séville.