

Université du Québec en Outaouais

**Standardisation des groupes de discussion sans leader :
l'impact de la nature des interventions et du nombre de participants**

par
Sabrina Poirier

Département de relations industrielles

Mémoire présenté dans le cadre du grade de
M. Sc. en relations industrielles et ressources humaines

Décembre 2005

Sommaire

Afin de sélectionner des employés possédant les caractéristiques essentielles propres à un poste à doter, il est nécessaire de choisir des moyens d'évaluation adaptés tant aux déterminants de cet emploi qu'aux variables organisationnelles de temps et de ressources financières et humaines disponibles pour y parvenir. Le groupe de discussion sans leader (GDSL) est un outil d'évaluation fréquemment utilisé puisqu'il démontre des avantages incontournables, par exemple en permettant d'évaluer simultanément plusieurs candidats à la fois, en plus d'offrir une économie de temps et de ressources impliquées. Par ailleurs, il permet de recréer un milieu social donnant naissance à des comportements de la part des candidats qu'une entrevue et un test écrit n'auront pas tendance à susciter. Toutefois, comme la dynamique de groupe varie d'une administration à l'autre en fonction des participants qui composent le groupe, le GDSL comporte une portion d'évaluation non standardisée qui risque d'affecter la qualité psychométrique des résultats qu'il produit. Il est donc légitime de se questionner sur l'effet de la nature des interventions et du nombre de participants à un GDSL sur les résultats d'une personne candidate à un processus de sélection. Par le biais d'une estimation de la fidélité par la méthode des formes parallèles, 53 sujets provenant d'organisations scolaires et prenant part à un véritable processus de sélection ont permis de mesurer la stabilité des résultats d'un GDSL à l'autre. En fonction de l'évaluation d'une série de huit compétences, l'élaboration de quatre groupes de discussion faisant varier tour à tour la nature des interventions (positives vs. négatives) et le nombre de participants (deux participants vs. quatre participants), contrôlés par le biais d'acteurs professionnels, a permis de démontrer que ces deux variables n'introduisent pas de façon systématique un biais préjudiciable pour un candidat prenant part à un processus de sélection.

Liste des tableaux

Tableau 1 Analyse comparative de diverses méthodes de sélection du personnel pour quatre critères d'évaluation.....	11
Tableau 2 Fréquence d'utilisation des différents exercices composant un centre d'évaluation.....	13
Tableau 3 Processus général d'élaboration d'instruments de mesure fondés sur la validation de contenu.....	23
Tableau 4 Évidence des dimensions observables dans le cadre de GDSL avec rôle assigné et sans rôle assigné.....	28
Tableau 5 Indices de validité critérielle des GDSL.....	38
Tableau 6 Coefficients de fidélité interexamineurs des GDSL.....	44
Tableau 7 Les groupes restreints comme un système de traitement de l'information.....	51
Tableau 8 Répartition des comportements dans un groupe selon trois pôles fonctionnels.....	60
Tableau 9 Les rôles fonctionnels dans les groupes restreints.....	62
Tableau 10 Les catégories d'interventions de Bales.....	65
Tableau 11 Échelle d'évaluation des compétences.....	71
Tableau 12 Exemple d'une compétence évaluée.....	79
Tableau 13 Schème expérimental.....	82
Tableau 14 Répartition des évaluateurs en fonction des candidats et des GDSL.....	83

Tableau 15	
Analyse descriptive des résultats en fonction de la nature des interventions.....	88
Tableau 16	
Analyse descriptive des résultats en fonction de la nature du nombre de participants.....	90
Tableau 17	
Coefficients de corrélation de Pearson en fonction de la nature des interventions.....	92
Tableau 18	
Coefficients de corrélation de Pearson en fonction du nombre de participants.....	94
Tableau 19	
Coefficients de fidélité inter-examineurs.....	95
Tableau 20	
Analyse de variance par les rangs.....	97
Tableau 21	
Test de Wilcoxon pour la compétence Mobilisation d'équipe.....	98
Tableau 22	
Test de Wilcoxon pour la compétence Orientation vers les résultats	98

Liste des figures

Figure 1	
Utilisation du GDSL dans le cadre d'un processus de sélection.....	69

Remerciements

Un merci tout spécial à Monsieur André Durivage. Pour son écoute, ses conseils et sa façon de transmettre son savoir, je le remercie d'avoir dirigé mes efforts vers l'accomplissement de ce mémoire.

Par ailleurs, une étude de cette envergure n'aurait pu se réaliser sans l'appui et l'implication d'une multitude de personnes. Je les remercie tous sincèrement de nous avoir appuyer dans cette aventure.

Finalement, merci à Jacques, à ma famille et à mes amis pour leurs encouragements et leur soutien jusqu'à la dernière ligne.

Sommaire	i
Liste des tableaux	ii
Liste des figures	iv
Remerciements	v
Introduction	1
PARTIE 1 Recension des écrits	3
Chapitre 1 Le GDSL : un outil d'évaluation en contexte de sélection	3
1.1 La nature du centre d'évaluation	6
1.2 Le groupe de discussion sans leader	13
1.2.1 La définition du GDSL	14
1.2.2 Les origines de la technique d'évaluation	14
1.2.3 Les principales caractéristiques	16
1.2.4 Les étapes du développement de l'exercice	22
1.2.5 L'évaluation des critères	26
Chapitre 2 Évaluer la qualité psychométrique d'un outil d'évaluation	33
2.1 La validité	34
2.1.1 La validité des centres d'évaluation	37
2.1.2 La validité des GDSL	38
2.2 La fidélité	39
2.2.1 La fidélité des centres d'évaluation	43
2.2.2 La fidélité des GDSL	43
2.3 Les considérations éthiques et légales	45
2.4 La qualité psychométrique des GDSL	46
Chapitre 3 La dynamique de groupe	48
3.1 La définition du concept de groupe	49
3.2 Les intrants	52
3.2.1 Les caractéristiques des membres	52
3.2.2 La tâche à accomplir	55
3.2.3 La taille du groupe	56
3.3 Les processus	59
3.3.1 Les comportements	59
3.3.2 Les rôles	61
3.3.3 Le climat	63
3.3.4 Les interactions	64
PARTIE 2 Question de recherche	67
Chapitre 4 Modèle théorique	67
4.1 Le modèle conceptuel et ses variables	68
4.1.1 Variable dépendante	70
4.1.2 Variables indépendantes	71
4.2 Hypothèses de recherche	72
4.3 Implications théoriques et empiriques	73
4.4 Stratégie de recherche	74
PARTIE 3 Méthodologie	76
Chapitre 5 Opérationnalisation des variables	76
5.1 Étape 1 : Finalités et spécification du domaine à mesurer	77
5.1.1 Finalités de l'instrument de mesure	77

5.1.2	Analyse et description de l'emploi	77
5.1.3	Spécification du domaine à mesurer	78
5.2	Étape 2 : Développement des instruments de mesure	79
5.3	Étape 3 : Implantation et enregistrements des données.....	81
5.3.1	Schème expérimental	81
5.3.2	Scénario des acteurs	84
5.3.3	Évaluation des compétences.....	85
5.3.4	Description de la population cible.....	85
5.3.5	Engagement déontologique et considérations éthiques.....	86
PARTIE 4	Analyses de données	87
Chapitre 6	Analyses statistiques.....	87
6.1	Statistiques descriptives	87
6.2	Comparaison des coefficients de corrélation.....	91
6.3	Fidélité inter-examineurs.....	95
6.4	Analyse de variance par les rangs	96
Chapitre 7	Discussion.....	100
7.1	Vérification des hypothèses	102
7.1.2	Hypothèses sur l'effet de la nature des interventions	102
7.1.2	Hypothèses sur l'effet du nombre de participants.....	104
7.1.3	Autres considérations.....	105
7.2	Contribution théorique et pratique	107
7.3	Forces et limites de la recherche	108
7.4	Pistes de recherche	110
Conclusion		113
Bibliographie		115

Introduction

Dans un contexte de sélection de personnel, le choix des outils d'évaluation revêt une très grande importance. Que ce soit pour respecter les contraintes administratives ou législatives, pour une représentation valide du rôle et des responsabilités de l'individu au sein de l'organisation ou pour la qualité psychométrique de l'instrument de mesure, différents facteurs doivent être considérés par le responsable d'un processus de sélection. Le groupe de discussion sans leader (GDSL) existe depuis plus de cinquante ans et continue d'être largement utilisé comme outil d'évaluation en contexte de sélection (Thorton, 1992). Il peut prendre différentes formes, porter sur différents thèmes et impliquer un nombre variable de participants. Essentiellement, cet exercice consiste à placer en interaction directe les personnes candidates à un poste afin qu'ils discutent d'une problématique et en arrivent à une solution commune (Balicco, 1997). Cette discussion permet alors à des évaluateurs d'observer des comportements reliés à des caractéristiques personnelles ou à des habiletés qui sont souvent difficilement observables par le biais de d'autres outils d'évaluation.

Bien que les GDSL offrent des avantages importants, ils soulèvent cependant des questions quant à leur validité. Par exemple, les différences individuelles entre les personnes candidates à un processus de sélection impliquent et suscitent des comportements, des rôles, des interactions et un climat de discussion différents de l'administration d'un GDSL à l'autre. Ces facteurs, tout comme le nombre de participants à l'intérieur du groupe, peuvent influencer en retour la dynamique de la discussion.

L'objectif du présent projet de recherche est donc d'évaluer l'impact de ce manque standardisation potentiel des GDSL sur les résultats des personnes candidates en situation réelle de sélection. De façon plus spécifique, ***est-ce que les résultats des personnes candidates aux groupes de discussion sans leader varient en fonction de la nature des interventions et du nombre de participants et ce, dans un contexte de sélection ?***

PARTIE 1 Recension des écrits

Chapitre 1 Le GDSL : un outil d'évaluation en contexte de sélection

Les organisations doivent désormais conjuguer avec une multitude de données nouvelles imposées par les profondes transformations de leur environnement. Les coûts importants des ressources humaines, la rapidité et la complexité des changements technologiques, la concurrence accrue, les méthodes de gestion de plus en plus décentralisées et les signes d'insatisfaction observés sur les lieux du travail suscitent une réflexion sur l'importance du potentiel humain et le choix des employés nécessaires au fonctionnement des organisations. Dans cette optique, l'acquisition d'un personnel compétent et l'affectation des employés à des postes où ils seront efficaces et satisfaits sont des activités qui peuvent contribuer tant au succès économique et social des organisations qu'à combler certains des besoins de la main-d'œuvre du XXI^e siècle (Pettersen, 2000).

En matière d'évaluation et de sélection de personnel, ces activités impliquent de recueillir de l'information sur les personnes ayant posé leur candidature à un poste dans le but de prendre une décision d'embauche ou de rejet. Pour les personnes candidates à un poste, la sélection est comparable à une course dans laquelle il faut franchir divers obstacles avec l'espoir d'arriver en première position. Pour les organisations, la gestion des activités de sélection consiste surtout à mettre en place des stratégies et des mécanismes qui les aideront à prédire le succès au travail et ce, avec le plus de précision possible (St-Onge, Audet, Haines et Petit, 1998). Afin d'atteindre cet objectif, il est nécessaire d'utiliser des instruments de sélection efficaces

qui vont permettre l'identification des différences sur le plan des compétences des personnes.

Plusieurs avantages sont reliés à la conduite d'un processus de sélection basé sur la mesure et l'évaluation des caractéristiques individuelles des personnes. En fait, selon Spencer et Spencer (1992), de tels processus standardisés permettent, d'une part, de contrer :

- **La faible performance ou productivité d'un employé :** la mesure préalable des compétences et des facteurs permettant de prédire la performance en emploi des futurs employés est une façon pour l'organisation de demeurer concurrentiel ou même d'obtenir un avantage concurrentiel;
- **Le taux élevé de roulement de personnel et le faible taux de rétention des employés :** un employé qui maîtrise les compétences nécessaires pour réussir en emploi fait bien son travail et n'a pas à craindre d'être congédié. De plus, des employés satisfaits sont moins portés à quitter leur travail;
- **Les longues courbes d'apprentissage dans le temps :** Les employés qui ont les compétences nécessaires pour occuper un emploi sont productifs plus rapidement et ont donc moins besoin d'une longue période de formation.

D'autre part, toujours selon Spencer et Spencer (1992), de tels processus permettent également :

- **De déterminer les besoins de formation à l'entrée :** il est possible de déterminer, en fonction de l'individu embauché et des compétences requises pour le poste, la formation qui devra lui être fournie à son entrée en poste.

- **De planifier la relève :** les gestionnaires ont davantage une vision à long terme des employés qui peuvent devenir les gestionnaires de demain et assurer la relève de l'organisation.

Cependant, prendre en considération les différences individuelles dans un processus de sélection n'est pas toujours simple et la mesure et l'évaluation de ces caractéristiques doit être réalisée d'une façon juste, équitable et valide, tout en respectant les contraintes législatives et administratives prévues dans la convention collective s'il y a lieu (Pettersen, 2000). Pour y parvenir, il existe différents outils d'évaluation et, parmi ces moyens, les plus couramment utilisés sont l'entrevue de sélection, les échantillons de travail, les tests d'aptitudes intellectuelles ou mécaniques, les tests de personnalité, d'intérêts, de connaissances et d'habiletés et les centres d'évaluation qui combinent différents exercices de simulation, tel que le GDSL.

Au-delà de la valeur scientifique d'un instrument de mesure, il s'agit pour une organisation de vérifier si l'utilisation d'un outil d'évaluation est justifiée par les enjeux de gestion qui s'y rattachent (Pettersen, 2000). D'une part, puisque les coûts des activités reliées au recrutement d'un bassin de nouvelles candidatures sont de plus en plus élevés pour une organisation, il est important que l'outil d'évaluation utilisé pour discriminer les personnes candidates donne des résultats prédictifs de leur rendement futur en emploi. D'autre part, une contestation légale peut obliger la démonstration de la qualité de l'outil d'évaluation et ce, dans le but d'attester du traitement équitable des personnes candidates au processus de sélection.

Aux fins de ce projet de mémoire, il est nécessaire de situer d'abord le contexte dans lequel le GDSL est le plus couramment utilisé, soit dans le cadre d'un centre d'évaluation. En fait, en exposant les paramètres du centre d'évaluation, il sera possible d'avoir une meilleure compréhension des origines et du cadre d'application initial du GDSL. Par la suite, la définition du GDSL, ses différents contours opérationnels et administratifs ainsi que ses avantages et inconvénients seront présentés afin de mieux comprendre les enjeux de cet outil d'évaluation en contexte de sélection.

1.1 La nature du centre d'évaluation

Le centre d'évaluation est principalement caractérisé par une technique d'évaluation multiple impliquant plusieurs évaluateurs indépendants, des échantillons de travail différents et une démarche structurée permettant d'observer, d'apprécier et d'effectuer des prédictions sur des comportements de gestion (Thorton et Byham, 1982, Gatewood et Field, 1994, *International Task Force on Assessment Center*, 2000). Dans cette optique, le centre d'évaluation consiste en une procédure détaillée avec plusieurs activités d'évaluation écrites et interactives où certaines parties des fonctions du poste à doter sont simulées. Les différents exercices composant le centre d'évaluation, ainsi que les problématiques à traiter par les personnes candidates, sont choisis et arrimés selon les réalités spécifiques d'un poste et visent à mesurer les caractéristiques requises pour l'emploi ou la fonction particulière.

Ce sont les Allemands qui ont été les premiers à appliquer l'approche de l'évaluation multiple au début des années 1930. Le concept s'est ensuite développé

pendant la Deuxième guerre mondiale, lorsque les Allemands, les Anglais et les Américains se sont servis des données sur l'observation et l'évaluation psychologique pour la sélection du personnel destiné aux forces armées et aux services secrets (Yan et Slivinski, 1976, Levy-Leboyer, 1996, Moulinier, 1998,). À cette époque, les psychologues cliniques et les psychiatres reposaient leur évaluation sur l'individu dans son ensemble, plutôt que selon ses qualités personnelles ou ses habiletés distinctes. De cette manière, les mesures psychologiques et les entrevues étaient très peu structurées et standardisées, et les données étaient souvent intégrées selon une intuition clinique (Highhouse, 2002).

La première application industrielle des centres d'évaluation remonte aux années 1950 avec le « *Management Progress Study* » de la compagnie AT&T (Thorton et Byham, 1982). Selon Levy-Leboyer (1996), l'étude longitudinale réalisée par Bray et ses collaborateurs (1966 et 1974) à l'*American Telephone and Telegraph Company* est certainement à l'origine du succès et de la multiplication des centres d'évaluation aux États-Unis. Pour Highhouse (2002), le centre d'évaluation est né de la nécessité d'évaluer les aspects non intellectuels de la performance humaine. Les exercices de type situationnel dans le cadre d'un centre d'évaluation ont également été élaborés pour apporter une plus grande standardisation autant par rapport aux conditions d'évaluation que pour l'attribution de notes (Flanagan, 1954 dans Highhouse, 2002).

Depuis, le centre d'évaluation est un outil largement mis à contribution dans le cadre de différentes activités reliées aux ressources humaines et demeure un outil de sélection bien connu et très utilisé (Motowildo, Dunette et Carter, 1990, Appelbaum, Harel et Shapiro, 1998, Lievens, 2002). C'est pourquoi, en 1975, un comité de travail international réunissant plusieurs spécialistes des centres d'évaluation a développé des

directives précises concernant l'application de cette démarche (*The Guidelines and Ethical Considerations for Assessment Center Operations*) (Cladwell, Thornton III et Gruys, 2003). Ces directives, révisées en 1989 et en 2000, visent surtout à assurer que les psychologues industriels et organisationnels, les professionnels en ressources humaines, les consultants ainsi que la haute direction ont une compréhension claire des procédures, des définitions et des concepts sur lesquels le centre d'évaluation est basé (*International Task Force on Assessment Center, 2000*).

Les centres d'évaluation servent habituellement à l'identification des aptitudes à la gestion dans le cadre de processus où l'organisation désire obtenir un profil clair d'habiletés ou de traits directement reliés à la réussite dans un poste de gestionnaire. Thornton et Byham (1982) ont établi une liste des raisons pour lesquelles les centres d'évaluation sont utilisés par les organisations :

- Diagnostique des besoins de formation et de développement;
- Diagnostique et prédiction des habiletés de gestion à l'intérieur d'une stratégie organisationnelle de changement;
- Évaluation de l'efficacité des programmes de formation;
- Évaluation du rendement des maisons d'enseignements;
- Évaluation de l'efficacité du recrutement et des activités de sélection collégiales;
- Évaluation des étudiants comme critère de graduation;
- Sélection et orientation de candidats à des postes de gestion supérieure ou à des postes sans gestion.

De façon générale, la démarche administrative d'évaluation du centre d'évaluation est plus ouverte que celle reliée à d'autres méthodes d'évaluation de personnel (Thorton et Byham, 1982). Selon Cosner et Baumgart (2000), la démarche d'évaluation suppose que chaque observateur surveille un ou deux participants puis compile ses observations en fonction d'une échelle de comportements attendus standardisée. Par la suite, chaque observateur trie et classe sous les critères appropriés ses observations et il attribue une cote à chaque critère. À la fin de la ou des journées d'évaluation, chaque observateur fait un compte rendu aux autres observateurs présentant une vue d'ensemble du déroulement de l'exercice, une description trait par trait des observations notées et un résumé concis de son évaluation. Les évaluateurs discutent alors de ces observations et en arrivent à un consensus sur les différents critères évalués. Le responsable du centre d'évaluation rédige ensuite un rapport détaillé des appréciations émises par les observateurs.

Selon Thorton et Byham (1982), l'efficacité des centres d'évaluation peut être résumée comme suit :

- Les exercices utilisés sont conçus spécifiquement pour mettre en évidence les capacités et les aptitudes spécifiques au poste pour lequel un groupe de candidats est évalué;
- Les exercices standardisés permettent aux observateurs d'évaluer les candidats dans des conditions relativement constantes afin d'effectuer des comparaisons valides;
- Les évaluateurs ne connaissent pas personnellement les candidats ce qui évite les biais découlant d'une implication émotionnelle;

- Les procédures forcent les évaluateurs à centrer leur attention sur les catégories de comportement qu'ils doivent observer pour évaluer un groupe de candidat;
- Les évaluateurs ont été formés spécifiquement à l'observation et à l'évaluation de ces catégories de comportement.

L'efficacité du centre d'évaluation peut également être expliquée par son aspect de simulation des fonctions du poste visé. En effet, plusieurs études ont démontré que le meilleur indicateur permettant de prédire des comportements futurs dans une situation d'emploi donnée sont les comportements adoptés en situation similaire (Motowildo et al., 1990). Dans ce contexte, les recherches tendent à démontrer que les centres d'évaluation permettent de prédire l'accès à des promotions à l'intérieur des organisations de même que le rendement en emploi dans environ 75 % des cas (Appelbaum et al., 1998). Finalement, Lievens (2002) avance que le principal avantage de l'approche par centre d'évaluation est que les participants obtiennent un rapport de rétroaction détaillé sur leurs forces et faiblesses.

Selon Pettersen (2000), pour mieux saisir la valeur d'un outil de sélection utilisé dans une situation donnée, quatre critères d'évaluation doivent être considérés, soit la validité, les coûts d'utilisation, la validité apparente et les implications légales. Puisqu'il est nécessaire d'analyser ces facteurs lors du choix d'une méthode d'évaluation en processus de sélection, le **Tableau 1** présenté à la page suivante est tiré de l'ouvrage de Pettersen (2000) et expose une analyse comparative de diverses méthodes de sélection du personnel selon ces critères. Cette compilation permet d'apprécier la valeur relative du centre d'évaluation en contexte de sélection.

TABLEAU 1
Analyse comparative de diverses méthodes de sélection du personnel
pour quatre critères d'évaluation ¹

Méthode de sélection	Validité critériée approximative	Coûts d'utilisation	Validité apparente	Facilité à être défendue légalement
Échantillons de travail	0,54	Modérés à élevés	Très élevée	Très élevée
Test d'aptitude mentale générale	0,51	Faibles	Faible	Faible à modérée
Entrevues structurées	0,51	Modérés à élevés	Élevée à très élevée	Élevée à très élevée
Examen de connaissances	0,48	Modérés	Très élevée	Très élevée
Entrevues non structurées	0,38	Faibles à modérés	Faible à modérée	Faible à modérée
Centres d'évaluation	0,37	Très élevés	Élevée à très élevée	Élevée à très élevée
Données biographiques	0,35	Modérés à élevés	Variable	-
Inventaire de personnalité	0,31	Faibles	Faible à modérée	Faible à modérée
Vérification de références	0,26	Faibles	Faible à modérée	-
Expérience	0,18	Faibles	Élevée	-
Inventaire des intérêts	0,10	Faibles	Modérée à élevée	-

A la lecture de ces informations, il est possible de constater que le coefficient de validité critériée du centre d'évaluation demeure peu élevé et que les coûts d'utilisation sont très élevés en comparaison avec d'autres méthodes d'évaluation. Malgré tout, puisque les participants et la haute direction valorisent cette démarche structurée et que le risque de devoir reprendre un processus en entier à la suite d'un recours légal est plutôt faible, il est peut être parfois très justifié d'utiliser le centre d'évaluation en contexte de sélection.

¹ Tiré de l'ouvrage : Pettersen, N. (2000). Évaluation du potentiel humain dans les organisations : Élaboration et validation d'instruments de mesure. Sainte-Foy, Presses de l'université du Québec, page 20.

En ce qui concerne les différentes activités utilisées dans le cadre de l'administration d'un centre d'évaluation, il n'y a pas deux centres qui utilisent exactement les mêmes exercices. Par contre, les points communs entre la plupart des centres d'évaluation sont les simulations (individuelles ou de groupe), les études de cas, les paniers de gestion, les tests psychométriques, les entrevues et les inventaires biographiques (Blaney, Hobon, Meade et Scodro, 1993, Cosner et Baumgart, 2000). Selon Joiner (2002), les tests de jugement situationnel (*Situational Judgement Tests*) constituent une nouvelle approche qui est de plus en plus populaire lors de l'administration d'un centre d'évaluation. Par le biais de ces exercices écrits, la personne candidate doit indiquer, parmi une liste d'actions présentées, la meilleure (ou la mauvaise) façon de réagir face à une mise en situation donnée.

À titre indicatif, le **Tableau 2** présenté à la page suivante démontre la fréquence d'utilisation des différents exercices lors de l'administration d'un centre d'évaluation. Cette liste a été développée à la suite d'une recension d'environ 500 centres d'évaluation avec lesquels les auteurs ont été associés (Thorton et Byham, 1982). Les résultats démontrent que le groupe de discussion sans leader (v.f. *Leaderless Group Discussion*) est utilisé, en moyenne, dans 65% des administrations d'un centre d'évaluation, donc plus d'une fois sur deux. Il est à noter que ces centres d'évaluation n'ont pas été seulement utilisés pour des activités de sélection de personnel et que les données ne sont pas aussi récentes qu'il serait souhaité, toutefois, l'étendue de l'échantillon représente tout de même une justification de la pertinence de la présente recherche.

TABLEAU 2
Fréquence d'utilisation des différents exercices composant un centre d'évaluation²

Exercices d'évaluation	% de la fréquence d'utilisation
Panier de gestion	95%
Groupe de discussion sans leader (avec rôle assigné)	85%
Exercice de simulation	75%
Groupe de discussion sans leader (sans rôle assigné)	45%
Exercice de planification	40%
Exercice d'analyse	35%
Jeu de gestion	10%
Entrevue sur les expériences passées	5%
Test d'aptitudes intellectuelles	2%
Examen de lecture	1%
Mathématique et arithmétique	1%
Personnalité	1%
Test de projections	1%

1.2 Le groupe de discussion sans leader

Même si le centre d'évaluation a démontré sa popularité et son caractère valide en sélection de personnel depuis les cinquante dernières années (Gaugler, Rosenthal, Thornton III et Bentson, 1987), il n'en demeure pas moins qu'il s'agit d'un outil d'évaluation très exigeant sous l'angle du temps ainsi que des ressources humaines et financières impliquées. En effet, la nature exhaustive d'un centre d'évaluation rend souvent les coûts liés à son implantation et à son opération difficiles à justifier au sein d'une organisation (O'Connell, Hattrup, Doverspike et Cober, 2002). De plus, les difficultés liées à la disponibilité de la main-d'œuvre ont suscité de la part des organisations une réflexion pour trouver des techniques qui réduisent le temps associé au recrutement et à la sélection du personnel (O'Connell et al., 2002). Dans ce cadre, une des possibilités qui s'offrent aux organisations afin de réduire le temps et les coûts

² Tiré de l'ouvrage : Thornton, G. et Byham, W. (1982). *Assessment Centers and Managerial Performance*. Florida, Academic Press inc., page 148.

reliés à l'administration d'un centre est d'utiliser les différents exercices, autant le groupe de discussion que les paniers de gestion et les simulations individuelles, de façon individuelle comme méthode d'évaluation et de sélection à part entière.

1.2.1 La définition du GDSDL

Le groupe de discussion sans leader (GDSDL) est un exercice collectif d'évaluation qui implique un nombre variable de participants (généralement de quatre à huit participants) qui doivent discuter entre eux et arriver à une décision de groupe à l'intérieur d'une période de temps spécifique (Balicco, 1997, Thorton et Byham, 1982). Le GDSDL peut prendre différentes formes et porter sur différents thèmes. Le terme sans leader fait référence au fait qu'il n'y a personne de nommé officiellement afin de diriger la prise de décision, d'animer ou de structurer la discussion. Les personnes candidates sont donc placées sur un pied d'égalité (Gatewood et Feild, 1994). La discussion et la contribution personnelle de chacune permettent d'observer plusieurs comportements reliés à leurs capacités et leurs qualités interpersonnelles selon les critères jugés importants pour le poste à doter.

1.2.2 Les origines de la technique d'évaluation

Selon Ansbacher (1951), cité dans Thorton et Byham (1982), la première utilisation du GDSDL remonte au programme d'évaluation militaire des Allemands. Cet auteur accorde la création de cette technique d'évaluation à J. B. Rieffert, directeur de la psychologie militaire allemande de 1920 à 1931. En fonction de ses écrits, une procédure d'évaluation basée sur un processus de table ronde, impliquant plusieurs participants, aurait été utilisée à partir de 1925 pour sélectionner les officiers de l'armée. Dans ce contexte d'évaluation, le GDSDL était surtout utilisé pour observer le

comportement envers les pairs et la résolution de problèmes. Puis, l'armée allemande aurait abandonné l'utilisation du GDSL vers la fin des années 1930, mais il semble que la marine allemande aurait modifié légèrement la technique d'évaluation collective et aurait continué à l'utiliser jusqu'à la fin de la Deuxième Guerre mondiale. À l'instar des responsables des processus de sélection de l'armée, ceux de la marine auraient cherché à introduire des sujets de discussion beaucoup plus controversés afin d'observer des comportements sociaux suscitant davantage de compromis, de conciliation et d'objectivité (Ansbacher, 1951 dans Thorton et Byham, 1982).

Pour Highhouse (2002), la méthode de groupe de discussion sans leader a été conçue par W.R. Bion en 1946, pionnier de l'approche anglo-analytique de l'évaluation individuelle. Bion et ses collègues n'acceptaient pas la notion voulant que les traits de personnalité soient des qualités constantes d'une personne qui existeraient indépendamment du contexte dans lequel la personne se trouve. Au lieu de cela, ils pensaient que la personnalité, organisée comme un tout, devait être évaluée dans des situations réelles où la configuration totale des traits était nécessaire pour réussir socialement. À partir de cette conception, les tests de groupe sans leader étaient plus orientés vers l'action. Par exemple, un groupe de six à huit candidats pouvait recevoir comme directive de construire un pont ou de déplacer un objet lourd à travers une série d'obstacles (Highhouse, 2002). Bion voyait dans ce type d'activité la production d'une pression pour le candidat qui devait tenter d'entretenir de bonnes relations interpersonnelles avec les autres participants et ce, tout en considérant ses propres intérêts dans l'accomplissement de la tâche (Bion, 1946 dans Highhouse, 2002).

D'un point de vue industriel, le GDSL a joué un rôle important très tôt dans plusieurs programmes d'évaluation. Entre 1948 et 1954, Bass a complété une série de

recherches démontrant en détail l'utilisation du GDSL dans plusieurs organisations de l'industrie anglaise, australienne et américaines (Bass, 1950, 1954). En fait, jusqu'en 1978, le GDSL était l'exercice le plus populaire lors de l'administration d'un centre d'évaluation (Thorton et Byham , 1982).

1.2.3 Les principales caractéristiques

L'administration d'un GDSL ressemble à un exercice de simulation ou une étude de cas, de type échantillon de travail, utilisés également dans le cadre de centres d'évaluation, à la différence près que le GDSL est le seul exercice d'évaluation de nature collective. Il est également possible de comparer le groupe de discussion à un exercice de présentation orale où toutes les personnes candidates doivent effectuer un exposé et faire valoir leur point de vue auprès de leurs collègues.

Selon Tixier (1974), il est possible d'utiliser un GDSL *libre* dans le cadre d'un processus d'évaluation. Dans cette optique, les personnes candidates doivent décider elles-mêmes du sujet dont ils débattront au cours de la période de discussion. Même si cette méthode d'évaluation peut comporter certains avantages, comme de faire ressortir les phénomènes de leadership, d'initiative et de prise de décision, il semble que cette méthode ne permettrait pas d'avoir, dans un minimum de temps, un éclairage précis sur le comportement et les différences individuelles de chacune des personnes candidates.

Il semble donc plus opportun de retenir l'administration d'un outil d'évaluation de nature situationnelle de type *étude de cas*, dans le cadre duquel les personnes candidates reçoivent tout d'abord de la documentation écrite sur une mise en situation fictive (Tixier, 1974). Les documents peuvent être de la correspondance provenant de

l'intérieur ou de l'extérieur de l'organisation, des sondages sur la satisfaction de la clientèle ou des ressources humaines, des dossiers sur des projets d'innovation ou de développement de nouveaux marchés, etc. Les personnes candidates disposent alors d'une période de préparation individuelle afin d'identifier le contexte, spécifier les enjeux et dégager les orientations principales du problème, selon la nature du thème soumis et les caractéristiques de la tâche à accomplir. Lors de cette période, elles se préparent donc à discuter avec les autres participants.

Le **temps alloué**, autant pour la préparation que pour la discussion, est variable et dépend de beaucoup d'autres facteurs comme de la complexité du sujet soumis, du nombre de critères à évaluer, des considérations pratiques ou financières de l'organisation et, finalement, de l'utilisation ou non de d'autres outils d'évaluation. Selon Gatewood et Fields (1994), dans la plupart des cas, une heure et demie à deux heures est la période de temps maximum qui sera allouée à l'administration de cet outil. Finalement, les considérations relatives à la période de temps allouée peuvent être également reliées au besoin de structure qui sera nécessaire pour arriver à une prise de décision commune (Wilson, 1996).

Le GDSL peut prendre différentes formes et le format privilégié influencera grandement la nature du thème soumis aux participants aux fins de la discussion. D'une part, la tâche à réaliser peut impliquer soit une interaction basée sur la **compétition** ou la **collaboration**. La compétition est souvent présente lorsque les participants ont à leur disposition des ressources limitées afin d'en arriver à un consensus ou qu'ils ont des rôles assignés afin de faire évoluer la discussion. Les problèmes impliquant une compétition entre les participants sont souvent caractérisés par le fait que les ressources allouées ne permettront pas de combler tous les désirs et intérêts des

participants (Gatewood et Fields, 1994). Par exemple, chaque participant peut avoir comme mandat d'obtenir une augmentation de salaire pour un de ses employés ou comme tâche d'obtenir une plus large partie du budget pour financer son service (Thorton et Byham, 1982). L'approche de collaboration est, à contrario, présente lorsque les participants n'ont pas de rôle assigné et qu'ils se rassemblent pour trouver une solution à une problématique commune. Ce type de groupe de discussion est surtout centré sur une approche de prise de décision. Un problème décrivant une situation où l'organisation doit élaborer un plan d'action pour répondre à une accusation de mesures discriminatoires dans l'embauche du nouveau personnel ou par rapport à un manque de contrôle sur des agents pollueurs seraient des exemples de discussions orientées vers des actions de collaboration (Gatewood et Fields, 1994).

De plus, les personnes candidates peuvent avoir des **rôles assignés ou non** par rapport à la tâche qu'elles s'appêtent à accomplir. Des rôles assignés sont présents dans une mise en situation quand chacun des participants reçoit de l'information spécifique sur un sujet qui demeure inconnu des autres participants. Cette information peut également diviser le groupe en deux camps qui doivent défendre des positions opposées et négocier un accord (Balicco, 1997; Tixier, 1974). Dans tous les cas, les informations reçues au préalable précisent le rôle de la personne candidate au sein du groupe et fournissent de l'information supplémentaire à propos de la direction ou des individus que la personne candidate doit représenter au cours du GDSL. Ces lignes directrices peuvent être utilisées par la personne candidate lorsqu'elle juge qu'elles influenceront les actions du groupe (Gatewood et Fields, 1994). Par exemple, les participants peuvent être mandatés pour défendre un projet, représenter les intérêts d'un groupe de personnes ou seulement mettre de l'avant une solution pré-déterminée par le concepteur du GDSL. En ce qui concerne les GDSL sans rôle assigné, la position des

participants lors de la discussion est personnelle. Cette forme de GDSDL ressemble à des comités organisationnels qui se réunissent pour discuter d'une problématique en tentant de mettre en œuvre de nouvelles procédures visant à la régulariser (Thornton et Byham, 1982).

En utilisant une autre typologie, Tixier (1974) mentionne que la combinaison d'un format compétitif avec rôle assigné peut être déterminée par l'appellation groupe de discussion **conflictuel**. Selon lui, ce modèle d'évaluation oblige les participants à mobiliser leur énergie lorsque le groupe est séparé en deux camps ou plus, ce qui peut représenter un certain avantage sur le plan des comportements stimulés. Cependant, l'observation de ces comportements devient difficile et il y a facilement confusion quant aux rôles de chaque personne candidate (Tixier, 1974).

Même si cette forme de groupe de discussion n'a pas été retenue dans le cadre du ce projet de mémoire, l'administration de cet exercice peut également être réalisée avec un **leader** assigné. Dans ce contexte, un participant possède le titre de dirigeant, donne au groupe une tâche à accomplir (Thornton, 1992) et structure la prise de décision. L'avantage de ce type d'exercice est qu'il simule le travail des gestionnaires responsables de la supervision d'un groupe de subalternes. Il permet par le fait même d'évaluer une variété de comportements reliés au leadership et à la communication. À propos des inconvénients, il serait très difficile, dans le cadre d'un processus de sélection, de permettre à tous les participants d'agir à titre de leader désigné. Pour assurer un traitement équitable, chaque participant devrait avoir ce rôle dans un exercice différent et c'est pour cette raison pratique que cette forme de groupe de discussion ne fera pas l'objet d'attention dans la présente recherche.

Tel que mentionné précédemment, les GDSDL peuvent porter sur différents **thèmes**. Le contenu de la mise en situation, le nombre de sous-problèmes spécifiques et la complexité du sujet varieront en fonction de la composition du groupe (Thorton et Byham, 1982). Le choix du thème est important, car les stimulus présentés ne font pas appel aux mêmes qualités et comportements. En effet, un thème très large oblige les participants à choisir une méthode de travail pour la résolution du problème alors qu'avec des règles du jeu plus précises, tous les participants ont la possibilité de s'exprimer beaucoup plus rapidement (Tixier, 1974). D'ailleurs, l'aspect directif de la consigne laisse moins de spontanéité à l'expression du leadership, mais permet une observation plus facile des comportements, dans la mesure où il existe déjà des références sur le thème (Tixier, 1974). À titre d'exemple, le thème peut impliquer un problème relié à la discipline du poste à doter qui est relativement simple ou porter sur un sujet abstrait associé à des jugements de valeurs comme la responsabilité corporative de l'organisation par rapport à la protection de l'environnement ou sur l'équité en emploi (Thorton et Byham, 1982). Selon Bernatchez (1982), l'essentiel est que l'orientation des membres du groupe soit suffisamment claire puisque, selon lui, l'absence de consignes claires requiert habituellement un degré de prise en charge trop élevé chez les participants. Finalement, toujours selon Bernatchez (1982), la nature du sujet à discuter ne doit pas être trop large et doit être concentrée et restreinte à des préoccupations qui demeurent gérables à l'intérieur d'une discussion de groupe. Pour identifier la complexité et l'étendue du groupe de discussion, un des points importants à considérer est le temps disponible pour le groupe de discussion (Wilson, 1996), mais aussi bien sûr le rôle et les responsabilités associés au poste à doter.

Dans la plupart des cas, les personnes candidates doivent en venir à un **consensus**. La notion de consensus signifie que tous les membres du groupe sont

d'accord sur la décision à adopter à la suite de la discussion. Comme Wilson (1996) l'expose, toutefois, dans le processus de prise de décision, on s'accorde pour dire que le consensus signifie que tous les membres du groupe sont fondamentalement d'accord avec le fait que la décision est acceptable. Cela ne veut pas nécessairement dire que tous les membres du groupe auraient nécessairement opté initialement pour cette décision. Dans le cadre d'un GDSL, le groupe est souvent encouragé à atteindre un consensus, puisqu'il s'agit d'une situation représentative de plusieurs emplois. Dans le même sens, lorsqu'une situation similaire au GDSL se présente en cours d'emploi dans les organisations, c'est souvent le groupe attribué pour l'élaboration de solutions qui sera également responsable de l'implantation de cette décision (Wilson, 1996). De cette manière, les participants au GDSL doivent à l'occasion produire un rapport écrit spécifiant les actions qui doivent être prises par l'organisation afin de faire face à la situation problématique présentée (Gatewood et Fields, 1994).

Dans ce contexte, le GDSL recrée un environnement dynamique qui stimule des comportements reliés à l'interaction entre le client et l'employé et de cette nature ce qui le rend intéressant comme moyen de sélection. En effet, chacun s'accorde maintenant à reconnaître que le personnel oeuvrant au service à la clientèle influence fortement, par leurs comportements, la qualité du service perçue par le client, et la compétitivité de l'entreprise. Dans le même sens, l'environnement des organisations est influencé par divers intervenants et certains emplois nécessitent une participation active à différents comités de travail au cours desquels l'employé devra faire face à des situations politiques délicates et difficiles qui nécessiteront de sa part une bonne interaction, situation typique d'un GDSL.

De plus, la supervision d'une équipe de travail exigera habituellement d'un employé de bonnes habiletés relationnelles ainsi qu'une facilité à initier des structures dans le cadre d'une discussion entre les membres de son équipe. Il est connu qu'un cadre doit animer des équipes de travail et défendre oralement une quantité appréciable de dossiers de tous genres (Bernatchez, 1982). Il est également largement admis que l'efficacité des gestionnaires réside relativement dans la capacité de soutenir verbalement des idées, principalement puisque les décisions se prennent de plus en plus en équipe dans les organisations. Dans ce contexte, il sera pertinent de développer un exercice de GDSL basé sur ces interactions représentatives et caractéristiques du poste à doter.

En somme, le format et la nature du thème qui seront choisis dans le cadre de l'administration d'un GDSL, en contexte de sélection, reposeront sur les informations qui seront recueillies lors des étapes reliées au développement de l'exercice. Tel que présenté dans la prochaine section, la réalité du poste à doter permettra au concepteur de choisir les conditions et les stimulus appropriés afin de permettre la mesure et l'évaluation des différences individuelles entre les personnes candidates.

1.2.4 Les étapes du développement de l'exercice

L'élaboration d'un instrument de mesure et d'évaluation, tel que l'exercice du GDSL, doit être appuyée sur un processus rigoureux et exhaustif. Aux fins de ce projet de recherche, le modèle de Pettersen (2000), c'est-à-dire le processus général d'élaboration d'instruments de mesure fondé sur la validation de contenu, a été retenu. Lorsque ces étapes sont réalisées minutieusement, l'outil d'évaluation est représentatif

du contenu de l'emploi et répondra aux différentes normes canadiennes et américaines en matière d'évaluation de personnel.

TABLEAU 3
Processus général d'élaboration d'instruments de mesure
fondés sur la validation de contenu³

PHASE I	
Finalités spécification du domaine à mesurer	
Étape 1	Finalités de l'instrument de mesure : préciser à quelles fins sera utilisé l'instrument de mesure (sélection, promotion, counselling d'emploi ou placement, formation, etc.)
Étape 2	Analyse et description de l'emploi : relever les éléments observables, non triviaux, pertinents tels que ce qui doit être effectué (éléments, tâches et responsabilités), ce qui doit en résulter (produits ou résultats), le niveau de rendement attendu, le contexte de l'emploi, l'équipement et les technologies utilisés et, finalement, les exigences requises du point de vue de l'organisation.
Étape 3	Spécification du domaine ou du sous-domaine à mesurer : choisir les éléments critiques ou pertinents de l'emploi à être représentés dans l'instrument, et les transposer en termes : <ol style="list-style-type: none"> 1. de comportements (éléments, tâches et responsabilités) 2. de résultats produits par ces comportements ou 3. de connaissances, d'habiletés, d'aptitudes ou d'autres caractéristiques personnelles nécessaires à ces comportements.
PHASE II	
Développement de l'instrument de mesure	
Étape 4	Conception de l'instrument dans son ensemble : en fonction des ressources et des contraintes de l'organisation, préciser le format de l'instrument, le type d'items, la durée et le nombre approximatif d'items, le mode de correction, le processus d'interprétation et l'usage de normes, le nombre de version de l'instrument, etc.
Étape 5	Création des items et élaboration des conditions d'application : rédiger les items afin de représenter le domaine à mesurer (défini à l'étape 3) tout en respectant les spécifications de l'instrument désiré (précisées à l'étape 4). Rédiger les autres éléments de l'instrument comme les directives, le mode d'enregistrement des réponses, etc.
Étape 6	Élaboration des outils d'évaluation : préparer la clé de correction notamment en identifiant les éléments de réponse attendus, déterminer la façon de calculer les scores et, s'il y a lieu, les normes utilisées. Dans le cas d'un instrument de mesure qui requiert un processus d'observation, d'interprétation ou d'évaluation, élaborer les outils nécessaires à ces opérations.

³ Tiré de l'ouvrage : Pettersen, N. (2000). Évaluation du potentiel humain dans les organisations : Élaboration et validation d'instruments de mesure. Sainte-Foy, Presses de l'université du Québec, page 218 et 219.

TABLEAU 3 (suite)
Processus général d'élaboration d'instruments de mesure
fondés sur la validation de contenu

Étape 7	Révision de la version expérimentale par des experts : soumettre l'instrument de mesure à des experts afin d'examiner la clarté et la pertinence du contenu.
Étape 8	Essai de la version expérimentale et contrôle des qualités métrologiques : faire passer l'instrument de mesure à un groupe témoin et s'il y a lieu, modifier les directives, les items, les outils d'évaluation ou toute autre composante de l'instrument. Les données recueillies peuvent aussi servir à l'examen des qualités métriques et à l'établissement de normes et de notes de passage.
PHASE III	Implantation de l'instrument de mesure
Étape 9	Rédaction des documents techniques : rédiger un document qui décrit les étapes de la construction de l'instrument de mesure et un manuel de procédures qui indique notamment les usages prévus, la manière de l'administrer, de le corriger et d'en interpréter les scores.
Étape 10	Implantation et suivi : planifier et organiser l'implantation de l'instrument, puis revoir périodiquement l'ensemble de l'instrument, ou lorsqu'il y a eu des changements dans la situation.

Comme l'objectif d'un processus de sélection de personnel est d'identifier les meilleurs individus pour performer dans un emploi spécifique à une organisation, l'information relative au poste à doter constitue un élément primordial. Dans cette optique, il est habituellement suggéré d'appuyer l'analyse de l'emploi sur plusieurs outils de collecte d'informations comme la création d'un groupe de travail, des entrevues individuelles avec des experts de contenu provenant de l'organisation, une revue de la documentation, l'observation, etc. Il est d'abord très important de recueillir de l'information sur l'ensemble des tâches, des responsabilités, des activités, des éléments, des comportements et des résultats qui se rattachent aux fonctions du poste à doter. Cette étape permettra d'identifier le domaine de contenu de l'outil d'évaluation et d'identifier les mesures relatives au domaine de performance relié à l'emploi.

La performance en emploi représente tous les comportements mesurables pertinents aux buts de l'organisation et qui est sous le contrôle d'un individu (Campbell, Gasser et Oswald, 1996 dans Hough et Oswald, 2000). Dans ce contexte, celle-ci est complexe, dynamique et multidimensionnelle et conséquemment, le système de sélection de personnel doit prédire les différences individuelles pour plusieurs types de performance en emploi (Hough et Oswald, 2000). De plus, la finalité d'un instrument de mesure utilisé dans un contexte d'évaluation ou de sélection est d'évaluer les caractéristiques individuelles sous-jacentes au rendement et difficilement modifiables à court terme par la formation et le perfectionnement. En fait, un tel outil psychométrique sera orienté vers les caractéristiques générales et relativement stables des individus. En utilisant l'information recueillie sur le poste et sur les données relatives à la performance en emploi pour ce type de fonctions, il sera plus facile d'identifier les caractéristiques requises par l'emploi. Comme pour le choix des critères à évaluer dans le cadre d'un centre d'évaluation, les caractéristiques déterminées doivent être spécifiques, observables et représenter une tâche reliée logiquement à l'emploi en question (Gatewood et Fields, 1994).

Une fois les dimensions et les tâches identifiées, la dernière étape, avant l'implantation de l'instrument de mesure, vise à créer un ou des problèmes situationnels représentatifs de ceux rencontrés par la personne en emploi. Comme pour le développement de n'importe quel autre outil d'évaluation, la situation doit être réaliste et représentative des fonctions du poste visé. Un exercice de validation et un pré-test doivent être conduits afin d'utiliser l'instrument dans le cadre d'un processus réel de sélection. Ces deux activités vont permettre d'apporter des modifications au contenu selon les commentaires des experts de contenu ou des sujets au pré-test.

1.2.5 L'évaluation des critères

Au cours d'un GDSL la personne candidate doit réagir sur le vif à des situations qui sont susceptibles de se présenter en emploi, au lieu d'exposer la façon dont elle se comporterait si telle ou telle situation se présentait. En effet, contrairement aux méthodes traditionnelles d'évaluation comme l'entrevue, le GDSL exige que les participants **démontrent** les comportements requis dans ces situations au lieu de les relater. Selon Smith, Gregg et Andrews (1990), les exercices de simulation comptent parmi les meilleurs outils de prédiction de la performance future au travail.

Tixier (1974) est d'avis que cet outil d'évaluation favorise notamment la projection du comportement social, par exemple :

- Il suscite des processus de confrontation, de tension, de déséquilibre qui réclament des efforts d'élucidation et d'adaptation;
- Il conduit les participants à vivre, à percevoir, à maîtriser les problèmes affectifs et fonctionnels mis en jeu par la situation où ils se trouvent engagés.

Il s'agit donc d'un instrument de mesure qui sert habituellement à évaluer les habiletés relationnelles des candidats, comme le leadership (Gatewood & Field, 1994; Thornton, 1992), la capacité à résoudre des problèmes ou à prendre des décisions. Le GDSL s'est avéré être un excellent facteur permettant de prédire le style interpersonnel de gestion (Craik, Ware, Kamp, O'Reilly III et Zedek, 2002). Gatewood et Fields (1994) mentionnent que le GDSL a été créé pour représenter les qualités de gestion qui sont

requis lors d'interactions avec un petit groupe d'individus dans le but de régler un problème commun avec succès, c'est-à-dire la persuasion, la persévérance, la résistance au stress et la gestion des conflits. Certains auteurs sont d'avis que les GDSL permettent aux participants d'exprimer leur intelligence (Tixier, 1974; Paulhus et Morgan, 1997).

Les GDSL peuvent également permettre de porter une appréciation sur l'habileté à contribuer en apportant des idées et la capacité à guider le processus de discussion (Harris, 1949 dans Thornton 1992). L'analyse de problèmes et la capacité à prendre des décisions peuvent également être évaluées (Bass, 1950 et 1954 dans Thornton 1992). Ce type d'exercice est souvent utilisé lors du recrutement de représentants commerciaux afin d'identifier un certain nombre de caractéristiques indispensables pour ce type de poste, telles que l'aisance verbale, la qualité de l'argumentation et les relations interpersonnelles (Balicco, 1997). Finalement, Cronbach (1970), dans Thornton et Byham (1982), mentionne que le GDSL peut permettre la mesure et l'évaluation de la prééminence de la personne, sa capacité à faciliter l'atteinte d'objectifs communs et sa sociabilité.

Une étude de Byham et Byham (1976) présentée dans Thornton et Byham (1982) présente une revue de plus de 1 000 centres d'évaluation utilisant des GDSL avec rôle assigné ou non afin d'identifier les dimensions observables par le biais de cette méthode d'évaluation. Pour parvenir à ces conclusions, les évaluateurs avaient comme consigne de porter une appréciation seulement sur les dimensions pour lesquelles ils avaient obtenu suffisamment d'informations dans le cadre de l'exercice. Le **Tableau 4** présenté à la page suivante expose le pourcentage pour lequel les exercices ont provoqué des comportements associés aux différentes dimensions évaluées.

TABLEAU 4
Évidence des dimensions observables dans le cadre de
GDSL avec rôle assigné et sans rôle assigné⁴

Dimensions reliées au poste		GDSL avec rôle assigné	GDSL sans rôle assigné
1.	Impact	90 %	90 %
2.	Capacité de communiquer oralement	95 %	95 %
3.	Capacité de présenter oralement	95 %	-
4.	Leadership	90 %	95 %
5.	Persuasion / Capacité à la vente	90 %	-
6.	Sensibilité	90 %	95 %
7.	Souplesse du comportement	90 %	-
8.	Ténacité	90 %	-
9.	Prise de risques	-	90 % *
10.	Initiative	90 %	95 %
11.	Jugement	-	90 % *
12.	Esprit de décision	-	90 % *

** La moyenne observée est de 90 % pour certains exercices et non pas pour l'ensemble des exercices d'évaluation.*

De manière opérationnelle, pendant que les personnes candidates sont placées autour d'une table ronde au milieu d'une salle de réunion pour discuter ensemble, des évaluateurs professionnels sont assis à l'extérieur du groupe ce qui leur permet d'observer et d'enregistrer les comportements des participants au GDSL, de façon discrète, sans interférer dans la dynamique du groupe (Gatewood et Field, 1994). Il peut arriver que le groupe contienne des observateurs anonymes que les autres participants considèrent aussi comme des candidats. Toutefois, les considérations éthiques reliées à cette forme d'évaluation limitent son utilisation dans le cadre de processus de sélection.

Comme dans le cas d'un centre d'évaluation, le rôle des évaluateurs présents lors d'un GDSL consiste à observer et à enregistrer les comportements des participants lors

⁴ Tiré de l'ouvrage : Byham, R.N. et Byham, W.C. (1976), Effectiveness of Assessment Center Exercises in Producing Behavior, *Assessment & Development*, March 1976, p. 9-10.

du déroulement de l'exercice (Lievens et Conway, 2001). En effet, comme l'expose la littérature sur les centres d'évaluation, les évaluateurs ont comme consigne dès le début du processus de retenir toute interprétation rapide, d'observer les candidats avec attention, et d'enregistrer leurs observations en ce qui concerne les comportements (Thorton, 1992). Selon Lievens et Conway (2001), cette attention portée à l'observation des comportements permet de rencontrer trois objectifs principaux, soit :

- Une évaluation plus précise puisqu'il est plus facile de déterminer une note sur une dimension pour laquelle on a recueilli plusieurs comportements.
- Une intégration des observations plus rapide pour les évaluateurs puisqu'il est possible de s'appuyer sur les notes des comportements pour défendre leur évaluation individuelle lors de la période d'intégration avec les évaluateurs;
- Des séances de rétroaction avec les personnes candidates plus détaillées puisque les comportements peuvent servir à illustrer leurs forces et faiblesses.

De façon générale, chaque évaluateur surveille un ou deux participants puis compile ses observations au fur à mesure qu'elles se manifestent. Par la suite, chaque observateur trie et classe ses observations sous les critères appropriés en fonction d'une échelle de comportements standardisée et il attribue une cote à chaque critère. À la fin de la ou des journées d'évaluation, chaque observateur fait un compte rendu aux autres évaluateurs. Ils discutent alors de ces observations et de leur note afin d'en arriver à un consensus sur les différents critères évalués. Selon Thorton et Byham (1982), l'échelle d'évaluation habituellement utilisée dans le cadre de ces exercices de simulation comporte habituellement six points (de zéro à cinq). Le responsable du

processus rédige finalement un rapport détaillé des appréciations émises par les observateurs et de la note finale obtenue par critère ce qui peut permettre l'élaboration de normes d'apprentissage et de perfectionnement.

Comme le souligne Thornton (1992), plus d'une personne est généralement impliquée dans l'évaluation des participants selon la méthode du centre d'évaluation. Les observations de plusieurs évaluateurs aident donc à assurer que certains biais n'influenceront pas les résultats finaux du processus d'évaluation. Le fait de considérer plusieurs points de vue indique que des jugements sont portés sur la compétence professionnelle globale des participants. Dans le cas de l'administration d'un GDSL de façon indépendante, il n'est pas possible de combiner l'information obtenue de d'autres sources. Toutefois, deux évaluateurs peuvent évaluer le même participant sur les mêmes dimensions afin de respecter ce principe de base.

D'après Tixier (1974), l'observation individuelle d'un participant est indissociable de l'observation du déroulement et de l'évolution de la discussion du groupe dans son ensemble. Ce constat implique donc que l'évaluateur devrait être en mesure d'identifier les phénomènes de groupe classiques, tels que l'autorité et l'influence, le réseau de communication, l'efficacité et la créativité. Pour Tixier (1974), ce sont les mécanismes de raisonnement des participants, l'état émotionnel du groupe, le niveau de collaboration, la capacité à résoudre des conflits et à atteindre les objectifs fixés par l'exercice qui fournissent en quelque sorte un cadre de référence pour l'évaluation des participants. Même si cette méthode d'évaluation n'est pas celle retenue dans le cadre du présent projet de recherche, la littérature sur la dynamique de groupe sera exposée dans la prochaine partie afin de valider si effectivement celle-ci a une influence sur l'évaluation individuelle d'une personne candidate à un processus de sélection.

Sur le plan des qualifications des évaluateurs, le Centre de formation et de consultation (1987) recommande l'utilisation de responsables hiérarchiques comme évaluateurs et des spécialistes en ressources humaines comme administrateurs de l'outil de mesure. Selon eux, les responsables hiérarchiques sont davantage informés des exigences de l'organisation et les spécialistes en ressources humaines possèdent les connaissances requises pour la sélection du personnel et l'utilisation des instruments (CFC, 1987). Lévy-Leboyer (1996) et Thorton et Byham (1982), quant à eux, stipulent que les évaluateurs ne devraient pas être en supervision directe avec les personnes candidates. Dans le même sens, une étude de Schmitt, Schneider et Cohen (1990) révèle que les évaluateurs qui connaissent personnellement les personnes candidates ne sont pas aussi objectifs dans leur évaluation que ceux qui ne connaissent pas les candidats. Les évaluateurs qui participent à un centre d'évaluation peuvent être des gestionnaires, le personnel provenant des ressources humaines de l'organisation, des consultants externes, des psychologues ou une variété d'autres individus, dans la mesure où l'ensemble de ces personnes sont familières avec le travail en question. En effet, selon Thornton (1992), les évaluateurs doivent être choisis selon leur niveau hiérarchique de gestion, pour leur connaissance technique et leur expertise concernant le poste à doter. Néanmoins, les recherches ont indiqué qu'une variété de différents types d'évaluateurs peuvent apporter des contributions valides à l'évaluation (Gaugler et al., 1987).

La formation des évaluateurs est primordiale. Le programme de formation porte habituellement sur les dimensions évaluées, les types de comportements observés dans chaque exercice, les normes pour l'évaluation des comportements par rapport aux exercices et, finalement, le processus d'intégration de l'information lors de la discussion

entre les évaluateurs (Thorton et Byham, 1982). Dans le même sens, les directives concernant l'opération des centres d'évaluation recommandent que les évaluateurs reçoivent une formation exhaustive sur le processus d'observation et de classification des critères et des comportements. Ils doivent également démontrer un habileté à noter adéquatement les candidats (Caldwell, 2003).

En somme, il est possible de conclure que l'utilisation du GDSL représente une approche pertinente au moment d'évaluer certaines qualités particulières associées à un poste en fonction des écrits scientifiques sur les centres d'évaluation. Cependant, comme tout outil de mesure et d'évaluation, l'utilisation du GDSL doit répondre à deux conditions de base, soit la validité et la fidélité. Dans le prochain chapitre, ces deux aspects fondamentaux seront examinés, mais plus particulièrement la notion de fidélité qui demeure au cœur même de cette recherche.

Chapitre 2 Évaluer la qualité psychométrique d'un outil d'évaluation

L'évaluation de la qualité psychométrique d'un instrument de mesure est une des étapes les plus importantes lors du développement et de l'utilisation d'un exercice en contexte de sélection. De façon générale, cette appréciation consiste à valider l'outil, à mesurer son degré de fidélité et à prendre en compte ses considérations éthiques et légales.

Pour y parvenir, des critères pour évaluer les tests, les pratiques de testing et les effets de leur utilisation ont été compilés et codifiés dans « *Standards for Educational and Psychological Testing* » (1999). Ces critères ont été élaborés grâce aux efforts de trois organisations - *American Educational Research Association (AERA)*, *American Psychological Association (APA)*, et *National Council on Measurement in Education (NCME)* – et fournissent des lignes directrices par rapport à l'évaluation des principaux aspects énoncés ci-haut. Ces critères sont d'autant plus importants qu'un outil d'évaluation dont la validité et la fidélité sont insuffisantes n'a pas sa place en gestion des ressources humaines. Non seulement celui-ci ne peut servir à guider les décisions organisationnelles, mais il peut également être nuisible et entraîner des conséquences néfastes, telles que des choix de candidats inadéquats, une évaluation inéquitable, des difficultés à se défendre en cas de grief ou de poursuite devant les tribunaux, sans parler du manque d'éthique à l'égard des personnes injustement traitées (Pettersen, 2000).

Pour rendre compte de la qualité psychométrique des GDSL, les concepts de validité et de fidélité seront définis dans les prochaines lignes. Le concept de fidélité, les

erreurs de mesure et les problèmes associés au manque de standardisation d'un outil d'évaluation seront revus et décrit en profondeur puisqu'ils se retrouvent à la base des questions de recherche de ce projet de mémoire. Pour compléter ce bagage théorique, des informations relatives aux indices de validité et de fidélité des centres d'évaluation et des GDSDL seront discutées. Les considérations éthiques et légales relatives à l'utilisation d'un moyen d'évaluation en contexte de sélection seront également traitées. Finalement, quelques lignes sur la qualité psychométrique des groupes de discussion sans leader permettront de situer de façon plus précise les questions de recherche.

2.1 La validité

D'un point de vue théorique, la validité est la qualité première et essentielle à considérer pour déterminer la qualité d'un instrument de mesure et il s'agit du principal aspect à considérer lorsque vient le temps de l'évaluer (Guion, 1998). La validité indique jusqu'à quel point l'instrument parvient à mesurer ce qu'il est censé mesurer ou à prédire ce qu'il est censé prédire (Pettersen, 2000). De façon plus spécifique, « *la validité évalue jusqu'à quel point les inférences faites à partir des résultats recueillis à un instrument de mesure sont exacts et fondés dans la situation particulière où elles sont appliquées*⁵» (Nunnally et Bernstein, 1994, Schmitt et Chan, 1998, *Standards*, 1999 dans Pettersen, 2000). Également, la validité se concentre sur tous les éléments de preuve de ce résultat, quelque soit la démarche de validation utilisée (*Standards* 1999 dans Pettersen, 2000).

⁵ Tiré de l'ouvrage : Pettersen, N. (2000). Évaluation du potentiel humain dans les organisations : Élaboration et validation d'instruments de mesure. Sainte-Foy, Presses de l'université du Québec, page 47.

Selon Pettersen (2000), l'efficacité des instruments de mesure dépend fondamentalement de la validité des inférences faites à partir des résultats obtenus au moyen de ces instruments. Si les inférences ne sont pas valides, les résultats ne peuvent pas servir à décrire le répondant (inférence descriptive) ni à prédire son comportement (inférence relationnelle). Pour vérifier l'indice de validité d'un outil d'évaluation, on propose d'utiliser trois stratégies de validation se rapportant à la validité de contenu, à la validité critérielle et à la validité de construit (Gatewood et Field, 1990, Ciasco, 1987).

Pour démontrer une **validité de contenu**, l'exercice doit refléter les types et le niveau de complexité des activités reliées à l'emploi. Le réalisme et la représentativité d'un outil d'évaluation sont donc vérifiés par le degré de relation entre son contenu (mises en situation, questions, tâches, procédures d'administration, de correction, etc.) et le contenu du poste visé (activités, tâches, rôle et responsabilités, etc.). Dans ce contexte, un outil d'évaluation est valide si son contenu est représentatif des tâches dont doit s'acquitter le candidat dans le poste visé (Gatewood et Fields, 1990) et si le contenu des réponses recherchées à l'outil est représentatif du domaine de performance requis par l'emploi (Thorton, 1992). La stratégie de validation basée sur le contenu permet de recueillir des résultats représentatifs du domaine à mesurer et de vérifier si ce domaine est pertinent à l'usage prévu des résultats. Les jugements de similarités entre le contenu de l'exercice et le poste sont les types d'évidences les plus utilisés pour démontrer la validité des exercices de simulation (Gaugler et al., 1987).

La **validité critérielle** sert à vérifier s'il existe une corrélation entre les résultats individuels obtenus à l'outil d'évaluation et un critère, par exemple le rendement au

travail. Cette stratégie vérifie jusqu'à quel point les différences individuelles, telles que mesurées par l'outil, permettent de prédire le comportement d'un individu au travail (Cascio, 1987). On peut parler dans cette optique de la validité concourante et de validité prédictive. La principale différence entre ces deux types est de nature temporelle (Marcoux, 1999). Dans le cadre d'une stratégie prédictive, la vérification de la validité se fera en comparant la relation entre les résultats des sujets à l'outil d'évaluation et leur performance future au travail, alors que dans le premier cas, les inférences sur la prédiction de la performance au travail du candidat peuvent être émises dès la première administration de l'instrument de mesure, le degré de validité ayant été obtenu lors de son développement. Cette stratégie prédictive émet l'hypothèse selon laquelle les employés les plus performants au travail auront de meilleurs résultats au test. Les écrits sur les méthodes de sélection du personnel accordent généralement l'importance à un indicateur spécifique de la validité : le coefficient de validité critériée. Celui-ci est orienté sur le degré auquel les procédures de sélection de personnel peuvent prédire adéquatement le rendement en emploi.

La **validité de construit** doit également être considérée dans l'appréciation d'un instrument de mesure. Celle-ci indique jusqu'à quel point un outil d'évaluation mesure le concept (ou le construit) qu'il est censé mesurer, par exemple des compétences. Par conséquent, la vérification de la validité de construit ou conceptuelle implique de définir si un outil d'évaluation mesure bien les compétences identifiées pour un poste (Standards, 1999).

2.1.1 La validité des centres d'évaluation

Comme nous l'avons traité dans le premier chapitre de ce projet de recherche, le centre d'évaluation a démontré au cours des cinquante dernières années qu'il représentait une procédure d'évaluation possédant un niveau de validité acceptable en sélection. Puisqu'un de ses facteurs d'efficacité réside dans son aspect de simulation du poste visé, le centre d'évaluation possède donc une forte validité apparente (Lévy-Leboyer, 1996, Pettersen, 2000). La validité prédictive moyenne des centres d'évaluation est de l'ordre de 0,40 (Schmitt, Gooding, Noe et Kirsch, 1984, Hunter et Hunter, 1984) et la recension des écrits sur les centres d'évaluation démontre que ceux-ci permettent de prédire avec assez de justesse le taux de promotion et d'avancement aux postes de gestion. L'ensemble des études de validation démontre également des indices de validité comparables et ce, peu importe le critère d'évaluation utilisé. Gaugler et al. (1987) ont publié une méta-analyse dans laquelle le coefficient de validité corrigé moyen est de 0,37 pour le centre d'évaluation (la corrélation non corrigée de 0,32). De plus, la validité est de 0,53 (corrélation non corrigée de 0,45) lorsque le critère est une évaluation du potentiel de gestion du candidat et de 0,36 (corrélation non corrigée de 0,31) lorsque le critère est une évaluation du rendement en emploi. Ces coefficients ont été corrigés pour tenir compte de l'erreur d'échantillonnage, de la restriction de l'étendue de l'échantillon et de l'infidélité du critère (Pettersen, 2000).

D'ailleurs, depuis les années 1980, la conclusion générale qui se dégage des écrits est que les centres d'évaluation ont une validité discriminante faible, c'est-à-dire qu'à l'intérieur d'un même exercice la distinction entre les critères évalués est floue et qu'il y a donc une forte corrélation entre les résultats aux différents critères, ainsi qu'une

faible validité convergente, c'est-à-dire qu'il y a peu de corrélations entre les résultats à des dimensions évaluées plusieurs fois par le biais d'exercices différents (Lievens, 2002).

2.1.2 La validité des GDSL

Bien que les GDSL soient utilisés dans la majorité des centres d'évaluation, il n'est pas possible de tirer des conclusions définitives sur leur niveau de validité prédictive (Smith et Robertson, 1993). En fait, même si les écrits présentent de nombreuses recherches sur la validité des centres d'évaluation, il est rare d'en trouver qui portent directement sur la validité de ses exercices spécifiques. Les seules indications précises trouvées à ce propos remontent aux années 1950, alors qu'une étude de Bass (1950 et 1954 dans Thorton et Byham, 1982) établissait la validité prédictive des GDSL entre 0,40 et 0,50.

Dans le cadre d'études sur la validité des centres d'évaluation, les auteurs Thorton et Byham (1982) ont établi les indices de validité critérielle présentés dans le **Tableau 5** spécifiquement pour le GDSL.

Tableau 5
Indices de validité critérielle des GDSL⁶

Exercice	Auteur	Validité critérielle
Groupe de discussion sans leader	Bray et Grant (1969)	0,60
	Wollowick et McNamara (1969)	0,25
	Bentz (1971)	0,37
	Hack (1974)	0,66

⁶ Tiré de l'ouvrage : Thorton, G. et Byham, W. (1982). *Assessment Centers and Managerial Performance*. Florida, Academic Press inc., page 173.

2.2 La fidélité

Les résultats obtenus par les personnes candidates à un processus de sélection doivent également être fiables ou fidèles. Le concept de fidélité concerne le degré d'exactitude des mesures et désigne la *précision* ou la *constance*, c'est-à-dire la stabilité des résultats (Pettersen, 2000). Dans ce contexte, une mesure fidèle est une mesure sans erreur et, en gestion de ressources humaines, ces erreurs de mesure qui peuvent affecter un instrument de mesure proviennent de multiples sources, soit : (Pettersen, 2000)

- Du candidat lui-même qui n'est pas toujours constant dans sa façon de répondre à un instrument de mesure ;
- De l'examineur dont le comportement est aussi influencé par plusieurs facteurs ;
- De la situation ou des conditions dans lesquelles est appliqué l'instrument.

Dans le cas du GDSL, la nature imprévisible de la dynamique du groupe amène potentiellement des erreurs de mesure provenant de la situation. En effet, de par sa nature spécifique, le GDSL est représenté par des conditions d'administration non standardisées d'une fois à l'autre. Ces conditions changeantes ou non conformes peuvent donc potentiellement altérer les résultats du candidat. Les principales caractéristiques d'une évaluation standardisée sont relatives au contenu (toutes les personnes candidates sont évaluées à partir de la même information), à l'information qui est recueillie selon les mêmes conditions à chaque fois que la mesure de sélection est

appliquée et par rapport aux règles pour l'attribution des notes qui sont spécifiées à l'avance et qui sont appliquées de la même façon à tous les participants.

Dans ce contexte, « *pour éviter que des conditions changeantes deviennent source d'instabilité, il suffit qu'elles soient scrupuleusement contrôlées et maintenues constantes pour tous les candidats à toutes les applications* »⁷. La standardisation est similaire au contrôle des variables dans le cadre d'une expérimentation et c'est ce principe qui a pour effet de contrôler directement l'erreur de mesure à sa source (Pettersen, 2000). Depuis le début des procédures de testing, la standardisation des conditions a été acceptée comme une bonne pratique et n'a jamais été remise en question dans l'histoire de la sélection du personnel (Guion, 1998). Il s'agit d'une garantie du traitement équitable entre l'ensemble des personnes candidates à un processus de sélection. Toutefois, selon Guion (1998), ce concept devrait faire l'objet d'une réévaluation :

*«Recent advances in psychometric theories suggest that the ideals of standardization may be better realized by individually tailoring sets of items to each applicant, perhaps using different numbers of items as well as different items for the different applicants; new method of administration by computer may call for varying de facto time limits by controlling response times for individual item.»*⁸

Pour évaluer la fidélité des résultats obtenus à un instrument de mesure et s'assurer qu'une mesure est exacte, on peut les comparer à d'autres mesures indépendantes du même objet. Si les résultats sont semblables, on conclut que la mesure est fiable et s'il y a des différences, alors il y a de l'erreur et c'est à l'aide du

⁷ Tiré de l'ouvrage : Pettersen, N. (2000). Évaluation du potentiel humain dans les organisations : Élaboration et validation d'instruments de mesure. Sainte-Foy, Presses de l'université du Québec, page 167.

⁸ Tiré de l'ouvrage : Guion, R.M. (1998). Assessment measurement, and prediction for personnel decisions. Mahwah, NJ, Lawrence Erlbaum Associates, p. 387.

coefficient de corrélation que l'on mesure dans quelle proportion ils varient (Pettersen, 2000). Voyons maintenant les différentes méthodes permettant d'estimer la fidélité.

D'une part, un instrument de mesure doit permettre à un observateur d'établir des mesures similaires d'un même sujet à des moments différents, toujours avec le même instrument. Il s'agit de la **fidélité test-retest** (Thiétart et coll., 1999). Cette méthode pour évaluer la fidélité d'un instrument de mesure consiste à faire passer aux mêmes personnes deux fois le même instrument à des moments différents et d'évaluer dans quelle mesure les résultats de la deuxième passation sont reliés à ceux de la première. Le degré de fidélité d'un test peut donc être mesuré en vérifiant la stabilité d'un test à reproduire les mêmes résultats sur une période de temps donnée (Pettersen, 2000).

D'autre part, les diverses parties d'un instrument de mesure doivent permettre de mesurer la même chose. Il s'agit de la **fidélité par consistance interne** (Guion, 1965 dans Pettersen, 2000). À partir de cette méthode, il s'agit en fait d'examiner les résultats obtenus par les mêmes personnes entre des parties équivalentes de l'examen à travers une seule passation. Plus les réponses des candidats à une partie de l'instrument sont reliées à celles des autres parties, plus les résultats sont fidèles.

Dans le cadre de ce projet de mémoire, la **fidélité par formes parallèles** et la **fidélité interexamineurs** sont les deux méthodes qui sont reliées plus particulièrement aux questionnements reliés à la qualité psychométrique des GDSDL. D'abord, la méthode des formes parallèles ou équivalentes est similaire à celle de la méthode du test-retest sauf que c'est une autre version de l'instrument de mesure qui est administrée lors de la deuxième passation. De façon plus précise, il s'agit d'administrer deux tests aux mêmes individus, la différence étant que le second test

n'est pas identique au premier : il s'agit d'une forme alternative élaborée pour mesurer la même caractéristique, mais selon une formulation différente (Thiétart et coll., 1999). Pour calculer la fidélité du GDSL par ce coefficient de corrélation, il s'agit donc d'élaborer des versions alternatives de cet instrument de mesure en faisant varier différentes conditions, comme la nature du thème, et de déterminer dans quelle proportion les résultats des personnes candidates sur les mêmes caractéristiques convergent entre eux.

D'un autre regard, de nombreux outils utilisés en gestion de ressources humaines nécessitent la participation de plusieurs examinateurs ou évaluateurs pour observer et porter une appréciation des personnes. Comme le jugement de l'évaluateur joue un rôle important dans ce contexte, un outil visant la mesure de certaines caractéristiques de la personne doit permettre à des observateurs différents de faire des mesures concordantes d'un même sujet avec le même instrument (Thiétart et coll., 1999). Pour ce faire, l'influence de l'examineur peut être évaluée en comparant deux ou plusieurs examinateurs lorsqu'ils évaluent la même personne et s'il y a concordance entre les examinateurs, les résultats sont fidèles (Pettersen, 2000). Pour calculer la fidélité par ce coefficient de corrélation, il s'agit donc de demander à des évaluateurs indépendants de corriger ou de porter une appréciation sur certaines caractéristiques spécifiques à partir du même bassin de personnes candidates et de déterminer dans quelle proportion les résultats convergent entre eux.

En conclusion, au cours d'une démarche de sélection, on ne devrait pas se fier à un instrument qui ne donne pas les mêmes résultats d'une mesure à l'autre dans des conditions similaires, mais il y a des contextes où le principe de la standardisation n'est pas si simple à appliquer comme c'est le cas pour le GDSL. D'un point de vue de

standardisation, la situation idéale impliquerait d'avoir seulement un groupe. Or, administrativement, les bassins de candidatures dépassent habituellement le nombre de personnes optimal pour un groupe de discussion. Pour évaluer un niveau de fidélité, il est possible de se référer à la recommandation de Gatewood et Field (1994) qui soutiennent que dans un contexte spécifique de sélection de personnel, le seuil minimal acceptable serait de 0,85 même s'ils trouvent davantage acceptable un seuil de 0,90 et plus.

2.2.1 La fidélité des centres d'évaluation

La vérification du degré de fidélité inter-interexamineurs est l'indice de fidélité le plus utilisé dans les centres d'évaluation (Marcoux, 1999). Selon Schneider et Schmitt (1992), l'effet de la forme de l'exercice constitue la principale source de variance lors de l'évaluation des participants alors que le contenu de l'exercice a un effet presque nul. Par contre, leurs résultats suggèrent que la performance de certains candidats est différente selon qu'ils sont évalués dans un contexte individuel ou en groupe. Les auteurs précisent qu'il est donc primordial que les exercices soient conçus de façon à être le plus près possible du type d'interactions que l'on retrouvera dans l'emploi cible.

2.2.2 La fidélité des GDSL

Thorton (1992) établit les coefficients de fidélité interexamineurs des GDSL présentés au **Tableau 6**.

Tableau 6
Coefficients de fidélité interexamineurs des GDSL⁹

Exercice	Auteur	Moyenne du niveau d'accord
Groupe de discussion sans leader	Bray et Grant (1966)	0,75
	Gatewood et al. (1990)	0,93
	Greenwood et McNamara (1967)	0,64
	Tziner et Dolan (1982)	0,83

La seule autre étude sur la standardisation des conditions d'administration du GDSL est celle de Gatewood, Thornton III et Hennessey (1990) et les résultats de cette étude sont mitigés. S'ils ont observé une fidélité interexamineurs autour de 0,93 lorsque des évaluateurs observaient des personnes candidates dans des groupes de discussion portant sur le même sujet, ils ont noté une fidélité de 0,66 à 0,88 lorsqu'il s'agissait d'évaluateurs différents et indépendants pour les mêmes sujets. Finalement, lorsque les conditions d'administration étaient très faibles (candidats évalués avec les mêmes juges dans des groupes de discussion différents et constitués de participants différents), ils ont observé un niveau de fidélité également très faible (variant de 0,35 à 0,62). En résumé, les résultats concernant la fidélité des groupes de discussion démontrent qu'elle varie en fonction du type d'erreur évalué. La fidélité interexamineurs est systématiquement supérieure aux autres types de fidélité (c'est-à-dire inter-groupe et inter-exercice).

Cependant, il est important de noter que les sujets de cette étude étaient, comme dans plusieurs autres recherches portant sur les groupes de discussion sans leader (Gibson, Randel et Earley, 2000; Stasser, Taylor et Hanna, 1999; Craik et al., 2002; Schneider et Schmitt, 1992), des étudiants et non des personnes candidates dans le cadre d'un processus réel de sélection. Il y a tout lieu de croire qu'un étudiant qui

⁹ Tiré de l'ouvrage : Thornton, G.C. III (1992). *Assessment Centers in Human Resource Management*. Reading, Mass; Don Mills, Ont., Addison-Wesley, page 125.

participe à une simulation sans conséquence pour lui n'agira pas nécessairement de la même manière qu'un candidat qui doit démontrer le maximum de sa performance afin d'obtenir l'emploi convoité.

2.3 Les considérations éthiques et légales

Les activités reliées aux ressources humaines sont encadrées par diverses lois et règlements. Au Québec, il y a la Charte canadienne des droits et libertés, la Charte des droits et libertés de la personne du Québec, le Code du travail, les contrats collectifs de travail et les nombreuses politiques gouvernementales en matière d'emploi (Pettersen, 2000). À cet égard, les outils d'évaluation utilisés en contexte de sélection peuvent porter atteinte aux droits fondamentaux de la personne ou certaines décisions d'embauche ou de rejet peuvent contrevenir à un article de la convention collective.

En ce sens, s'il y a un cadre législatif qui s'applique et qu'il y a contestation légale par rapport à n'importe quel outil d'évaluation utilisé, il faudra répondre aux allégations des appelants. Pettersen (2000) énonce qu'il y a une augmentation croissante des procédures judiciaires en matière de gestion des ressources humaines. Devant les tribunaux et les organismes qui édictent les lignes directrices par rapport aux instruments de mesure, la validation de contenu est considérée comme une démarche acceptable et suffisante. Selon Boucher et Doyon (1991), les exercices de simulation, qu'ils soient de nature individuelle ou collective, réduisent de beaucoup les possibilités de poursuite pour discrimination ou intrusion dans la vie privée puisqu'ils sont en fait des répliques de la situation qu'aura à vivre la personne en emploi. Bien que chaque cas soit d'espèce, certains outils seront plus faciles à défendre que d'autres dus à la qualité

psychométrique qui leur est propre. Dans le contexte du GDSDL, jusqu'à maintenant, son degré de fidélité a suscité de nombreuses questions.

Certaines responsabilités éthiques doivent également être respectées par les membres d'un groupe de discussion. Toujours selon Wilson (1996), chaque personne doit s'engager et être déterminé à donner le meilleur d'elle-même, s'engager à jouer juste, à écouter attentivement et à participer pleinement et, finalement à prendre un rôle de participant analyste.

2.4 La qualité psychométrique des GDSDL

De façon générale, les GDSDL créent un environnement dynamique qui génèrent des interactions tant négatives que positives qui permettent aux évaluateurs de recueillir de nombreuses observations qui n'apparaissent habituellement pas lors de l'administration de d'autres types de moyens d'évaluation. Les évaluateurs ont donc l'opportunité d'examiner la façon dont les personnes candidates seraient les plus propices à réagir lors de situations sociales variées et difficiles, situations en principe typiques de l'emploi à combler.

Cependant, des interrogations demeurent sur la qualité des inférences produites et sur la constance des résultats obtenus à un groupe de discussion sans leader en situation réelle de sélection. En fait, comme il l'a été exposé, le GDSDL implique une certaine portion d'évaluation non standardisée d'une administration à une autre et ce, en raison des dynamiques de nature sociale, mais aussi psychologiques qui sont générées lors des interactions entre les participants. L'équivalence d'un GDSDL peut également être affectée par la présence de toutes les personnes candidates au processus de

sélection convoquées à la journée d'évaluation, aspect difficilement contrôlable par les responsables d'un concours.

En somme, puisque la composition des GDSDL risque de varier et que les énoncés formulés par un individu engendrent des réactions et des interventions différentes d'une administration à l'autre, ce manque potentiel de standardisation d'une administration à l'autre signifie que les évaluateurs peuvent être amenés à se questionner si les comportements qu'ils observent sont réellement en fonction de l'individu ou en fonction de la dynamique du groupe. Pour avancer les conclusions appropriées et vérifier si la dynamique d'un GDSDL peut influencer en retour les résultats d'un processus de sélection, il est nécessaire de parcourir les écrits sur la notion de dynamique de groupe et, de cette manière vérifier exactement quelles sont les sources d'erreur associées aux GDSDL. Comme la nature de l'instrument est fortement associée à la notion de dynamique de groupe, examinons donc les connaissances dans ce domaine.

Chapitre 3 La dynamique de groupe

Le GDSL recrée une situation de travail représentative de diverses fonctions des emplois d'aujourd'hui, mais également un milieu social où des personnes doivent interagir ensemble sur une problématique particulière. Il génère donc différentes interactions, telles que des confrontations, des alliances, des luttes pour le leadership (Cragan et Wight, 1986) ou des demandes d'éclaircissement, de l'écoute et des encouragements (Parker, 1994) Même s'il est possible d'effectuer certains parallèles avec la théorie sur les groupes restreints (*small group research*) et les méthodes de groupes de discussion expérimentales, cette technique en contexte de sélection possède certaines caractéristiques distinctes. D'une part, les GDSL ont une durée très limitée dans le temps et le contexte est généralement celui d'une mise en situation fictive, à l'instar de comités organisationnels, par exemple, où les membres sont impliqués par leur expérience, leurs connaissances ou émotionnellement. Il est clair que les groupes qui travaillent ensemble sur une longue période vont présenter certaines spécificités que l'on ne retrouve pas dans les rassemblements temporaires, comme le GDSL. Certaines normes de conduite, par exemple, vont commencer à s'établir, et la façon dont les individus se sentent les uns vis-à-vis des autres affectera leur comportement face aux objectifs du groupe et entre eux. Comme le GDSL est un groupe éphémère, l'observation des interactions est limitée à cette tranche de vie : « *c'est l'aspect superficiel de la technique* » (Tixier, 1974).

Par ailleurs, il existe une certaine compétition entre les membres du groupe, même si la tâche à accomplir peut impliquer que les gens collaborent afin d'en arriver à un consensus. Au-delà du fait que la personne candidate a été informée qu'elle sera

évaluée en groupe et qu'elle a accepté de « *jouer le jeu* », elle sera tentée de se faire valoir et de défendre ardemment sa position dans le groupe pour « *remporter le concours* », ce qui amènera une certaine divergence d'intérêts entre les participants. Ce contexte est inhérent à la conscience et à l'attitude que les membres auront entre eux et il est important de le mettre en perspective avec l'ensemble des théories s'appliquant à la notion de groupe.

En gardant ces particularités à l'esprit, les prochaines lignes auront pour objectif de présenter les principaux fondements des études sur la dynamique de groupe afin de vérifier l'impact potentiel d'une évaluation collective sur les résultats individuels d'une personne candidate à un processus de sélection.

3.1 La définition du concept de groupe

Tout d'abord, le ***groupe de discussion*** fait référence à une ou plusieurs réunions d'un petit groupe qui communique, face à face, afin de réaliser un objectif commun et accomplir une tâche de groupe (Bormann, 1975). Suivant cette définition, les concepts clés qui y sont associés sont celui de ***petit groupe*** (ou groupe restreint) et de ***communication***.

Adler et Towne (1991) présente le groupe comme un petit ensemble de personnes qui interagissent, habituellement face à face, pendant un certain temps, dans le but d'atteindre un objectif précis. Pour Schein (1965), le groupe représente un ensemble formé d'un nombre quelconque de personnes qui interagissent les unes avec les autres, qui ont psychologiquement conscience les unes des autres et qui se considèrent comme un groupe. Selon Wilson (1996), une des caractéristiques

importantes d'un groupe est l'influence mutuelle exercée par les interactions et les comportements individuels des membres, ce qui confère à chaque groupe un caractère unique. Dans une définition suffisamment large, le groupe sera défini comme un ensemble de personnes, interdépendantes dans la poursuite en commun d'un ou plusieurs objectifs (Serruys, 1973).

Même si les écrits tendent à suggérer que deux personnes ne puissent constituer un groupe, pour Krech et Crutchfield (1952), dans Serruys (1973), il suffit que deux personnes aient entre elles des relations psychologiques et réciproques pour qu'il y ait un groupe. Selon la perspective de Wilson (1996), trois personnes, au minimum, forment un groupe à cause de la capacité de deux d'entre elles à exercer une influence sur l'autre. Toutefois, ce même auteur est d'avis que la théorie s'appliquant à un groupe peut également s'appliquer à deux personnes interagissant ensemble.

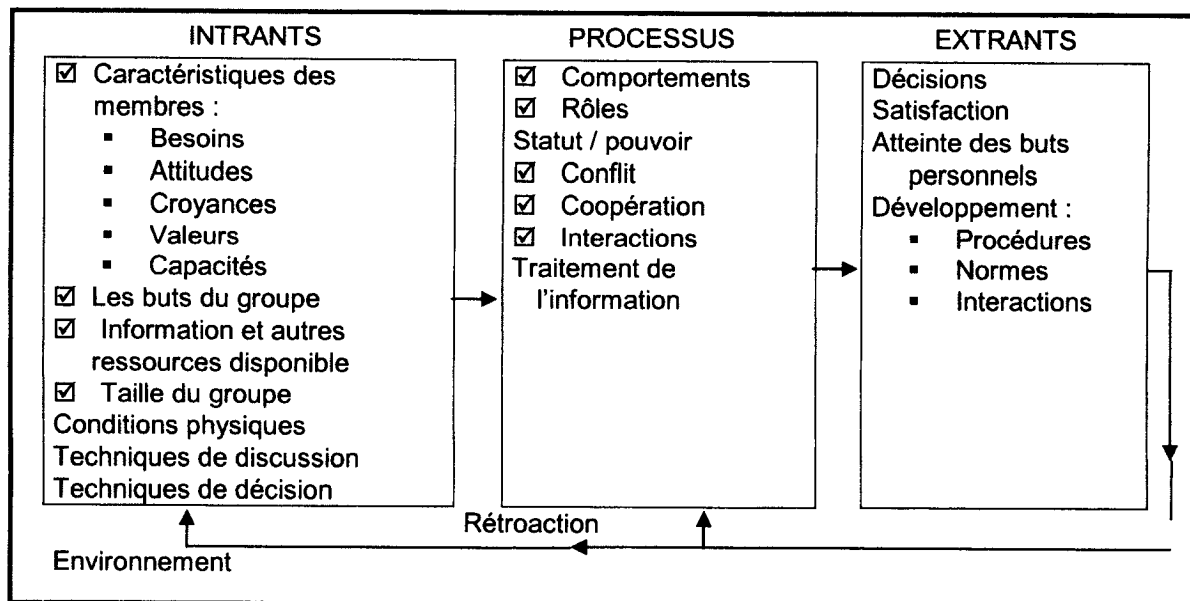
Plusieurs typologies relatives au concept de groupe sont présentes dans les écrits scientifiques. Celle exposée par Serruys (1973) est sans aucun doute la plus complète :

- Groupe primaire et groupe secondaire;
- Groupe formel et groupe informel;
- Groupe permanent et groupe éphémère;
- Groupe libre et groupe obligatoire;
- Groupe de travail et groupe psychique;
- Groupe de réunion (ci-inclus groupe d'information et de négociation, panel, symposium, séminaire, etc.).

Suivant ces différents types de groupe, cet auteur a classé le GDSL comme instrument de sélection et outil d'appréciation individuelle comme un **groupe de réunion**. Ce type de groupe a comme activité principale de réaliser ses objectifs essentiellement au moyen d'un échange de communications verbales.

En fonction de ces fondements théoriques, le modèle de Wilson (1996), représenté par le **Tableau 7**, expose les éléments à définir. Une approche systémique paraît appropriée pour comprendre la dynamique de groupe qui prend forme à chaque administration d'un GDSL. Les intrants ainsi que les processus de groupe relatifs à la période de discussion de groupe seront analysés, alors que les extrants seront traités en cours de route au lieu de faire l'objet d'une discussion séparée. Pour chacun de ses aspects, des liens avec les caractéristiques propres aux GDSL seront effectués.

TABLEAU 7
Les groupes restreints comme un système de traitement de l'information¹⁰



¹⁰ Tiré de l'ouvrage : Wilson, G.L. (1996). Group in context: Leadership and participation in small groups. New-York, McGraw-Hill inc., page 24.

3.2 Les intrants

À partir des concepts suggérés par Wilson (1996), seuls les intrants susceptibles de diminuer la standardisation lors de l'administration d'un GDSL seront analysés. En premier lieu, il va de soi que les caractéristiques individuelles des personnes candidates sont un aspect important à considérer. Par la suite, les buts du groupe ainsi que l'information et les autres ressources disponibles, éléments relatifs à la tâche à accomplir, pourront également faire varier la dynamique de groupe. Finalement, le nombre de participants au groupe de discussions est aussi un facteur susceptible d'apporter des interactions différentes d'une administration à l'autre. En ce qui concerne les conditions physiques, telles que la température, le bruit, l'endroit où a lieu la discussion, elles sont relativement faciles à contrôler par les responsables d'un processus de sélection et demeurent généralement constantes d'une administration à l'autre. Les techniques relatives à la discussion et à la prise de décision ne seront pas discutées dans le cadre de ce projet de recherche, puisqu'elles seront habituellement suggérées par le concepteur et font plutôt référence à la tâche à accomplir (par exemple réunion de remue-méninges où la décision doit être prise par consensus).

3.2.1 Les caractéristiques des membres

Les groupes sont d'abord des réalités composées d'individus. Lewin affirmait que le groupe est influencé par les besoins personnels de ses membres tandis que la personne membre est affectée par les standards du groupe (cité dans Richard, 1995). La citation suivante de Serruys (1973) rend compte de cette prémisse de base :

« Les différentes particularités psychophysiques de l'individu sont structurées et organisées en un ensemble qui met en œuvre, avec son caractère unique et original, l'adaptation au milieu. C'est cette organisation qui constitue la personne. Elle n'est pas statique mais dynamique; elle réagit créativement de manière constante à la situation environnante qui lui est sensible. »¹¹

D'un premier regard, chaque membre du groupe a besoin d'être considéré comme une personne à part entière selon ses particularités uniques. Il désire développer ses potentialités et s'accomplir personnellement, mais il a également besoin d'entretenir des relations avec autrui (Serruys, 1973). En ce sens, ces besoins peuvent être parfois contradictoires (besoin de rester soi-même vs. besoin d'approbation). La façon dont les personnes agiront et traduiront leurs besoins pourra en retour influencer le déroulement de la discussion. En fait, les motifs, la conscience de soi et les traits de personnalité de chaque individu l'amènera à se comporter d'une telle façon afin d'influencer les autres membres et chaque groupe de discussion deviendra une expérience unique. De plus, les croyances et les valeurs de chaque individu affecteront la culture du groupe. Finalement ses capacités, comme sa facilité à s'exprimer oralement, pourront agir sur la progression des travaux du groupe. Selon Field et House (1995), il faut également considérer les caractéristiques démographiques d'ordre moins complexe, mais qui peuvent également avoir un impact sur le processus de groupe, telles que l'âge, le sexe, le statut socio-économique, l'éducation ainsi que l'expérience de travail et de vie.

La composition d'un groupe est généralement définie comme la somme des caractéristiques individuelles de ses membres et s'explique par l'homogénéité ou l'hétérogénéité des membres qu'il réunit. Tel que Dolan, Lamoureux et Gosselin (1996)

¹¹ Tiré de l'ouvrage : Serruys, J.W. (1973). Initiation à la dynamique de groupe. Bruxelles, Les éditions vie ouvrière, page 25.

l'exposent, chaque caractéristique liée à la composition du groupe est susceptible d'apporter des données différentes au processus de groupe.

*« Dans un groupe dit **homogène**, [propriété d'un groupe reposant sur des similitudes psychologiques, sociologiques, économiques, spirituelles ou culturelles entre les membres] la compatibilité des besoins, des motivations ainsi que des personnalités sont synonymes d'une plus grande efficacité, et ce parce que la coopération et la communication y sont particulièrement favorisées. Toutefois, il est important de retenir que l'homogénéité du groupe n'est pas nécessairement synonyme d'efficacité. En effet, si plusieurs membres du groupe possèdent des traits de personnalité qui les incitent à exercer fortement leur leadership, le groupe sera composé de plusieurs chefs et de peu d'indiens et il en résultera une plus faible productivité.*

*Dans un groupe **hétérogène**, [propriété d'un groupe reposant sur des différences psychologiques, sociologiques, économiques, spirituelles ou culturelles entre les membres] la variation dans les caractéristiques individuelles intensifie la qualité de résolution de problèmes éventuels en raison de la stimulation interactive des individus. Par contre, l'hétérogénéité peut engendrer des situations conflictuelles qui auront pour effet de nuire à l'épanouissement du groupe.»¹²*

L'homogénéité du groupe tend à le rendre fonctionnel plus rapidement en favorisant les communications et la coopération. Des normes de comportement auront tendance à se mettre en place plus facilement et les membres du groupe seront généralement plus satisfaits de la progression des travaux (Bormann, 1975). Lors de l'administration d'un GDSL, puisque le temps d'interaction est souvent très limité, l'homogénéité des participants, relativement aux qualifications et à l'éducation, sera importante afin d'atteindre les objectifs d'évaluation (Bass 1954 dans Thorton et Byham, 1982, Tixier, 1974, Lévy-Leboyer, 1996). En effet, l'efficacité de l'outil d'évaluation sera probablement affectée si ces caractéristiques individuelles de base sont très différentes entre les personnes candidates et ce degré d'homogénéité est facilement contrôlable par le responsable du processus de sélection lors de l'étape de la pré-sélection des candidatures. Toutefois, Bass (1954), dans Thorton et Byham (1982), mentionne que le GDSL pourrait être inefficace si les capacités des participants sont très homogènes.

¹² Tiré de l'ouvrage : Dolan, S.L., Lamoureux, G., Gosselin, E. (1996). Psychologie du travail et des organisations. Montréal, Gaétan Morin Éditeur ltée, page 234.

Selon Tixier (1974), il est difficile de rendre un groupe de discussion homogène et certaines connaissances sur la problématique présentée peuvent avantager l'un ou l'autre des participants. Dans un groupe très hétérogène, l'évolution des travaux peut être ralentie et le groupe peut se séparer en coalitions selon les besoins et les caractéristiques de ses membres. Cependant, selon Field et House (1995), l'hétérogénéité d'un groupe peut aider ce dernier à comprendre différentes perspectives et façons d'accomplir la tâche du groupe. De façon résumée, à court terme, l'efficacité d'un GDSL ne sera pas favorisée par des différences individuelles très prononcées entre les membres du groupe.

En somme, la réalité des groupes répond aux lois des individus qui lui sont propres et il est possible de voir le groupe comme une addition d'individus (Serruys, 1973). Selon cette optique, certains auteurs s'inspirent de l'approche psychologique individuelle qui met en évidence la primauté de la personne et stipule qu'un ensemble d'idées et de pensées n'existe que chez les individus. Mais, les individus ne sont pas qu'eux-mêmes, ils sont façonnés par le groupe. Le comportement d'un individu isolé n'est pas nécessairement le même lorsqu'il travaille au sein d'un groupe. C'est selon cette perspective qu'une approche psychosociale sera retenue lors de l'exploration des processus du groupe.

3.2.2 La tâche à accomplir

Les buts du groupe ainsi que l'information et les autres ressources disponibles sont susceptibles de faire varier le processus du groupe. À cet effet, Dolan et al. (1996) mentionnent que :

*« Plus une tâche est complexe, plus les membres du groupe prennent de temps à l'accomplir et plus les probabilités d'en arriver à un consensus sont faibles. Plus la tâche requiert d'informations diversifiées et complexes, plus le risque d'erreur augmente à cause des difficultés de coordination et d'organisation de ces informations ».*¹³

Lors de l'administration d'un GDSDL ces facteurs sont relatifs en grande partie à l'intention du concepteur. Afin de diminuer l'effet de la complexité de la tâche, il serait pertinent que les buts du groupe soient clairement spécifiés dans la mise en situation. Par ailleurs, le thème ne devrait pas privilégier une personne candidate pour ses connaissances du sujet et devrait être de nature fictive. Dans ce contexte, il sera important que l'information requise pour participer à la discussion soit la même pour toutes les personnes candidates et soit spécifiée à l'avance dans la documentation soumise. Au niveau de la complexité, il est essentiel de donner une tâche appropriée à la hauteur des qualifications des participants au GDSDL.

3.2.3 La taille du groupe

Le nombre de participants dans un groupe est un sujet qui a fait l'objet de plusieurs écrits. Des avantages et des désavantages ont été associés aux plus grands groupes. D'après Renz et Greg (2000), une augmentation du nombre de participants permettra de stimuler les interactions et de faciliter la résolution des conflits entre les membres du groupe. Dans cet ordre d'idées, les études indiquent que la performance du groupe et la prise de décision sont généralement améliorées avec l'ajout de membres selon la contribution qu'ils apportent au déroulement de la discussion, mais chaque nouveau membre apporte un gain de performance de plus en plus petit (Hare, 1962). Barker, Wahlers et Watson (2001) ajoutent que les groupes plus grands permettent un meilleur échange d'information en vue de résoudre un problème. Il est donc légitime de

¹³ Tiré de l'ouvrage : Dolan et al. (1996), op. cit. page 265.

penser que « *l'augmentation du nombre d'individus enrichit naturellement la somme d'informations disponibles* » et « *permet d'avoir une gamme d'opinions plus hétérogènes* », mais qu'au-delà d'un certain seuil « *l'addition d'un individu n'ajoute plus rien, car statistiquement un groupe ne peut produire qu'un nombre déterminé de solutions et d'idées originales* » (Amado et Guittet, 1997). Dans le cadre d'un groupe de discussion, Pettersen (2000) mentionne que le fait d'augmenter le nombre de participants augmente la probabilité d'y retrouver un échantillon représentatif, et donc plus comparable d'un groupe à l'autre.

En ce qui concerne les désavantages, plusieurs auteurs rappellent qu'une augmentation du nombre de participants peut diminuer considérablement le temps d'intervention d'une personne et que cette diminution n'est pas distribuée de façon égale entre les membres du groupe (Pavitt et Curtis, 1990; Barker et al, 2001; Renz et Greg, 2000). Bales et ses collaborateurs (1951) ont en effet observé que le pourcentage de temps d'intervention des personnes moins volubiles passait de 26% à 8% lorsque les groupes étaient constitués de trois ou huit personnes, alors qu'il ne diminuait que de 4% (de 42% à 38%) dans le cas des personnes plus volubiles. Selon Pavitt et Curtis (1990), cet écart important entre le temps d'intervention des participants peut affecter la structure perçue du pouvoir détenu par chacun des membres. À cet effet, Wilson (1996) mentionne qu'un groupe comprenant un nombre de participants très large aura besoin de davantage de structure afin de permettre à tous les participants de s'exprimer ainsi qu'une prise de décision rapide, les communications étant plus difficiles que dans un groupe restreint. De plus, Carron et Spink (1995) indiquent que dans les plus grands groupes, les personnes qui prennent une position de leadership ont tendance à adopter une attitude plus autocratique. Finalement, Pavitt et Curtis (1990) ainsi que Tedeschi et Lindskold (1976) mentionnent que les plus grands groupes sont plus propices à la

création de coalitions visant à augmenter le pouvoir respectif des participants. En somme, plus les groupes sont grands, plus ils auront tendance à se diviser en sous-groupes et à avoir des non-participants (Hare, 1981 dans Field et House, 1995).

Le nombre de personnes dans un groupe est une variable qui affecte sans contredit la quantité et la qualité des communications entre les individus ainsi que la nature des relations interpersonnelles que les membres développeront entre eux (Krech et Crutchfield (1948) dans Hare, 1962). Selon Dolan et ses collègues (1996), dans un groupe de trois à cinq personnes, les échanges d'idées et les interactions sont plus faciles ce qui amène une plus grande satisfaction des membres. C'est dans les grands groupes que les individus manifestent le plus rapidement leur désintérêt et leur mécontentement. En résumé, plus il y a de participants dans le groupe, plus les discussions tourneront autour de quelques personnes et plus les interactions de groupe déclineront (Bormann, 1975).

Pour Hare (1962), un groupe de discussion exige au moins trois personnes et au plus de douze à quinze. Selon cet auteur, le nombre de participants optimal est de cinq personnes. Toutefois, il met en perspective les avantages et désavantages des grands groupes selon le type de sujet à discuter. De cette façon, la recherche d'une solution logique serait mieux conduite par un petit groupe d'environ trois personnes. Le groupe créatif, où la résolution du problème demande plusieurs solutions différentes, serait optimal avec environ six personnes. L'examen critique d'un problème, une situation dans laquelle il est nécessaire de confronter des opinions et des points de vue différents, devrait être assuré au mieux par un ensemble d'une douzaine de personnes. Bien entendu, ces chiffres ne sont présentés qu'à titre indicatif.

3.3 Les processus

Comme il l'a été démontré, les processus des GDSL sont donc influencés par plusieurs variables internes à la composition du groupe et à la nature de la tâche à accomplir. À l'inverse des exercices écrits ou même des entrevues, les GDSL utilisent les mécanismes de la psychologie sociale pour apprécier le comportement individuel des participants (Tixier, 1974) et les processus qui animent la dynamique du groupe trouvent leur fondement dans ces théories. Il s'agit du psychologue Kurt Lewin qui a popularisé le terme dynamique de groupe (Serruys, 1973). Pour cet auteur, un groupe adopte une forme déterminée par l'équilibre des forces, des tensions et par le « *champ perceptif* » des individus (Amado et Guittet, 1997). Ainsi, pour celui-ci, les groupes doivent s'appréhender comme des totalités dynamiques qui résultent des interactions entre les membres.

En fonction des éléments portant sur les écrits sur la psychosociale des groupes et la typologie de Wilson (1996), les comportements, les rôles, le climat ainsi que les interactions seront discutés comme processus ayant lieu dans le cadre d'un groupe de discussion. Puisque les participants à un GDSL sont placés sur un pied d'égalité, les écrits portant sur le statut et le pouvoir des membres d'un groupe (voir Tableau 5) ne feront pas l'objet d'une attention particulière aux fins de la présente discussion.

3.3.1 Les comportements

Benne et Sheats (1948) ont identifié trois catégories de comportements, soient ceux associés à l'atteinte des résultats, ceux reliés à l'identité du groupe et ceux touchant aux besoins individuels des membres. Les comportements reliés à la tâche consistent à définir et à résoudre les problèmes communs de la discussion. Les

fonctions reliées à l'identité du groupe visent la construction, l'autoformation du groupe ainsi que le maintien des comportements centrés sur le groupe. Finalement, d'autres comportements sont uniquement centrés sur les besoins individuels des membres et s'opposent à la formation et au maintien du groupe (Richard, 1995). Le **Tableau 8** présenté à la page suivante dresse l'inventaire de l'ensemble de ces comportements identifiés par Benne et Sheats (1948).

TABLEAU 8
Répartition des comportements dans un groupe selon trois pôles fonctionnels¹⁴

Les comportements reliés à la tâche	
1.	Amorce, suggère ou propose des idées nouvelles relatives aux buts, à la définition du problème à des solutions.
2.	Demande des informations en vue de clarifier d'obtenir des faits.
3.	Demande des opinions sur les faits en vue de clarifier les valeurs.
4.	Donne des informations, des faits.
5.	Offre des opinions personnelles, ses croyances concernant les suggestions, ce que le groupe devrait considérer comme valeurs.
6.	Coordonne des idées données ou explicite par des exemples ce qui a été suggéré.
7.	Orienté, définit la position du groupe, synthétise.
8.	Évalue, manifeste un esprit critique.
Les comportements reliés à la construction et au maintien du groupe	
1.	Encourage, félicite, donne son accord, montre de la chaleur, de la solidarité dans ses attitudes envers les autres membres.
2.	Harmonise les différences entre les participants, résout les tensions en plaisantant.
3.	Facilite la participation des autres, ouvre les communications.
4.	Propose un idéal à atteindre comme groupe.
5.	Observe le groupe, fait des commentaires sur sa marche.
6.	Marche dans le mouvement du groupe, adhère volontairement ou passivement, écoute les idées des autres.

¹⁴ Tiré de l'ouvrage : Richard, B. (1995). Psychologie des groupes restreints. Québec, Presses Inter-Universitaires, page 68.

TABLEAU 8 (suite)

Répartition des comportements dans un groupe selon trois pôles fonctionnels

Les comportements reliés aux besoins individuels des membres	
1.	Attaque le groupe ou ce qu'il fait, déprécie autrui.
2.	Bloque, fait de l'obstruction systématique, est toujours contre.
3.	Cherche à mobiliser l'attention sur soi en se vantant, en racontant ses réussites personnelles.
4.	Utilise le groupe comme auditoire pour exprimer son idéologie, ses sentiments personnels, se confesse lui-même au groupe.
5.	Manque d'intérêt, d'implication, s'exprimant sous forme de cynisme, de bouffonnerie.
6.	Demande constamment de la sympathie, soit par insécurité, soit par mésestime personnelle.
7.	Domine le groupe, impose ses idées, sa supériorité sous forme de vantardise, d'affirmation de sa supériorité.

La présence ou l'absence de participants adoptant un ou plusieurs de ces comportements peut avoir un effet critique sur l'efficacité du groupe (Wilson, 1996). Lorsque l'ensemble de ces comportements sont présents, le groupe, ainsi que ses membres, sont généralement productifs et peuvent atteindre les objectifs recherchés. L'absence d'un ou plusieurs de ces types de comportements peut, au contraire, diminuer l'efficacité globale et donc avoir un impact sur le rendement d'un, de plusieurs ou même de tous les participants. Dans le même ordre d'idées, certains comportements reliés aux besoins individuels des membres, habituellement improductifs, pourront exercer un impact négatif sur le groupe et influencer, par conséquent, le rendement d'un ou de plusieurs participants à un GDSDL.

3.3.2 Les rôles

Benne et Sheats (1948) ont également précisé qu'il existait vingt-neuf rôles découlant de l'ensemble de ces comportements. Le **Tableau 9** présente ces rôles, toujours en fonction des trois pôles fonctionnels identifiés précédemment. Il est à noter que l'ensemble des termes ont été traduits librement à partir de l'information présentée dans l'ouvrage de Wilson (1996).

TABLEAU 9
Les rôles fonctionnels dans les groupes restreints¹⁵

Les rôles reliés à la tâches			
1.	Initiateur / contributif	2.	Chercheur d'informations
3.	Donneur d'informations	4.	Chercheur d'opinions
5.	Donneur d'opinions	6.	Élaboration / clarification
7.	Coordonnateur	8.	Poseur de diagnostics
9.	Orienteur / synthétiseur	10.	Énergique
11.	Assistant aux procédures	12.	Secrétaire / preneur de notes
13.	Évaluateur / critique		
Les rôles reliés à la construction et au maintien du groupe			
14.	Supporteur / encourageant	15.	Harmoniste
16.	Libérateur de tensions	17.	Chercheur de compromis
18.	Vigil des réseaux de communication	19.	Expressif
20.	Procédurier	21.	Suiveur
Les rôles reliés aux besoins individuels des membres			
22.	Bloqueur	23.	Agresseur
24.	Déserteur	25.	Dominateur
26.	Narcissique	27.	Confesseur
28.	Clown	29.	Prêcheur

En résumé, un groupe restreint est composé d'une culture particulière avec des rôles qui reflètent les points de vue des membres à propos des besoins du groupe, des talents ainsi que des préférences de ses membres (Hare, 1962). Les rôles sont un ensemble de comportements que les membres adopteront en fonction des attentes des autres membres du groupe (Wilson, 1996). Tout comme pour les comportements, ces rôles peuvent être tant positifs que négatifs et auront une influence sur le processus de groupe et son efficacité.

¹⁵ Tiré de l'ouvrage : Wilson, G.L. (1996). Group in context: Leadership and participation in small groups. New-York, McGraw-Hill inc., pages 195-201.

3.3.3 Le climat

Les processus de compétition et de collaboration entre les membres affectent le climat qui à son tour influence le comportement de ses membres (Barker et al., 2001). Le climat fait référence aux conditions de fonctionnement, de communication et de travail, faites au groupe par ses membres et par le milieu institutionnel dans lequel s'insère le groupe (Richard, 1995). Par exemple, Gibb (1961) a démontré que la présence de certains comportements, tels que démontrer de la supériorité ou tenter de manipuler les autres dans un groupe, pouvaient produire un état défensif alors que d'autres comportements, par exemple la résolution de problèmes ou l'égalité entre les membres, contribuaient à diminuer cet état. Alderton et Frey (1986) ont également remarqué que les membres d'un groupe devenaient plus défensifs lorsqu'il y avait des différences d'opinion.

Bien que le format du GDSL spécifiera à l'avance si le groupe de discussion évoluera dans un contexte de collaboration ou de compétition, il est pertinent d'exposer les caractéristiques propres à ces deux situations énoncées par Deutsch (1949) dans Richard (1995). Selon lui, la compétition provoque :

- L'insécurité personnelle, l'anxiété;
- Les tensions entre les membres;
- Les rivalités entre les personnes;
- L'inhibition de la communication;
- La suspicion et la méfiance entre les candidats;
- Un rendement moyen faible;

- Des remarques agressives;
- Très peu de satisfaction dans le groupe.

Toujours selon Dueutsch (1949), la situation de collaboration favorise :

- La cohésion dans le groupe;
- La coordination des efforts;
- La subdivision de l'activité;
- Le sens de l'obligation entre les membres;
- Une pression à produire;
- La diversité dans la quantité des contributions;
- Une attention portée à la personne des membres;
- La compréhension mutuelle;
- Des rapports amicaux;
- Une communication assurée entre les membres;
- Une production de meilleure qualité.

3.3.4 Les interactions

Dans les écrits récents, les interactions entre les personnes sont certainement les exemples les plus fréquemment utilisés pour décrire le groupe et son action (Adler et Towne 1991). À travers les écrits scientifiques, la grille d'analyse élaborée en 1950 par Bales (dans Richard, 1995) s'avère la plus complète afin de comprendre les interventions des participants. Tout d'abord, une intervention peut se situer à différents niveaux, soit d'information, de formulation d'opinion ou évaluation, expression d'une direction ou d'une suggestion et finalement elle reflète un sentiment ou une tension.

Selon Bales, toutes ces interventions peuvent être classées par rapport à six types de problèmes constamment présents dans un groupe, soit de communication, d'évaluation, de contrôle, de décision, de tension et d'interprétation.

Le **Tableau 10** expose cette vaste gamme d'interventions. Il est à noter que les interactions n° 1 à n° 3 font partie de ce que Bales appelle la zone socio-affective positive, les interactions n° 4 à n° 9 à l'aire des tâches socio-opératoires et finalement, les interactions n° 10 à n° 12 à la zone socio-affective négative. À partir des informations recueillies, Bales a formulé l'hypothèse selon laquelle les processus d'interaction reproduisent toujours les mêmes phases, car les échanges entre les individus passent toujours par des séquences identiques, soit la phase d'information, d'évaluation, d'influence et de contrôle, enfin la phase de décision.

TABLEAU 10
Les catégories d'interventions de Bales¹⁶

	Interventions	Problèmes
Réactions socio-affectives positives	1. de solidarité, d'encouragement	d'interprétation
	2. de détente, de satisfaction	de tension
	3. d'approbation, d'accord	de décision
Réactions à la tâche	4. d'apport de suggestion	de contrôle
	5. d'apport d'opinion	d'évaluation
	6. d'apport d'information, d'orientation	de communication
	7. de demande d'information, d'orientation	de communication
	8. de demande d'opinion, d'évaluation	d'évaluation
	9. de demande de suggestion	de contrôle
Réactions socio-affectives négatives	10. de désapprobation passive, de rejet	de décision
	11. de tension	de tension
	12. d'antagonisme	d'interprétation

¹⁶ Tiré de l'ouvrage : Richard, B. (1995). Psychologie des groupes restreints. Québec, Presses Inter-Universitaires, page 80.

Les interactions et les comportements de chaque individu composant le groupe peuvent donc influencer la dynamique et la progression des travaux du GDSL. Thornton (1992) indique que le climat et le ton de la discussion peuvent être différents d'un groupe de discussion à l'autre, allant de très animé à plutôt tranquille et soumis, dépendamment de la composition et de l'humeur du groupe. Gatewood et al. (1990) et Pettersen (2000) sont d'avis que le manque de standardisation en ce qui a trait aux différences individuelles des participants (par exemple, plus ou moins agressifs, collaborateurs, analytiques, etc.) peut influencer le comportement des candidats et ainsi affecter la fidélité des résultats. Il y a donc lieu de croire que la dynamique entourant la discussion, tout comme le temps d'intervention de chacune des personnes candidates, seront également influencés par le nombre de participants au GDSL.

En somme, les écrits présentent quatre facteurs qui peuvent contribuer de façon importante à diminuer la standardisation des groupes de discussion sans leader et qui doivent être pris en considération au moment d'évaluer les personnes candidates en situation de groupe. Ces facteurs concernent la nature du thème (la problématique) proposé dans la mise en situation (Gatewood et al., 1990), le format du GDSL (c'est-à-dire de collaboration ou de compétition) (Gatewood et al., 1990; Thornton, 1992), le nombre de participants au groupe (Pettersen, 2000) et les différences individuelles des participants quant à leurs comportements (Gatewood et al., 1990; Thornton, 1992; Pettersen, 2000). Dans le cadre de ce mémoire, il s'avère impossible de contrôler l'ensemble de ces facteurs. Cependant, on peut tenter de faire varier certains d'entre eux et contrôler les autres afin d'examiner leur effet sur la performance des personnes candidates dans une situation d'évaluation comprenant un enjeu de sélection.

PARTIE 2 Question de recherche

Chapitre 4 Modèle théorique

Il semble légitime que le déroulement d'un GDSL soit influencé par plusieurs facteurs relatifs aux intrants de ce groupe ainsi qu'au processus qui a lieu lors de cette réunion. Par conséquent, ce problème risque d'affecter la qualité psychométrique de cet instrument de mesure en contexte de sélection, plus particulièrement le degré de constance et de précision des résultats qu'il produit. La question est donc de savoir si les résultats d'une personne candidate sont influencés par ce manque de standardisation à propos des conditions d'administration. En effet, s'il est normal qu'un participant agisse différemment en fonction de la situation dans laquelle il se trouve, on ignore cependant si les résultats qu'il obtient par rapport à l'évaluation de sa performance au GDSL seront affectés négativement ou positivement.

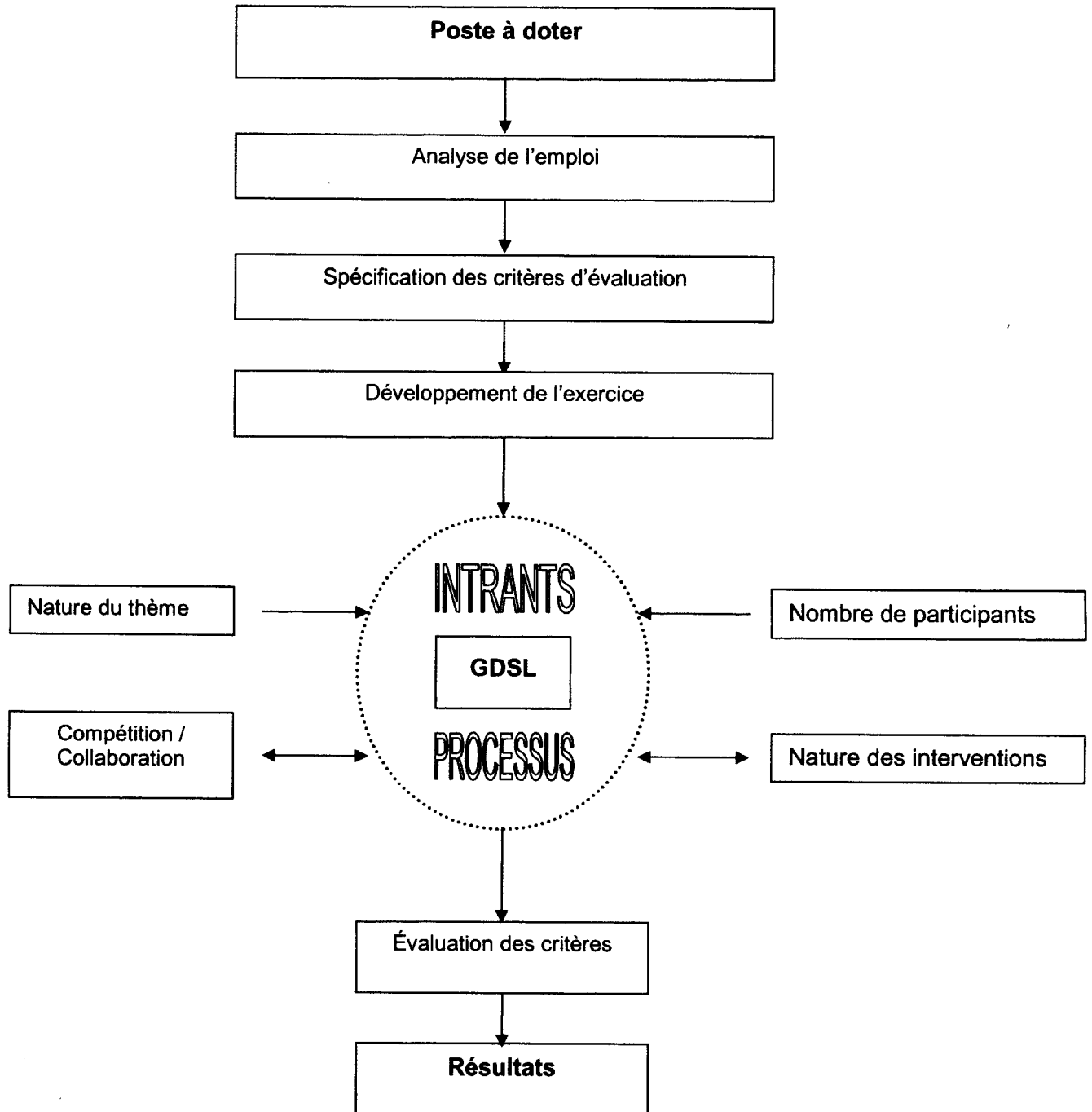
L'objectif du présent projet de recherche est donc d'évaluer l'impact de ce manque de standardisation potentiel des GDSL sur les résultats des personnes candidates en situation réelle de sélection. Dans cette optique, selon les quatre facteurs qui peuvent contribuer de façon importante à diminuer la standardisation des groupes de discussion sans leader présentés à la suite de la recension des écrits, soit la nature du thème, le format du GDSL (collaboration vs. compétition), le nombre de participants au groupe et la nature des interventions des participants, seuls les deux derniers facteurs seront manipulés et les deux autres seront contrôlés.

Cette recherche permettra donc de vérifier *si les résultats des sujets aux groupes de discussion sans leader varient en fonction de la nature des interventions et du nombre de participants et ce, dans un contexte de sélection.*

4.1 Le modèle conceptuel et ses variables

De façon générale, la **Figure 1** présentée à la page suivante désigne le déroulement logique de l'utilisation du GDSL dans le cadre d'un processus de sélection. Ce modèle a été retenu en premier lieu afin d'identifier les facteurs susceptibles de diminuer la fidélité des résultats obtenus au GDSL dans une perspective globale.

Figure 1
Utilisation du GDSL dans le cadre d'un processus de sélection



4.1.1 Variable dépendante

La variable dépendante de ce projet de recherche est représentée par les résultats obtenus par les personnes candidates aux différents critères d'évaluation pertinents pour le poste. Dans le cadre d'un GDSL, l'attribution des résultats est établie par des évaluateurs qui ont été formés spécifiquement sur l'observation et la notation de comportements reliés aux critères d'évaluation retenus. La variable dépendante varie donc en fonction des comportements des personnes candidates et des observations des évaluateurs. Tel que précisé au premier chapitre de cette recherche, chaque évaluateur surveille généralement un ou deux participants puis compile ses observations au fur à mesure qu'elles se manifestent. Par la suite, chaque observateur trie et classe ses observations sous les critères appropriés en fonction d'une échelle de comportements standardisée et il attribue une cote à chaque critère. À la fin de la ou des journées d'évaluation, chaque observateur fait un compte rendu aux autres observateurs, présentant une vue d'ensemble du déroulement de l'exercice, une description par comportement des observations notées et un résumé concis de son évaluation.

L'échelle d'évaluation des résultats varie habituellement sur un continuum de un à cinq et celle présentée au **Tableau 11** a été validée à plusieurs reprises, notamment dans le cadre de processus de sélection au sein de différents ministères fédéraux et provinciaux.

TABLEAU 11
Échelle d'évaluation des compétences¹⁷

5	Excellent Le candidat répond aux principaux aspects et critères de l'évaluation et ses réponses sont appropriées. Le candidat a surpassé les attentes.
4	Très bon Le candidat répond à la plupart des critères de l'évaluation et aucune lacune majeure dans les aspects évalués a été notée. Le candidat démontre une performance au-dessus de la moyenne.
3	Bon Quelques-uns des critères principaux et secondaires de l'évaluation ont été soulevés. Le candidat démontre quelques lacunes mineures au niveau du critère évalué.
2	Faible Le candidat omet des aspects et des critères de l'évaluation importants. Le candidat affiche quelques lacunes majeures et éprouve des problèmes.
1	Médiocre Le candidat omet la plupart des aspects évalués et affiche plusieurs lacunes majeures. Il est incapable de répondre ou ses réponses sont inappropriées.

4.1.2 Variables indépendantes

À la lumière des écrits et selon la question de recherche retenue, les variables indépendantes de ce modèle sont donc la nature des interventions auxquelles la personne candidate doit faire face et le nombre de participants au groupe. Ces variables indépendantes ont été choisies étant donné l'influence prépondérante qu'elles peuvent avoir sur les résultats aux groupes de discussion (Gatewood et al., 1990, Thorton, 1992 et Pettersen, 2000).

En ce qui concerne la nature des interventions, un continuum contribution positive vs. négative, ce que Bales (1950) appelle la zone socio-affective positive vs. zone socio-affective négative, permettra une meilleure classification des comportements, des rôles et des interactions des individus. Les comportements et les rôles positifs seront donc reliés à la construction et au maintien du groupe, alors que les comportements et les rôles négatifs seront reliés aux besoins individuels des membres.

¹⁷ Tiré du site web : Agriculture et agroalimentaire Canada, adresse URL http://www.agr.gc.ca/hr-rh/index_f.php?s1=cc-ce&page=adapt

Conséquemment, cette division permettra de circonscrire avec plus de précision l'influence de la nature des interventions sur la variable dépendante.

Par rapport à la taille du groupe, Pettersen (2000) mentionne que le fait d'augmenter le nombre de participants augmente la probabilité d'y retrouver un échantillon représentatif et, partant, plus comparable d'un groupe à l'autre. En fait, puisque certaines études tendent à démontrer que deux personnes ne peuvent constituer un groupe et que les conditions de standardisation sont très faibles avec deux personnes (chaque participant est évalué dans des conditions différentes), il est pertinent de comparer les résultats d'une personne à la suite de sa participation à un GDSL de deux individus et à un autre de plus grande taille.

4.2 Hypothèses de recherche

À ce stade, il est nécessaire de présenter les hypothèses énoncées en fonction des deux conditions expérimentales qui façonnent la question de recherche.

H1 : Hypothèses sur l'effet de la nature des interventions sur les résultats

H₁₀ : Les résultats des personnes candidates à un GDSL sur une même compétence sont exempts de variation peu importe la nature des interventions auxquelles ces personnes sont soumises.

H₁₁ : Les résultats des personnes candidates à un GDSL sur une même compétence varient en raison de la nature des interventions auxquelles ces personnes sont soumises.

H2 : Hypothèses sur l'effet du nombre de participants sur les résultats

H2₀ : Les résultats des personnes candidates à un GDSL sur une même compétence sont exempts de variation peu importe le nombre de personnes prenant part à l'exercice.

H2₁ : Les résultats des personnes candidates à un GDSL sur une même compétence varient en raison du nombre de personnes prenant part à l'exercice.

Les hypothèses nulles **H1₀** et **H2₀** seront rejetées et les hypothèses alternatives **H1₁** et **H2₁** seront acceptées si, d'une part, la nature des interventions à l'intérieur d'un GDSL provoque un effet significatif sur les résultats des personnes candidates à un GDSL et si, d'autre part, le nombre de personnes candidates à une GDSL provoque un effet significatif sur les résultats des personnes candidates. Pour ces hypothèses, le seuil de signification statistique sera représenté par un niveau de probabilité inférieur ou égal à 0,05.

4.3 Implications théoriques et empiriques

Les conclusions de ce projet de recherche auront de nombreuses implications tant au plan théorique qu'empirique. De façon plus spécifique, il sera possible de vérifier si l'introduction des variables indépendantes amène un biais susceptible de faire en sorte que les personnes candidates récoltent de meilleures ou de pires résultats selon que le groupe est plus grand ou constitué de candidats ayant des comportements plus ou moins positifs. D'un point de vue pratique, ce projet de recherche permettra de vérifier si une carence de standardisation pour ces deux facteurs peut être préjudiciable

pour les candidats à un processus de sélection ou conduire à de mauvaises décisions organisationnelles en matière de sélection de personnel.

4.4 Stratégie de recherche

Selon le collectif de Thiétart (1999), « *lorsque l'on se préoccupe de la fiabilité, il s'agit de s'assurer que si l'on mesure le même objet ou le même phénomène avec le même instrument de mesure, on obtient des résultats les plus similaires possibles* »¹⁸. Pour cela, il est nécessaire de calculer les corrélations entre des mesures reproduites d'un même objet obtenues par un même instrument. Cette reproduction peut être effectuée dans le temps (différentes mesures à des moments différents) ou par des individus différents (différents observateurs, différents codeurs). Dans le cadre de ce projet de recherche, toutefois, l'importance sera accordée à l'estimation de la fidélité par formes équivalentes (différentes mesures) et de la fidélité interexamineurs (différents codeurs).

L'objectif étant de mesurer l'influence de certains facteurs sur la variation des résultats des participants à un GDSL et de vérifier s'ils représentent une source d'erreur de mesure, le moyen privilégié pour prouver qu'une variable quelconque est la cause d'une autre variable demeure la méthode expérimentale. En effet, l'expérimentation désigne le dispositif par lequel le chercheur manipule les variables et observe les effets de cette manipulation sur d'autres variables (Thiétart et coll., 1999).

¹⁸ Tiré de l'ouvrage : Thiétart, R.A. et coll. (1999). Méthodes de recherche en management, Paris, Dunod, page 265.

Alors afin de vérifier ***si les résultats des sujets aux groupes de discussion sans leader varient en fonction de la nature des interventions des participants et du nombre de participants au groupe et ce, dans un contexte de sélection***, il est donc nécessaire de soumettre les sujets à une évaluation semblable et comparer les résultats selon ces deux pôles, soit à des moments différents et par des observateurs différents. L'opérationnalisation de ces variables et le schème expérimental associé à cette recherche seront présentés en détail dans la prochaine partie.

PARTIE 3 Méthodologie

Chapitre 5 Opérationnalisation des variables

Pour rendre les variables opérationnelles, la méthode des formes alternatives a été retenue, c'est-à-dire administrer quatre instruments de même nature aux même individus, mais portant sur des sujets différents (Thiétart et coll., 1999). Afin de recréer une situation similaire à celle observée en contexte de sélection, des démarches ont dû être entreprises avec différentes organisations afin d'avoir accès à des sujets participant à un véritable processus de sélection par le biais d'une évaluation des compétences. Comme un conseil scolaire francophone de la région était à la recherche de moyens d'évaluation afin de sélectionner des candidats en vue de l'obtention d'un poste de direction d'école et ainsi identifier leurs forces et faiblesses. Par la suite, trois autres commissions scolaires de la région se sont jointes à la recherche en soumettant également leurs candidats à une évaluation par le biais des différents GDSL.

Ces différents groupe de discussion sans leader ont été élaborés selon le processus général d'élaboration d'instruments de mesure fondés sur la validation de contenu (Pettersen, 2000) et ce, selon l'information recueillie auprès d'un comité de travail réunissant dix intervenants du milieu scolaire (surintendants, directeurs et conseillers en ressources humaines). L'idée principale, tant du point de vue de la recherche que dans le cadre des processus de sélection, était d'identifier les compétences essentielles à un poste de direction d'école, de développer des mises en situation représentatives de l'emploi qui permettrait d'évaluer ces compétences et, finalement, de soumettre les personnes candidates à une évaluation par le biais

d'instruments de mesure présentant différentes conditions. Mais voyons quelles étapes ont été réalisées afin d'y parvenir.

5.1 Étape 1 : Finalités et spécification du domaine à mesurer

5.1.1 Finalités de l'instrument de mesure

Dans un contexte de planification stratégique de la relève et selon l'approche par compétences, les instruments de mesure élaborés visaient à évaluer des candidats et identifier leurs forces et faiblesses dans le but de sélectionner ceux démontrant le meilleur potentiel pour accéder à un poste de direction scolaire et ce, sur la base de leurs résultats aux différents instruments de mesure.

5.1.2 Analyse et description de l'emploi

Lors de cette étape cruciale, il s'agissait de bien comprendre le contexte dans lequel s'inscrit le poste pour lequel les participants devront être sélectionnés. L'étape d'analyse basée sur l'approche par compétences implique autant une compréhension globale de la mission, de l'environnement ainsi que de la direction stratégique privilégiée par l'organisation (Durivage, 2002). Ces analyses doivent tenir compte du niveau du poste (importance des responsabilités, des pouvoirs et de l'obligation de rendre compte), du rôle (activités et responsabilités majeures et mineures), du stade de développement organisationnel (cycle de vie) et des variables contextuelles (par exemple, les valeurs de l'organisation) (Slivinski et Miles, 1996 dans Durivage, 2002).

Cette étape a été réalisée à partir d'une analyse de tâches, d'une revue de la documentation pertinente (planification stratégique, rapport annuel, etc.) ainsi qu'à partir

des informations transmises par les intervenants du comité de travail lors de groupes de discussion.

5.1.3 Spécification du domaine à mesurer

Après une analyse rigoureuse de l'emploi, il est alors possible d'identifier les compétences nécessaires à la réalisation du poste ou de la fonction. L'identification des compétences à partir des responsabilités identifiées se fait habituellement à partir d'un dictionnaire ou d'un inventaire qui comporte un nombre limité de compétences. L'identification des compétences et des connaissances requises doit se faire en étroite collaboration avec les personnes qui connaissent bien la fonction, soit les gestionnaires et les employés en place. Cette étape a donc été complétée à la suite de rencontres avec deux groupes de travail, soit un composé des supérieurs du poste et un autre composé de titulaires des postes.

En somme huit compétences ont été identifiées par les membres des groupes de travail comme étant essentielles pour offrir un rendement supérieur en emploi, soit :

- | | |
|------------------------------------|------------------------------------|
| (1) Sens de l'environnement | (5) Capacité d'analyse |
| (2) Prise de décision | (6) Orientation vers les résultats |
| (3) Communication interpersonnelle | (7) Flexibilité |
| (4) Mobilisation d'équipe | (8) Communication orale |

Pour chacune de ces compétences, une définition et des indicateurs comportementaux, au nombre de sept en moyenne par compétence, ont été conçus (voir exemple présenté ci-dessous au **Tableau 12**). Les compétences et leurs indicateurs ont été soumis et approuvés par la direction générale de chacune des quatre commissions scolaires.

Tableau 12
Exemple d'une compétence évaluée

Relations interpersonnelles
Capacité à établir et à entretenir, avec autrui, des rapports contribuant à l'atteinte des objectifs de l'organisation, en favorisant le respect, l'écoute et la compréhension mutuelle.
▪ Est capable de discerner les besoins des autres et de décoder leurs opinions et leurs pensées
▪ Sait résoudre des situations interpersonnelles délicates et complexes
▪ Ses relations interpersonnelles sont teintées de courtoisie et de respect en ce qui concerne les opinions et les positions des autres
▪ Sait établir rapidement des relations positives et constructives
▪ Est reconnu pour sa grande capacité à écouter les autres
▪ Collabore avec les autres
▪ Démonstre une attitude positive et ce, malgré des circonstances parfois difficiles
▪ Représente un élément positif dans une équipe
▪ Est réfléchi et prudent dans ses actions et ses affirmations
▪ Réagit avec tact et diplomatie en tout temps

5.2 Étape 2 : Développement des instruments de mesure

Lors de différents groupe de discussion, le comité de travail a permis de relever les éléments observables et pertinents reliés à l'emploi, tels que ce qui doit être effectué (éléments, tâches et responsabilités), ce qui doit en résulter (produits ou résultats), le niveau de rendement attendu, le contexte de l'emploi, et, finalement, les exigences requises du point de vue de l'organisation. Par ailleurs, ces rencontres ont également permis d'obtenir des incidents critiques des postes à doter (des situations courantes, des difficultés rencontrées, les meilleures façon de régler certains problèmes, les comportements à adopter, etc.). En fonction des ressources et des contraintes

financières et logistiques provenant autant des organisations scolaires que de celles associées à la recherche, il a été nécessaire de préciser le format des instruments de mesure, le type de mise en situation privilégié, la durée et le nombre approximatif de problématique, le mode de correction, le processus d'interprétation et l'usage de normes et, finalement, le nombre de version des instruments.

En somme, quatre GDLS ont été élaborés et chacune des mises en situation portaient sur un thème différent, soit : (1) un sondage de satisfaction auprès du personnel, (2) une plainte d'un parent, (3) un problème de pédagogie et (4) l'intégration de technologies informatiques. Les mises en situation comportaient une problématique de travail réaliste pour un poste de directeur d'école et présentaient un degré de complexité et de difficulté jugé équivalent par le comité de travail. En effet, dans tous les cas, celui-ci a validé autant le contenu des instruments de mesure que leur format.

Afin de contrôler un aspect relatif au processus et susceptible de diminuer la standardisation de l'outil, dans tous les GDLS, le format privilégié a été un climat de collaboration. Les candidats devaient résoudre la problématique et en arriver à des propositions communes, sans se voir imposer de restrictions qui pourraient les mettre en rivalité. L'objectif ultime de la discussion était donc de parvenir à un consensus sur les solutions à privilégier afin de résoudre les problématiques de chacune des mises en situation. De plus, les candidats étaient avisés qu'ils n'étaient pas en compétition les uns avec les autres et que leur candidature serait retenue, pour l'étape suivante, à la condition que leurs résultats individuels atteignent la note de passage (trois sur une échelle en cinq points).

En ce qui a trait aux directives reçues par les personnes candidates, ces dernières spécifiaient que le contexte et les informations contenues dans chacune des mises en situation étaient de nature fictive. Même s'il est possible que certains renseignements s'apparentent ou soient même identiques à des situations vécues au sein de l'organisation du participant, celui-ci devait demeurer attentif au contenu des mises en situation et utiliser l'information fournie. De plus, la personne candidate était avisée qu'il s'agit de son comportement individuel et de sa propre contribution aux mises en situation qui sont évalués par le biais des mises en situation.

Le déroulement de l'exercice a été divisé en deux parties. La première étape, de nature individuelle, consistait à prendre connaissance du matériel de base et se préparer à la discussion de groupe. Pour chacune des mises en situation, les personnes candidates disposaient d'une période de **vingt (20) minutes** afin de se préparer à échanger avec eux sur leur position et leurs conclusions à l'égard de la problématique. Au cours de la deuxième étape, les participants devaient discuter du sujet avec le groupe de candidats selon un période de temps établie selon le nombre de participants.

5.3 Étape 3 : Implantation et enregistrements des données

5.3.1 Schème expérimental

Tel que présenté au **Tableau 13**, le schème expérimental de ce projet de recherche spécifie les composantes de l'expérience à mener et combine systématiquement les variables indépendantes « **nombre de participant** » et « **nature des interventions des autres participants** ». Chaque GDSL a donc varié en fonction

de la contribution positive ou négative apportée par un des membres du groupe et le nombre de personnes participant à la discussion. Le nombre de participants a varié entre deux et quatre personnes. Quant à la nature des interventions des autres participants, elle a été manipulée grâce au concours d'acteurs qui apportaient une contribution positive ou négative au déroulement de la discussion (voir section suivante).

TABLEAU 13
Schème expérimental

NATURE DES INTERVENTIONS DES AUTRES PARTICIPANTS	NOMBRE DE PARTICIPANTS	
	Deux participants	Quatre participants
Positive	GDLS 1	GDLS 3
Négative	GDLS 2	GDLS 4

Dans chaque GDLS, il n'y avait qu'un seul candidat évalué alors que les autres participants étaient tous des acteurs qui jouaient un rôle établi à l'avance. Dans le cas des GDLS comportant deux personnes, l'acteur adoptait soit des comportements positifs (GDLS 1), soit des comportements négatifs (GDLS 2). Dans le cas des GDLS à quatre personnes, un seul des acteurs exprimait les comportements positifs (GDLS 3) ou négatifs (GDLS 4) selon le cas, tandis que les deux autres manifestaient des comportements neutres. Au cours de chacun des GDLS, un évaluateur différent prenait en note les comportements manifestés par les personnes candidates et procédait à une évaluation par rapport aux huit compétences critiques dans l'emploi visé.

Chaque candidat a participé aux quatre GDSL et a donc été évalué à quatre reprises. Pour contrôler l'effet qu'aurait pu avoir le thème de la mise en situation sur l'évaluation du candidat, chacun des quatre thèmes a été réparti de manière à ce que, par exemple pour la mise en situation « sondage de satisfaction auprès du personnel », la moitié des candidat(e)s aient eu à faire face à un participant positif alors que l'autre moitié des candidat(e)s aient côtoyé un participant négatif.

Quant à la durée des GDSL, elle a été ajustée au prorata du nombre de participants : vingt (20) minutes pour les GDSL de deux personnes et quarante (40) minutes pour celles de quatre personnes.

Finalement, les évaluateurs ont été répartis par alternance pour éviter toute contamination (voir **Tableau 14** présenté ci-dessous).

TABLEAU 14
Répartition des évaluateurs en fonction
des candidats et des GDSL

	GDSL 1	GDSL 2	GDSL 3	GDSL 4
Candidat 1	Évaluateur 1	Évaluateur 2	Évaluateur 3	Évaluateur 4
Candidat 2	Évaluateur 4	Évaluateur 1	Évaluateur 2	Évaluateur 3
Candidat 3	Évaluateur 3	Évaluateur 4	Évaluateur 1	Évaluateur 2
Candidat 4	Évaluateur 2	Évaluateur 3	Évaluateur 4	Évaluateur 1
Candidat 5	Évaluateur 1	Évaluateur 2	Évaluateur 3	Évaluateur 4
Et ainsi de suite...	Évaluateur 4	Évaluateur 1	Évaluateur 2	Évaluateur 3

5.3.2 Scénario des acteurs

Au niveau des scénarios des acteurs, chaque thème est associé à des arguments spécifiques à formuler. En effet, la dynamique créée à l'intérieur de chacun des groupes de discussion est basée sur des normes très spécifiques comme par exemple, le nombre d'interventions à effectuer, la nature des arguments à présenter, etc. D'ailleurs, chaque intervention est reliée à la définition des différents rôles dans un groupe tiré de l'ouvrage de Wilson (1996).

De façon générale, les comportements neutres consistaient en quatre interventions de nature administrative : (1) demander des informations, (2) récapituler, (3) jouer le rôle de secrétaire ou (4) suggérer un mode de fonctionnement. Les comportements positifs comprenaient cinq types d'interventions de nature relationnelle : (1) inciter à la collaboration, (2) demander l'opinion des autres participants, (3) demander aux autres d'élaborer sur les idées qu'ils présentaient, (4) donner des critiques positives et (5) rechercher le consensus. Finalement, les comportements négatifs étaient basés sur cinq types d'intervention également de nature relationnelle : (1) rejeter les idées présentées par les autres, (2) confronter les participants par la critique, (3) imposer des façons de procéder, (4) rechercher à se valoriser et (5) afficher un humour déplacé. Tous les acteurs ont reçu une formation d'une journée, jusqu'à ce qu'ils maîtrisent leurs rôles.

De plus, l'ordre de présentation des groupes de discussion ainsi que les thèmes abordés ont été contrebalancés. En d'autres mots, le thème portant sur la qualité du service à la clientèle sera discuté un nombre de fois égal dans le groupe

comportant une personne apportant une contribution négative et dans le groupe comportant une personne apportant une contribution positive.

5.3.3 Évaluation des compétences

La série de huit compétences évaluées est demeurée la même dans chacun des groupes de discussion. Chaque compétence comportait une définition et des indicateurs comportementaux de même que des éléments associés à la présentation d'idées ou d'opinions de la part des candidats. Sur le plan de la démarche d'évaluation, au cours de chaque GDSL, un évaluateur prenait en note les comportements manifestés par le candidat. Une fois l'exercice terminé, il procédait à une évaluation individuelle et indépendante sur chacune des huit compétences en attribuant une note selon une échelle de un à cinq et ce, en fonction des indicateurs comportementaux observés. Puis à la fin de la journée, les évaluateurs se rencontraient afin de discuter de cette évaluation et d'en venir à un consensus sur le résultat réel de la personne candidate. Même si tous les évaluateurs possédaient déjà une grande expérience pour ce genre d'évaluation (au-delà de 200 évaluations par évaluateurs réparties sur une période de deux à quatre ans), ils ont tous reçu une formation d'une journée complète spécifiquement pour cette recherche.

5.3.4 Description de la population cible

Au total, 53 personnes ont accepté sur une base volontaire que leurs résultats soient utilisés dans le cadre de la présente recherche. Par conséquent, l'échantillon est constitué de 19 hommes et de 34 femmes.

5.3.5 Engagement déontologique et considérations éthiques

Cette recherche a été effectuée sous la supervision du Département de relations industrielles de l'Université du Québec en Outaouais. Par conséquent, chaque étape a été conduite en fonction du code de déontologie et d'éthique de l'UQO et du réseau des Universités du Québec.

Chaque participant a accepté sur une base volontaire que ses résultats soient compilés sous forme de données et utilisés dans le cadre d'une étude universitaire. En effet, un formulaire de consentement a été distribué à chacun d'eux au début de la séance d'évaluation afin qu'ils y apposent leur signature s'ils consentaient à ce que leurs résultats d'évaluation soient traités de façon anonyme. Les sujets ont donc été mis au courant des modalités de l'étude et des objectifs de la recherche et, par conséquent le chercheur a obtenu un consentement éclairé de leur part.

Les enjeux sur la confidentialité de la recherche et les sources de données auront été préservées tout au cours de la recherche puisque dans tous les cas, les organisations scolaires n'ont reçu que les résultats des candidats aux exercices faisant partie de l'évaluation normale (cinq des huit compétences critiques). Finalement, les candidats étaient au courant qu'ils étaient en relation avec des acteurs dans le cadre des GDSL et ce, pour des considérations éthiques.

PARTIE 4 Analyses de données

Chapitre 6 Analyses statistiques

6.1 Statistiques descriptives

À la page suivante, le **Tableau 15** présente les statistiques descriptives de l'évaluation des compétences des personnes candidates en fonction de la nature des interventions auxquelles elles ont été soumises lors de leur évaluation alors que le **Tableau 16** indique celles obtenues en fonction du nombre de participants aux GDSL. La première colonne de ces tableaux représente les compétences ayant fait l'objet d'une évaluation et le résultat global (moyenne des résultats aux compétences évaluées) et la deuxième colonne précise le nombre de sujets utilisés pour effectuer les calculs statistiques. Les colonnes suivantes indiquent respectivement la moyenne, l'écart-type, la variance, la kurtose et l'asymétrie qui sont présentés pour chacune des conditions, c'est-à-dire interventions positives vs. interventions négatives au **Tableau 15** et deux participants vs. quatre participants au **Tableau 16**. Il est à noter que pour chacune des variables, ces statistiques sont calculées à partir de la moyenne des résultats obtenus aux deux groupes qui présentaient la même condition expérimentale (par exemple, le résultat moyen des GDSL n^o 1 et n^o 3 pour le contexte positif).

Tableau 15
Analyse descriptive des résultats en fonction de la nature des interventions

Compétences évaluées	N	Interventions positives GDSL n° 1 et n°3					Interventions négatives GDSL n° 2 et n°4				
		Moyenne	Écart type	Variance	Kurtose	Asymétrie	Moyenne	Écart type	Variance	Kurtose	Asymétrie
Sens de l'environnement		2,9	0,57	0,33	0,01	-0,61	3,0	0,76	0,58	-0,59	0,12
Prise de décision		3,0	0,53	0,28	-0,53	-0,49	2,9	0,69	0,48	-0,59	0,41
Relations interpersonnelles		3,3	0,52	0,27	0,07	0,04	3,2	0,72	0,52	0,78	-0,32
Mobilisation d'équipe		2,8	0,56	0,32	-0,35	0,15	2,8	0,73	0,53	-0,30	0,32
Communication orale		3,1	0,47	0,22	1,87	-0,48	3,1	0,61	0,37	-0,18	-0,22
Capacité d'analyse		2,8	0,61	0,37	-0,53	0,21	2,7	0,74	0,54	-1,03	0,21
Orientation vers les résultats		3,1	0,62	0,39	-0,22	-0,27	3,1	0,75	0,56	0,16	-0,01
Flexibilité		3,0	0,51	0,26	0,23	-0,30	2,9	0,62	0,38	1,53	-0,13
Résultat global		3,0	0,57	0,32	-0,18	-0,71	3,0	0,72	0,51	-0,30	-0,23

D'une part, on constate que la nature des interventions n'a pas provoqué de distinction majeure sur le résultat pour chacune des compétences puisque la différence entre les moyennes de chacune d'entre elles est d'au plus 0,1 sur l'échelle en cinq points. D'ailleurs, le résultat global pour les deux conditions est semblable ce qui témoigne également du peu de différence obtenue dans les résultats selon que les personnes candidates étaient confrontées à des interventions positives ou négatives. Il en va de même pour trois compétences sur huit, c'est-à-dire *Mobilisation d'équipe*, *Communication orale* et *Orientation vers les résultats* qui ont des moyennes semblables selon les deux conditions. En ce qui concerne les cinq autres compétences, le sens de cette légère différence entre les moyennes n'est pas systématique puisque les résultats à quatre compétences se sont avérés plus élevés dans le cadre des interventions

positives, mais pour la compétence *Sens de l'environnement*, c'est dans un contexte négatif que la moyenne était la plus élevée.

Sous le regard de la courbe normale de distribution des résultats, on remarque à partir de l'indice de kurtose que la majorité des compétences sont représentées par une courbe pratiquement normale à l'exception de la compétence *Communication orale* dans un contexte positif qui en est loin ($k=1,87$). En regardant de plus près ces statistiques à partir de leur distribution en pourcentage, il apparaît que seulement environ 10% des candidats ont obtenu un score de 2,5 / 5 ou moins pour cette compétence. Il y a également la compétence *Flexibilité* en contexte négatif qui présente un kurtose élevé ($k=1,53$). Toujours à partir de la distribution de ces résultats en pourcentage, on constate qu'il y a seulement 15% des candidats qui ont obtenu un résultat de 3,5 / 5 et plus. Par ailleurs, sur le plan de l'indice d'asymétrie, la compétence qui retient le plus d'attention sous l'angle de la condition positive est *Sens de l'environnement* avec un indice de -0,61. De plus près, cette statistique révèle en fait qu'il y a plus de 25% des personnes candidates évaluées qui ont obtenu un score moyen de 3 / 5 pour cette compétence ce qui rend la courbe plutôt asymétrique. Finalement, sous l'angle de la condition négative, c'est la compétence *Prise de décision* qui présente une courbe pour le moins asymétrique avec 70% des personnes candidates évaluées qui ont obtenu un score de 3 / 5 et moins.

Tableau 16
Analyse descriptive des résultats en fonction du nombre de participants

Compétences évaluées	N	2 participants GDSL n° 1 et n°2					4 participants GDSL n° 3 et n°4				
		Moyenne	Écart type	Variance	Kurtose	Asymétrie	Moyenne	Écart type	Variance	Kurtose	Asymétrie
Sens de l'environnement		3,0	0,64	0,33	-0,31	-0,45	2,9	0,57	0,41	-0,15	-0,18
Prise de décision		3,0	0,53	0,28	-0,35	-0,44	2,8	0,60	0,36	-0,55	0,27
Relations interpersonnelles		3,3	0,61	0,37	-0,01	-0,45	3,2	0,56	0,31	0,31	0,13
Mobilisation d'équipe		2,9	0,54	0,29	-0,84	0,04	2,7	0,64	0,41	-0,43	0,08
Communication orale		3,2	0,50	0,26	0,28	-0,18	3,1	0,50	0,25	-0,37	0,56
Capacité d'analyse		2,7	0,62	0,38	-0,81	0,11	2,9	0,65	0,42	-0,62	0,28
Orientation vers les résultats		3,2	0,63	0,40	-0,38	-0,31	3,0	0,65	0,42	-0,30	-0,26
Flexibilité		2,9	0,54	0,29	0,05	-0,02	2,9	0,49	0,24	0,60	0,25
Résultat global		3,0	0,60	0,36	-0,41	-0,45	2,9	0,59	0,35	-0,23	-0,70

En ce qui concerne la variable *Nombre de participants*, on observe d'abord une différence de 0,2 points pour les moyennes de quatre des huit compétences, c'est-à-dire *Prise de décision*, *Mobilisation d'équipe*, *Capacité d'analyse* et *Orientation vers les résultats*. Toutefois, le sens de cette différence n'est pas systématique, puisque pour trois compétences le résultat est plus élevé lorsque la personne candidate était dans un groupe de discussion de deux participants, alors pour une compétence (*Capacité d'analyse*), c'est en fonction d'un groupe de discussion de quatre participants que les résultats sont plus élevés. Par ailleurs, on remarque que les résultats ont tendance à être légèrement plus faibles dans les groupes de quatre participants, ce qui n'est toutefois pas suffisant pour donner lieu à une grande différence au niveau du résultat global (0,1 point d'écart en faveur d'un groupe de discussion constitué de deux personnes). On constate donc que le nombre de participants n'a pas donné lieu à des

distinctions majeures dans le résultat moyen des personnes candidats aux groupes de discussion sans leader. D'ailleurs, pour cette condition, l'écart-type et la variance sont pratiquement semblables pour chacune des compétences et les résultats, en moyenne, présentent une variation se situant autour de 0,35 sur 5.

Finalement, sur le plan de la courbe normale de distribution des résultats, il convient également de noter qu'il y a seulement 7,5 % des candidats qui ont obtenu un score de 2,5 / 5 ou moins pour la compétence *Communication orale* dans une situation de groupe constitué de 2 participants ($k=0,28$). De plus, on constate que pour la compétence *Flexibilité* ($k=0,60$) dans un groupe de discussion comprenant 4 participants, 80 % des candidats ont obtenu un résultat moyen se situant entre 2,5 / 5 et 3,3 / 5 laissant les deux extrémités de la courbe pratiquement vides. Finalement, l'indice d'asymétrie de la compétence *Communication orale* dans un groupe de 4 participants démontre en fait que seulement 15% des personnes candidates évaluées ont obtenu un score supérieur ou égal à 3,5 / 5.

6.2 Comparaison des coefficients de corrélation

Au-delà des statistiques descriptives, même si les deux variables ne semblent pas amener de distinction sur les résultats moyens, ces derniers peuvent cependant cacher des différences individuelles. En effet, les résultats à chaque compétence compilés pour le groupe de participants ne garantit pas qu'il n'y a pas de différence significative sur les résultats propres à chaque individu. Par exemple, le fait que la présence d'un participant ayant des comportements positifs ou négatifs n'ait pas d'impact sur les résultats du groupe ne signifie pas qu'une personne candidate en particulier ne soit pas affectée par cette condition.

Par conséquent, les hypothèses émises au départ misaient sur la présence ou non d'une variation significative de la variable dépendante (résultats d'une personne candidate en contexte de sélection) en fonction de la manipulation des variables indépendantes (nature des interventions et nombre de participants). Pour arriver à vérifier ces hypothèses et déterminer l'influence des conditions expérimentales sur les résultats individuels, le **Tableau 17** et le **Tableau 18** présentent d'abord de façon résumée les corrélations obtenues. Il est à noter que ces corrélations sont calculées à partir du résultat individuel moyen des 53 participants aux conditions expérimentales prises deux à deux (par exemple, le résultat moyen des GDSL n° 1 et n° 3 est corrélé avec le résultat moyen des GDSL n° 2 et n°4).

Tableau 17
Coefficients de corrélation de Pearson en fonction de la nature des interventions

Compétences évaluées	N	Corrélation <i>r</i>
Sens de l'environnement		0,22
Prise de décision		0,24
Relations interpersonnelles		0,33*
Mobilisation d'équipe		0,32*
Communication orale		0,19
Capacité d'analyse		0,36**
Orientation vers les résultats		0,26
Flexibilité		0,16
Résultat global		0,24

** Corrélation significative au niveau 0.01

* Corrélation significative au niveau 0.05

D'une part, les corrélations apparaissant au **Tableau 17** fluctuent substantiellement pour chaque compétence et se situent entre 0,16 et 0,36, donc de faibles à modérées, et elles se sont avérées significatives pour moins de la moitié des compétences évaluées. En effet, seulement trois compétences sur huit ont obtenu une corrélation significative, c'est-à-dire les compétences *Relations interpersonnelles*, *Mobilisation d'équipe* et *Capacité d'analyse*. Dans ce contexte, il ressort qu'une personne démontrant de bonnes relations interpersonnelles a été en mesure de le démontrer tant en présence d'interventions positives que négatives et il en va de même pour un participant démontrant une capacité à mobiliser efficacement une équipe ainsi qu'à analyser adéquatement une situation. En effet, puisque les corrélations à ces trois compétences sont significatives, c'est dire qu'il y a une relation entre les résultats obtenus à ces compétences en présence d'interventions positives et négatives. En revanche, comme ces relations se manifestent de faible à modérément, la force de cette corrélation ne permet pas de prédire que la manifestation de la compétence pourrait se reproduire dans les mêmes proportions d'un groupe de discussion à l'autre. En ce qui concerne les cinq autres compétences évaluées, les corrélations apparaissent donc comme étant aléatoires.

En fonction de la deuxième condition expérimentale, l'impact du nombre de participants composant un GDSL, le **Tableau 18** présente les corrélations entre les résultats obtenus par les personnes candidates.

Tableau 18
Coefficients de corrélation de Pearson en fonction du nombre de participants

Compétences évaluées	N	Corrélation <i>r</i>
Sens de l'environnement		0,49**
Prise de décision		0,45**
Relations interpersonnelles		0,52**
Mobilisation d'équipe		0,57**
Communication orale		0,38**
Capacité d'analyse		0,54**
Orientation vers les résultats		0,45**
Flexibilité		0,43**
Résultat global		0,54**

** Corrélation significative au niveau 0.01

* Corrélation significative au niveau 0.05

Les corrélations apparaissant au **Tableau 18** fluctuent substantiellement pour chaque compétence et se situent entre 0,38 et 0,57, on peut donc parler d'un degré d'homogénéité modérée dans la relation entre les résultats à un groupe de deux participants vs. un groupe de quatre participants. Par ailleurs, les corrélations se sont avérées significatives pour l'ensemble des huit compétences de même que pour le résultat global. Il en ressort donc qu'une personne démontrant, par exemple, une bonne capacité à lire son environnement a été en mesure de le démontrer tant en faisant partie d'un groupe de discussion comprenant une autre personne qu'avec trois autres participants. En effet, puisque les corrélations à l'ensemble des compétences sont significatives, c'est dire qu'il y a une relation entre les résultats obtenus à ces compétences dans les deux situations présentées. La force de cette relation ne permet pas de prédire que la manifestation de la compétence pourrait se reproduire dans les

mêmes proportions d'un groupe de discussion à l'autre ou dans d'autres conditions (dans des groupes de six ou huit participants, par exemple). On peut néanmoins ajouter que ces résultats s'apparentent à ceux de Gatewood et al. (1990), c'est-à-dire des corrélations de 0,35 à 0,62 dans le cas où les sujets participaient à des groupes de discussion portant sur des thèmes différents.

6.3 Fidélité inter-examineurs

Afin d'évaluer si les fluctuations des corrélations et des moyennes à chacune des compétences sont l'effet des conditions expérimentales ou simplement un manque de fidélité, les coefficients de fidélité inter-examineurs sont présentés au **Tableau 19**. Dans ce cas-ci, la fidélité inter-examineurs a été obtenue en calculant la moyenne des coefficients de corrélations entre les quatre groupes de discussion (GDSL n⁰ 1 et n⁰2, GDSL n⁰ 1 et n⁰3, GDSL n⁰ 1 et n⁰4, GDSL n⁰ 2 et n⁰3, GDSL n⁰ 2 et n⁰ 4 et GDSL n⁰ 3 et n⁰4).

Tableau 19
Coefficients de fidélité inter-examineurs

Compétences évaluées	N	Fidélité inter-examineurs
Sens de l'environnement		0,28
Prise de décision		0,31
Relations interpersonnelles		0,35
Mobilisation d'équipe		0,33
Communication orale		0,26
Capacité d'analyse		0,29
Orientation vers les résultats		0,25
Flexibilité		0,23
Résultat global		0,29

Dans l'ensemble, on observe que la relation entre les scores attribués par chaque évaluateur s'étend de faible à modérée pour se situer entre 0,23 et 0,35. Sur le plan de la signification des corrélations, la *Mobilisation d'équipe* s'est avérée la plus significative avec une moyenne de cinq corrélations significatives sur six. Par la suite, on retrouve la compétence *Sens de l'environnement et les Relations interpersonnelles* avec quatre corrélations significatives sur six. Pour les cinq autres compétences, trois corrélations ou moins se sont avérées significatives. En somme, le degré d'homogénéité dans les scores attribués par les évaluateurs s'étend de faible à modéré, à part pour la *Mobilisation d'équipe* où le rapport entre les évaluations est plus fort. En ce qui concerne la force de cette relation, un coefficient de fidélité se situant entre 0,23 et 0,35 indique en fait une faible convergence entre les résultats attribués par les examinateurs ayant pris part au processus d'évaluation des candidats.

6.4 Analyse de variance par les rangs

D'un dernier regard, comme l'échantillon de données est plus important que 30 sujets et qu'il est apparu une certaine asymétrie dans la distribution normale des résultats, une analyse de variance par les rangs à mesures répétées a été conduite afin d'obtenir davantage d'informations sur l'effet des deux conditions expérimentales. Le **Tableau 20** démontre donc les résultats obtenus, d'une part, à la moyenne des rangs pour les quatre scénarios du schème expérimental, en fonction du Test de Friedman (quatre participants en contexte positif – 4P, quatre participants en contexte négatif – 4N, deux participants en contexte positif – 2P et, finalement, deux participants en contexte négatif– 2N). Dans ce contexte, le seuil de probabilité qui sera retenu comme significatif sera au niveau $p \leq 0,05$.

D'autre part, ce tableau présente le CHI^2 ainsi que le seuil de probabilité statistiques obtenus à chacun des compétences évaluées. Finalement, pour les compétences qui ont démontré un effet significatif, un Test de Wilcoxon a ensuite effectué une comparaison entre le scénario positif vs. négatif et le scénario deux vs. quatre participants afin de déterminer quel facteur avait eu une influence sur ces résultats.

Tableau 20
Analyse de variance par les rangs

Compétences évaluées	N	4P	4N	2P	2N	Chi ²	Probabilité	Contraste
Sens de l'environnement		2,4	2,5	2,6	2,5	1,312	0,726	
Prise de décision		2,4	2,4	2,8	2,4	5,705	0,127	
Relations interpersonnelles		2,4	2,2	2,8	2,6	7,521	0,057	
Mobilisation d'équipe		2,3	2,3	2,9	2,4	9,140	0,027	Voir Tableau 21
Communication orale		2,5	2,3	2,7	2,5	2,809	0,422	
Capacité d'analyse		2,6	2,6	2,5	2,3	3,893	0,273	
Orientation vers les résultats		2,1	2,4	2,8	2,8	10,78	0,013	Voir Tableau 22
Flexibilité		2,7	2,4	2,4	2,6	1,915	0,590	

Le **Tableau 21** et le **Tableau 22** démontrent, pour les deux compétences ont obtenu un seuil de probabilité statistiques significatif (*Mobilisation d'équipe* et *l'Orientation vers les résultats*), les résultats obtenus par le biais du Test de Wilcoxon afin de déterminer laquelle des deux conditions expérimentale a provoqué une variation au niveau des résultats des personnes candidates.

Tableau 21
Test de Wilcoxon pour la compétence Mobilisation d'équipe

Mobilisation d'équipe	4 part. Négatif 4 part. Positif	2 part. Positif 4 part. Positif	2 part. Négatif 2 part. Positif	2 part. Négatif 4 part. Négatif
Z	-0,421	-1,769	-1,712	-0,926
Probabilité	0,674	0,077	0,087	0,355

Tableau 22
Test de Wilcoxon pour la compétence Orientation vers les résultats

Orientation vers les résultats	4 part. Négatif 4 part. Positif	2 part. Positif 4 part. Positif	2 part. Négatif 2 part. Positif	2 part. Négatif 4 part. Négatif
Z	-0,925	-2,625	-1,035	-0,652
Probabilité	0,355	0,009	0,301	0,514

En regardant les résultats obtenus au Test de Wilcoxon, on remarque que pour la compétence *Mobilisation d'équipe*, il semble que la variation ne s'est pas avérée assez importante pour ressortir avec ce test statistique qui a une approche plus conservatrice. Par conséquent, il est possible de conclure que la variation qui a été observée pour cette compétence n'est pas l'effet d'une condition plus que l'autre étant donné que les probabilités présentées ci-haut ne sont pas statistiquement significatives au niveau $p \leq 0,05$.

En ce qui concerne la compétence *Orientation vers les résultats*, le Test de Wilcoxon a démontré un contraste significatif en fonction de l'impact du nombre de participants. En effet, en contexte positif, la variation a été significative selon que les candidats étaient placés en interaction directe avec trois autres participants ou seulement avec une autre personne. Il est difficile de déterminer une raison conceptuelle qui pourrait valider le fait qu'il serait systématiquement plus facile d'obtenir de meilleurs résultats sur cette compétence et ce, en contexte positif. Peut-être faut-il aussi considérer l'effet du hasard qui aurait pu jouer à travers toutes les comparaisons effectuées.

En somme, l'analyse de variance des moyennes par les rangs permet de conclure d'une part, que la nature des interventions n'a pas produit de variation significative sur les résultats des personnes candidates à un GDSL sur une même compétence. De plus, la compétence *Mobilisation d'équipe* est ressorti comme présentant une variation significative en fonction du Test de Friedman, mais en voulant déterminer quelle condition expérimentale avait sollicité cette variation, le Test de Wilcoxon n'a pas démontré de contraste significatif. Finalement, le nombre de participants a produit une variation significative pour une compétence sur huit (*Orientation vers les résultats*).

Chapitre 7 Discussion

L'objectif premier de cette recherche était de vérifier si les résultats des sujets aux groupes de discussion sans leader varient en fonction du nombre de participants au groupe et de la nature des interventions des participants et ce, dans un contexte de sélection de personnel. Cette préoccupation est née du fait qu'il est nécessaire pour une organisation d'utiliser des instruments de sélection efficaces permettant de porter un éclairage précis et constant sur le plan des compétences des personnes évaluées. Or, en simulant les fonctions de plusieurs postes d'aujourd'hui, le GDSL recrée un environnement dynamique qui représente une approche pertinente lorsque vient le moment d'évaluer certaines caractéristiques particulières, seulement si la nature imprévisible de cette dynamique n'introduit pas des erreurs de mesure en présentant des conditions changeantes d'une administration à l'autre de façon à amener un biais préjudiciable pour une personne candidate à un processus de sélection.

En effet, les écrits ont présenté quatre facteurs qui peuvent contribuer de façon importante à diminuer la standardisation des groupes de discussion sans leader et qui doivent être pris en considération au moment d'évaluer les personnes candidates en situation de groupe. Ces facteurs concernent la nature du thème (Gatewood et al., 1990), le format du GDSL (c'est-à-dire de collaboration ou de compétition) (Gatewood et al., 1990; Thorton, 1992), le nombre de participants au groupe (Pettersen, 2000) et les différences individuelles des participants quant à leurs interventions (Gatewood et al., 1990; Thorton, 1992; Pettersen, 2000). Aux fins de cette recherche, seuls les deux derniers facteurs ont été manipulés afin d'en vérifier l'impact sur les résultats d'une personne candidate prenant part à un processus de sélection.

En manipulant les deux variables *Nature des interventions* et *Nombre de participants*, la première préoccupation de ce mémoire était de vérifier s'il pouvait être dommageable pour un candidat d'être évalué à travers une discussion ou une autre. Parce que s'il est normal qu'un participant réagisse différemment en fonction de la situation dans laquelle il se trouve, il est impossible de dire si ses résultats sont affectés négativement en fonction de la nature des interventions auxquels il a été soumis lors de son évaluation ou en fonction du nombre de personnes avec lesquels il devait discuter. L'impact de la réponse à cette question est majeur puisqu'une mauvaise évaluation des compétences des participants à un processus de sélection peut non seulement conduire à décisions organisationnelles inéquitables (par exemple, inégalité des chances de parvenir à une promotion), mais également à des fâcheuses décisions d'affaires en rejetant des candidatures de personnes maîtrisant les compétences requises pour un poste, mais défavorisées par la dynamique ayant eu lieu au cours de l'évaluation, ou inversement en choisissant des gens ne possédant pas les compétences critiques pour l'emploi.

D'un point de vue théorique, il semble logique que la performance d'une personne candidate évaluée par le biais de différents GDSDL demeure stable d'un groupe à l'autre. En fait, quelle que soit la composition du groupe, c'est-à-dire les comportements démontrés par les participants ou le climat de la discussion, les résultats d'un candidat à une même compétence devraient demeurer invariables (H1). Il semble également logique que la performance d'une personne candidate évaluée par le biais de différents GDSDL demeure stable d'un groupe à l'autre quel que soit le nombre de personnes participant à la discussion, c'est-à-dire que ses résultats à une même compétence devraient demeurer invariables (H2).

7.1 Vérification des hypothèses

Sous le regard des deux hypothèses énoncées, **H1** sur l'effet de nature des interventions et **H2** sur l'effet du nombre de participants, dans l'ensemble, les résultats obtenus indiquent que les deux variables n'ont pas d'effet systématique sur la performance d'un candidat à un GDSL. En effet, les résultats moyens sont demeurés pratiquement similaires pour l'ensemble des compétences quelle que soit la condition expérimentale mises en cause dans les groupes de discussion. Par conséquent, ces variables n'ont pas introduit de biais d'emblée préjudiciable pour les candidats qui auraient pu faire en sorte qu'un candidat récolte automatiquement de meilleurs ou de pires résultats selon la composition de groupe.

Par contre, comme les résultats moyens ont été pratiquement similaires pour la série des huit compétences, il faut garder à l'esprit qu'il y a pu y avoir un effet de halo sur les évaluations des personnes candidates. Afin de déterminer quel a pu être l'apport de cet effet et répondre à ce questionnement, il aurait fallu effectuer un rapport de corrélation supplémentaire entre chaque compétence et chaque évaluateur de façon à déterminer si cette erreur d'évaluation a pu avoir un impact sur les résultats des personnes candidates. Comme il ne s'agissait pas spécifiquement de l'objet de cette recherche, cette question demeure en plan pour l'instant.

7.1.2 Hypothèses sur l'effet de la nature des interventions

H1₀ : Les résultats des personnes candidates à un GDSL sur une même compétence sont exempts de variation peu importe la nature des interventions auxquelles ces personnes sont soumises.

H1₁ : Les résultats des personnes candidates à un GDSL sur une même compétence varient en raison de la nature des interventions auxquelles ces personnes sont soumises.

En revanche, les résultats moyens ont caché certaines différences individuelles. En ce qui concerne la nature des interventions, avec seulement quatre corrélations significatives sur neuf et des coefficients se situant entre 0,16 et 0,37, la fidélité des résultats obtenus entre les groupes de discussion donnant lieu à des interventions positives et ceux constitués d'interventions négatives demeure faible et aléatoire. Une des explications plausibles est que les résultats ont varié dans une mesure beaucoup plus large dans le cadre des interventions négatives. En effet, les évaluations pourraient avoir eu tendance à être plus tranchée dans un contexte négatif puisque la variance a atteint presque 0,6 ce qui n'est pas négligeable sur une échelle en cinq points. Malgré tout, l'analyse de variance des moyennes par les rangs a démontré que la variable *Nature des interventions* ne provoquait pas de variation significative sur les résultats des personnes candidates à un GDSL sur une même compétence.

En somme, comme la comparaison des moyennes de rangs a permis d'observer que les différences n'étaient pas statistiquement significatives, l'hypothèse **H1₀** est acceptée puisque les résultats des personnes candidates à un GDSL sur une même compétence sont exempts de variation peu importe la nature des interventions auxquelles ces personnes sont soumises.

7.1.2 Hypothèses sur l'effet du nombre de participants

H₂₀ : Les résultats des personnes candidates à un GDSDL sur une même compétence sont exempts de variation peu importe le nombre de personnes prenant part à l'exercice.

H₂₁ : Les résultats des personnes candidates à un GDSDL sur une même compétence varient en raison du nombre de personnes prenant part à l'exercice.

En ce qui concerne le nombre de participants (**H₂**), la carence de standardisation dans ce sens ne serait pas d'emblée préjudiciable pour un candidat. Avec des corrélations significatives se situant entre 0,38 et 0,57, une relation existe entre les résultats d'un participant évalué dans un groupe composé d'un nombre de participants plus grand ou plus petit. De plus, la vérification de la différence des moyennes de rangs a permis d'observer des variations significatives pour seulement deux compétences sur huit, mais lorsque les analyses statistiques ont été poussées plus loin, il est clairement ressorti que le nombre de participants a provoqué une variation significative pour une seule compétence (*Orientation vers les résultats*).

En somme, l'hypothèse nulle **H₂₀** est rejetée et l'hypothèse alternative **H₂₁** est acceptée partiellement puisque les résultats des personnes candidates à un GDSDL, pour une compétence sur la série de huit ayant fait l'objet d'une évaluation, varient en raison du nombre de personnes ayant pris part à l'exercice.

7.1.3 Autres considérations

Il ne faut cependant pas conclure hâtivement qu'il n'est pas important de contrôler les variables indépendantes en cause dans ce mémoire lors de la planification des séances d'évaluation même si seulement une variation significative a été observé seulement pour une compétence. D'une part, ces variables n'ont été étudiées seulement pour quelques valeurs. Rien ne prouve que, par exemple, des groupes de plus grande taille (six ou huit participants) n'auraient pas eu d'effet significatif sur les résultats à toutes les compétences évaluées. D'autre part, les résultats obtenus ne contredisent pas le fait qu'augmenter le nombre de participants augmente également la possibilité d'y retrouver un échantillon davantage représentatif. En somme, les conclusions de cette recherche a pu démontrer principalement que le nombre de participants avaient un effet sur la compétence *Mobilisation d'équipe* et *Orientation vers les résultats* en situation de groupe impliquant 2 participants. Cependant, il convient de remarquer que cette différence significative sur les résultats a été en faveur du candidat. En effet, force est d'admettre que la dynamique s'installant avec une seule autre personne permet plus facilement de démontrer des comportements adéquats liés à la mobilisation d'une équipe et à l'orientation vers les résultats.

En ce qui concerne les compétences qui ont démontré une distribution de résultats qui ne concordent pas avec une normalité parfaite, il y a lieu d'envisager différentes explications. D'une part, en ce qui concerne la compétence *Communication orale*, il pourrait avoir une erreur d'échantillonnage dans le sens où comme tous les candidats étaient des professionnels provenant du monde scolaire, ils pouvaient tous posséder une facilité à s'exprimer de manière convaincante et articulée dû à leur grande expérience dans le domaine de la communication auprès de différents auditeurs. Il

serait plausible d'avancer également que les candidats n'ont pas obtenu des scores faibles à cette compétence en situation positive et lorsqu'ils discutaient seulement avec une autre personne parce que ces deux conditions favorisent incontestablement des échanges de qualité. Pour ce qui est de la compétence *Flexibilité*, soit qu'il s'agit également d'une erreur d'échantillonnage parce que les personnes candidates sont, par déformation, plus souvent en contrôle de la discussion, soit les conditions d'administration de l'outil a eu un effet sur les résultats à cette compétence en particulier, dans le sens où il s'est avéré difficile pour les candidats de s'adapter en situation négative et dans un groupe de quatre participants.

Finalement, comme les corrélations obtenues pour les deux variables sont assez faibles pour invoquer autre chose qu'un manque de fidélité par rapport à la situation, il était important de considérer d'autres sources d'erreurs potentiellement importantes (Cronbach, 1990 dans Pettersen, 2000). Comme l'examineur peut également être une source d'erreur importante qui pourrait avoir affecté la relation entre les résultats obtenus aux différents groupes de discussion, l'indice de fidélité inter-examineurs a également été calculé et cela a permis de mettre en lumière que la convergence entre les résultats des évaluateurs était plutôt faible. D'ailleurs, l'effet de halo a également été soulevé comme étant une source d'erreur potentielle qui pourrait avoir affecté les évaluations des candidats. En somme, cela ne permet pas de dire avec certitude que les conditions expérimentales sont les auteurs de cette relation faible entre les résultats obtenus aux quatre groupes de discussion, mais peut-être que le travail des évaluateurs y est également pour quelque chose.

Pour juger de la valeur des coefficients de fidélité obtenus par la méthode des formes équivalentes, il n'y a pas de seuil généralement approuvé par les experts

concernant le seuil minimum d'un coefficient de fidélité, mais il doit être le plus élevé possible (Pettersen, 2000). Toutefois, comme Pettersen (2000) le fait remarquer, cette affirmation ne tient pas compte de la réalité puisque la plupart des coefficients se situent entre 0,95 et 0,70, et parfois même à des niveaux aussi faibles que 0,50. Comme la pratique est l'art du possible, on doit également considérer des aspects plus contextuels (Guion 1998 et Nunnally et Bernstein, 1994 dans Pettersen, 2000). D'un point de vue pratique, les conclusions de ce mémoire concernant le nombre de participants sont donc très intéressantes d'autant plus que la recherche a été effectuée auprès de candidats participant à un véritable processus de sélection.

7.2 Contribution théorique et pratique

Dans ce contexte, la contribution théorique de cette recherche a été de démontrer qu'en contexte réel de sélection, la composition d'un GDSDL, c'est-à-dire la nature des interventions et le nombre de participants, ne provoque pas de variation significative et systématique sur les résultats des personnes candidates et ce, sur une même compétence. Par conséquent, de ce point de vue, ces deux variables n'introduisent pas de biais susceptibles de faire en sorte qu'une personne candidate récoltera de meilleurs ou de pires résultats selon les personnes avec lesquelles elle aura eu à discuter au cours de la séance d'évaluation.

En ce qui concerne l'aspect pratique de cette recherche, il y a bon nombre d'organisations qui effectuent présentement des processus de sélection en utilisant le GDSDL comme moyen d'évaluation. Par contre, jusqu'à présent, en situation de contestations légales sur la fidélité des résultats produits par cet outil, aucune d'entre elles n'avaient accès à des données permettant de confirmer ou de réfuter les

allégations d'un appelant qui aurait estimé avoir été lésé dans l'évaluation de ces compétences par le biais du GDSL.

7.3 Forces et limites de la recherche

D'un premier regard, il convient de noter que la méthodologie rigoureuse et l'expérimentation scientifique qui a été appliquée en contexte réel de sélection font de ce mémoire une recherche unique comparée à la plupart des autres recherches sur le sujet qui sont habituellement réalisés avec la participation d'étudiants. En effet, il est remarquable que quatre (4) organisations provenant du monde scolaire aient collaboré à toutes les étapes de l'expérimentation (élaboration et validation du profil de compétences, élaboration et validation des mises en situation, etc.) et qu'elles ont par la suite acceptée que leurs candidats soient évalués et sélectionnés par le biais du processus. D'ailleurs, l'implication de ces organisations a sans aucun doute permis de manipuler les conditions expérimentales par le biais d'acteurs professionnels possédant une très vaste expérience dans l'apprentissage de textes et du jeu en tant que tel, ce qui a amené un avantage incontournable pour vérifier l'impact de la nature des interventions. En résumé, il est possible de certifier qu'une méthodologie très stricte a été suivie tant pour le développement et l'implantation des outils que pour les règles d'administration des outils d'évaluation.

Bien que cette expérimentation ait été réalisée et standardisée dans les meilleures conditions possibles, quelques éléments auraient pu être bonifiés afin d'améliorer les conclusions de cette recherche. Premièrement, les évaluateurs ont varié tout au cours de la recherche ce qui fait qu'environ dix (10) évaluateurs différents ont été impliqués dans l'évaluation des sujets. Bien que l'ensemble des ressources impliquées

possédaient une formation académique et professionnelle en ressources humaines, un minimum de deux ans d'expérience dans l'évaluation de candidats en situation réelle de sélection, elles ont tous reçu une formation sur les mises en situation et les règles de cotation en cause dans le processus. Même si cette formation a toujours été donnée par la même personne, dans le cadre d'une prochaine expérimentation, l'idéal serait d'avoir la participation des quatre mêmes évaluateurs tout au cours du processus, de manière à ce qu'ils soient tous formés sur le processus au même moment et de la même façon.

Deuxièmement, il convient de noter que le nombre important de compétences à évaluer par chaque évaluateur (huit compétences par exercice) était très élevé si l'on compare à ce que les écrits avancent (ratio de 1,3 compétence par exercice dans le cadre d'un centre d'évaluation) et les indicateurs de comportements utilisés sont demeuré plutôt généraux au lieu d'être reliés spécifiquement à chacune des mises en situation comme il l'est suggéré pour améliorer la qualité des mesures (Lievens et Conway, 2001).

Troisièmement, il aurait été souhaitable d'effectuer un enregistrement vidéo de chacun des groupes de discussion de façon à donner la chance aux évaluateurs de re-visionner certains moments de la simulation et d'améliorer leur compréhension de la performance du candidat. Il semble que cette façon de faire ait contribué à augmenter la validité des résultats de candidats participants à un centre d'évaluation opérationnel (Ryan, Daum, Mauman, Grisez, Mattimore et Nalodka, 1995 dans Lievens, 2002).

En outre, pour des considérations éthiques les candidats connaissaient l'identité du candidat en face de lui, c'est-à-dire qu'il savait qu'il s'agissait d'un acteur. Par

conséquent, il se peut que certains sujets aient pu être influencé par cet élément, même si la qualité du jeu de ces acteurs professionnels ne laissait souvent aucun doute sur la véracité de leurs propos et la crédibilité de leur rôle.

Bien que cette situation ait été très complexe à gérer à cause du nombre considérable de ressources financières et humaines qui y étaient associées (administrateurs, évaluateurs, acteurs, intervenants des organisations, personnes candidates, etc.), l'étude a été poursuivie puisqu'elle avait la chance d'avoir pour sujet des candidats en situation réelle de sélection en comparaison avec la plupart des études sur les outils d'évaluation dont les sujets sont des étudiants n'ayant aucun enjeu dans le processus.

7.4 Pistes de recherche

Comme l'on possède peu de connaissances à propos du groupe de discussion sans leader, il serait intéressant de considérer d'autres avenues de recherche, en s'inspirant entre autre des écrits associés aux centres d'évaluation. À ce sujet, Lievens et Klimoski (2001) ont répertorié deux modèles pour aider à comprendre de quelle manière les évaluateurs et le processus de cotation influencent la qualité des mesures prises par le centre d'évaluation. Par conséquent, pour améliorer la validité des résultats des candidats à un centre d'évaluation, il s'agit de simplifier la façon de coter, d'une part, et, d'autre part, d'utiliser des évaluateurs d'expérience (Lievens et Klimoski, 2001).

En ce qui concerne le premier modèle, les études tendent à démontrer qu'il serait plus efficace de limiter le nombre de compétences mesurées à un ratio de 1,3

compétence par exercice (Gaugler et Thornton, 1989, Maher, 1990). De plus, il serait plus efficace de coter des dimensions aux concepts distincts (Kleinmann, Exler, Kuptsch et Köler, 1995). Par exemple, si la compétence *Vision stratégique* figure parmi celles évaluées dans le cadre de l'exercice, il serait souhaité d'éviter des compétences comme *Sens de l'environnement* ou *Perspicacité en affaire* qui font appel à des concepts similaires. Selon d'autres écrits, il serait également appréciable de spécialiser l'exercice de façon à ce que seulement une dimension de la personne soit évaluée (Robie, Adam, Osburn, Morris et Etchegaray, 2000). Par exemple, si l'on recherche de l'information sur la capacité d'une personne à communiquer efficacement, il serait donc souhaitable dans ce cas de ne pas donner comme tâche aux évaluateurs d'observer également des comportements reliés à la *Capacité d'analyse* dans le cadre du même exercice. De plus, les auteurs suggèrent de donner des listes de comportements associés aux compétences qui se rapporteront spécifiquement à la problématique de l'exercice (Donahue, 1997 et Reilly, 1990). Finalement, il faudrait donner la chance aux évaluateurs de visionner à nouveau les enregistrements vidéos des exercices afin qu'ils puissent approfondir leur compréhension de la performance du candidat (Ryan et al., 1995).

Le deuxième modèle de Lievens et Klimoski (2001) repose sur le principe de l'évaluateur expert. Sous cet angle, il semble que des différences majeures sur la qualité des évaluations ont été observées entre des évaluateurs novices et experts (Chi, Glaser et Farr, 1988). D'ailleurs, dans une étude de Lievens (2002) sur l'effet des facteurs reliés à l'évaluateur et ceux reliés aux candidats sur la validité de construit des centres d'évaluation, trois échantillons d'évaluateurs (psychologues industriels ou organisationnels, gestionnaires et étudiants) ont démontré qu'il n'y avait pas de preuve concernant la validité discriminante et convergente des évaluations faites par des

étudiants. Cette étude a également mis de l'avant que la validité de construit du centre d'évaluation était affectée dans une très large mesure par une faible fidélité inter-examineur.

Finalement, il y a également lieu de se questionner sur l'effet de l'exercice sur la performance du candidat afin de mieux comprendre la qualité des mesures prises à partir des exercices d'un centre d'évaluation. En effet, certains chercheurs ont soutenu que le centre d'évaluation faisait ressortir l'habileté des candidats à réagir dans certaines situations plutôt que d'évaluer des compétences (Sackett et Dreher, 1982, Craig, 1987, Howard, 1997, Sackett et Tuzinski, 2001). Par conséquent, le concept évalué par cet instrument serait davantage la capacité d'un individu à répondre à son environnement que des compétences en soi ce qui amènerait des performances très différentes d'un exercice à l'autre.

À ce sujet, l'étude de Lievens (2002) a également mis de l'avant qu'il y avait une très grande inconsistance dans les performances des candidats à travers les exercices d'un même centre d'évaluation ce qui affecte grandement la validité de construit de cet outil d'évaluation et de sélection de personnel. Dans ce contexte, il serait également intéressant de considérer d'autres pistes de recherche vers les conditions d'une situation qui vont amener les candidats à ajuster leurs comportements.

En somme, il serait donc tout à fait pertinent d'appliquer ces pistes de recherche reliées à l'avancement des travaux sur les centres d'évaluation au cas particulier du groupe de discussion sans leader et d'effectuer une nouvelle expérimentation en tenant compte de ces différents facteurs.

Conclusion

Le groupe de discussion sans leader est principalement utilisé par les professionnels en ressources humaines comme instrument de mesure en contexte de sélection. Il présente des avantages comparés à d'autres méthodes d'évaluation, mais la recension des écrits a démontré que ses conditions d'administration pouvaient potentiellement amener une diminution de la standardisation et en retour influencer les résultats d'une personne candidate à un processus de sélection.

Dans le cadre de ce mémoire, un schème expérimental rigoureux a été créé afin de vérifier si deux facteurs contribuant à diminuer la standardisation du GDSL pouvaient potentiellement créer un biais préjudiciable pour des candidats évalués par cet instrument de mesure. La méthode de la fidélité par formes parallèles a permis de recueillir des données sur l'impact de la nature des interventions et du nombre de participants sur les résultats d'un candidat en contexte réel de sélection. En effet, avec la participation de sujets provenant d'organisations du monde scolaire de la région (des candidats prenant part à un véritable processus de sélection), l'élaboration de quatre GDSL différents a permis d'évaluer une série de huit compétences.

De façon générale, les résultats aux analyses statistiques ont démontré que ni la nature des interventions ni le nombre de participants n'avaient un impact directement préjudiciable sur les résultats des personnes candidates aux compétences évaluées. En effet, les résultats moyens sont demeurés pratiquement les mêmes d'un groupe à l'autre et leur variation n'a jamais été systématique, ce qui est très encourageant d'un point de vue pratique puisque cela veut dire qu'un candidat ne récoltera pas

automatiquement de meilleurs ou de pires résultats selon les comportements des autres participants ou le nombre de personne ayant pris part à la discussion.

Cependant, que ce soit sous l'angle de la nature des interventions ou du nombre de participants, il n'a malheureusement pas été prouvé statistiquement qu'il y avait une forte relation entre les résultats obtenus par la manipulation des quatre conditions expérimentales. Néanmoins, la nature des interventions n'a pas provoqué de différence significative pour la série de huit compétences évaluées et le nombre de participants a créé une différence significative pour seulement une des huit compétences évaluées, c'est-à-dire l'*Orientation vers les résultats*. Finalement, il est important de mettre en perspective que la fidélité entre les évaluateurs s'est avérée plutôt faible.

L'intérêt premier de cette recherche était de présenter aux organisations en quête d'un outil efficace des considérations scientifiques sur la précision des résultats obtenus par des participants à un GDSL. Toutefois, il s'avère que d'autres études seront nécessaires particulièrement en ce qui concerne l'impact du nombre de participants. Dans le cadre d'une prochaine expérimentation, il serait intéressant d'appliquer l'ensemble des pistes de recherche énoncées ci-haut concernant la simplification de la manière de coter et les évaluateurs experts afin de faire avancer encore davantage la recherche sur la validité et la fidélité des GDSL.

Bibliographie

- Adler, R., Towne, N. (1991). *Communication et interactions*. Laval, Éditions Études Vivantes.
- Alderton, S.M., Frey, L.R. (1986). Argumentation in small group decision making. Dans Hirokawa, R.Y., Marshall, S.P. *Communication and Group Decision Making*. California, Sage, pp. 157-173.
- Amado G., Guittet, S. (1997). *Dynamique des communications dans les groupes*. Paris, Armand Colin, 213 p.
- American Educational Research Association, American Psychology Association, et National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, D.C., American Psychological Association.
- Appelbaum, S., Harel, V., Shapiro, B. (1998). The developmental assessment centre: the next generation. *Career Development International*. Bradford. Vol. 3, N° 1; pg. 5.
- Atkins, P. et Wood, R. (2002). Self-versus others' ratings as predictors of assessment center ratings : validation evidence for 360-degree feedback programs. *Personnel Psychology*, 55, 4, p. 871.
- Bales, R.F. (1950). *Interaction Process Analysis*. Cambridge, Mass., Addison-Wesley.
- Bales, R.F., Strodtbeck, F.L., Mills, T.M. et Roseborough, M.E., (1951). Channels of communication in small groups, *American sociological Review*, Vol. 16, pp. 461-468.
- Balocco, C. (1997). *Les méthodes d'évaluation en ressources humaines : la fin des marchands de certitude*. Paris, Éditions d'Organisation, 342 p.
- Barker, L.L., Wahlers, K.J., Watson, K.W. (2001). *Groups in process – An introduction to small group communication*. Boston, Allyn and Bacon.
- Bass, B.M. (1950). The leaderless group discussion, *Personnel Psychology*, 3, pp. 17-32.
- Bass, B.M. (1954). The leaderless group discussion, *Psychological Bulletin*, 51, pp. 462-492.
- Benne, K.D., Sheats, P. (1948). Functional roles of group members, *Journal of Social Issues*, 4(2), pp. 41-49.
- Bernatchez, J-C. (1982). *La sélection du personnel par simulation*. Montréal, les Éditions Agence d'ARC.
- Bion, W.R. (1946). The leaderless group project. *Bulletin of the Menninger Clinic*, 10, pp. 77-81.

Blaney, D., Hobson, C., Meade, M.E., Scodro, J. (1993) The assessment center: Evaluating managerial potential. *Nursing Management*. Chicago: Vol. 24, N° 2; pp.54-57.

Bormann, E.G. (1975). *Discussion and Group Methods : Theory and Practices*. New-York, Harper & Row Publishers, 395 p.

Boucher, D. et Doyon C. (1991). *La sélection du personnel, pour trouver l'excellence*. Québec, Les presses du management.

Bray, D.W. et Grant D.L. (1966). The assessment center in the measurement of potential for potential for business management, *Psychological Monographs*, 80, (17, no. 625), pp. 1-27.

Bray, D.W., Campbell, R.J. et Grant D.L. (1974). *Formative years in business: A long-term AT&T study of managerial lives*, New York, Wiley.

Byham, W.C. (1977). Assessor selection and training. Dans J.L. Moses et W.C. Byham (Eds.), *Applying the assessment center method*. New York, Pergamon Press.

Byham, W., (1986). *The Assessment Center Method and Methodology: New Applications and Technologies*. Monograph VII, Development Dimensions International.

Byham, R.N. et Byham, W.C. (1976), Effectiveness of Assessment Center Exercises in Producing Behavior, *Assessment & Development*, March 1976, pp. 9-10.

Campbell, J.P., Gasser, M.B. et Oswald, F.L. (1996). The substantive nature of job performance. *Individual Differences and Behavior in Organizations*. San Francisco, Jossey-Bass, pp. 258-299.

Caldwell, C., Thorton III, G.C., Gruys, M.L. (2003). Ten Classic Assessment Center Errors : Challenges to Selection Validity. *Public Personnel Management*. Spring, 32, 1, pg 73.

Carron, A.V. et Spink, K.S. (1995). The group size-cohesion relationship in minimal groups. *Small Group Research*, 26, 86-105.

Chi, M.T.H., Glaser, R. et Farr, M.J. (1988). *The nature of expertise*. Hillsdale, New Jersey : Erlbaum.

Ciasco, W.F. (1987). *Applied psychology in Personnel Management*. Englewood Cliffs, Prentice-Hall inc.

Ciascio, W.F. et Sweet, D.H. (1989). *Human resource planning, employment, and placement*. Washington, D.C., Bureau of National Affairs.

Centre de formation et de consultation (1987). *L'Appréciation du personnel par simulation*. Ste-Foy, Les presses de l'Université Laval.

Cosner, T.L. et Baumgart, W.C. (2000). An Effective Assessment Center Program : Essential Components. *FBI Law Enforcement Bulletin*, 69, 6, pg 1.

Cragan, J.F. et Wright, D.W. (1986). *Communication in small group discussions*. St-Paul, West Publishing Compagny.

Craig, J.R. (1987). Person characteristic versus role congruency explanations for assessment centre ratings, *Academy of Management Journal*, Vol. 30, No. 4, pp.817-826.

Craik, K.H., Ware, A.P., Kamp, J., O'Reilly III, C., Staw, B., Zedeck, S. (2002). Explorations of construct validity in a combined managerial and personality assessment programme. *Journal of Occupational and Organizational Psychology*, 75, pp. 171-193.

Cronbach, L.J. (1990.). *Essentials of psychological testing (5e éd. rev.)*. New-York, Harper Collins.

Dolan, S.L., Lamoureux, G., Gosselin, E. (1996). *Psychologie du travail et des organisations*. Montréal, Gaétan Morin Éditeur Itée, 500 p.

Dueutsch, M. (1949). An experimental study of the effects of cooperation and competition on group success. *Human Relations*, 2, pp. 199-231.

Durivage, A. (2002). *La gestion des compétences et le processus de sélection du Mouvement Desjardins*. Document de recherche 2002-2, Département de Relations industrielles, Université du Québec en Outaouais. Monographie publiée.

Field, R.H.G., House, R.J. (1995). *Human Behavior in organizations: a Canadian Perspective*. Scarborough, Prentice Hall Canada Inc., 706 p.

Flanagan, J.C. (1954). Some considerations in the development of situation tests. *Personnel Psychology*, 7, pp. 461-469.

Gatewood, R.D. et Field, H. S. (1994). *Human Resource Selection*. Forth Worth: The Dryden Press.

Gatewood, R.D., Thorton III, G. C., Hennessey, H. (1990). Reliability of exercise ratings in the leaderless group discussion. *Journal of Occupational Psychology*, Vol. 63, pp. 331-342.

Gaugler, B.B., Rosenthal, D.B., Thornton, G.C. III et Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, Vol. 72, pp. 493-511.

Gaugler, B.B. et Thornton, G.C. (1989). Number of assessment centre dimensions as a determinant of assessor accuracy. *Journal of Applied Psychology*, Vol. 74, pp. 611-618.

Gibb, J.R. (1961). Defensive communication. *Journal of Communication*, Vol. 11, pp. 142-148.

Gibson, C.B., Randel, A.E., Earley, C. (2000). Understanding group efficacy: An empirical test of multiple assessment methods. *Group and Organization Management*, 25 (1), pp. 67-97.

Guion, R.M. (1965). *Personnel testing*. New York, McGraw-Hill.

- Guion, R.M. (1998). Assessment measurement, and prediction for personnel decisions. Mahwah, NJ, Lawrence Erlbaum Associates, 690 p.
- Hare, P.A. (1962). Handbook of Small Group Research. New-York, The Free Press, 512 p.
- Hare, P. A. (1981). Group size. *American Behavioral Scientist*, 24, pp. 695-708.
- Harris, H. (1949). The group approach to leadership testing. London, Routledge and Paul.
- Herriot, P., Chalmers, C., Wingrove, J. (1985). Group decision making in an assessment center. *Journal of Occupational Psychology*, 58, pp. 309-323.
- Highhouse, S. (2002). Assessing the candidate as a whole: a historical and critical analysis of individual psychological assessment for personnel decision making. *Personnel Psychology*, 55, pp. 363-396.
- Hough, L.M. et Oswald, F.L. (2000). Personnel Selection: Looking Toward the Future – Remembering the Past. *Annual Review of Psychology*, 51, pp. 631-663.
- Howard, A. (1997). A reassessment of the assessment center challenges for the 21st century, *Journal of Social and Behavioral Personality*, vol. 12, pp.13-52.
- Hunter, J.E. et Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance, *Psychological Bulletin*, 96, pp. 72-98.
- International Task Force on Assessment Center (2000). The Guidelines and Ethical Considerations for Assessment Center Operations. *Public Personnel Management*, Vol. 29, No 3, pp. 315-331.
- Joiner, D. (2002). Assessment Centers : What's New ?. *Public Personnel Management*, 31, 2, pg. 179.
- Kleinmann, M., Exler, C., Kuptsch, C. et Köler, O. (1995). Independence and observability of dimensions as moderators of construct validity in the assessment center, *Organization Psychologie*, Vol. 39, pp. 22-28.
- Krech, D., Crutchfield, R. (1948). Theory and problems of social psychology. New-York, McGraw-Hill.
- Krech, D., Crutchfield, R. (1952). Théories et problèmes de la psychologie sociale, Paris, P.U.F.
- Lévy-Leboyer, C. (1996). Évaluation du personnel: quels objectifs ? Quelles méthodes ? 2e éd., Paris. Éditions d'Organisation, 247 p.
- Lievens, F. (2002) Trying to Understand the Different pieces of the Construct Validity of Assessment Centers : An examination of Assessor and Assessee Effects. *Journal of Applied Psychology*, Vol. 87, No. 4, pp. 675-686.

Lievens, F., Conway, J. M. (2001). Dimension and exercise variance in assessment center scores – A large-scale evaluation of multitrait-multimethod studies. *Journal of Applied Psychology*, 86 (6), pp. 1202-1222.

Lievens, F., De Fruyt, F. et Van Dam, K. (2001). Assessor's use of personality traits in descriptions of assessment centre candidates : A five-factor model perspective. *Journal of Occupational and Organisational Psychology*, 74, pp. 623-636.

Lievens, F. et Klimoski, R.J. (2001). Understanding the assessment centre process : Where are we now ? dans C.L. Copper et I.T. Robertson (Eds), *International Review of Industrial and Organizational Psychology* (Vol. 16, pp. 245-286). Chichester, England: Wiley.

Maher, P.T. (1990). How many dimensions are enough? Paper presented at the International Congress on the Assessment Centre Method, Orange, CA.

Marcoux, A. (1999). L'exercice du panier de gestion : que mesure-t-il exactement ? Mémoire de maîtrise, Département de relations industrielles, Université du Québec à Hull, 148 p.

Motowildo, S., Dunnette, M., Carter, G., (1990). An Alternative Selection Procedure : The low-Fidelity Simulation. *Journal of Applied Psychology*, 75 (6), pp. 640-647.

Moulinier, R. (1998). L'évaluation du personnel : pour la démocratie dans l'entreprise. Paris, Top Éditions, 175 p.

Nunnally, J.C. et Bernstein, I.H. (1994). *Psychometric Theory*. (3e éd. Rev.) New-York, McGraw-Hill.

O'Connell, M.S., Hattrup, K., Doverspike, D. et Cober, Al. (2002). The Validity of "Mini" Simulations for Mexican Retail Salespeople. *Journal of Business and Psychology*, 16 (4), pp. 593-599.

Parker, G.M. (1994). *Cross-functional teams : Working with allies, enemies or other strangers*. San Francisco, Jossey-Bass.

Paulhus, D.L. et Morgan, K.L. (1997). Perceptions of intelligence in leaderless groups: The Dynamic effects of shyness and Acquaintance. *Journal of personality and Social Psychology*, 72 (3), pp. 581-591.

Pavitt, C., Curtis, E. (1990). *Small group discussion : A theoretical approach*. Arizona, Gorsuch Scarisbrick Publishers.

Petterson, N. (2000). *Évaluation du potentiel humain dans les organisations : Élaboration et validation d'instruments de mesure*. Sainte-Foy, Presses de l'université du Québec, 374 p.

Renz, M.A., Greg, J.B. (2000). *Effective Small Group Communication in Theory and Practice*. Boston, Allyn and Bacon.

Richard, B. (1995). *Psychologie des groupes restreints*. Québec, Presses Inter-Universitaires, 138 p.

Sackett, P.R. et Dreher, G.F. (1982). Constructs and assessment center dimensions : Some troubling empirical findings, *Journal of Applied Psychology*, Vol. 67, pp. 401-410.

Sackett, P.R. et Tuzinski, K. (2001). The role of dimensions and exercices in assessment centre judgements, M. London Ed., New Jersey, Elbaum.

Schein, E.H. (1965). *Organizational Psychology*. New Jersey, Prentice-Hall, 67 p.

Schmitt, N. et Chan, D. (1998). *Personnel selection*. Thousand Oaks, CA, Sage.

Schmitt, N., Gooding, R.Z., Noe, R.A., Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, pp. 407-422.

Schmitt, N., Schneider, J.R., Cohen, S.A. (1990). Factors Affecting the Validity of a Regionally Administred Assessment Center. *Personnel Psychology*, Vol. 43, No. 1, pp. 1-13.

Schneider, J.R., Schmitt, N. (1992). An exercise design approach to understanding assessment centre dimension and exercise constructs. *Journal of Applied Psychology*, 77 (1), pp. 32-41.

Serruys, J.W. (1973). *Initiation à la dynamique de groupe*. Bruxelles, Les éditions vie ouvrière, 215 p.

Smith, G., Gregg, M., Andrews, D. (1990). *Savoir recruter, sélection et évaluation : une nouvelle approche*. Paris, Éditions Eyrolles, p. 51-88.

Smith, M., Robertson, I.T. (1993). *The theory and practice of systematic personnel selection*. Hampshire, Angleterre, Macmillan, 329 p.

Spencer, L.M. et Spencer, S.M. (1992). *Competence at Work, Models for Superior Performance*. New York, John Wiley and Sons.

Stasser, G., Taylor, L.A., Hanna, C. (1999). Information sampling in structured and unstructured discussion of three- and six-person groups. *Journal of Personality and Social Psychology*, 57 (1), pp. 67-78.

St-Onge, S., Audet, M., Haines, V. et Petit A. (1998). *Relever les défis de la gestion des ressources humaines*. Boucherville Gaëtan Morin Éditeur.

Tedeschi, J.T. et Lindskold, S. (1976). *Social Psychology*. New York, Wiley.

Thiétart, R.A. et coll. (1999). *Méthodes de recherche en management*, Paris, Dunod, 535 p.

Thornton, G.C. III (1992). *Assessment Centers in Human Resource Management*. Reading, Mass; Don Mills, Ont., Addison-Wesley. 276 p.

Thorton, G. et Byham, W. (1982). *Assessment Centers and Managerial Performance*. Florida, Academic Press inc., 458 p.

Tixier, J. (1974). Le groupe de discussion : une méthode active pour la sélection des cadres. Paris, Hommes et techniques, 109 p.

Wilson, G.L. (1996). Group in context: Leadership and participation in small groups. New-York, McGraw-Hill inc., 387 p.

Yan T.R. et Slivinski, L.R. (1976). A History of the Assessment Centre Method in the military. Ottawa, Commission de la fonction publique.

Site Web

Agriculture et agroalimentaire Canada
www.agr.gc.ca/hr-rh