

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN PSYCHOLOGIE

PAR
NATHALY JUTRAS

APPLICATION D'UNE PROCÉDURE D'ÉLABORATION D'INSTRUMENTS
DE SÉLECTION BASÉS SUR LA VALIDITÉ DE CONTENU:
CONSTRUCTION D'UN « IN-BASKET »

JUIN 1998

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

Application d'une procédure d'élaboration d'instruments de sélection

basée sur la validité de contenu :

construction d'un « in-basket »

Peu d'instruments valides élaborés sur la base de la validité de contenu sont disponibles sur le marché. Cette recherche vise donc à élaborer un « in-basket » en appliquant une procédure d'élaboration standard. Après avoir analysé diverses procédures d'élaboration, l'approche de Pettersen (1997) a été choisie pour être appliquée étape par étape, avec toute la rigueur propre à la psychométrie. Ceci permettra de faire des recommandations sur le procédé d'élaboration en fonction du degré de validité de contenu du test ainsi élaboré. Le test « in-basket » est adapté aux fonctions d'un gestionnaire de projet, type de gestion appelée à être la solution à la complexité organisationnelle moderne. Il est ensuite validé auprès d'un échantillon de 43 gestionnaires de projets par la méthode de Lawshe (1975). Les résultats indiquent un faible coefficient de validité de contenu. Des suggestions en vue d'améliorer le procédé d'élaboration découlent de ces résultats et la procédure de validation est réévaluée.

Remerciements

Merci à Normand Pettersen, mon directeur de recherche, pour son expertise, sa patience ainsi que pour avoir cru en moi.

TABLE DES MATIÈRES

CHAPITRE I : INTRODUCTION	1
CHAPITRE II : CONTEXTE THÉORIQUE	8
1. Validité de contenu	9
1.1. Définition	9
1.2. La validité apparente : définition	11
2. Historique du développement des simulations et plus spécifiquement du « in-basket »	12
2.1. Les simulations	13
2.2. Le « in-basket »	14
3. La procédure d'élaboration d'un instrument de sélection basé sur la validité de contenu	18
3.1. Les composantes de l'instrument de mesure	18
3.1.1. L'instrument lui-même	18
3.1.2. Conditions d'application	19
3.1.3. Le processus d'observation et d'évaluation	20
3.2. Différentes procédures d'élaboration	20
3.2.1. Society for Industrial and Organizational Psychology (SIOP)	21
3.2.2. Gatewood et Feild	25
3.2.3. Mussio et Smith	28
3.2.4. Pettersen	31
4. Simulations et « in-basket »	33
4.1. Description	34
4.2. Forces et faiblesses	36
CHAPITRE III : ÉLABORATION DE L'INSTRUMENT (PROCESSUS GÉNÉRAL D'ÉLABORATION)	40
1. Étape 1 : Finalités de l'instrument de mesure	41
2. Étape 2 : Analyse et description de l'emploi	43
3. Étape 3 : Spécification du domaine ou du sous domaine de contenu de l'emploi à mesurer	52
4. Étape 4 : Conception de l'instrument dans son ensemble	80
5. Étape 5 : Création des items et échantillonnage	81

5.1.	Contexte de la simulation et documents de support	82
5.2.	Rédaction des items	84
6.	Étape 6 : Révision de la version expérimentale par des experts	89
 CHAPITRE IV : RÉSULTATS CONCERNANT LA VALIDITÉ DE CONTENU DE L'INSTRUMENT		 91
1.	Les composantes de l'instrument de mesure	92
1.1.	Les directives	92
1.2.	Le mode de présentation des stimuli	93
1.3.	Mode utilisé pour les réponses	94
1.4.	Les conditions d'administration	94
1.5.	Les items	95
 CHAPITRE V : DISCUSSION		 100
 CHAPITRE VI : CONCLUSION		 121
 RÉFÉRENCES		 125
 APPENDICES		
Appendice A :	Définition des dimensions (Yukl, 1987)	132
Appendice B :	Les catégories de comportements du PMBOK	143
Appendice C :	Les consignes aux candidats pour l'instrument de mesure « in-basket »	147
Appendice D :	Le contexte de la simulation	152
Appendice E :	Les documents de support	157
Appendice F :	Le protocole de cueillette des incidents critiques	166
Appendice G :	Les items du « in-basket »	180
Appendice H :	Les consignes aux experts	206
Appendice I :	Les consignes aux candidats (test de Lawshe)	208
Appendice J :	Les résultats des coefficients de validité de contenu (sujets regroupés par domaine d'expertise)	217

N.B. Les appendices C, D, E et G ont été retirées de la présente version afin d'assurer la confidentialité de l'instrument de mesure.

Liste des tableaux

Tableau

1	Tâches, rôles et responsabilités du gestionnaire de projet	51
2	Aperçu des différentes taxonomies de la gestion comparées par Yukl (1987)	57
3	Comparaison entre la taxonomie proposée par le PMBOK (1996) et celle de Yukl (1987)	77
4	Matrice (dimensions x domaines d'application)	79
5	Compilation entre les domaines définis par le PMBOK et la comparaison des incidents critiques	86
6	Distribution des items en fonction des dimensions	91
7	Caractéristiques des sujets	99
8	Résultats de la compilation des réponses des 43 sujets (coefficients de validité de contenu)	100

CHAPITRE I : INTRODUCTION

La société d'hier valorisait la terre, le travail physique, les matières premières et le capital. L'industrie était axée sur le travail à la chaîne et la production de masse, amenant par le fait même une tendance à la déresponsabilisation. La révolution industrielle et l'école classique de gestion (taylorisme et bureaucratisme) ont apporté avec elles l'uniformisation des tâches et, ainsi, l'individu est devenu un maillon interchangeable de la chaîne d'opérations. Les travailleurs ne possédaient alors qu'une toute petite partie des moyens de production.

Aujourd'hui, ces paramètres ont subi une profonde mutation. Le savoir est devenu le plus puissant instrument d'accroissement de l'industrie. L'individu, par ses connaissances, dispose maintenant d'une part des moyens de production. Plus son savoir s'avère spécialisé, critique, décisif, moins le travailleur est interchangeable.

La tendance amorcée entraîne une élimination des cadres intermédiaires, des entreprises en constante réorganisation, une augmentation des besoins d'expertise spécialisée, la nécessité d'une vision globale, la destruction de la

structure pyramidale. Les valeurs essentielles sont l'autonomie, l'initiative, l'imputabilité. Le héros de notre temps est l'innovateur : il associe le savoir imaginaire à l'action. Les unités organisationnelles se multiplient, le volume de l'information à créer et à communiquer s'accroît, on favorise la formation continue, la responsabilisation de chacun (Peters, 1993 ; Serieyx, 1993 ; Toffler, 1991).

Dans la mesure où les individus deviennent indispensables, le recrutement et la sélection revêtent un caractère névralgique pour l'entreprise, surtout en ce qui a trait au personnel cadre. En effet, une mauvaise sélection de personnel de direction peut limiter l'efficacité de l'organisation dans son entier. Il n'est donc pas exagéré de dire qu'une des clefs du succès d'une entreprise réside dans la qualité de son processus de sélection (Werther et al., 1990).

Pour l'entreprise, la sélection efficace signifie un meilleur rendement. Pour l'employé, cela sous-entend un plus grand confort, une meilleure estime de soi reliée à la satisfaction du travail bien fait et, par conséquent, moins de stress. Ainsi, pour l'employeur comme pour l'employé, une sélection adéquate est primordiale.

Ayant pour but principal le mieux-être de l'individu et, par extension, de la collectivité, le psychologue est naturellement amené à s'intéresser à la sélection. De plus, il dispose de moyens, de techniques et de points de vue permettant une

analyse des éléments qui constituent les diverses situations de sélection (Duhamel, 1992). (1) Sa formation en psychométrie, (2) sa propension à l'analyse et au diagnostic ainsi que (3) sa rigueur en recherche font du psychologue un agent de première importance au sein des processus de sélection.

Différents instruments de sélection sont disponibles sur le marché. Plusieurs tests psychométriques ont été développés, validés et ont fait la preuve de leur utilité. L'expertise en psychométrie est à la fine pointe : tests objectifs d'intelligence et d'aptitudes, de personnalité, d'intérêts, ainsi que plusieurs tests projectifs.

Par contre, comme le disait Dunnette (1994), la plupart de ces tests ont été conçus pour mesurer des variables psychologiques appartenant à des typologies d'aptitudes, de traits de personnalité ou autres, qui ne proviennent pas du monde du travail et de l'univers des emplois. En conséquence, il est difficile de relier directement, sans inférence logique ou par déduction, la performance à ces tests psychométriques (ex., perception spatiale, stabilité émotionnelle, extraversion, etc.) aux comportements souhaités dans un travail donné (identification de problèmes en gestion, résolution de conflits interpersonnels, accueil de la clientèle, etc.).

Des instruments de mesures élaborés spécifiquement en fonction du monde du travail et de ses propres typologies en termes de connaissances ou de comportements existent. Ce sont des instruments construits suivant une démarche

de validité de contenu et ayant pour objectif d'être représentatifs des tâches dans un emploi à exécuter. D'ailleurs, tel que le mentionne Cronbach (1970), la validation de contenu est une approche particulièrement importante pour tout outil mesurant des compétences acquises par formation ou par expérience. Selon Thornton (1992), les tests élaborés sur la base de la validité de contenu font partie des instruments les plus efficaces pour la sélection du personnel cadre.

Depuis une vingtaine d'années, la popularité d'instruments représentatifs des tâches à mesurer amène une recrudescence de l'intérêt envers la validation de contenu (Pettersen, 1995). L'augmentation de la judiciarisation en matière de gestion des ressources humaines, particulièrement en ce qui a trait à la sélection et à la promotion du personnel, peut expliquer en partie cet intérêt (Pettersen, 1995). Les entreprises sont de plus en plus conscientes de la nécessité de démontrer hors de tout doute raisonnable que les instruments sur lesquels elles fondent leurs décisions sont tout à fait en accord avec les chartes, le code du travail ou toute autre politique gouvernementale, et ne présentent aucune discrimination.

Malheureusement, la plupart de ces instruments de mesure maison sont élaborés par des gestionnaires de personnel, dans un délai restreint et avec peu de connaissances en psychométrie. Chacun y va de son propre processus d'élaboration, intuitivement, se fiant à son jugement et à sa connaissance du

domaine. Pourtant, la validité de contenu d'un instrument est assurée par la rigueur de son élaboration et ce, à chacune des étapes. Si ce procédé d'élaboration n'est pas standardisé, n'a pas été appliqué rigoureusement afin d'en assurer la valeur, il est difficile de prétendre à la validité de l'instrument. En effet, force est de constater que les quelques spécimens disponibles sur le marché ne présentent pas nécessairement la rigueur scientifique proposée par la psychologie.

Pourtant, la valeur d'une conclusion ou d'une décision ne peut être meilleure que les données sur lesquelles elle est appuyée (DeVellis, 1991). Ainsi, il y a donc de nombreuses circonstances où le gestionnaire de personnel serait plus avisé d'adopter une démarche rigoureuse et fondée sur des données objectives (Pettersen, 1995). L'obtention de ces données objectives suppose de s'appuyer sur un patron, une ligne de conduite, un processus d'élaboration comportant des étapes bien définies et adapté aux instruments basés sur la validité de contenu. Cela pourrait permettre d'assurer une plus grande rigueur à l'élaboration de ces dits instruments. Ce processus devrait être général et pourrait ainsi s'adapter à différents tests (examens de connaissances, tests de type échantillon de travail et mises en situation, etc.).

L'objectif de la présente recherche est d'appliquer et d'adapter une procédure d'élaboration d'instrument de mesure basée sur le contenu et développée spécifiquement pour faire la sélection du personnel. Cette procédure,

ainsi appliquée, pourrait éventuellement devenir une procédure type à l'usage des psychologues ou d'autres spécialistes oeuvrant en sélection de personnel.

Pour ce faire, un instrument de sélection sera élaboré en suivant cette dite procédure. L'application du processus d'élaboration et les résultats obtenus pourront permettre d'entériner le processus ou de proposer des suggestions afin de l'améliorer.

Les stratégies de validation par le contenu en fonction de l'emploi regroupent plusieurs types de tests, parmi lesquels on retrouve les simulations. N'importe lequel de ces test aurait pu être envisagé afin d'appliquer le processus d'élaboration. Par contre, à l'intérieur des simulations, le test « in-basket » est parmi les instruments les plus valides pour la sélection (Robertson et Iles, 1988). Il apparaît donc très intéressant d'élaborer un « in-basket » à l'intérieur de cette recherche.

L'univers de contenu qui sera mesuré par le « in-basket » est la gestion de projet. Comme c'est une forme d'organisation à la fois souple et changeante, la gestion de projet représente une réponse prometteuse au grand défi de la complexité organisationnelle moderne (Bennis, 1989). La gestion de projet s'impose comme une véritable forme d'organisation et de gestion dans toutes les sphères d'activités. Dans un contexte de mondialisation où la performance est

essentielle à la survie des organisations, la gestion de projet est en train de devenir la règle plus que l'exception (Pettersen, Fabi et Garand, 1997). Le test « in-basket » ici développé servira non seulement à appliquer et éventuellement à fournir des informations pertinentes sur le procédé d'élaboration, mais pourra aussi être utilisé par les nombreuses organisations pratiquant la gestion de projet.

CHAPITRE II : CONTEXTE THÉORIQUE

La première partie de ce chapitre présente la notion de validité de contenu, cette notion étant très souvent à la base de la construction de tests de sélection. Ensuite, un des avantages découlant de la validité de contenu, la validité apparente, sera précisé. La deuxième partie comprend un résumé historique du développement des instruments basés sur la validité de contenu et, plus spécifiquement, du « in-basket ». La troisième partie comporte deux volets : la description des composantes de l'instrument de mesure, ainsi que l'étude des différents modes d'élaboration de tests basés sur la stratégie de validation. Un procédé d'élaboration sera retenu pour être appliqué dans le cadre de cette recherche. En quatrième partie, le terme « simulation », contexte plus large à l'intérieur duquel s'inscrit le « in-basket », sera expliqué, puis l'accent sera mis sur le « in-basket » lui-même. Dans la même section, afin de bien cadrer le « in-basket » dans le monde de la sélection, il importe d'en faire ressortir les forces et les faiblesses en le comparant aux autres méthodes de sélection. Plusieurs recherches faisant l'analyse critique du « in-basket » seront donc présentées.

1. Validité de Contenu

1.1 Définition

La validité de contenu est un concept basé sur la représentativité du test par rapport à ce qu'il est censé mesurer (Anastasi, 1982). Dans le contexte de sélection de personnel, elle réfère au degré selon lequel les résultats du test peuvent être représentatifs du rendement dans un domaine explicitement défini (Commission de la Fonction Publique, 1982). La validité de contenu représente le niveau d'adéquation entre les caractéristiques ou les compétences mesurées par l'instrument et celles nécessaires à la performance dans l'emploi réel. La question fondamentale concernant la validité de contenu est donc de savoir si les éléments qui composent le test sont représentatifs de l'emploi. Ainsi, un test d'épellation et de multiplication est valide, par définition, s'il comporte des problèmes d'épellation et de multiplication. Par contre, la solution n'est pas aussi simple qu'elle le paraît. Une des difficultés est d'obtenir un échantillon adéquat de l'univers à mesurer. L'univers doit être systématiquement analysé afin de s'assurer que les aspects majeurs de cet univers de contenu soient représentés à l'intérieur du test, et ce, dans des proportions justes par rapport à l'emploi (Anastasi, 1982 ; *Equal Employment Opportunity Commission*, 1978). L'élaboration d'un test basé sur la validité de contenu implique donc de définir un échantillon représentatif des connaissances, traits ou habiletés que la performance à l'emploi exige. Cela nécessite aussi la consultation d'experts du domaine de contenu. L'opinion de ces

experts devra avoir préséance sur toute approche quantitative (Gatewood et Feild, 1994). En effet, la validité de contenu d'un instrument de sélection repose d'abord et avant tout sur le jugement des experts du domaine (Société Industrielle et Organisationnelle de Psychologie, 1987).

Un instrument de mesure est un moyen de recueillir un échantillon de réponses ou de comportements appartenant à un domaine préalablement défini (Guion, 1965). De plus, tel que le mentionnent Tiffin et Mc Cormick (1967), la validité d'un instrument est sa capacité à atteindre les objectifs qui lui sont assignés. Conséquemment, la validité est un attribut des scores obtenus et non de l'instrument lui-même (Cronbach, 1971 ; Guion, 1977 ; Messick, 1974). Par contre, même si la validité est un attribut des scores, la démonstration de cette dite validité, suivant une approche de validation de contenu, repose sur l'analyse de l'instrument de mesure et de ses composantes. La représentativité des composantes de l'instrument par rapport au domaine à mesurer sera donc les prémisses de la validation des inférences, que l'on fera à partir des réponses des sujets. Ainsi, on doit s'assurer de la validité de ces composantes pour pouvoir prétendre à la validité des résultats obtenus à partir de l'instrument. « Il faut donc une méthode rationnelle, systématique et faisant preuve de discernement qui indique de quelle manière les items du test concordent avec les qualités requises et comment ces qualités sont également liées aux tâches à exécuter » (Commission de la Fonction Publique, 1982). La validité de contenu d'un test est

donc assurée par la rigueur de son élaboration, suivant une logique et un rationnel éprouvés et ce, à chacune des étapes.

1.2 Validité apparente

La validité apparente découle de la validité de contenu mais, contrairement à celle-ci, la validité apparente n'est pas garante de la validité du test proprement dite. En fait, si le test a une grande validité de contenu par rapport aux tâches à mesurer, il a habituellement une grande validité apparente. Par contre, la réciproque n'est pas nécessairement vraie. Ainsi, la validité apparente est un concept important mais limité. Elle n'indique pas ce que le test mesure en réalité mais bien ce que le test « paraît » mesurer (Anastasi, 1997). En fait, ce que l'item paraît mesurer peut différer considérablement de ce qu'il mesure en réalité, puisque cela peut dépendre des conditions dans lequel il est mesuré et corrigé. Des changements dans le temps, dans les consignes ou dans le processus de correction peuvent faire varier considérablement la mesure (Anastasi, 1982).

Parmi les simulations, le « in-basket » présente une des meilleures validités apparentes (Robertson et Iles 1988). « La validité apparente est généralement considérée comme une qualité souhaitable de l'instrument plutôt que comme un indice technique de validité » (Anastasi, 1988, Cronbach, 1970 ; Tziner, Jeanrie, Cusson, 1993). Elle est donc en quelque sorte un outil « marketing », si l'on peut

dire, puisque si elle ne garantit pas la validité du test comme telle, elle est un atout important pour convaincre les principaux acteurs de l'intérêt de l'instrument de mesure pour la sélection. La validité apparente est un attrait pour les employeurs. De plus, elle suscite l'intérêt du candidat, le sensibilise aux exigences de la tâche et a souvent pour effet de rendre ses attentes plus réalistes. Elle facilite également l'image d'une évaluation objective permettant ainsi au candidat d'accepter plus facilement une évaluation négative (Pettersen, 1993). La validité apparente est aussi un atout intéressant dans des causes de grief et de contestation puisqu'elle facilite la défense de l'instrument (Kleiman et Faley, 1985).

Par contre, il est important de retenir que la validité apparente n'est pas suffisante à elle seule pour justifier l'utilisation d'un instrument de mesure. Elle est utile, mais ne remplace pas la véritable validité de contenu de l'instrument. Le travail du chercheur doit donc se centrer sur la validité de contenu de l'instrument, la validité apparente en étant une conséquence probable.

2. Historique du développement des instruments de sélection

L'historique du développement des simulations et du « in-basket » permet de mieux en comprendre le niveau de connaissance actuel, ainsi que les lacunes qui demeurent à combler. Il apparaît donc intéressant d'en faire la présentation afin d'éclairer la démarche subséquente.

2.1 Les simulations

Au début du siècle, tenant lieu de simulations, de simples tests de performance, où l'individu faisait étalage de ses connaissances, étaient utilisés pour mesurer les différences individuelles des candidats (Thornton et Byham, 1982). En 1930, des simulations plus près de celles que l'on connaît aujourd'hui ont fait leur apparition pour la sélection d'officiers allemands. Lors de la Deuxième Guerre mondiale, les développements en matière de sélection se sont multipliés. En 1938, l'*Office of Strategic Service* développe les premiers programmes de sélection par simulation aux États-Unis. Entre 1943 et 1945, plusieurs psychologues et psychiatres participent à des recherches dans le but de développer des simulations valides. Un peu après la Deuxième Guerre, le *Veterans Administration (V.A.)* finance un vaste programme de recherche en sélection. Cette recherche dirigée par Kelly et Fiske (1951) met en évidence la nécessité de perfectionner les simulations (cité dans Thornton et Byham, 1982). La simple simulation d'un échantillon du travail n'est pas suffisante, les résultats en termes de validité sont décevants. Les simulations deviennent un sujet de recherche de plus en plus investigué. Les groupes de recherche se multiplient : *Institute for Personality Assessment and Research*, Stern, Stein et Bloom (1956), Menninger (1958) sont autant de chercheurs qui font rapidement évoluer les simulations (cités dans Thornton et Byham, 1982).

2.2 Le « in-basket »

En 1952, Fredericksen de l'*Educational Testing Service (ETS)* développe le premier « in-basket » pour la formation de pilotes (*Air Force Personnel and Training Research Center*). Une étude minutieuse auprès des instructeurs a été effectuée dans le but d'identifier les besoins des aspirants officiers en termes de formation. Douze catégories de comportements ont été mises en évidence dont six concernaient des comportements typiquement individuels, plus propices à ce type d'évaluation. Quatre de ces catégories ont été choisies pour s'inscrire à l'intérieur du test en développement parce qu'elles incluaient des comportements routiniers correspondant aux procédures standards d'opération.

Il était important de développer un test qui ressemblerait le plus possible au travail réel d'un officier de l'*Air Force*. Des entrevues avec des officiers ont révélé que la majorité du travail solitaire était centré sur le contenu des paniers « In and Out ». Ces paniers se trouvent sur le bureau de tout officier ou gestionnaire et contiennent des notes de service, rapports, et autres documents qui sont déposés quotidiennement. Il fut donc convenu d'élaborer un test situationnel qui serait le plus réaliste possible, en utilisant comme items les éléments de ces paniers, d'où l'appellation « in-basket ».

La première critique des officiers de l'*Air Force* portait sur le fait qu'il n'était pas raisonnable de prendre des décisions à partir des items à cause de l'information trop restreinte qu'ils contenaient, alors qu'une bonne prise de décision dépendait de plusieurs facteurs inconnus de la situation. Il paraissait donc nécessaire de fournir un entraînement adéquat au candidat, comprenant la description de son poste, de l'organisation qui l'employait, de l'histoire de l'entreprise, de sa mission, de l'équipement dont il bénéficiait, de l'organigramme de son unité, etc. Tout cet arsenal demandait beaucoup plus de temps qu'il était prévu au départ. Le compromis final fut que le test devait inclure des directives mettant le candidat en situation de remplacer son prédécesseur d'urgence, sans bénéficier de toute la formation habituelle, tout en lui fournissant le plus d'informations possible. D'où la création des documents de support que le candidat pourrait examiner avant la passation.

Le « in-basket » ainsi préparé plaçait le candidat dans quatre rôles successifs à l'intérieur de la même organisation. Chacun de ces quatre tests a été administré au cours d'une période de deux heures, quatre jours durant. À l'intérieur de chacun des rôles, le sujet bénéficiait du contenu d'un panier de gestion (lettres, mémos, lettres préparées par un subordonné et autre matériel similaire). Des documents d'aide ont été attachés à plusieurs items démontrant les actions déjà entreprises.

Le test lui-même a satisfait les exigences de l'*Air Force*. Cependant, des faiblesses ont été observées dans le processus de correction. Des recherches subséquentes se sont donc penchées sur la possibilité d'intégrer des dimensions ou des catégories afin de regrouper les items selon le contenu mesuré et ainsi, en arriver à une plus grande rigueur.

Air American Telephone and Telegraph Company a travaillé avec le personnel d'ETS afin de construire un « in-basket » utilisable pour un poste de cadre moyen. Celui-ci a été utilisé à grande échelle pour l'entraînement et la promotion du personnel cadre. D'autres compagnies comme *IBM*, *Dayton Rubber Cie* et *The Boeing Company* ont suivi. *The Port of New York Authority* et *Proctor and Gamble* ont développé leur propre programme d'entraînement. L'Université du Michigan (Secteur Relations Industrielles) a construit une série de « in-basket » destinée à être vendue aux entreprises. Un exercice « in-basket » utilisé pour les cours de management a également été développé.

Les « in-basket » ont été largement utilisés pour la formation, à cause de leur grande validité apparente et de leur acceptation par les candidats. Par contre, l'utilisation pour la sélection implique plus qu'une simple validité apparente. On cherche donc à maximiser sa validité de contenu et à assurer ses qualités métrologiques (Fredericksen, Jensen, Beaton, 1972).

Les études concernant la validité de contenu deviennent de plus en plus courantes (Dunnette, 1971 ; Howard, 1974 ; Bray, 1982 ; Robertson et Kandola, 1982 ; Turnage et Muchinsky, 1984 ; Hakstian, Woolsey et Schroeder, 1986). Des recherches visant à établir une méthodologie de construction de tests de sélection se multiplient (Cronbach, 1970 ; Anastasi, 1988 ; Society for Industrial and Organizational Psychology, 1987). Plusieurs analyses et méta-analyses sont effectuées concernant les instruments de sélection de type simulations (Bray et Grant, 1966 ; Moses et Byham, 1976 ; Hunter et Hunter, 1984 ; Gaugler, Rosenthal, Thornton III et Bentson, 1987).

Aujourd'hui, des méthodes d'élaboration sont disponibles (Society for Industrial and Organizational Psychology, 1987 ; Pettersen, 1995). De plus, plusieurs études de validité comparant les résultats à des simulations avec l'avancement de carrière confirment la valeur de ces instruments (Bray et Grant, 1966 ; Thornton et Byham, 1982 ; Tziner, 1990). Les études critiquant les simulations sont nombreuses, les améliorations possibles sont mises en évidence; le chemin est tracé.

3. Procédure d'élaboration d'un instrument de sélection basé sur la validité de contenu

3.1 *Les composantes de l'instrument de mesure*

Il a été mentionné précédemment que la démonstration de la validité de l'instrument de mesure repose sur la validité de ses éléments. Or, les éléments d'un instrument de mesure peuvent être regroupés en trois grandes composantes: l'instrument lui-même, ses conditions d'application ainsi que le processus d'observation et d'évaluation (Pettersen, 1995).

3.1.1 *L'instrument lui-même*

La première composante est l'instrument en soi, elle comprend : les items, les questions, les situations ou les tâches, en fait, les stimuli auxquels les candidats sont appelés à réagir. Ces stimuli doivent être représentatifs du domaine à mesurer. « Dans le cas d'un instrument de sélection, ce domaine est défini par rapport à l'emploi, de sorte que les éléments de contenu d'un instrument de mesure doivent refléter les situations qui sont effectivement rencontrées dans l'emploi » (Pettersen, 1995). Une bonne analyse de l'emploi est primordiale. Celle-ci éliminera tout ce qui est trop spécifique, mais tiendra compte de suffisamment de domaines de décision avec lesquels le candidat aura à composer pour permettre une certaine généralisation du test (Society of Industrial and Organizational Psychology, 1987). Tout item n'étant pas directement lié à l'emploi,

comportant une difficulté trop élevée pour les besoins de l'emploi ou relié à une tâche trop occasionnelle de l'emploi sera éliminé. De plus, la forme du stimulus devra être tout aussi représentative. Cronbach (1970) mentionne à titre d'exemple le fait qu'un sujet puisse échouer un item non pas par son manque de connaissances, mais par une difficulté de langage dans la formulation de la question. Ainsi, l'item ne serait pas représentatif si une telle qualité du langage n'était pas nécessaire à l'efficacité dans l'emploi en question.

3.1.2 *Les conditions d'application*

Cette deuxième composante de l'instrument de mesure comprend les directives, les conditions d'administration, le mode de présentation des stimuli et le mode utilisé pour les réponses. Ainsi, les conditions présentées dans le test se doivent de présenter adéquatement ce qui se passe dans l'emploi afin de s'assurer que la mesure reflète véritablement la performance que le candidat fournirait réellement dans l'emploi.

Il est important que les directives et les conditions d'application reflètent le domaine de l'emploi à mesurer. Les candidats doivent être soumis à des facteurs temporels, spatiaux et psychologiques les plus près possibles de l'emploi réel. Par exemple, un test mettant l'accent sur la vitesse n'est justifié que si celle-ci est une dimension importante de l'emploi (Cronbach, 1970).

Les différents stimuli, de même que les réponses, doivent être présentés d'une façon semblable à l'emploi (Tziner, Jeanrie et Cusson, 1993). Par exemple, si un poste implique des interactions avec les autres, il serait plus pertinent que les stimuli soient présentés à travers des interactions verbales que par écrit, tout comme effectuer concrètement une soudure serait plus pertinent qu'énumérer les étapes à suivre pour souder, etc.

3.1.3 Le processus d'observation et d'évaluation

La troisième composante concerne le processus d'observation et d'évaluation. Les dimensions évaluées se doivent d'être représentatives de l'emploi à mesurer. La grille de correction doit être basée sur des solutions reconnues comme étant valables par rapport à la situation réelle de travail. La compilation des scores, s'il y a lieu, ne doit pas fausser les résultats par une pondération inadéquate.

3.2 Différentes procédures d'élaboration

Afin de documenter le choix d'une procédure d'élaboration d'instrument de sélection basé sur la validité de contenu, plusieurs modèles ont été envisagés. Ces différentes procédures visent à circonscrire le plus rigoureusement possible le domaine de contenu de l'emploi mesuré.

3.2.1 *Society for Industrial and Organizational Psychology (SIOP)*

La Société Industrielle et Organisationnelle de psychologie présente une méthode standard ainsi que divers conseils pour l'élaboration de tests basés sur la validité de contenu (1987). Les étapes d'élaboration proposées sont les suivantes :

La définition du domaine de contenu et l'échantillonnage

La définition de l'échantillon de travail doit être basée sur une compréhension du travail à effectuer et des besoins organisationnels. Le domaine de contenu ne doit pas nécessairement couvrir le travail à faire en entier. Par contre, l'échantillon doit être représentatif. L'investigateur se doit d'expliquer ce qui est inclus dans le domaine de contenu, les raisons pour lesquelles ces tâches sont incluses et pour lesquelles certaines parties n'ont pas été retenues.

Les circonstances spéciales

Si certaines tâches nécessitent des connaissances qui peuvent varier, ou qui peuvent être interprétées différemment dépendamment des situations, ces aspects de l'emploi ne doivent pas être inclus dans la mesure.

L'analyse minutieuse et précise de l'emploi

La justification d'un test élaboré sur la base de la validité de contenu est la similarité entre le contenu du test et le domaine de contenu de l'emploi. La

définition du domaine de contenu doit être faite à partir d'une analyse des tâches, des activités et des responsabilités relatives à l'emploi mesuré. Les habiletés, les connaissances, les traits et les autres caractéristiques personnelles nécessaires à la performance à l'emploi doivent être bien définies. Pour ce faire, l'analyse du travail doit être basée sur le jugement d'experts du domaine.

L'expérience à l'emploi

Le domaine de contenu doit être défini en des termes permettant d'évaluer ce que le candidat doit savoir faire ou connaître à la base, en excluant toute tâche qui nécessiterait un entraînement spécifique. Ainsi, il importe de définir toute connaissance ou habileté que le candidat doit maîtriser avant d'exercer véritablement l'emploi. Évidemment, la délimitation est mince entre la contribution ou l'apport de l'employé versus le résultat après l'entraînement à l'emploi. Malgré tout, il est important de s'assurer que la généralisation du test est suffisamment grande pour éviter que tous les aspects qui nécessitent un entraînement spécifique soit éliminés de la mesure. Par contre, on doit s'assurer que le candidat possède les connaissances ou les habiletés qui lui permettront de s'adapter et d'apprendre afin de fournir une performance adéquate pour le poste.

Les activités et tâches fréquentes ou critiques à l'emploi

Les aspects triviaux de l'emploi ne doivent pas intervenir dans la mesure. Par contre, un type d'activité non fréquent mais critique pour l'emploi peut être inclus dans le test.

Les propriétés de la mesure

Lorsque c'est possible, il importe de pré-tester la procédure de sélection. Cela pourra permettre d'ajuster le test. Par exemple, les parties du test qui ne contribuent pas à la variance devraient être éliminées. Par contre, il importe que le choix des items qui seront retenus se fasse en rapport avec leur degré de représentativité du contenu de l'emploi avant toute considération reliée à l'analyse d'items. Il en est de même pour la corrélation entre les items. En effet, une extrême corrélation doit être évitée mais malgré tout, le critère principal d'inclusion ou d'exclusion demeure le degré de représentativité de l'item par rapport au domaine de contenu.

La fidélité présente des aspects particuliers lorsqu'il s'agit d'instruments de mesure pour la sélection. Des variations en ce qui concerne la correction peuvent apparaître : un désir de minimiser le temps de passation peut amener une variance dans les résultats si la passation est trop courte ou si l'échantillon de travail est trop petit. Les effets de la pratique ou de la fatigue des candidats peuvent aussi amener une fluctuation dans les résultats.

La grille de correction doit tenir compte de réponses reconnues dans le domaine. En effet, une réponse acceptée signifie qu'on peut raisonnablement attendre cette réponse dans toute circonstance.

Quant à l'interprétation des résultats, elle dépend des propriétés de la mesure. En effet, si l'instrument échantillonne une part substantielle ou importante du domaine de contenu, il permettra une discrimination adéquate des candidats. Cependant, si l'instrument est construit dans une optique de formation plutôt que pour la sélection, les examinateurs évalueront en fonction du potentiel des répondants et dans ce cas, un score critique serait plus approprié. Un tel type d'évaluation est aussi approprié dans des cas où les conditions de la mesure ne peuvent être identiques à l'emploi à cause de limitations dans la procédure.

Les experts reconnus

Les experts ayant contribué à n'importe laquelle des étapes d'élaboration du test doivent être qualifiés et leur expertise du domaine doit être reconnue. Les personnes qui auront été consultées pour définir le domaine de contenu, pour le développement des items ou pour l'évaluation du test doivent être choisies minutieusement. De plus, elles doivent connaître et doivent être informées précisément de leur rôle.

Quoique la Society for Industrial and Organizational Psychology (1987) suggère ces étapes d'élaboration, le processus peut être un peu flou et laisse place à l'interprétation. En effet, ce procédé contient beaucoup plus de recommandations que d'étapes d'élaboration clairement définies. Il pourrait être difficile de trancher, à savoir quelle étape franchir avant une autre. Cette imprécision pourrait rendre les résultats difficiles à interpréter puisqu'il serait ardu de déterminer où se situent les failles dans l'élaboration : cette étape aurait-elle dû être effectuée avant celle-ci? À quelle étape le jugement d'experts aurait-il dû intervenir? Néanmoins, ces recommandations devront être prises en considération lors de l'élaboration de l'instrument.

3.2.2 Gatewood et Feild (1994)

Ces deux auteurs proposent une méthode d'élaboration générale pour tout instrument de sélection élaboré sur la base de la validité de contenu, en voici les différentes étapes :

L'analyse de poste

L'analyse de poste est au coeur de toute démarche d'élaboration d'instrument de sélection. La qualité de cette analyse est l'ingrédient essentiel de la validité de l'instrument. Les résultats de l'analyse de tâche permettent de définir le domaine de contenu ; en alliant le domaine de contenu avec le contenu de la procédure de sélection, la validité est établie. Cette analyse doit comprendre :

- Une description des tâches de l'emploi.
- Le degré d'importance de ces tâches ou le niveau critique.
- La spécification des connaissances, des traits et des habiletés que requiert la performance aux tâches considérées comme critiques.
- L'évaluation du degré de l'importance des tâches ou des connaissances, des traits ou des habiletés essentiels à la performance doit comporter :
(1) une définition opérationnelle de chacune des connaissances, des traits ou des habiletés (CTH), (2) une description des relations entre chacune des CTH et la tâche correspondante, (3) une description de la complexité ou de la difficulté de chacune des CTH, (4) une indication expliquant la nécessité de chacune des CTH en relation avec la performance à l'emploi.

La sélection des experts

L'application de stratégies de validité de contenu nécessite le jugement d'experts. Habituellement, les experts sont choisis parce qu'ils occupent déjà le poste mesuré ou qu'ils occupent un échelon supérieur (superviseurs du poste visé). Ces experts doivent être des individus aptes à procurer des informations précises sur les tâches à effectuer et sur les connaissances, habiletés ou traits à posséder afin de performer à l'intérieur de ces dites tâches. De plus, il est essentiel que ces juges soient bien sélectionnés, informés et entraînés face aux extrants attendus d'eux.

La spécification du contenu de la mesure

Une fois que le domaine de contenu a bien été défini, les items, questions ou autre contenu de l'instrument de mesure doivent être spécifiés. Il importe à cette étape que les items, questions ou autres soient représentatifs de la définition du domaine de contenu. Les connaissances, traits ou habiletés doivent être mesurés en proportion de leur importance dans le domaine de contenu. En d'autres mots, l'instrument doit présenter des stimuli représentatifs de l'univers à échantillonner. Les experts doivent se positionner afin de juger cette représentativité. L'inclusion ou l'exclusion d'un item se fera à partir de leur jugement.

Ce procédé d'élaboration est plus concret que celui présenté par la SIOP et les étapes sont assez bien définies. Par contre, les sous-étapes ne sont pas explicites : le processus est regroupé en trois larges phases, ce qui laisse place à l'interprétation. Il serait intéressant d'utiliser une procédure plus précise, plus exhaustive, laissant moins de marge de manoeuvre pour le chercheur. À l'intérieur d'une démarche de construction de test élaborée sur la base de la validité de contenu, plusieurs décisions doivent être prises. Le jugement entre en considération à plusieurs endroits. Cette complexité, inhérente à ce type de construction, complique déjà la démonstration de la validité du test. Il est donc inutile d'ajouter un degré de difficulté en laissant place au jugement dans le procédé d'élaboration lui-même.

3.2.3 *Mussio et Smith (1979)*

Ces deux auteurs, dans un ouvrage sur la sélection, proposent différentes étapes pour l'élaboration d'un procédé de sélection.

La sélection du format de l'instrument

Il importe de définir le type d'instrument qui permettra de mesurer le plus adéquatement possible les tâches de l'emploi. Cela sous-entend que plus le test est une approximation précise de l'emploi, plus la qualité des inférences faites à partir de l'instrument sera grande.

Le type de communication et le niveau de langage

Le test doit être construit de façon à choisir le type de communication approprié pour l'emploi. En effet, il importe de définir le type de communication adéquat, le niveau de langage, la pertinence de termes techniques ou de codes relatifs à l'emploi. Les candidats doivent être en mesure de comprendre les instructions du test en fonction de leurs connaissances. Le test ne doit pas comporter des difficultés, dans sa présentation, qui ne sont pas présentes à l'emploi. La façon dont l'instrument est présenté doit être représentative de l'emploi mesuré. Par exemple, si le poste nécessite une parfaite connaissance de l'anglais, l'instrument pourrait être présenté dans cette langue ; si les interactions humaines sont courantes, le test aurait peut-être avantage à simuler de véritables échanges.

Les items doivent présenter les qualités suivantes :

- Utiliser un niveau de langage approprié.
- Rédiger les instructions et les items dans un langage le plus clair et le plus concis possible.
- Regrouper les items concernant le même contenu à l'intérieur de la grille de correction.
- Utiliser un type de contenu par question.
- Éviter la formulation négative qui pourrait ajouter une complexité de compréhension inutile.

La durée de la passation

Le test doit être représentatif des contingences de l'emploi. Ainsi, sa durée doit présenter des facteurs qui sont semblables à l'emploi, comme le degré d'anxiété et la durée nécessaire pour se familiariser avec le matériel.

La note de passage

La note de passage doit être établie avec logique. Les candidats ne possédant pas les habiletés ou les connaissances nécessaires afin de satisfaire les exigences de base de l'emploi devraient être éliminés à partir de la note qui a été établie.

La révision du contenu de l'instrument

Le contenu doit être revu par un comité d'experts. Ce comité est composé de trois personnes ou plus qui possèdent une connaissance et/ou une expertise étendue du domaine. Chacun des experts doit examiner le contenu indépendamment de l'avis des autres. Ils doivent être bien informés de la démarche et de leurs responsabilités dans l'établissement de la validité de l'instrument.

Il est suggéré de regrouper les items en trois catégories, selon les recommandations des experts du domaine (1) les items inacceptables, qui doivent être rejetés, (2) les items acceptables et (3) les items qui requièrent des modifications mineures pour être acceptables. Les items classés dans la deuxième catégorie (acceptables) doivent être reconnus comme tels par tous les experts et ce, unanimement. En ce qui a trait aux items de la troisième catégorie (corrections mineures), ceux-ci devront être présentés une deuxième fois aux experts, une fois les modifications apportées, avant d'être inclus dans le test.

Dans les faits, certaines étapes se devraient d'être inversées ou ajoutées. En effet, il serait difficile d'établir la durée limite de la passation sans l'avis des experts du domaine. De même, établir la note de passage sans faire de pré-test serait très arbitraire. Certains points peuvent être précisés ainsi à première vue. Par contre, tout changement dans la procédure implique des réajustements à

toutes les étapes de construction, chacune étant intimement liée aux autres. Il est donc difficile de proposer des changements et de les appliquer ainsi à l'aveuglette. Seule l'application d'une procédure d'élaboration étape par étape permet d'identifier ces lacunes et, possiblement, de proposer des ajustements à posteriori. Cette procédure présente toutefois des avenues ou des recommandations intéressantes qui devront être intégrées afin d'enrichir le procédé d'élaboration qui sera choisi.

3.2.4 *Pettersen (1997)*

Pettersen, après avoir analysé les démarches proposées dans la littérature, a élaboré un processus segmenté en étapes et en sous-étapes.

Phase I : Finalités et spécification du domaine à mesurer

Étape 1. Finalités de l'instrument de mesure

Préciser à quelles fins sera utilisé l'instrument de mesure.

Étape 2. Analyse et description de l'emploi

Identifier les éléments observables, le contexte de l'emploi, le rendement espéré, les habiletés spécifiques, techniques.

Étape 3. Spécification du domaine ou sous-domaine de contenu de l'emploi à mesurer

Déterminer les éléments critiques de l'emploi afin de les représenter dans l'instrument.

Phase II : Développement de l'instrument de mesure

Étape 4. Conception de l'instrument dans son ensemble

Préciser le format, la durée, le mode de correction.

Étape 5. Création des items et échantillonnage

Rédiger les items en fonction des critères qui ont été précisés aux étapes 3 et 4.

Étape 6. Élaboration des outils d'observation et d'évaluation

Identification des réponses attendues, transformation en termes de scores.

Étape 7. Révision de la version expérimentale par des experts

Vérifier auprès d'experts du domaine la clarté et la pertinence du contenu de l'instrument.

Étape 8. Pré-test de la version expérimentale

Administrer l'instrument auprès d'un groupe témoin. Modifications de certaines des composantes de l'instrument, si nécessaire.

Phase III : Implantation et suivi

Étape 9. Rédaction du manuel technique

Décrire de façon détaillée les différentes étapes de construction, la manière d'administrer l'instrument, de le corriger.

Étape 10. Révisions périodiques

Révision des différentes étapes périodiquement ou lorsque la situation évolue.

Cette procédure comprend des étapes décrites avec précision, la démarche est exhaustive. Cette précision augmente la probabilité d'assurer la validité de l'instrument ainsi conçu. Chacune des étapes sont inter-reliées et s'inscrivent dans un processus logique. De plus, cette démarche est conforme aux recommandations et conseils présentés antérieurement dans cette section. Cette procédure d'élaboration sera donc retenue pour l'élaboration de l'instrument. Les étapes seront présentées plus en détails à l'intérieur du chapitre de méthodologie.

4. Simulations et « in-basket »

Les différents instruments de mesure pour la sélection élaborés sur la base de la validité de contenu par rapport aux tâches à mesurer sont les simulations ou les mises en situation, les examens de connaissances et les échantillons de travail. Dans le cadre de la présente recherche, un « in-basket » sera élaboré afin d'appliquer la procédure d'élaboration, un tel type de test s'inscrit dans la catégorie des simulations. Voici donc les définitions et quelques critiques concernant les simulations et plus spécifiquement, le « in-basket » :

4.1 Description

Les simulations ou mises en situation sont des instruments de mesure basés sur la participation des candidats qui doivent utiliser des comportements typiques requis par le poste visé (Rodger, 1988). La mise en situation place les candidats dans une situation contrôlée qui représente une approximation de la réalité et dans laquelle ils doivent réagir. Lopez (1966) présume qu'ils reproduiront leurs comportements habituels, permettant ainsi de prédire leur performance future. La mesure de la performance ou des comportements typiques de l'emploi se présentera sous différentes formes, selon l'emploi à mesurer.

Les différents types de simulations peuvent être regroupé en deux catégories : simulations de groupe et simulations individuelles.

Les simulations de groupe permettent habituellement d'observer les comportements sociaux de l'individu. Les principaux types de simulations de groupe sont :

- la discussion avec ou sans leader,
- la discussion avec ou sans rôles assignés,
- la discussion dans un climat de compétition ou de collaboration,
- ainsi que les jeux d'entreprise.

(Cook, 1988 ; Rodger, 1988).

Les simulations individuelles regroupent, quant à elles :

- la plupart des tests de type « papier-crayon »,
- les jeux de rôles,
- ainsi que différents types d'entrevues.

(Cook, 1988 ; Rodger, 1988).

L'exercice du courrier ou « in-basket » s'inscrit dans la catégorie des simulations individuelles comme étant un test de type « papier-crayon ». Des différentes mises en situation (ou simulations), celui-ci semble être le test le plus couramment utilisé. En effet, lorsque l'on a recours à une simulation, le « in-basket » est utilisé dans 85% des cas (Thornton, 1992). L'exercice du courrier ou « in-basket » confronte le candidat à une situation de travail hypothétique où il doit prendre des décisions, écrire des lettres, prévoir des réunions, etc., à partir de mémos, de messages téléphoniques et d'autres documents provenant de son panier de gestion (Lopez 1966 ; Thornton 1992). Ces documents présentent des problèmes réalistes qui suscitent des réponses exprimées en termes de comportements. Le cumul de ces réponses fournit une base d'analyse des qualités démontrées lors d'actions ou de décisions administratives (Lopez 1966 ; Moreau 1983). La durée relativement courte de l'administration d'un « in-basket », de une à trois heures selon le cas, pousse l'individu à établir des priorités relativement aux problèmes à résoudre (Fredericksen, Saunders et Wand 1957, cité dans Lopez 1966).

4.2 Forces et faiblesses

La simulation constitue un outil des plus efficaces pour la sélection et la promotion du personnel cadre (Thornton, 1992). La validité prédictive des tests situationnels a été largement éprouvée (Gaugler, Rosenthal, Thornton & Bentson, 1987 ; Huck, 1977 ; Thornton et Byham 1982 ; Tziner, 1990). Les conclusions de ces recherches sont très positives. Les résultats à des tests situationnels ont été comparés à des critères tels : la progression de carrière, l'avancement salarial, les évaluations du superviseur et les évaluations de potentiel. Selon ces mêmes auteurs, un coefficient de corrélation médian variant autour de 0.40 entre les résultats aux simulations et les critères externes a été obtenu.

Les simulations ont été comparées à différents modes de sélection dans une étude de Thornton et Byham (1982). Les simulations ont été comparées aux évaluations de superviseurs, aux tests d'intelligence, aux tests d'intérêt et de personnalité, aux évaluations cliniques, aux entrevues et curriculums. Les résultats indiquent que les simulations sont plus appréciées des candidats et des employeurs et que leur coefficient de validité est semblable aux tests d'intelligence, d'intérêt et de personnalité.

Cette même étude affirme que les entrevues et les évaluations de rendement sont efficaces lorsqu'il y a peu de candidats. Par contre, lorsque le

nombre de candidats augmente ou lorsque l'emploi postulé est différent du précédent, les simulations apportent plus d'informations. De plus, les évaluations de rendement permettent d'éliminer les candidats indésirables, mais ne permettent pas d'obtenir suffisamment d'informations pour différencier un bon candidat d'un autre, d'autant plus que les évaluations ne proviennent pas de la même personne.

Pour ce qui est des tests psychométriques, Thornton et Byham (1982) recommandent leur usage afin de produire une première élimination des candidats. En effet, l'administration et la correction de tels tests étant plus standardisées, ils sont plus faciles à utiliser que les autres méthodes lorsque le nombre de candidats est très élevé.

Une évaluation clinique s'est révélée profitable seulement lorsque le nombre de postulants est très restreint. De plus, les conditions d'évaluation doivent être très bien contrôlées.

En conclusion, les simulations sont des procédures plus onéreuses, mais fournissent plus d'informations et sont plus facilement acceptées par les candidats et les employeurs (Thornton et Byham, 1982).

Par contre, les résultats aux simulations semblent corrélérer plus fortement avec le potentiel promotionnel du candidat qu'avec des indicateurs objectifs de sa

compétence ou de son efficacité (Turnage et Muchinsky, 1984). Est-ce le système promotionnel qui est déficient ou les tests situationnels qui ne mesurent pas la bonne variable? Des recherches de Schmidt, Noe, Meritt et Fitzgerald (1984) ont démontré que les corrélations entre les différentes évaluations d'un employé, réalisées par des superviseurs différents, sont faibles. On peut donc supposer que l'évaluation du rendement étant subjective, elle n'est pas une variable suffisamment fiable comme critère de comparaison pour déterminer la validité des tests.

Parmi les simulations, le « in-basket » a obtenu la meilleure corrélation entre les résultats au test et les promotions obtenues huit ans plus tard (Byham et Thornton, 1970). La validité de pronostic et de diagnostic du « in-basket » est très bonne (Cook, 1988), habituellement plus consistante que d'autres outils d'évaluation (Sackett, 1982). Le niveau de biais négatifs par rapport aux différentes minorités est très restreint (Robertson et Iles 1988). De plus, c'est aussi le test qui mesure le plus de dimensions (Gaugler et al. 1990).

Selon Bray et Grant (1966), parmi les simulations, le « in-basket » s'avère l'un des instruments présentant la plus grande validité prédictive. Le « in-basket », comme apport particulier aux centres d'évaluation, aide à poser un diagnostic sur les participants en comparant les prises de décisions obtenues lors de la simulation avec un modèle idéal de fonctionnement (Moreau, 1983). Si

quelqu'un peut traiter avec succès les problèmes administratifs qui lui sont présentés dans le test « in-basket », nous pouvons supposer, comme le suggèrent Hemphill, Griffiths et Fredericksen (1962), qu'il possède les habiletés d'un bon administrateur et que ces mêmes habiletés se transposent dans la vie réelle.

Selon Fredericksen, Jensen et Beaton (1972), le « in-basket » procure des mesures comportementales de la personnalité, comme la tendance à prendre des décisions, à être informel, à utiliser la discussion comme méthode de résolution de problème, à interagir avec les supérieurs ou à retarder indûment les décisions.

CHAPITRE III : ÉLABORATION DE L'INSTRUMENT

(PROCESSUS GÉNÉRAL D'ÉLABORATION)

Dans la section précédente, il a été démontré que la validité de contenu d'un instrument de sélection réside dans la rigueur de son élaboration. Le développement de ces tests au cours des cinquante dernières années a permis d'identifier les lacunes et les forces des simulations. Différents auteurs ont proposé des modes d'élaboration. Chacun de ces modes présente des forces et des faiblesses. Le processus retenu pour la présente recherche est celui de Pettersen (1995), car il présente des caractéristiques intéressantes: il est complet, exhaustif et précis ; la démarche à suivre est explicite et laisse peu de place à l'improvisation. À chacune des étapes d'élaboration, il importe de démontrer la rigueur dans les choix et les décisions qui seront prises. Chacune de ces étapes devra contribuer à la validité du test et s'appuyer sur une base solide afin d'assurer le passage à l'étape subséquente. Ainsi, et seulement ainsi, l'on pourra prétendre, dans une certaine mesure, à la validité de contenu de l'instrument. L'application de la démarche permettra ensuite de vérifier la pertinence de chacune des étapes, la justesse de la séquence de ces dites étapes ainsi que la précision du processus. Des recommandations pourront ensuite être faites en conséquence des résultats obtenus.

Comme l'objectif de cette recherche n'est pas la production d'un instrument comme tel, l'instrument ne comportera pas de solutionnaire, cette étape comprenant à elle seule une démarche tout aussi complexe que l'élaboration de l'instrument lui-même.

Le présent chapitre sera divisé en fonction des étapes d'élaboration, soit les étapes no 1, 2, 3, 4, 5 et 7. Chacune des sections, ou étapes, comportera l'explication de la démarche ainsi que le rationnel qui la sous-tend. Voici un rappel des étapes de construction du test, soit les divisions du chapitre.

- Étape 1. Finalités de l'instrument de mesure
- Étape 2. Analyse et description de l'emploi
- Étape 3. Spécification du domaine ou sous-domaine de contenu de l'emploi à mesurer
- Étape 4. Conception de l'instrument dans son ensemble
- Étape 5. Création des items et échantillonnage
- Étape 7. Révision de la version expérimentale par des experts

1. Étape 1 : Finalité de l'instrument de mesure

Il importe tout d'abord de préciser les objectifs de l'instrument en termes d'utilisation. Selon qu'il s'agisse de sélection, de formation ou de promotion, les

contingences de passation et de correction seront différentes et la façon de conduire l'analyse de l'emploi pourra varier (Pettersen, 1997). À cet égard, une étude ayant pour but d'expliquer le grand nombre de taxonomies que l'on retrouve en gestion mentionne justement que ces variations dans les taxonomies sont dues aux différents angles selon lesquels le poste est analysé (Yukl, 1987).

Tel que mentionné précédemment, le « in-basket » à élaborer sera conçu pour la sélection de personnel cadre, en l'occurrence des gestionnaires de projet. Il importe d'être conscient de l'implication du choix de l'utilisation du test dans un contexte de sélection. En effet, un instrument conçu en ce sens se doit d'être robuste à la falsification des réponses et facile à défendre en cas de contestation (Pettersen, 1995). De plus, il devra tenir compte de la période d'apprentissage normale à l'emploi, afin de mesurer seulement des variables ne nécessitant pas un entraînement spécifique (SIOP, 1987). Il est essentiel d'être congruent en ce qui a trait au niveau de généralisation du test lors de la construction. Il importe donc de retenir le fait que le test doit être utilisable, généralisable à la gestion de projet en général. Compte tenu de ce choix, le test devra échantillonner des tâches communes à tous les gestionnaires de projet, et non pas à un poste en particulier, comme la gestion d'un chantier de construction ou d'un projet informatique, par exemple. On doit donc retenir l'importance du choix des finalités de l'instrument de mesure, puisque ce choix induit plusieurs implications, conséquences, balises, qui accompagneront la démarche du début à la fin.

2. Étape 2 : Analyse et description de l'emploi

« Le concept d'analyse d'emploi réfère à tout processus de collecte d'informations servant à décrire un emploi » (Harvey, 1991). Cette étape est essentielle puisqu'elle constitue le canevas de la construction du test (SIOP, 1987). Le but est d'arriver à préciser le plus clairement possible l'ensemble des éléments d'un poste de travail (SIOP, 1987). D'abord, l'analyse de l'emploi ne doit pas se limiter à observer le quotidien du gestionnaire sur une courte période. Ainsi, le « in-basket » devra échantillonner une année complète de travail et non pas seulement une partie de la réalité de tous les jours (Lopez, 1966). De plus, comme la performance peut varier en fonction des diverses situations (Sackett et Wilson (1982), il est essentiel d'échantillonner toutes les situations importantes de l'emploi visé. Une partie du « in-basket » devra donc comporter des problèmes qui ne se présentent pas couramment, mais qui sont tout de même représentatifs de l'emploi sur une longue période. Il importe aussi de procéder à l'analyse de l'emploi afin d'éliminer tout ce qui est trop spécifique, tout en tenant compte de suffisamment de domaines de décision pour permettre une certaine généralisation du test (SIOP, 1987). La définition du poste devra être basée sur une bonne compréhension de l'emploi, tenir compte des besoins de l'organisation et du marché et comprendre toutes les dimensions du poste (SIOP 1987).

Souvent appelée analyse de tâches ou analyse de postes, l'analyse des emplois est un domaine de la psychologie industrielle et de la gestion des ressources humaines fort développé. Plusieurs techniques de collectes de données ont été mises de l'avant. Le but de cette recherche n'étant pas centré sur l'analyse de l'emploi, il n'apparaît pas important d'en énumérer toutes les variantes ainsi que leurs définitions et procédures. Mais notons, entre autres, l'analyse fonctionnelle (Fine et Wiley, 1971 ; Tziner, Jeanrie et Cusson, 1993), l'analyse des processus de travail (Mc Cormick, Jeaneret et Mecham, 1972 ; Tziner, Jeanrie et Cusson, 1993) ainsi que l'analyse des situations de travail (ou incidents critiques) (Flanagan, 1954).

En ce qui a trait à la gestion et à la gestion de projet plus spécifiquement, plusieurs études ont été faites afin d'en définir et d'en circonscrire les tâches. La cueillette de données primaires n'est donc pas nécessaire. L'envergure des données secondaires disponibles permet de s'appuyer sur ces différentes recherches et de s'en inspirer afin de définir la gestion de projet.

La gestion de projet se définit par des méthodes ou techniques créées pour la conception, l'analyse et la conduite d'activités temporaires fortement irréversibles, non répétitives, réalisées sous contraintes de temps et de ressources (Declerc, Eymery et Crener, 1980 ; Pettersen, 1989).

À la base, la gestion de projet est un domaine très près de la gestion plus traditionnelle. Elle s'est différenciée de la gestion fonctionnelle à cause de certains besoins du marché. En effet, la gestion de projet répond avant tout à une nouvelle philosophie dans l'entreprise (Ménard, 1995).

La disparition des marchés et de la production de masse, la demande pour une qualité supérieure et une réponse rapide et la fragmentation des marchés sont autant de facteurs qui ont amené une complexité croissante de la gestion (Peters, 1988). En fonction de ces nouvelles contingences, il est nécessaire d'adopter une philosophie de management par projet (Serieyx, 1993). Toute entreprise ne sachant pas s'adapter à ces nouvelles exigences aura peine à survivre dans le monde qui se crée (Ménard, 1995).

Cette nouvelle philosophie, qui distingue la gestion de projet de la gestion plus traditionnelle, est basée sur deux postulats principaux :

- (1) les caractéristiques particulières d'un projet impliquent une gestion différente de celle proposée pour une activité continue
- (2) une approche générique peut être utilisée pour tout projet.

De plus, cette philosophie de la gestion suppose : (1) une culture particulière, (2) un mode d'organisation particulier, (3) un ensemble de techniques

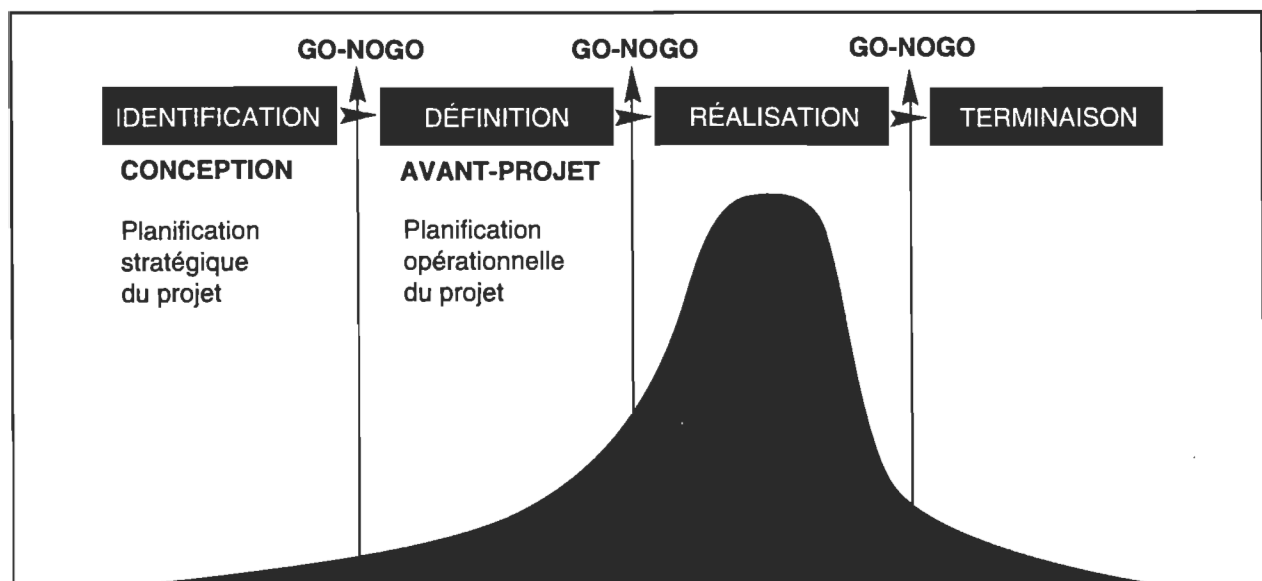
et d'outils pour gérer de façon efficiente et efficace le cycle de vie des projets (Ménard, 1995).

Puisque le concept de projet est ce qui caractérise principalement la gestion de projet, il importe d'en définir clairement le sens. « Un projet est un ensemble d'activités interdépendantes menant à la livraison d'un produit ou d'un service clairement identifié et généralement dans un contexte de temps et de ressources limité » (Ménard, 1995). Les principaux domaines où le concept de projet est utilisé sont : la construction, l'ingénierie, la recherche et le développement de nouveaux produits, l'informatique, les programmes sociaux, les études de faisabilité, le milieu organisationnel, l'organisation d'un événement (festival, jeu), le cinéma et la télévision. Le concept de projet est caractérisé par sa durée limitée, par l'assujettissement à diverses contraintes de performance, de qualité, de délais et de coûts, par un cycle de vie dynamique, par l'implication de nombreux intervenants, ainsi que par un contexte d'incertitude.

Un projet est caractérisé par sa nouveauté, son unicité. De fait, un projet implique généralement de faire quelque chose de nouveau, qui n'a pas été entrepris exactement de la même façon ou dans le même contexte qu'auparavant. Par contre, il est évident que le degré de nouveauté ou d'unicité peut varier considérablement d'un projet à l'autre. Il est assujéti à une durée limitée, il est de nature temporaire. La durée du projet peut être relativement courte (une

construction, par exemple) ou très longue, dans le cas de méga-projets (tels l'ouverture d'un hôpital ou un développement urbain). Un projet se caractérise par un cycle de vie dynamique, généralement représenté en quatre phases : l'identification (préciser l'objectif), la définition (planification), la réalisation et la terminaison (livraison du produit et fermeture administrative du projet). Le schéma du cycle de vie d'un projet est présenté à la figure 1 (Ménard, 1995).

Figure 1



Extrait de Ménard, P. (1995). *La gestion de projet*, Ordre des comptables licenciés du Québec. p.8.

La partie ombragée correspond au taux d'utilisation des ressources. Elle augmente progressivement pour atteindre son apogée à la phase de réalisation.

La définition de la gestion de projet impose le cadre dans lequel l'instrument s'inscrira. Ce cadre, ou l'univers à échantillonner se doit d'être précisé. Suite à la

définition, voyons donc les différentes fonctions, les différentes tâches ou les différents rôles spécifiques du gestionnaire de projet.

Selon Cleland & King (1988), les fonctions principales du manager de projet sont la planification et l'exécution. La dimension planification inclut les concepts de planification à long terme et la prise de décision stratégique. Le terme exécution comprend des sous-fonctions, comme l'organisation, la motivation et le contrôle (Pettersen, 1989).

Dans le domaine de la construction, O'Brien (1975) identifie trois principales tâches attribuées au gestionnaire de projet. On y retrouve : (1) le développement d'un plan de projet ainsi que son implantation, (2) la coordination de l'équipe de projet, fonction reliée à ses capacités de leadership et (3) l'utilisation des outils de gestion et de communication. Également dans ce domaine, Mueller (1979) dénombre huit tâches principales dévolues au gestionnaire de projet (1) la supervision, (2) l'estimation et l'offre de service, (3) la planification et l'inventaire, (4) l'acquisition, (5) les communications écrites, (6) l'expédition, (7) les contrats administratifs et (8) l'inspection. De plus, il mentionne dix autres tâches plus secondaires, en ordre d'importance : (9) la gestion des plans et devis, (10) la négociation de contrats, (11) les ventes et les relations publiques, (12) la gestion de la sécurité, (13) la comptabilité des coûts, (14) la gestion des équipements, (15)

la structuration des tâches, (16) l'approbation de la paie, (17) les comptes à payer et (18) les comptes à recevoir (Pettersen, 1989).

S'inscrivant dans l'approche systémique, Ménard (1995) propose cinq rôles principaux : (1) le stratège, (2) le pilote, (3) le négociateur, (4) le leader et (5) le dépanneur. Le stratège effectuera la gestion des interfaces externes, prendra des décisions stratégiques et politiques et sera responsable de la gestion des risques. Le rôle de pilote fait appel à des dimensions comme la planification, l'organisation et le contrôle. Son mandat de négociateur touche ce qui a trait à l'acquisition des ressources, aux approvisionnements, aux mandats et aux contrats. L'aspect « leadership » implique la gestion de l'équipe, la gestion du personnel, la gestion des conflits et les communications. Le rôle de dépanneur concerne la gestion des problèmes imprévus.

Adams, Barndt et Martin (1979) déterminent quant à eux cinq rôles principaux qui caractérisent le gestionnaire de projet : (1) l'intégrateur, (2) le communicateur, (3) le leader d'équipe, (4) le preneur de décisions et (5) le responsable du climat de travail.

Les membres de la firme de consultation *Technology Management Group of Pugh-Roberts Associates* ont observé que le succès d'un projet reposait sur cinq dimensions. La conception, l'entrepreneurship, la direction (planification et

coordination), la recherche d'information ainsi que le parrainage (guider les membres moins expérimentés de l'équipe) sont ses cinq rôles principaux. À ceci, ils ajoutent la résolution de problèmes de nature technique (Pettersen, 1989).

Shanks (1983), quant à lui, mentionne quatre aspects fondamentaux à la gestion de projet : (1) la gestion des résultats, (2) la gestion du travail, (3) les ressources humaines et matérielles et (4) la gestion des crises (cité dans Pettersen, 1989).

Spitz (1982) regroupe les activités des gestionnaires en sept catégories : (1) la planification, (2) l'identification et l'acquisition des ressources, (3) la coordination, (4) la recherche d'information, (5) la communication, (6) le contrôle et (7) le catalyseur ou leadership.

Il n'est pas simple de faire la synthèse de ces publications. En effet, chaque auteur a tendance à développer sa propre grille d'analyse. On parle de tâches, de fonctions ou de rôles et la méthode d'analyse diffère. De plus, il s'agit ici d'une description générique. L'importance des fonctions pourra varier d'un secteur d'activité à un autre, ainsi que d'un mode d'organisation à un autre. D'ailleurs, comme le mentionne Pettersen (1989), « la diversité des cadres d'analyse est tellement grande qu'il est pratiquement impensable de réunir en une image cohérente des pièces qui n'ont pas été conçues pour s'emboîter les unes dans les

autres ». Voici un tableau comparatif (tableau 1), un résumé de Pettersen (1989) qui a repris les principaux travaux des auteurs ayant identifié les tâches du gestionnaire de projet. Ce tableau, si ce n'est une synthèse intégrative, permettra à tout le moins d'avoir une vue d'ensemble des tâches, rôles et responsabilités du gestionnaire de projet.

Tableau 1

Tâches, rôles et responsabilités du gestionnaire de projet

Ménard (1979)	Adams et al. (1979)	Roberts et Fusfeld (1981)	Martin (1978)	Shanks (1983)	Wilemon et Cicdero (1970)	Spitz (1982)
Stratège et Ag. de liaison	Communicat.	Conception Résolution de problème Entrepre- neurship Recherche d'information	Représentant Rel. avec la direction	Gestion des résultats Gestion du travail		Planification Recherche d'information Communicat
Pilote	Preneur de décisions	Direction	Preneur de décisions	Gestion des ressources Gestion du risque et des crises	Contraintes organisation. Risque	Identification et acquis des ressources
Coordinateur	Intégrateur	Parrainage	Intégrateur		Équilibre technique versus managérial Relations interperson.	Coordination Contrôle Catalyseur
	Responsable du climat					

Extrait de Pettersen, N. (1989). *Que font les gestionnaires de projet*, Gestion 2000, 5, page 102.

Cette démarche aura permis de bien circonscrire les tâches, rôles ou responsabilités du gestionnaire de projet et permettra de soutenir l'étape subséquente, soit la définition du domaine à mesurer. Les recherches qui ont inspiré la définition des rôles ou des tâches sont rigoureuses et bien construites. En effet, elles s'inspirent pour la plupart de connaissances reconnues du domaine. Entre autres, Ménard (1979) est reconnu comme une sommité d'envergure nationale en gestion de projet. Certains ont complété leur étude par une recherche empirique, c'est le cas de Spitz (1982) qui, dans une thèse de doctorat, a compilé les résultats de 20 entrevues auprès de spécialistes, experts du domaine. S'inspirer de recherches rigoureuses permet d'assurer la validité de cette étape, de consolider la démarche et d'offrir les assises nécessaires afin de spécifier le domaine et le sous-domaine de contenu à mesurer.

3. Étape 3 : Spécification du domaine ou du sous-domaine de contenu de l'emploi à mesurer

Les rôles, tâches qui caractérisent son univers, ont été mis en évidence lors de l'étape précédente, mais quelles sont les principales dimensions qui définissent son univers? La définition des dimensions du domaine de contenu à mesurer permettra de s'assurer de bien échantillonner le domaine de contenu.

L'analyse de tâche offre plusieurs indications à ce sujet, mais il s'agit ici d'en arriver à une définition conceptuelle du domaine ou du sous-domaine à mesurer, à savoir une définition faite de mots et de concepts (Kerlinger, 1973). Dans un contexte de sélection, il a été établi que le domaine de contenu représenté dans l'instrument doit être la performance ou les connaissances ou habiletés nécessaires à cette performance (Equal Employment Opportunity Commission et al., 1978 ; SIOP, 1987). La démarche de la définition du domaine de contenu doit donc permettre l'obtention de dimensions présentant des définitions opérationnelles qui permettront ensuite de transposer ces habiletés sous-jacentes en stimuli.

Une recherche sérieuse ainsi qu'une recension complète des écrits ont été effectuées par Yukl en 1987. Cette recherche ne concerne malheureusement pas la gestion de projet spécifiquement. Yukl a défini les dimensions qui caractérisent le gestionnaire plus général, ou gestionnaire fonctionnel. Ces données sont-elles applicables? Le niveau de généralisation peut-il être adéquat?

Pettersen (1989) mentionne que les activités présentées dans les écrits sur la gestion fonctionnelle et sur la gestion de projet présentent plusieurs concordances et même s'apparentent de façon étonnante, pour une grande partie des tâches recensées. Pourtant, il semble acquis pour les spécialistes en gestion de projet, de par le contexte particulier, que leurs activités diffèrent de celles de

leurs collègues en gestion fonctionnelle. Par contre, la seule étude effectuée comparant directement des gestionnaires de projet à des gestionnaire fonctionnels conclut que les différences sont beaucoup moindres que ce qui aurait pu être attendu (Irwig, 1979). Ainsi, cette différence, attribuable au contexte, pourrait être représentée par le type d'activité dans laquelle la dimension s'illustre, soit les domaines dans lesquels la dimension s'applique.

Afin d'assurer la rigueur dans le choix des dimensions, il apparaît donc approprié d'utiliser la grille de Yukl décrivant la gestion fonctionnelle, puisque celle-ci offre un niveau d'approfondissement supérieur à toute autre recherche en gestion de projet. Ainsi, la grille offrira un niveau de généralisation intéressant et nous verrons subséquemment comment s'assurer que le domaine ainsi défini soit représentatif de la gestion de projet, en illustrant adéquatement le contexte particulier de ce type de gestion.

Tout d'abord, voici les conclusions de la recherche de Yukl (1987) sur les dimensions de la gestion. Ensuite, un sous-ensemble sera défini pour les besoins du « in-basket ».

Yukl (1987) s'est penché sur neuf études présentant des taxonomies différentes. Le nombre de dimensions varie, les différentes catégories se recoupent quelquefois, des taxonomies sont plus spécifiques que d'autres et

divisent les dimensions en sous-catégories. Certaines dimensions sont présentes chez certains auteurs et absentes chez d'autres.

Certaines différences sont attribuables à trois principales causes (Yukl, 1987). Premièrement, le fait que les catégories ou dimensions dans lesquelles sont regroupés les comportements, soient des abstractions en soi. Ces catégories n'existent pas de façon tangible et peuvent être interprétées de façon différente, étant des construits avant tout théoriques. Selon l'angle pour lequel la taxonomie sera développée (formation, sélection, etc.), le focus ne sera pas fait au même endroit. Évidemment, selon que l'on s'intéresse aux dimensions qui illustrent le mieux la performance, ou celles qui permettent de mieux mesurer l'apprentissage, les résultats risquent de différer.

Deuxièmement, le niveau de généralité des catégories peut différer, en ce sens qu'une catégorie plus globale peut en inclure deux ou trois et ne peut donc être comparée à une sous-catégorie d'un autre auteur qui aura choisi de séparer les dimensions de façon plus précise. La difficulté sera de s'assurer de l'indépendance de ces différentes dimensions si la catégorie est trop étroite ou exiguë.

Troisièmement, les différences peuvent être attribuables à la méthode utilisée pour développer la taxonomie. Elles peuvent être développées notamment

par une analyse factorielle, par le jugement d'experts ou par l'approche hypothético-déductive.

Il importe donc, dans le choix des dimensions que l'on jugera pertinentes pour l'élaboration du test, de s'assurer : (1) qu'elles auront été développées pour un contexte semblable, c'est-à-dire la sélection, (2) avec une précision suffisamment grande pour bien définir la dimension, mais sans être trop restreintes afin qu'il y ait vraiment indépendance des critères, (3) et que la méthode de cueillette et d'analyse des données pour le développement de la taxonomie soit conforme à la démarche rigoureuse ici proposée.

Les neuf taxonomies faisant partie de la recension sont le résultat de méthodologies de recherches différentes (analyse factorielle, méthode hypothético-déductive) (tableau 2). De plus, l'objectif de la recherche peut différer. Yukl (1987) a donc dû tenir compte de ces variantes afin d'établir une taxonomie intégrative exempte de contradictions et uniforme quant à l'objectif de la mesure.

Tableau 2

Aperçu des différentes taxonomies de la gestion comparées par Yukl (1987)

Auteurs et dates	Nombre de catégories	Objectif	Approche
Bowers & Seashore (1966)	4	Les comportements efficaces du leader	Théorie hypothético-déductive
House & Mitchell (1974)	4	Les comportements efficaces du leader	Théorie hypothético-déductive
Luthans & Lockwood (1984)	12	Classifier les comportements observables	Classification par des juges
Mahoney, Jerdee, & Carroll (1963)	8	Décrire les exigences administratives	Théorie hypothético-déductive
Mintzberg (1973)	10	Classifier les comportements observables	Classification par des juges
Morse & Wagner (1978)	6	Les comportements efficaces du leader	Analyse factorielle
Page (1985)	10	Décrire les exigences administratives	Analyse factorielle
Stogdill (1963)	12	Les comportements efficaces du leader	Théorie hypothético-déductive
Yukl (1983)	22	Les comportements efficaces du leader	Analyse factorielle

Extrait de YUKL, G.A. (1987). *A new taxonomy for integrating diverse perspectives on managerial behavior*. Working Paper, SUNYA., p. 23

Cinq des taxonomies avaient pour but premier de classer les comportements qui pourraient potentiellement expliquer la performance administrative à travers le leadership. Trois de ces taxonomies ont été développées par la méthode hypothético-déductive. Stogdill (1963) propose une taxonomie du leadership afin d'expliquer les comportements de groupe du

gestionnaire. Bowers and Seashore (1966) proposent une taxonomie du leadership avec quatre catégories basées sur les théories de leadership élaborées par une recherche de l'Université du Michigan. House et Mitchell (1974) présentent une taxonomie du leadership avec quatre catégories de comportements basée sur leur théorie révisée du leadership. Les deux autres taxonomies ont été développées par analyse factorielle soit celle de Morse et Wagner (1978) et celle de Yukl (1983). Morse et Wagner (1978) proposent une taxonomie composée de six dimensions. Quant à Yukl, sur une période de huit ans, ses collègues et lui produisent six études d'analyse factorielle, à partir de questionnaires administrés à des gestionnaires et à leurs subordonnés. Vingt-deux facteurs orthogonaux ont été mis en évidence (Yukl & Nemeroff, 1979 ; Yukl & Van Fleet, 1982 ; Yukl, 1983).

En contraste avec ces cinq taxonomies, deux autres classifications de comportements ont été développées pour coder les observations sur les comportements, afin de fournir une description approfondie du comportement administratif. Ces taxonomies sont des classements basés sur le jugement et leur première source d'information est l'observation de gestionnaires plutôt que les descriptions obtenues par des questionnaires. Mintzberg (1973) a développé une taxonomie comprenant dix catégories de comportements appelés les rôles administratifs (*managerial roles*). Luthans et Lockwood (1984) ont développé une taxonomie du travail administratif regroupant douze catégories de comportements.

Les deux dernières taxonomies ont été élaborées afin de décrire les responsabilités des gestionnaires soit celles de Mahoney, Jerdee et Carroll (1963) et Page (1985). Les catégories de comportements comportent de l'information à propos des traits nécessaires afin de performer comme gestionnaire. Ce type de taxonomie procure une autre perspective pour décrire le travail administratif (Carroll & Gillen, 1987) et est utile afin de comparer les différents types de gestionnaires en termes des traits requis à la performance et du temps accordé à différentes activités. Cela permet de classifier les fonctions administratives (Fayol, 1949 ; Urwick, 1952).

Voici donc les treize catégories ou dimensions regroupant les comportements identifiés par : (1) Yukl (1983), (2) Morse et Wagner (1978), (3) Bowers et Seashore (1966), (4) House et Mitchell (1974), (5) Luthans et Lockwood (1984), (6) Mintzberg (1973), (7) Mahoney et al. (1963), (8) Page (1985) et (9) Stogdill (1963). Cette taxonomie intégrative a ensuite été validée selon diverses méthodes rigoureuses, comme il sera expliqué subséquemment.

Dans cette section, seront présentées les dimensions ainsi que leurs justifications et pertinence. Les définitions sont présentées en appendice afin de ne pas alourdir le texte. (Appendice A)

Information, communication (Informing)

L'importance de cette dimension a été notée pour la première fois par Likert (Yukl, 1987). Plus récemment, Mintzberg (1973) en a fait mention, présentant le gestionnaire comme un centre nerveux de la communication au sein de l'organisation. Pour certains auteurs, tels Mahoney et al. (1965), cette dimension s'intègre dans une autre catégorie, en l'occurrence, la supervision. Luthans et Lockwood (1984), pour leur part, incluent la réception et la communication d'information dans une même catégorie, ce qui implique aussi une capacité de discriminer l'information pertinente.

Malgré que cette dimension ne soit pas présentée avec la même importance selon l'auteur, tous s'entendent pour mentionner que la communication est une dimension importante de la gestion. Le nom de la dimension pourra varier (par exemple, échange d'information de routine, Luthans et Lockwood, 1984), ainsi que le fait qu'elle fasse partie d'une catégorie distincte ou qu'elle soit incluse à l'intérieur d'une autre définition.

Résolution de problème et gestion de crises (Problem solving and crisis management)

Mintzberg (1973) mentionne que le stress organisationnel est centré sur les requêtes demandant une assistance immédiate. Répondre à cette pression par une résolution efficace des problèmes est donc une dimension importante de

l'efficacité organisationnelle (Stewart, 1976 cité dans Yukl, 1987). Cependant, les gestionnaires trop préoccupés par la résolution à court terme de problèmes, perdent de l'efficacité à cause de leur vision trop proximale et n'ont plus suffisamment de temps ou de perspective pour la planification. Les problèmes se présentent donc la plupart du temps en situation d'urgence, étant donné leur mauvaise anticipation (Yukl, 1987).

Toutes les études mentionnent l'importance de la dimension résolution de problèmes et gestion des crises. Pour Morse et Wagner (1978), cette catégorie est appelée « résolution stratégique de problèmes » (« strategic problem solving »). Mintzberg (1973) la nomme « perturbations dans la prise en charge » (« disturbance handler role »). Luthans et Lockwood (1984) la mentionnent sous l'appellation de « résolution de problèmes et prise de décision » (« problem solving and decision making ») et, pour Page, (Yukl, 1987) cette catégorie se nomme « prise de décision » (« decision making »). Pour certains autres auteurs (Stogdill, 1963 ; Bowers and Seashore, 1966 ; Mahoney et al., 1965), il n'y a pas d'équivalent direct, mais la dimension « résolution de problèmes » est incluse à l'intérieur d'autres catégories.

Planification et organisation (Planning and organisation)

La planification est rarement un comportement compris dans un seul épisode. Elle sous-tend une série d'épisodes séparés, quoique reliés sur une

période de semaines ou de mois. Par ses stratégies à long terme, la planification et organisation regroupe des comportements plus difficilement observables (Snyder & Glueck, 1980). Malgré tout, il y a certains aspects de cette catégorie de comportements qui peuvent être observés. Par exemple, la rédaction de plans, la préparation de budgets écrits, le développement d'horaires et l'organisation de réunions afin de formuler objectifs et stratégies.

L'importance de la planification et de l'organisation en tant que dimensions administratives est depuis longtemps reconnue dans la littérature, comme le mentionne Drucker (1974) (cité dans Yukl, 1987). Plus récemment, Carroll et Gillen (1987) en soulignaient la pertinence comme mesure de la performance administrative. La plupart des taxonomies comparées ont une ou deux catégories qui incluent la planification. Morse et Wagner (1978) nomment cette dimension « organisation et coordination », Luthans et Lockwood (1984) la mentionnent sous l'expression « planification et coordination », et Mahoney et al. (1965) la décrivent sous l'appellation « planification ». Page (1985) divise ce comportement en deux catégories, une nommée la « planification stratégique à long terme » et une deuxième appelée « planification et organisation ». Cette distinction entre planification stratégique et planification opérationnelle est utile, mais à l'intérieur d'une taxonomie moins générique. Dans le cas qui nous occupe, cette distinction peut être utile à la compréhension, mais non essentielle à la mesure. Dans la taxonomie de Yukl, les facteurs innovation, planification et organisation sont

présentés comme deux catégories séparées. Par contre, des recherches subséquentes (Yukl & Van Fleet, 1982 ; Yukl & Clemence, 1983) démontrent que l'innovation est un élément inhérent à la planification et à la résolution de problèmes plutôt qu'à une catégorie distincte.

Consultation et délégation (Consultation and delegation)

Comme la dimension support, la dimension consultation et délégation a fait l'objet de multiples recherches. Ces différentes études démontrent que le partage des pouvoirs est un comportement important de l'efficacité du gestionnaire (Locke & Sweiger, 1979 ; Miller & Monge, 1986 ; Schweiger & Leana, 1985, tous cités dans Yukl, 1987). Par contre, encourager la participation des subordonnés aux prises de décision induirait une plus grande satisfaction au travail pour ceux-ci, mais aurait peu d'incidence sur la performance du gestionnaire comme telle. Selon certaines conditions, ces deux aspects pourraient être inclus à l'intérieur d'une même dimension (Yukl, 1987).

Malgré l'évidence de l'apport de la consultation et de la délégation sur la performance du gestionnaire, certains auteurs ne traitent pas cette catégorie séparément dans leur taxonomie. Stogdill (1963) inclut la délégation dans une catégorie nommée tolérance à la liberté, mais n'implique pas l'aspect consultation. Pour House et Mitchell (1974), la consultation est incluse dans une catégorie appelée participation au leadership, mais ne comprend pas la partie délégation.

Quant à Yukl (1983), la taxonomie qu'il présente comporte deux catégories pour couvrir ces aspects, la délégation (appelée délégation-autonomie) ainsi que la consultation (appelée participation aux décisions). Pour Mintzberg (1973), la consultation et la délégation sont impliquées dans la dimension rôle du leader. Pour Luthans et Lockwood (1984), quant à eux, ces comportements sont une partie de la catégorie motivation et renforcement.

La délégation peut être considérée à prime abord comme un concept relié directement au partage des pouvoirs. Par contre, dans la pratique, elle implique une série d'incidents qui peuvent aussi avoir trait à d'autres catégories. Par exemple, décider quelles responsabilités doivent être déléguées à quelles personnes implique la dimension planification. Sélectionner les tâches pouvant aider le développement de tel potentiel chez un employé est directement lié à la catégorie développement de carrière. Les problèmes utilisés pour mesurer la dimension consultation et délégation se doivent donc d'être très spécifiquement choisis afin d'éviter un chevauchement des dimensions dans la mesure.

Clarifier les rôles et objectifs (Clarifying roles and objectives)

L'importance de la clarification des rôles pour l'efficacité du gestionnaire est clairement démontrée par une recherche de Locke, Shaw, Saari et Latham (1981) (Yukl, 1987). Cette recherche a mis en évidence le fait que plus la performance attendue des subordonnés est clairement définie, meilleure est la coopération et

meilleurs sont les résultats de ces dits subordonnés. Early, Wojnaroski et Prest (1987) (cité dans Yukl, 1987), quant à eux, spécifient que la performance des subordonnés est meilleure quand les attentes sont bien clarifiées, puisque cela permet de guider leurs efforts dans des activités productives et de développer leur volonté de s'améliorer en fonction d'attentes explicites.

La clarification des rôles est la seule catégorie de cette taxonomie qui ne présente aucun équivalent direct dans les neuf taxonomies étudiées. Par contre, elle est impliquée comme étant une part d'autres catégories. Pour Stogdill (1963), elle est incluse dans la mise en place de la structure, pour Bowers et Seashore (1966), elle fait partie de la facilitation du travail. Yukl (1983) présente une distinction entre la clarification des rôles et la communication des objectifs ou des attentes. Pour Mintzberg (1973), c'est une part du rôle du leader. Page (1985) (Yukl, 1987) ainsi que Mahoney et al. (1965) en parlent comme d'une partie de la dimension supervision. Cependant, le chevauchement entre les dimensions devient trop évident, comme le démontre une recherche avec incidents critiques (Yukl & Clemence, 1973 ; Yukl & Van Fleet, 1982). Il serait donc plus approprié de traiter cette dimension de façon distincte.

Suivi et contrôle (« Monitoring operations »)

La dimension suivi et contrôle est un concept distinct de la planification et de la résolution de problèmes, mais elle y est étroitement liée. Elle procure

l'information nécessaire pour la planification et la résolution de problèmes. De plus, elle apporte de l'information afin de permettre au gestionnaire d'évaluer adéquatement ses subordonnés, de leur offrir les critiques ou les félicitations en lien avec leur rendement. D'ailleurs, de récentes recherches soulignent l'importance de cette dimension comme étant très reliée à la performance du gestionnaire (Komaki, 1986).

Pourtant, l'importance de cette dimension n'est pas reconnue dans la littérature concernant le leadership et n'est pas représentée dans les taxonomies de Stogdill (1963), Bowers et Seashore (1966), House et Mitchell (1974) ou Morse et Wagner (1978). Quatre des autres taxonomies en reconnaissent l'importance et présentent une catégorie séparée sur le suivi et contrôle. L'appellation peut varier, mais la définition demeure la même (Page, 1985). Yukl (1987), l'inclut à l'intérieur de deux autres catégories : la supervision (« monitoring ») et le contrôle (« controlling »).

Le suivi et contrôle est facilité par le développement de plans détaillés, de d'horaires, d'étapes intermédiaires vers la réalisation du projet et d'indicateurs concrets afin de mesurer l'avancement et la performance. Ces exemples sont donc des comportements observables utiles à la mesure afin d'évaluer cette dimension chez le candidat.

Motivation (« Motivating task commitment »)

L'importance de la motivation sur l'efficacité administrative se reflète aussi dans le fait que chacune des taxonomies a développé une catégorie comportant au moins un des aspects de cette dimension. Pour Stogdill (1974), c'est l'accent sur la production, pour Bowers et Seashore (1966) c'est l'accent sur les buts. Dans la taxonomie de House et Mitchell (1974), il s'agit du leadership orienté sur la réussite et pour Morse et Wagner (1978), cette dimension est appelée motivation et prise en main des conflits. Yukl (1983), quant à lui, a mis en évidence deux catégories qui correspondent à la dimension motivation : l'inspiration et l'accent sur la performance.

Les trois autres taxonomies ont inclus la motivation à l'intérieur d'autres dimensions. Pour Mintzberg (1973), cet aspect fait partie du rôle du leader, pour Luthans et Lockwood (1984), cela fait partie de la dimension renforcement et motivation et pour Page (1985) et Mahoney et al. (1963), cette dimension est citée comme étant une part de la supervision.

Support (Supporting)

Cette dimension a fait l'objet de bien plus de recherches que n'importe quel autre comportement du gestionnaire. Cet aspect important de la gestion a été manipulé en laboratoire, mesuré par des questionnaires, des observations, des entrevues et une cueillette d'incidents critiques. Toutes ces études ont démontré

que le support influence positivement l'attitude des subordonnés et leur performance (Bass, 1981 ; Yukl, 1981).

Une catégorie de comportement semblable au support se retrouve dans la majorité des taxonomies développées afin d'expliquer l'efficacité d'un gestionnaire. Le support est de façon globale l'équivalent de ce que Stogdill (1963) et Yukl (1983) appellent la considération, et House et Mitchell (1974), le leadership supportant. Par contre, les taxonomies développées pour décrire le travail du gestionnaire n'incluent aucune catégorie séparée qui soit équivalente au support. Pour Mintzberg (1973), ce type de comportement est présumé inclus dans la dimension du rôle du leader. Pour Luthans et Lockwood (1984), il fait probablement partie de la catégorie socialisation (Yukl, 1987).

Reconnaissance et récompenses (Recognizing & rewarding)

Malgré que l'importance de cette dimension ait été clairement démontrée, peu de taxonomies en font mention en tant que catégorie distincte. Pour Page (1985), cette dimension est incluse dans la catégorie supervision, Mintzberg (1973) en fait mention comme un aspect du rôle du leader, pour Mahoney et al. (1963), elle apparaît comme une sous-catégorie de supervision et évaluation.

Dans les taxonomies de Stogdill (1963), Bowers et Seashore (1966), House et Mitchell (1974) ou Morse et Wagner (1978), cette dimension n'est nullement

mentionnée. Yukl (1983), quant à lui, présente deux facteurs séparés : compliments et reconnaissance ainsi que récompenses. Cependant, la justification de combiner ces deux aspects à l'intérieur d'une même dimension a été clairement démontrée dans une recherche subséquente sur les incidents critiques (Yukl & Clemence, 1984).

Pour Luthans et Lockwood (1984), la reconnaissance et les récompenses apparaissent comme des aspects de la motivation et renforcement. L'utilisation de renforcements positifs est effectivement une part de la motivation des efforts des subordonnés. Par contre, il y a de bonnes raisons de considérer ce comportement comme étant une dimension distincte. La reconnaissance et les récompenses sont des comportements résultant de la démonstration de compétence ou d'un comportement désirable du subordonné. La motivation, quant à elle, sous-tend une attente d'influencer l'employé à fournir le travail désiré. Ainsi, l'un est un comportement résultant d'une performance du subordonné (reconnaissance et récompenses) tandis que l'autre (la motivation) tend à induire une performance. De façon encore plus spécifique, ces deux comportements se différencient par leurs valeurs sous-jacentes. La motivation est clairement orientée sur la tâche, tandis que ce qui sous-tend la reconnaissance et les récompenses serait de nature plus relationnelle.

Harmonie et travail d'équipe (Harmonizing)

L'importance de cette coopération, de l'identification au travail de l'équipe et de la résolution constructive de conflits sont des thèmes centraux dans la littérature sur les organisations (Peters & Waterman, 1982). De plus, ces recherches fournissent d'importants éléments qui sous-tendent l'évidence que l'harmonie et le travail d'équipe est grandement relié à l'efficacité administrative.

Quatre des taxonomies étudiées dans cette recherche présentent des dimensions exclusives concernant l'harmonie, seule l'appellation varie. Stogdill (1963) nomme cette catégorie l'intégration, Bowers et Seashore (1966) l'appellent la facilitation à l'interaction, Luthans et Lockwood (1984) en font mention sous le terme de conflits administratifs. Dans la taxonomie développée par Yukl (1983), cette catégorie de comportements est divisée en deux dimensions : la facilitation à l'interaction et les conflits administratifs. Par contre, plusieurs exemples de comportements dans cette analyse factorielle sont grandement alourdis par de doubles appellations qui n'ajoutent pas nécessairement à la précision de la mesure, tel que le mentionne d'ailleurs une étude subséquente du même auteur (Yukl, 1987). Pour les autres taxonomies, l'importance de cette dimension est mise en évidence mais elle est incluse à l'intérieur d'autres dimensions. Par exemple, Morse et Wagner (1978) l'incorporent dans motivation et prise en charge des conflits, Mintzberg (1973) la définit à l'intérieur de la dimension du rôle du leader, etc.

Développement de carrière (Developping)

Les recherches empiriques concernant les effets d'un tel type d'encouragement sont limitées. Par contre, il est largement accepté, dans la littérature concernant la gestion, que le fait de soutenir le développement des aptitudes des subordonnés est grandement relié à la performance du gestionnaire et de son département (Yukl, 1987 ; Bradford & Cohen, 1984). D'ailleurs, la littérature concernant l'entraînement et la formation des subordonnés soutient cette assertion (Wexley, 1984).

Trois des taxonomies étudiées présentent des catégories directement équivalentes. Morse et Wagner (1978) désignent cette dimension sous le nom de « croissance et développement », Luthans et Lockwood (1984) l'appellent « entraînement et développement ». Quant à Yukl (1983), comme pour bien des catégories, elle est présentée en deux sous-dimensions : entraînement et « counseling de carrière ».

Dans trois des autres taxonomies, cette dimension est incluse à l'intérieur d'autres concepts. Pour Mintzberg (1973), elle fait partie du « rôle du leader », pour Page (1985), la « supervision » et pour Bowers et Seashore (1966), c'est une part de la dimension « facilitation du travail ».

La taxonomie de Mahoney et al. (1963), qui se veut une description du travail administratif, ne présente pas d'équivalent. Par contre, on retrouve certaines correspondances à l'intérieur des catégories évaluation et direction du personnel.

Représentation (Representing)

L'importance de la représentation pour l'efficacité du gestionnaire est supportée par des recherches sur les relations latérales du gestionnaire (Kanter, 1983 ; Kotter, 1982 ; Sayles, 1979 ; Stewart, 1967 ; Strauss, 1962 dans Yukl, 1987).

La catégorie « représentation » n'apparaît pas dans les taxonomies du leadership proposées par Bowers et Seashore (1966) et House et Mitchell (1974), puisque ceux-ci se sont intéressés au comportement du leader à l'égard de ses subordonnés.

Dans la plupart des autres taxonomies, on peut retrouver une équivalence. Yukl (1983), tout comme Page (1985) (cité dans Yukl, 1987), la mentionne telle quelle. Stogdill (1963) présente lui aussi une dimension de ce nom. Par contre, sa définition en est plus restreinte, puisque la dite dimension ne fait appel qu'aux comportements de représentation auprès d'autres entreprises (le délégué de l'entreprise, en fait).

La description du travail du gestionnaire de Mahoney et al. (1963) comporte deux catégories qui correspondent chacune à une partie : la « représentation » et la « négociation ». Pour Luthans et Lockwood (1984), tous les comportements du gestionnaire, incluant des relations avec l'extérieur, sont regroupés dans la catégorie « interaction avec l'extérieur » ; cette dimension est donc plus large que celle de la « représentation » comme telle.

Inter-relations départementales (Interfacing)

L'importance de construire une coopération entre les divers départements d'une entreprise est très bien démontrée dans une recherche de Kanter (1983) et Kotter (1982) (cités dans Yukl, 1987). De plus, Peters et Waterman démontrent l'importance d'être près des préoccupations des clients en étant en contact étroit avec les autres départements du début à la fin de la chaîne de production.

Certaines taxonomies mettent plus d'emphasis sur les relations avec les subordonnés au détriment des relations inter-départementales, c'est le cas pour les taxonomies de Bowers et Seashore (1966), House et Mitchell (1974) ainsi que Stogdill (1974). Par contre, des taxonomies telles que celles présentées par Morse et Wagner (1978), notent l'importance de cette catégorie de comportements, appelée « gestion des ressources de l'environnement », pour Page (1985) et Mahoney (1965), elle se nomme « coordination ».

Valeur de la taxonomie intégrative de Yukl (1987)

Comme il a été mentionné précédemment, la taxonomie intégrative de Yukl (1987) a largement été éprouvée. C'est une démarche de très grande envergure, inspirée de recherches rigoureuses (neuf taxonomies préalablement citées), desquelles Yukl a intégré toutes les données et les informations, pour ensuite entreprendre une recherche empirique importante. Les démarches entreprises afin de valider la taxonomie intégrative sont décrites ci-dessous (Yukl, 1987) :

Trois études subséquentes ont été menées afin de vérifier la validité de contenu. En 1984, en 1985 et en 1986 respectivement, 32, 24 et 36 étudiants au MBA ont servi de juges afin de classifier les comportements dans chacune des catégories correspondantes. La représentativité, la pertinence des catégories de comportements a été vérifié. On a demandé à 24 gestionnaires et à 130 de leurs subordonnés de déterminer l'importance de chacun des comportements nécessaires pour effectuer de façon efficace le travail de gestionnaire (1984).

De plus, le degré de consistance interne a été évalué. On doit s'attendre à ce qu'un item corrèle plus fortement avec un autre item de sa catégorie qu'avec un item d'une catégorie différente (Yukl, 1987). Le coefficient Alpha de Cronbach a été utilisé pour calculer la consistance interne pour trois échantillons de sujets. Une autre preuve de la valeur de la taxonomie est la démonstration que la mesure est stable dans le temps. Cette qualité est appelée fidélité. Celle-ci a été vérifiée

en administrant le questionnaire sur l'importance des comportements au même échantillon de sujets, en deux différentes occasions. La corrélation entre les résultats des répondants a été calculée en comparant les résultats de la première administration et ceux de la deuxième. Trois études de ce type ont été effectuées à différents intervalles de temps.

Cette taxonomie intégrative de Yukl (1987) résume les exigences de la gestion, les habiletés et les dimensions sous-jacentes à l'efficacité du gestionnaire. Ces dimensions permettront de s'assurer de couvrir l'univers du gestionnaire et le domaine de contenu de l'emploi à mesurer.

Dans le contexte d'un « in-basket », quatre à huit dimensions sont habituellement mesurées (Pettersen, 1995). Il n'est pas réaliste ni habituel d'en mesurer plus d'où la nécessité de ne retenir qu'un sous-ensemble (sous-domaine) du contenu. Ce sous-domaine n'aura évidemment pas la prétention d'être représentatif et de mesurer tout le contenu. Par contre, ce sous-domaine devra être une représentation du sous-domaine visé, soit les dimensions retenues.

Des dimensions comme la délégation, la planification et l'organisation, le contrôle managérial et le jugement peuvent être mesurées par un tel type d'instrument (Thornton, 1992). Un tableau de Byham et Byham (1976) présente le « in-basket » comme étant l'instrument, parmi les simulations, qui mesure le plus de dimensions. Les dimensions mesurées dans plus de 95% des « in-basket »

sont les suivantes : communication écrite, sensibilité, initiative, planification et organisation, délégation, contrôle managérial, analyse, jugement et esprit de décision.

Les dimensions de Yukl (1987) qui seront mesurées sont les suivantes : (1) information et communication, (2) résolution de problèmes et gestion de crises, (3) planification et organisation, (4) consultation et délégation, (5) clarification des rôles et objectifs, (6) suivi et contrôle. (7) harmonie et travail d'équipe et (8) inter-relations départementales. Ces dimensions qui font toutes appel à une habileté ou une compétence sous-jacente à la performance en gestion, constituent le sous-domaine de contenu visé par le « in-basket ».

Évidemment, ces compétences sont génériques et ont été définies par Yukl (1987) pour des postes de gestion en général. Or, le « in-basket » à élaborer vise une forme spécifique, soit la gestion de projet. Même si les compétences retenues sont suffisamment générales pour s'appliquer à la gestion de projet, ce contexte d'application est tout de même différent de la gestion en général et comporte des spécificités (voir « analyse et description de l'emploi »). Il faudra donc que ce contexte particulier soit représenté dans les items.

Afin de définir ce contexte particulier le « *Project Managerial Body of Knowledge* » (PMBOK) est tout à fait adéquat. Comme son nom l'indique, le

PMBOK définit un corpus de connaissances, ou de compétences à maîtriser pour prétendre au titre de « *Project Manager Professional* » (*PMP*).

Ce corpus de connaissances propre à la gestion de projet est constitué de: (1) la planification du projet, (2) la définition en séquences des différentes étapes du projet, (3) les délais, (4) le budget, (5) le contrôle de qualité, (6) les relations humaines, (7) la communication, (8) la gestion des risques et impondérables, et (9) les fournitures. Les définitions sont présentées en appendice afin d'alléger le texte (Appendice B).

Comme les dimensions sont représentées par la taxonomie intégrative de Yukl (1987), il importe d'éviter une certaine redondance. Les dimensions préalablement sélectionnées seront donc comparées aux catégories du *PMBOK* (tableau 3, ci-dessous).

Tableau 3

Comparaison entre la taxonomie proposée par le *PMBOK* (1996) et celle de Yukl (1987)

<i>PMBOK</i>	Yukl (1987)
(1) Planification	Planification et organisation
(2) Définition en séquences...	Planification et organisation
(7) Communication	Information et communication
(8) Gestion du risque et des imp.	Planif. et org., rés. de problèmes

Cette redondance entre les deux taxonomies confirme la pertinence de la taxonomie intégrative de Yukl (1987) pour une application en gestion de projet. Les cinq autres catégories présentées par le *PMBOK* constituent davantage des domaines d'intervention, ou domaine d'application de la gestion de projet, d'où leur utilité à définir le « contexte d'application ».

En effet, une dimension fait appel à une habileté sous-jacente à la mesure de la performance, tandis que les délais, le budget, la qualité, les ressources humaines et matérielles caractérisent, quant à eux, des contextes d'application.

La sélection des items implique la construction d'une table de spécification (Hopkins et Antes 1987). Ainsi une matrice sera formée comportant deux axes, en abscisse les dimensions et en ordonnée le contexte d'application propre à la gestion de projet. L'abscisse de la matrice ainsi formée présentera donc les huit dimensions retenues (provenant de la taxonomie intégrative de Yukl, 1987) afin d'être mesurées par le « in-basket ». En ordonnée, les domaines d'application proposées par le *PMBOK* (cinq) permettront de concrétiser les dimensions dans des domaines d'application (ou contextes d'application) représentatifs de la gestion de projet.

La matrice est composée de quarante cellules obtenues par le croisement des différentes dimensions de la gestion et des différents domaines d'application (tableau 4).

Tableau 4

Matrice (dimensions x domaines d'application)

Dimens. ---- Domaine	Inform. et Commun.	Résol. de prob. et gestion des crises	Planif. et Organ.	Consult. et Délég.	Clarific. des rôles et objectifs	Suivi et Contrôle	Harmonie et travail d'équipe	Inter- relations départ.
Délais								
Budget								
Humain								
Fournit. et mat.								
Qualité								

Néanmoins, il importe de retenir que les domaines d'application assurent la représentativité du contexte gestion de projet mais que les dimensions définissent le domaine de contenu à mesurer. Ainsi, les domaines d'application sont retenus afin d'assurer la représentativité des items dans un contexte de projet et permettront une distribution des items dans des objets d'application différents.

À cet égard, la matrice ainsi formée n'est pas véritablement une table de spécification puisque seules les dimensions seront mesurées. En effet, les domaines, tel que mentionné précédemment, sont des « objets » permettant de caractériser en quoi les dimensions s'appliquent. En résumé, ils sont utiles à la composition des items mais sont exempts de la mesure.

De plus, il importe de préciser que toutes les cellules de la matrice ne seront pas couvertes. En effet, la distribution des items assure que chaque dimension ainsi que chaque domaine d'application est couvert mais les dimensions ne s'appliquent pas nécessairement dans chacun des domaines.

4. Étape 4 : Conception de l'instrument dans son ensemble

Il s'agit ici de déterminer le format et la durée de l'instrument. Il a été mentionné précédemment que le test est une simulation de type papier-crayon. Afin de placer le sujet devant l'obligation de prioriser certains items, la durée de la passation ne devra pas permettre au candidat de terminer entièrement l'exercice. Déjà en 1937, des auteurs s'étaient penchés sur les qualités nécessaires à la gestion, Chester Barnard (cité dans Lopez, 1966) mentionne que la caractéristique essentielle d'un bon gestionnaire réside dans sa capacité décisionnelle. Selon lui, « l'art de la prise de décision consiste à ne pas décider des questions qui ne sont pas immédiatement pertinentes, ne pas décider prématurément, ne pas prendre de décisions qui ne peuvent être appliquées et ne pas prendre de décisions que d'autres peuvent prendre ».

La passation du test auprès d'un échantillon de gestionnaires de projet permettra de déterminer la durée nécessaire. Cette partie sera réalisée à posteriori, de même que la grille de correction.

5. Étape 5 : Création des items et échantillonnage

L'exercice « in-basket » remis aux candidats comprend quatre éléments : des consignes en vue de la passation, la description de la situation spécifique simulée, des documents de support et des items.

Les consignes ont pour but d'expliquer au candidat le déroulement de l'exercice ainsi que les différentes règles auxquelles il devra se soumettre. Les directives ne font pas partie de la section chronométrée du test.

Le test comme tel débute avec la description de la situation simulée. Les documents de support, quant à eux, viendront ajouter de l'information nécessaire sur la situation ainsi que des outils de travail (par exemple, le nom des divers employés, leur fonction respective, etc.). Les problèmes à résoudre, ou les items, sont présentés sous forme de lettres, de mémos, de rapports, d'appels téléphoniques (écrits). Au nombre de 15 à 30 les items doivent être distribués d'un même coup au candidat, soit en même temps que le contexte de la simulation (description de la situation simulée) et les documents de support. (Pettersen, 1991)

Les consignes de l'instrument, les instructions au candidat sont présentées en appendice (appendice C). Les instructions sont inspirées d'un modèle présenté par Pettersen (1991). De plus, plusieurs tests disponibles sur le marché ont été

étudiés permettant ainsi de s'assurer une certaine conformité avec les exercices « in-basket » les plus reconnus.

5.1. Contexte de la simulation et documents de support

Il s'agit ici de déterminer l'environnement, le type d'entreprise qui sera mise en situation. Il importe d'assurer une certaine généralisation. Par exemple, choisir un milieu où le projet serait trop spécifique, notamment un forage pourrait nuire à une utilisation plus étendue du test. Le projet doit donc comporter des stimuli variés afin d'être généralisable mais ne doit pas nécessiter des connaissances spécifiques.

Pour la rédaction des documents de support, il importe de bien définir l'entreprise, le projet, sa philosophie, ses objectifs, au sein desquels le candidat évoluera. Ce dernier pourra ainsi éclairer ses décisions par une meilleure connaissance de l'environnement. Les documents de support sont donc élaborés de façon concise mais explicite, afin de fournir une information préalable à la passation et, ainsi, permettre de maximiser la pertinence des décisions que le candidat pourra prendre lors de la simulation.

Étant donné ces considérations, le projet choisi est celui de la construction et du démarrage d'un Hôpital de Jour. Tout d'abord, le milieu hospitalier permettait une vaste étendue de possibilités quant aux items. De plus, ce milieu ne

demandait pas, de la part du gestionnaire, une expertise trop spécialisée. Le contenu des items devra être ajusté à ce niveau d'expertise et contribuera ainsi à la généralisation du test.

Afin de bien cadrer le milieu dans la simulation, des entrevues ont été effectuées avec le directeur général (M. Nagi Habashi) d'un hôpital en soins prolongés. Les démarches entreprises auprès de cet expert sont les suivantes :

- Explication du déroulement de la construction du test.
- Clarification des attentes.
- Validation de la table de spécification (Afin de s'assurer qu'elle soit effectivement représentative de la gestion de projet et du milieu simulé)

L'hôpital approché (Centre d'hébergement et de soins longues durées Frédérick Georges Hériot) projetait d'ouvrir un Hôpital de Jour. Il nous offrait ainsi la possibilité de représenter la réalité de façon encore plus concrète, en nous permettant d'utiliser directement un projet existant. Le directeur général a eu l'obligeance de permettre la consultation des documents concernant ce projet, soit le budget de l'hôpital, le rapport annuel, l'organigramme, la convention collective, les buts et objectifs de l'Hôpital de Jour ainsi que la planification du projet.

Le contexte de la simulation ainsi que les documents de support ont été élaborés à la lumière de ces informations. Ils sont présentés en appendice (Appendices D et E).

5.2. Rédaction des items

Des incidents critiques ont été utilisés afin de rédiger les items. Les incidents critiques sont définis comme étant des comportements qui, dans des circonstances données, se sont avérés particulièrement efficaces ou inefficaces. Cette méthode a été développée par Flanagan (1954). Elle est utilisée dans un contexte de cueillette de données pour l'analyse de poste, mais aussi afin de fournir les éléments pertinents au développement des items d'un test basé sur la validité de contenu (Tziner, Jeanrie et Cusson, 1993). Ils sont souvent produits sous forme d'anecdotes relatives à un emploi en particulier (Tziner, Jeanrie et Cusson, 1993). De plus, la notion de critique introduit le fait que cet incident a eu un impact, une conséquence directement observable telle qu'une baisse de la productivité ou un climat de travail tendu, par exemple. Un incident critique est donc tout à fait approprié à l'élaboration d'un instrument basé sur la validité de contenu. Cette cueillette d'information (les incidents critiques) doit se faire selon une procédure bien explicite afin d'être valable. Flanagan (1954) en définit les paramètres de la façon suivante :

- (1) La personne qui rapporte l'incident doit décrire la circonstance dans laquelle l'événement s'est produit.
- (2) Elle doit décrire précisément le comportement qui en a résulté
- (3) Elle doit exprimer clairement en quoi ce comportement était approprié ou non et pourquoi.

Les incidents critiques ont été amassés auprès des étudiants inscrits à la maîtrise en gestion de projet par Messieurs Pettersen et Grandmont, professeurs au département des Sciences de la Gestion et de l'Économie. Ces étudiants, pour être admis à la maîtrise, devaient avoir un minimum de deux ans d'expérience en gestion de projet, soit à titre de gestionnaire soit à titre de professionnel. Dans le cadre d'un cours, ils avaient pour mandat de composer dix incidents critiques, cinq positifs et cinq négatifs. Une banque de 285 incidents critiques a ainsi été recueillie. Les directives qui leurs ont été données sont présentées en appendice (voir Appendice F). On remarquera qu'aucune directive ne mentionne les dimensions sélectionnées. En effet, le sous-domaine de contenu n'avait pas encore été défini au moment de la collecte des incidents.

Une classification sommaire obtenue par la compilation de ces dits incidents critiques (285 incidents) selon leurs domaines respectifs a été effectuée. Les six domaines qui ressortent le plus dans la compilation des incidents critiques sont : (1) les délais, (2) le budget, (3) les liens avec la direction et les pairs, (4) les relations avec les subordonnés, (5) le matériel et les fournitures et (6) la planification. Certains autres domaines étaient représentés, mais le petit nombre d'incidents (15 et moins par domaine) ne permettait pas de les considérer comme étant des domaines critiques de la gestion de projet. Les définitions de chacun de ces domaines sont assez évidentes et, puisqu'elles sont redondantes avec les domaines du *PMBOK* (1996) préalablement définis, il apparaît peu utile de les définir à nouveau.

Le tableau 5 présente les concordances entre les domaines définis par le *PMBOK* et les incidents critiques amassés.

Tableau 5

Comparaison entre les domaines définis par le *PMBOK* (1996)
et la compilation des incidents critiques

<i>PMBOK</i> (1996)	Compilation des incidents critiques (285 incidents provenant de 30 questionnaires de projet)
Planification du projet Définition en séquences des différentes étapes du projet Délais Budget Contrôle de qualité Ressources humaines Communication Risques et impondérables Fournitures	Planification Délais Budget Liens avec la direction et les pairs Relations avec les subordonnés Matériel et fournitures

Cette comparaison est convaincante. En effet, il apparaît clairement que la plupart des domaines du *PMBOK* se recoupe avec les domaines déterminés à partir de la compilation des incidents critiques.

Cependant, comme peu d'incidents critiques portaient sur le contrôle de la qualité, ce domaine ne pourra apparaître dans le présent « in-basket ». Compte tenu du fait que les contextes d'application sont utilisés à titre représentatif et sont exempts de la mesure, refaire une cueillette d'incidents afin d'en amasser de nouveaux concernant le contrôle de la qualité en particulier, n'apparaît pas essentielle.

Une fois cette classification par domaine d'application effectuée, un nouveau classement a été fait afin de rédiger les items. Premièrement, les incidents critiques ont été classés selon leur niveau d'adéquation avec la situation simulée. En effet, seuls les incidents les plus adéquats ont été retenus, et ce, en tenant compte de la matrice préalablement définie. Comme les participants, ayant composé ces items provenaient de milieux différents (construction, mines, usines, centres hospitaliers, projets internationaux, etc.), plusieurs incidents n'étaient pas adéquats pour la mise en situation et ne pouvaient être transposés en items. En effet, certains étaient trop spécifiques à leurs domaines d'intervention ou comportaient des décisions techniques impliquant une connaissance spécialisée. Lors de ce premier classement, 70 incidents critiques ont été retenus.

Deuxièmement, ces 70 incidents critiques ont été classés en trois catégories :

(1) Incidents correspondant aux dimensions préalablement définies et les situations décrites peuvent être utilisées telles quelles dans la situation simulée.

Seuls les noms doivent être changés. Incidents à transposition directe, 32 items ont été classés dans cette catégorie.

(2) Incidents correspondant aux dimensions préalablement définies, mais les situations décrites sont légèrement différentes de la situation simulée. Des changements mineurs sont nécessaires afin d'ajuster les incidents à la situation simulée. Les rôles des différents intervenants de la situation doivent être changés. Transposition indirecte, 15 items ont été classés dans cette catégorie.

(3) Incidents correspondant aux dimensions préalablement définies, mais les situations présentées ne sont pas suffisamment complexes, la réponse étant implicite dans la situation décrite. Deux incidents peuvent être utilisés pour composer le même item afin d'enrichir le problème à résoudre ou la situation. Transposition indirecte, 23 items ont été classés dans cette catégorie.

Les incidents critiques de la première catégorie ont été utilisés en priorité. Il a été impossible de couvrir l'ensemble de la matrice seulement avec les incidents de cette catégorie. En effet, parmi les 32 items, plusieurs couvraient les mêmes cellules. Ainsi, 15 incidents critiques ont été transposés en items.

Les incidents critiques de la deuxième catégorie ont été distribués dans la matrice. Cinq de ces incidents ont été retenus en fonction de leur complémentarité, par rapport à la table de spécification, avec les items composés à partir des incidents de la première catégorie.

Vingt items sont ainsi construits. La dimension résolution de problèmes et résolution de crises est peu représentée. Les incidents critiques de la troisième catégorie ont donc été classés selon la dimension qu'ils mesuraient et deux incidents critiques ont été retenus afin de former un item mesurant la dimension résolution de problème.

Les items de l'instrument ainsi conçus sont présentés en appendice (Appendice G).

6. Étape 7 : Révision de la version expérimentale par des experts

Comme il a été précédemment mentionné, la justification d'un test basé sur la validité de contenu pour la sélection est le degré de similarité entre les items et le travail réel. La définition de l'univers à mesurer devra donc tenir compte de l'avis de spécialistes dans ce domaine (SIOP, 1987).

La démarche concernant la matrice et la transposition des incidents critiques en items a donc été validée auprès de trois spécialistes de la gestion de projet. Ces spécialistes possèdent tous un diplôme doctoral en gestion, enseignent dans diverses universités du réseau de l'Université du Québec et ont chacun cumulé vingt années d'expérience en gestion de projet.

Tout d'abord, ces trois experts ont fait la lecture des documents de support et des consignes et ont suggéré des correctifs afin d'améliorer la clarté et la représentativité de la situation simulée par rapport au domaine de contenu.

Puis, ils ont classé les items selon chacune des dimensions. Le pré-classement effectué lors de la composition des items ne leur a pas été remis afin de ne pas contaminer leur jugement. Ces résultats sont présentés au tableau 6 de la façon suivante : chacun des items en abscisse est représenté par un « X » vis à vis de la dimension (en ordonnée) qu'il mesure. L'on remarquera qu'un item peut mesurer plus d'une dimension. Il est à noter que les experts ont été unanimes concernant ce classement.

Ensuite, il leur a été demandé de juger de la clarté, de la représentativité générale par rapport au domaine de la gestion de projet, de l'ambiguïté et du degré de difficulté des items.

Il serait difficile et fastidieux d'énumérer ici les correctifs suggérés par les experts mais précisons que la version finale présentée en appendice (Appendice G) comprend ces corrections. Leurs commentaires ont été unanimes et assurent la représentativité des items par rapport à la gestion de projet. Une copie des consignes adressées aux experts est présentée en appendice (Appendice H).

Tableau 6

Distribution des items en fonction des dimensions

Dimens. ----- Items	Inform. et Commun.	Résol. de prob. et gestion des crises	Planif. et Organ.	Consult. et Déleg.	Clarific. des rôles et objectifs	Suivi et Contrôle	Harmonie et travail d'équipe	Inter- relations départ.
Item 1	X					X	X	
Item 2	X	X	X				X	
Item 3								X
Item 4	X		X		X			
Item 5						X		X
Item 6				X		X		
Item 7						X		
Item 8	X			X				X
Item 9	X				X		X	
Item 10			X		X		X	
Item 11				X		X		X
Item 12			X			X		X
Item 13			X					X
Item 14			X					X
Item 15			X		X		X	X
Item 16						X		
Item 17		X	X	X				X
Item 18			X					X
Item 19								
Item 20	X		X		X			
Item 21	X		X				X	X
n total = 53	7	2	11	4	5	7	6	11

X : Item représente la dimension en abscisse

CHAPITRE IV : RÉSULTATS CONCERNANT LA VALIDITÉ DE CONTENU DE L'INSTRUMENT

La démonstration de la validité de contenu de l'instrument repose sur la validité de ses composantes (Pettersen, 1997). Comme, il a été maintes fois répété tout au long de ce texte, la validité du test est assurée par la rigueur de son élaboration. Revoyons donc, à posteriori, la représentativité de chacune des composantes de l'instrument de mesure par rapport au domaine de contenu à mesurer.

Ensuite, les résultats d'une étude de validité proposée par Lawshe (1975) afin de vérifier la validité de contenu des items sont présentés et expliqués accompagnés des tableaux de données les plus utiles à la compréhension du lecteur.

1. Les composantes de l'instrument de mesure

1.1 *Les directives*

Le candidat est placé dans une situation fictive où il doit remplacer son prédécesseur et prendre des décisions avec de l'information restreinte, une nouvelle équipe de collègues et de subordonnés. Même si cette situation est une

fiction nécessaire à la mise en situation, elle est tout de même très semblable à la réalité d'un gestionnaire de projet. Celui-ci est appelé à travailler avec des équipes formées pour les besoins du projet, avec des délais restreints et, souvent, son rôle à l'intérieur de l'entreprise débute et se termine avec le projet, tout comme la situation fictive du test.

De plus, le « in-basket » devient d'autant plus approprié dans un contexte de gestion de projet, puisque les directives imposent de remplacer son prédécesseur de façon inopinée, d'être en relation avec du nouveau personnel et de vivre des problèmes sortis du contexte de l'entreprise en général avec en prime une contrainte de temps. Tout à fait les contraintes que peut rencontrer un gestionnaire de projet à cause du concept de projet en soi.

1.2 *Le mode de présentation des stimuli*

Il a été mis en évidence que les mémos, messages téléphoniques et autres (items du « in-basket ») étaient le médium le plus représentatif de la tâche du gestionnaire. Ces items provenant directement du panier de gestion, ils sont présentés de la même façon qu'ils apparaissent au gestionnaire, c'est-à-dire par écrit (notes de secrétaires, notes à l'agenda, notes de collègues, employés, supérieurs, etc.). Tziner, Jeanrie et Cusson (1993) mentionnent que quelle que soit la représentativité des items, s'ils ne sont pas présentés de la même façon qu'à l'emploi, le test sera moins valide. Ainsi, il importe de présenter des items

écrits, si nécessaire, ou des discussions de groupe si plus approprié. Le travail décisionnel que le gestionnaire fait à partir de son panier de gestion est une tâche solitaire. Le mode de présentation est donc adéquat. Par contre, il aura à communiquer ses décisions à son équipe mais, dans bien des cas, pas nécessairement par écrit (voir mode de présentation des réponses).

1.3 Mode utilisé pour les réponses

Évidemment, le gestionnaire prévoit ses interventions par écrit mais, subséquemment, il pourra avoir à les communiquer verbalement à son équipe, à prévoir des réunions ou des entrevues. Cette partie n'est pas directement mesurée par le test puisque les réponses du candidat sont recueillies par écrit. Ainsi, la qualité de ses interactions, le mode de communication qu'il établit ne peut faire partie de la mesure. D'ailleurs le domaine de contenu visé ne comprend ni la communication orale, ni les relations interpersonnelles. Il est possible de compléter le test par une entrevue où le candidat aura à débattre ses réponses. Mais l'élaboration et la validation d'une telle entrevue comme élément de mesure pourraient faire l'objet d'une recherche subséquente.

1.4 Les conditions d'administration

La prise de décision est au centre du « in-basket ». De plus, la contrainte de temps pousse le candidat à remettre à plus tard certaines décisions, à déléguer et à régler les situations urgentes. Cronbach (1970) mentionne qu'un test qui met

l'accent sur la rapidité d'exécution n'est justifié que si la vitesse est une dimension importante de l'emploi. Le test apparaît donc tout à fait approprié afin de déterminer le choix optimal d'un gestionnaire de projet, les contraintes du test étant tout à fait semblables aux contraintes de l'emploi.

1.5. Les items

Les processus d'élaboration de tests basés sur la validité de contenu posent un problème quant à la démonstration de cette dite validité, quand il s'agit d'utiliser ce processus de validation pour les tests de sélection. Lawshe (1975) est le seul à proposer un test afin de valider les instruments de sélection élaboré sur la base de la validité de contenu. Ce test consiste à questionner un échantillon d'individus étant à même de juger de l'emploi mesuré. En l'occurrence, des gestionnaires de projet (43 sujets), étudiants à la Maîtrise en gestion de projet ont répondu au test de Lawshe (1975) présentée ci-dessous :

Est-ce que l'habileté (ou connaissance) mesurée par cet item est

- Essentielle
- Utile mais non essentielle, ou
- Non nécessaire

à la performance à l'emploi.

Une copie des consignes remises aux 43 sujets est présentée à l'Appendice I.

Chacune des réponses des sujets ou candidats est compilée et le nombre d'individus ayant répondu « essentielle » sera additionnée pour chacun des items. Ce nombre (essentielle) moins le nombre de sujets total divisé par deux ($43/2$) sera ensuite divisé par le nombre total de sujets (43). La réponse obtenue est un coefficient de validité pour chaque item. Ainsi, lorsque la moitié des sujets indiquent « essentielle » et l'autre moitié « utile mais non essentielle » ou « non nécessaire », le coefficient est de zéro. Lorsque plus de la moitié des répondants indiquent « essentielle », le coefficient se situe entre zéro et un (0,00 et 1,00). À l'inverse, lorsque plus de la moitié des répondants indiquent « utile mais non essentielle » ou « non nécessaire », le coefficient se situe entre zéro et moins un (0,00 et -1,00).

$$\frac{n_{\text{individus ayant répondu « essentielle »}} - n_{\text{total}} / 2}{n_{\text{total}}}$$

Les données ainsi obtenues permettent, une fois compilées, de faire une moyenne des coefficients des items et ainsi déterminer le coefficient de validité du test en entier. Des tables sont fournies par Lawshe (1975), indiquant le niveau significatif du coefficient de validité avec une marge d'erreur de 5 pour cent (0,05). Pour quarante sujets, le coefficient doit être égal ou supérieur à 0,29 pour que l'item soit jugé valide.

Pour chacun des sujets, l'âge, le sexe, le nombre d'années d'expérience en gestion de projet ainsi que leur domaine d'expertise (l'ingénierie, la construction,

etc.) ont été compilés. (voir tableau 7). La moyenne du nombre d'années d'expérience des sujets en gestion est de 3,5 années. La moyenne du nombre d'années d'expérience des sujets en gestion de projet est de 3,3 années.

Pour chacun des items, les sujets ont répondu à la question de Lawshe (1975) (voir section 2.6). Le tableau 8 présente le ratio de validité de contenu ainsi obtenu. Plusieurs coefficients sont négatifs et peu se situent au-dessus du seuil de signification qui est de 0,29. En effet, seuls les items 1 et 4 seraient considérés comme valides selon le seuil de signification, leurs ratios étant respectivement de 0,32 et 0,41.

Ensuite, étant donné le faible coefficient de validité des items et la grande variance entre les répondants, une compilation des réponses des sujets séparés selon leurs domaines d'expertise, a été effectuée. Ainsi, on pourra possiblement réduire la variance entre les répondants. De plus, il est raisonnable d'avancer l'hypothèse que les ratios de validité de contenu soient plus élevés pour les domaines d'expertise semblables au contexte du « in-basket ». En effet, cette démarche vise à vérifier l'hypothèse que les faibles ratios de validité de contenu soient attribuables aux contradictions entre les sujets, contradictions qui pourraient être possiblement évitées en les regroupant par domaine d'expertise. De fait, les résultats de cette classification ont démontrés une tendance beaucoup plus forte, les sujets se positionnant de façon plus marquée dans le même sens, les ratios

sont plus tranchés. En effet, on retrouve des résultats de -1,00, 0,00, 0,50. De plus, pour certains domaines d'expertise, les coefficients de validité sont beaucoup plus élevés que les résultats de tous les sujets confondus. Par contre, compte tenu du nombre restreint de sujets, puisque séparés selon leurs domaine d'expertise, le seuil de signification doit être augmenté. Ceci nous ramène à la problématique de départ, soit très peu de résultats significatifs, en d'autres termes, très peu d'items jugés valides. Comme les résultats se présentent sur plusieurs tableaux, afin de ne pas alourdir le texte ces résultats sont présentés en appendice (appendice J).

Tableau 7
Caractéristiques des sujets

Répondants	Âge	Sexe	Expérience en gestion	Expérience en gestion de projets	Domaine d'expertise
1	33 ans	Masculin	7 ans	6 ans	Informatique
2	27 ans	Féminin	2 ans	1 an	Construction
3	26 ans	Féminin	0 an	0 an	Biologie environnementale
4	42 ans	Masculin	2 ans	0 an	Construction
5	25 ans	Féminin	0 an	0,5 an	Ingénierie
6	47 ans	Masculin	10 ans	0 an	Services
7	32 ans	Masculin	8 ans	0 an	Construction
8	33 ans	Féminin	5 ans	5 ans	Architecture - construction
9	24 ans	Féminin	0 an	0,5 an	Ingénierie
10	34 ans	Masculin	8 ans	0 an	Services financiers
11	34 ans	Masculin	0 an	3 ans	Ingénierie
12	28 ans	Féminin	5 ans	0 an	Non spécifié
13	24 ans	Féminin	0,5 an	1 an	Ingénierie
14	30 ans	Masculin	3 ans	3 ans	Génie municipal
15	41 ans	Masculin	20 ans	6 ans	Enseignement
16	27 ans	Masculin	10 ans	2 ans	Militaire
17	26 ans	Masculin	0 an	0 an	Non spécifié
18	32 ans	Masculin	4 ans	5 ans	Non spécifié
19	36 ans	Masculin	0 an	11 ans	Ingénierie
20	36 ans	Féminin	3 ans	0 an	Comptabilité
21	51 ans	Masculin	20 ans	20 ans	Enseignement, consultation
22	31 ans	Masculin	4 ans	3 ans	Ingénierie
23	39 ans	Masculin	0 an	12 ans	Construction
24	50 ans	Masculin	0 an	18 ans	Construction
25	38 ans	Féminin	3 ans	3 ans	Promotion, prévention
26	47 ans	Masculin	0 an	7 ans	Non spécifié
27	24 ans	Masculin	1 an	1 an	Ingénierie
28	31 ans	Masculin	1 an	2 ans	Ingénierie
29	33 ans	Masculin	0 an	0 an	Sécurité au travail
30	31 ans	Masculin	0 an	1 an	Protection civile
31	30 ans	Masculin	0 an	0 an	Éducation
32	34 ans	Masculin	3 ans	6 ans	Ingénierie
33	46 ans	Masculin	23 ans	3 ans	Santé
34	38 ans	Masculin	0 an	0 an	Ingénierie
35	27 ans	Masculin	4 ans	1 an	Construction - informatique
36	27 ans	Masculin	1 an	1 an	Informatique
37	40 ans	Féminin	0 an	0 an	Ingénierie
38	27 ans	Masculin	7 ans	7 ans	Ingénierie
39	27 ans	Féminin	1 an	1 an	Ingénierie
40	30 ans	Masculin	0 an	2 ans	Administration
41	28 ans	Masculin	5 ans	5 ans	Construction - informatique
42	33 ans	Féminin	0 an	0 an	Sciences sociales
43	30 ans	Féminin	3 ans	3 ans	Ingénierie

Tableau 8

Résultats de la compilation des réponses des 43 sujets

Coefficient de validité de contenu

Item 1	0,32	Item 12	0,23
Item 2	-0,41	Item 13	0,27
Item 3	-0,45	Item 14	-0,36
Item 4	0,41	Item 15	-0,05
Item 5	0,05	Item 16	-0,32
Item 6	0,00	Item 17	0,23
Item 7	-0,64	Item 18	-0,64
Item 8	-0,23	Item 19	-0,41
Item 9	-0,09	Item 20	-0,59
Item 10	-0,23	Item 21	-0,09
Item 11	-0,36		

Les items en caractères gras présentent des coefficients de validité de contenu au desuus du seuil de signification , soit 0,29.

CHAPITRE V : DISCUSSION

Le domaine de contenu de l'emploi a été circonscrit par une recension des écrits dans le domaine. De la taxonomie intégrative de Yukl (1987), huit dimensions ont été retenues. Ces dimensions composent l'abscisse de la matrice de l'univers de l'emploi à mesurer. Afin d'illustrer les items dans un contexte caractéristique de gestion de projet, les domaines du *PMBOK* ont été utilisés en ordonnée. Ils sont présents à titre indicatif afin de s'assurer de représenter le contexte particulier de la contrainte, temps, budget, etc., de la gestion de projet. Par contre, il importe de retenir que ce sont les dimensions, et non les domaines qui représentent le squelette de l'instrument. Cette matrice comporte 32 cellules et les items sont distribués à l'intérieur de cette dite matrice afin de couvrir l'ensemble de l'univers à mesurer.

Des rencontres ont été effectuées afin de créer une entreprise fictive, représentative de la réalité, ainsi que des documents de support expliquant le poste du candidat dans la mise en situation. Afin de représenter la gestion de projet de façon la plus générale possible, le contexte retenu est celui de la construction et de la mise en place des services d'un Hôpital de Jour rattaché à un centre hospitalier déjà existant. Ce contexte semble suffisamment général et

peut permettre l'inclusion d'items concernant les normes, les relations inter-départementales, la gestion du personnel, la construction et les services. Les termes techniques des hôpitaux ont été éliminés, autant que possible, afin que le gestionnaire ne soit pas limité par des connaissances autres que celles rattachées directement à son domaine.

Les items sont ensuite composés. Ils s'inscrivent dans le contexte de l'Hôpital de Jour et sont reliés aux documents de support. Les items ont été composés à partir d'incidents critiques (285 incidents) recueillis auprès de gestionnaires de projet ayant plus de deux années d'expérience dans le domaine. Vingt et un items ont été retenus afin d'être inclus dans le test.

Trois experts en gestion de projet, ayant vingt années d'expérience et plus, ont revu les documents de support, les items et la matrice afin de donner leur appréciation. Ils ont jugé de la pertinence et de la clarté des items et des documents de support. Ils ont permis d'identifier des pistes de réponse aux items, ont jugé de la représentativité et du degré de difficulté des items par rapport au domaine de contenu. De plus, afin de s'assurer que chacun des items mesure bien ce qu'il est sensé mesuré, ils ont classé les différents items en se positionnant sur la dimension mesurée par l'item.

Suite à cette démarche, les items ont été présentés à 43 gestionnaires de projet, afin que ceux-ci répondent au test de Lawshe (1975), soit juger si les items sont essentiels ou non afin de mesurer la performance à l'emploi. Après la compilation statistique, seuls deux items ont un coefficient de validité significatif selon la table de Lawshe (1975). Ce constat amène plusieurs questionnements, autant sur le procédé d'élaboration et la rigueur avec laquelle il a été suivi, que sur la pertinence de l'utilisation du test de Lawshe (1975) dans le contexte.

Premièrement le choix des dimensions est appuyé sur une recherche d'envergure (celle de Yukl, 1987). Deuxièmement des experts du domaine ont jugés les items et les ont classés selon les dimensions qu'ils mesurent, assurant ainsi qu'ils mesurent bien la dimension qu'ils sont sensés mesurer. Troisièmement, plusieurs sujets ont jugés de la validité des items, sans tenir compte des dimensions cette fois. Comme chaque étape est imbriquée dans la précédente, les résultats décevants des coefficients de validité de contenu (obtenus par le test de Lawshe) invalidant les items, ces résultats invalident-ils aussi la pertinence des dimensions retenues? Doit-on remettre en cause toute la démarche? Est-ce que les résultats de 43 sujets invalident tout le processus? Si oui, quels sont les correctifs suggérés ; sinon, que signifient ces résultats?

Possiblement parce que la gestion de projet est une nouvelle approche en gestion et implique une philosophie particulière, il semble qu'il n'y ait pas

nécessairement consensus sur ce qu'est la gestion de projet ou sur ce qui est essentiel afin de mesurer la performance des gestionnaires de projet.

En effet, si l'on compare les résultats des trois experts qui sont des formateurs en la matière, qui ont, eux aussi, répondu au questionnaire de Lawshe, ils s'accordent pour dire que les items en grande majorité sont valides. L'instrument a été construit afin d'être généralisable à tous les domaines de la gestion de projet, ce qui est donc congruent puisque des experts du domaine général s'accordent sur sa validité, soit les trois experts en formation.

Par contre, l'échantillon de quarante-trois sujets était formé d'étudiants en gestion de projet de diverses provenances (informatique, ingénierie, etc.). Ceux-ci considèrent 19 des 21 items de l'instrument comme étant non essentiels (ou invalides selon le coefficient de validité de Lawshe, 1975). Serait-ce que ces items ne correspondent pas à la spécialisation des sujets, en fait à leurs domaines respectifs? Les items sont-ils trop généraux pour être considérés comme essentiels à la mesure d'un domaine particulier?

De plus, il est intéressant de revenir sur la formulation de la question de Lawshe (1975). Le terme « essentiel » est très tranché. Si un candidat échouait un item considéré comme essentiel, cela supposerait-il qu'il serait inapte à occuper l'emploi? La Société Industrielle et Organisationnelle de Psychologie (1987)

mentionne que les aspects critiques de l'emploi doivent être représentés à l'intérieur de l'instrument. Quelle est la définition exacte de « critique »? Une habileté doit-elle être considérée comme essentielle à la mesure de la performance pour être critique? Flanagan (1954), qui a grandement contribué à populariser le terme « critique » par sa méthode des incidents critiques (méthode d'analyse d'emploi décrite à l'étape 3), le définit comme étant un élément qui a un effet important sur la réalisation des objectifs de l'emploi étudié. Cet effet peut être direct ou indirect, à court, moyen ou long terme, mais il doit être suffisamment important pour avoir un impact significatif (Pettersen, 1995). Afin de déterminer si l'impact est significatif, cela suppose l'application d'un jugement ou d'une échelle de valeur. Ce jugement, qui peut déjà être difficile à trancher, devient carrément arbitraire si l'on se penche sur les différences entre les termes « critique », « significatif », « important » ou « essentiel ». Doit-on supposer que « utile » est comparable à « important » ou que « critique » est synonyme « d'essentiel »? Ou encore, peut-être que « critique » se situe dans une zone entre « essentiel » et « utile ». Si l'on creuse les nuances, les résultats peuvent devenir difficiles à interpréter.

La formule du coefficient de validité de Lawshe (1975) tient compte uniquement de la cote « essentielle » :

$$\frac{n_{\text{individus ayant répondu « essentielle »}} - n_{\text{total}} / 2}{n_{\text{total}}}$$

Ainsi, un item où la moitié des sujets auraient répondu qu'il est essentiel et où l'autre moitié l'aurait jugé utile mais non essentiel obtiendrait un coefficient de validité de zéro.

Un autre item pourrait lui aussi avoir le coefficient zéro malgré que les répondants l'aient jugé tout à fait différemment. Par exemple, la moitié des sujets auraient pu indiquer essentiel, l'autre moitié non essentiel et l'item obtiendrait le coefficient zéro. Pourtant, le degré de validité de ces deux items devrait être différent. Peut-on considérer qu'un item jugé essentiel par la moitié des répondants et jugé utile par l'autre moitié des répondants, doive être retiré de l'instrument? Son coefficient de zéro indique que cela devrait pourtant être la cas.

Un autre aspect peut être questionné, étant donné que le cadre de référence, c'est-à-dire la matrice, n'est pas fourni au sujet. Le jugement « essentiel ou non » implique d'inférer l'habileté sous-jacente mesurée par l'item. On peut donc supposer que si le candidat perçoit à sa façon l'habileté mesurée, peut-être différemment de la dimension réellement évaluée, son jugement sur la validité de l'item peut être remise en cause. En effet, les candidats utilisés pour évaluer la validité de l'instrument n'ont pas un profil différent de candidats qui pourront être évalués par le test. Les candidats évalués n'obtiendront pas tous une note parfaite. Les erreurs de jugement qu'ils feraient supposent qu'ils avanceraient une stratégie de résolution de l'item qui serait inadéquate. Par exemple, si la dimension

mesurée par l'item est la planification et que le candidat délègue la tâche, il échouera l'item. On peut supposer que cet échec implique qu'il n'aura pas perçu la nécessité de planifier en réponse au problème présenté. Qu'arrive-t-il si ce même candidat évalue la validité de l'item? Peut-il le juger non essentiel dans sa dimension délégation alors qu'il mesurait la dimension planification? On peut sûrement en émettre l'hypothèse... À tout le moins, on peut se questionner sur l'adéquacité de l'échantillon. En effet, les sujets sont des étudiants, possédant chacun un minimum d'expérience mais ne sont pas des experts du domaine.

De plus, cette démarche entre en contradiction avec le *Equal Employment Opportunity Commission* (1978), qui mentionne que les experts qui jugeront de la validité d'un instrument devront être informés de toute la démarche de construction, des critères d'inclusion ou d'exclusion des items, du cadre de référence. Par contre, la logique de Lawshe (1975) est tout aussi compréhensible: faire table rase, placer les sujets devant les items uniquement, afin qu'ils jugent directement la validité de l'item, sans tenir compte de quoi que ce soit d'autre que les stimuli. Malgré sa logique, cette démarche est-elle réalisable? À tout le moins, le risque est grand, puisqu'un jugement négatif (un coefficient de validité faible) devient difficile à interpréter. On peut se demander, en toute logique si l'item est jugé non valide par rapport au cadre de référence trop spécialisé des sujets, par erreur ou autre. Cette difficulté d'interprétation complique aussi la possibilité d'apporter des ajustements, ne sachant pas à quel niveau l'item est déficient.

La variance élevée entre les répondants peut indiquer que la question est floue, qu'elle laisse une trop grande place à l'interprétation. La spécifier impliquerait peut-être de fournir le cadre de référence, c'est-à-dire la matrice. Mais cette spécification pourrait peut-être invalider le test de Lawshe, puisque sa prémisse est de recueillir le jugement sans a priori. Une autre explication à la variance pourrait-elle être que le concept de gestion de projet en lui-même est flou? Ou encore que le terme « essentiel » porte à confusion et devrait être clairement défini?

Avec la cueillette des données, il serait intéressant d'avoir des informations complémentaires auprès des sujets. Est-ce que le faible coefficient de validité d'un item remet en cause uniquement sa pertinence ou l'habileté qui était mesurée? En fait, est-ce le type de stimuli présenté pour mesurer le suivi et contrôle, par exemple, ou la pertinence de l'inclusion de la dimension elle-même qui est remise en question? Si c'est l'item qui est remis en question, est-il inadéquat dans toutes les dimensions qu'il mesure (quelquefois deux ou trois) ou seulement pour une des dimensions? Ces informations seraient utiles afin de restructurer les items déficients ou de les rejeter, le cas échéant. Faute de précisions, des coefficients de validité faibles impliquent de refaire la démarche depuis le début, ne sachant pas où est l'erreur. L'addition d'un questionnaire complémentaire pourrait permettre de recueillir ces informations. De plus, répondre à ce questionnaire exigerait de

la part du candidat de pousser sa réflexion, et ainsi, ajouter une plus grande rigueur à son jugement.

Il est malheureux, en fin de processus, de retrouver autant d'hypothèses, de questions. Par contre, ces questionnements ne pouvaient être soulevés sans faire la démarche puisque le test de Lawshe est peu utilisé.

Répondre à ces questionnements, ces hypothèses, impliquerait de refaire la démarche en la précisant. Ces incertitudes démontrent l'importance d'affiner le procédé d'élaboration, de préciser les étapes, d'ajouter des sous-étapes afin qu'une élaboration éventuelle ne présente pas toutes ces hypothèse d'interprétation. Évidemment, le faible coefficient de validité de l'instrument ne peut être attribué entièrement à la démarche. Par contre, comme la validité d'un instrument de sélection basé sur la validité de contenu est assise sur la rigueur de son élaboration, plus le procédé sera précis, exhaustif, exempt de marge de manoeuvre, plus la validité de l'instrument risque d'augmenter.

Recommandations au processus d'élaboration

Voici donc les recommandations au processus qui sont proposées compte tenu de la faiblesse des résultats. Les recommandations débutent à l'étape 4 et reprennent le fil conducteur à partir de cette étape afin de compléter la démarche proposée par Pettersen.

Phase II : Développement de l'instrument de mesure

Étape 4 : Validation de la matrice

Un minimum de trois experts devront juger de la représentativité de la matrice pour couvrir l'univers de l'emploi, et ce, de façon individuelle.

- Fournir un résumé de la démarche.
- Indiquer le type d'instrument de mesure et les limites de celui-ci.
- Présenter la matrice ainsi que la justification des choix de dimensions et domaines.
- Fournir les définitions correspondantes

Consignes :

- Prendre connaissance des documents
- Pour chaque dimension et domaine, indiquer : « essentiel, utile mais non essentiel ou non nécessaire à la mesure de la performance à l'emploi ».
- Ajouter toute dimension ou domaine qui vous semble manquant afin de couvrir l'ensemble de l'univers de contenu. Tenir compte des limites de l'instrument.

Toute dimension ou domaine qui sera retenu devra être approuvé unanimement par les juges. Dans le cas contraire, une deuxième évaluation sera prévue, où les juges échangeront sur leurs divergences et devront atteindre un consensus.

Avantages de la démarche :

Évite la difficulté d'interprétation des résultats, puisque chaque étape est validée séparément. Il n'est plus possible de se demander si c'est l'item ou l'habileté qu'il mesure qui est invalide. De plus, il apparaît plus intéressant de juger si la dimension est plus essentielle à la mesure que l'item, puisqu'une fois que la dimension aura été validée, la seule question concernant l'item sera de juger s'il évalue correctement la dimension qu'il est sensé mesurer.

Dans le cas où l'instrument serait considéré comme non valide à la fin du processus, la démarche n'aurait pas à être reprise du début, puisque chaque étape aura été assurée.

Les experts seront plus informés de la démarche puisqu'ils l'accompagneront étape par étape et seront ainsi plus à même de juger. Il sera possible de réajuster ou de réorienter le test à n'importe laquelle des étapes.

Étape 5 :

À noter que toutes les simulations ne comportent pas nécessairement des documents de support, cette étape est donc appropriée dans le contexte de l'élaboration d'un « in-basket » plus particulièrement.

a) Élaboration des documents de support

Compte tenu de la matrice, un environnement approprié sera choisi afin de

permettre la mesure de toutes les dimensions. Des rencontres avec des gens en poste permettront de créer une entreprise fictive la plus représentative possible de la réalité. Les documents de support devront inclure toutes les informations nécessaires à la résolution des problèmes (outre les connaissances que le candidat doit posséder).

b) Validation des documents de support

Les experts précédemment consultés devront juger de la clarté et de la pertinence des informations fournies dans les documents de support. Chacune des recommandation des experts sera ajoutée au texte.

Avantages de la démarche :

De même que pour l'étape précédente, la validation séparée de chacune des démarches entreprises permettra de préciser les erreurs et d'assurer chacun des paliers de l'élaboration.

De plus, cette validation évitera une démarche subséquente, puisque des résultats trop faibles lors de la passation pourraient indiquer une faiblesse au niveau des informations fournies ou encore un degré de difficulté trop élevé. Cette validation préalable assurera que les documents de support ne sont pas en cause.

Étape 6 : Cueillette des incidents critiques

Un échantillon de spécialistes du domaine mesuré, titulaire en poste et leurs supérieurs immédiats (si possible moitié-moitié) (Lawshe, 1975) devront être consultés afin de créer des incidents critiques en vue de la création des items.

- Fournir un résumé de la démarche préalablement entreprise et des buts de celle-ci.
- Indiquer le type d'instrument de mesure, le type de candidat auquel il s'adresse et les limites de celui-ci.
- Présenter la matrice définissant l'univers, soit les dimensions et les domaines (s'il y a lieu) ainsi que toutes les définitions correspondantes.
- Présenter les documents de support.
- Fournir la définition d'incidents critiques.

Consignes :

- Prendre connaissance des documents fournis.
- Composer le plus d'incidents critiques possible.
- Caractéristiques des incidents critiques :
ils sont pertinents, compte tenu du type d'instrument en développement ;
ils sont représentatifs de la matrice préalablement définie ;
la moitié des incidents a un impact positif et l'autre moitié a un impact négatif;

ces incidents se sont réellement produits, soit vous en avez vous-même vécu l'expérience, soit ils vous ont été rapportés; ces incidents peuvent être transposés dans l'entreprise fictive préalablement définie.

Avantages de la démarche :

Préciser la matrice, les documents de support et le type d'instrument élaboré aux candidats qui produiront les incidents critiques permettra d'éviter que des incidents soient inadéquats et doivent être éliminés. Lorsque certains incidents sont éliminés, seul le hasard permet de supposer que les incidents retenus sont représentatifs de l'ensemble des incidents recueillis. De plus, le fait que des supérieurs immédiats du poste concerné soient mis à contribution, assure que plusieurs angles de vision soient représentés dans les incidents critiques.

Étape 6 : Jugement des incidents critiques

Les incidents les plus intéressants seront sélectionnés par les experts préalablement consultés.

Avantages de la démarche :

Cette étape permettra d'éviter l'inadéquacité de certains incidents, de s'assurer qu'ils sont critiques. Cette étape pourrait être réalisée par le psychométricien mais le fait que ce soient les experts du domaine qui en fassent

la vérification devient en quelque sorte une étape préalable à la validation des items, assurant ainsi le moins de marge d'erreur possible.

En effet, premièrement, les sujets ayant composés les items ne sont pas reconnus comme des experts du domaine. Deuxièmement, un seul sujet est d'avis que cet incident est essentiel (celui qui l'a composé), alors que dans un test de validation comme celui de Lawshe, plusieurs sujets évaluent l'item. Cette étape de pré-validation, si l'on peut dire, assurera que la qualité des incidents critiques n'est pas en cause si certains items sont jugés invalides. Ainsi, seule la transposition des incidents critiques en items pourra être questionnée.

Étape 7 : Transposition des incidents critiques

Les incidents critiques seront transposés en items en fonction des documents de support et seront ainsi ajustés à l'entreprise fictive.

Avantages de la démarche :

Le fait que plusieurs items soient disponibles permettra une élimination qui pourra assurer l'inclusion des items les plus pertinents.

Étape 8 : Jugement des items par les experts

Afin de structurer le jugement des experts et de fournir des indications claires, deux méthodes sont proposées par Mussio et Smith (1979) : (1) jugement

de la pertinence, de l'exactitude et de la clarté des items, (2) jugement du degré de difficulté des items. Chacun des experts préalablement consultés devra juger des items individuellement.

1. Pertinence, exactitude et clarté des items
 - 1.1 Est-ce que l'information présentée dans l'item est pertinente par rapport aux tâches de l'emploi?
 - 1.2 Est-ce que la situation est présentée de façon juste?
 - 1.3 Que pensez-vous que puisse être la meilleure solution possible par rapport à la situation présentée (réponse)?
 - 1.4 Y a-t-il d'autres solutions de rechange que vous considèreriez comme acceptables?
 - 1.5 Quelles sont les solutions de rechange plausibles que vous considèreriez comme inacceptables?
 - 1.6 Est-ce que l'information est présentée clairement?
 - 1.7 Avez-vous des suggestions afin d'améliorer la méthode de présentation de l'item?
 - 1.8 Croyez-vous que la présentation puisse tromper ou confondre les candidats?

Les experts, travaillant individuellement, devront classer les items en trois catégories. La première catégorie comprendra les items inacceptables ou rejetés,

la deuxième catégorie sera formée des items acceptables ou retenus et la troisième catégorie comportera les items jugés acceptables mais nécessitant des corrections mineures. Tous les items conservés dans le premier groupe devront avoir été choisis unanimement par les juges. (Mussio et Smith, 1979)

Les items jugés acceptables (items du deuxième groupe) seront assignés dans les cellules adéquates de la matrice par les experts. En fonction de l'univers à couvrir, certains items de la troisième catégorie pourront être intégrés afin d'assurer de couvrir l'ensemble de la matrice. Ces items du troisième groupe ayant été retenus, des corrections devront être apportées, puis les experts devront les revoir afin de juger, compte tenu des ajustements, si ces items peuvent être classés dans la première catégorie. Ces jugements doivent faire l'unanimité. Dans le doute, l'item est rejeté ou corrigé une deuxième fois.

2. Degré de difficulté des items

Les items retenus dans la matrice seront cotés en fonction de leur degré d'importance par rapport au domaine à mesurer. Cette cote pourra être utilisée afin de déterminer la note de passage (Figure 2 à la page suivante).

Pertinence des items

SVP indiquer par un crochet (✓) votre évaluation de l'item dans la section appropriée

Numéro de l'item	Le candidat doit posséder cette connaissance ou cette habileté pour performer à l'emploi		Le candidat possédant cette connaissance ou cette habileté pourrait mieux performer qu'un employé moyen		Un candidat possédant cette connaissance ou habileté pourrait performer de façon supérieure
	1	1,5	2	2,5	3
Item no 1.					
Item no 2.					
Item no 3.					
Item no 4.					
Item no 5.					
Item no 6.					
Item no 7.					
etc.					

La moyenne des cotes de chacun des items sera compilée. Les items ayant une cote inférieure à 2 sont considérés comme des items de base tandis que les items ayant obtenu une cote de 2 et plus sont considérés comme avancés.

Avantages de la démarche :

La précision dans les consignes aux experts permet de centrer leur attention aux endroits désirés. De plus, la standardisation du questionnaire permet de recueillir des résultats quantifiables plutôt que seulement des commentaires qualitatifs. Ces résultats pourront plus facilement être comparés à une analyse de validation plus poussée (par exemple celle de Lawshe, 1975), auprès d'un plus grand nombre de sujets et ainsi permettre d'avancer des explications plus assurées si une différence importante est constatée dans les résultats.

En résumé :

Toutes les démarches subséquentes s'appuient sur la validité de l'étape qui la précède. Ainsi tous les paliers de décision sont assurés et multiples hypothèses sont éliminées lors de l'interprétation des résultats. De plus, toute la démarche est construite à partir de la matrice. Une fois celle-ci validée, elle est le canevas de toutes les étapes ultérieures. En fait, chaque étape devient un sous-ensemble de la matrice. La démarche se déroule donc du plus général au plus spécifique, et aucun élément est à l'extérieur des limites préalablement définies. Ainsi, en aucun

moment dans la démarche, une information pourra remettre en question des étapes préalables.

Possiblement que la réalisation de toutes ces étapes n'est pas toujours raisonnable ni même possible, compte tenu des contraintes de délai et de budget que les entreprises rencontrent lorsqu'elles élaborent des instruments de mesure. Par contre, tel que précédemment cité : « ces considérations pragmatiques ne peuvent en rien ébranler le fait indéniable que la valeur d'une conclusion ou d'une décision ne peut être meilleure que les données sur lesquelles elle est appuyée » (DeVellis, 1991). La validité de contenu de l'instrument est assurée par la rigueur de son élaboration, ainsi, ajouter de la rigueur, de la précision à un processus déjà bien construit, ne peut qu'améliorer la valeur de l'instrument ainsi élaboré.

CHAPITRE VI : CONCLUSION

L'objectif principal de la présente recherche était d'appliquer un processus d'élaboration d'instrument basé sur la validité de contenu et de proposer s'il y a lieu des ajustements, des recommandations. Cette recherche ne présentait pas d'hypothèse, exception faite de l'attente, plus ou moins sous-entendue, d'obtenir un test valide en suivant rigoureusement les étapes d'élaboration. Le sujet est large à couvrir et semé d'embûches; chaque étape d'élaboration suscite de nouvelles informations, une nouvelle documentation.

Au cours de cette démarche, plusieurs difficultés sont survenues. À mesure que l'instrument évoluait, des informations s'ajoutaient et, quelquefois, remettaient en question ou indiquaient que des réajustements devaient être appliqués dans les étapes précédentes. Cet état de fait a grandement compliqué l'élaboration puisque cela a impliqué plusieurs pas en avant et plusieurs pas en arrière. De plus, étant donné que les résultats du jugement des experts et ceux de la procédure de validation de Lawshe ont été somme toute contradictoires, il a été ardu de déterminer où se situaient les failles dans l'élaboration. Les recommandations au processus concernent surtout ces aspects. Assurer l'avancement à chaque étape et ainsi éviter tout recul par différentes vérifications pour mieux déterminer à quel niveau les ajustement sont nécessaires.

Effectivement, les résultats de l'analyse statistique ont été surprenants, voire alarmants. Chacune des étapes est abondamment documentée, s'appuie sur des auteurs, des experts reconnus et est ajustée pas à pas au domaine de contenu. On pouvait donc espérer un produit (soit l'instrument de mesure) représentatif du contenu à mesurer malgré qu'il semble évident en fin de processus que les experts auraient dû être consultés plus fréquemment, comme il est d'ailleurs mentionné dans la discussion (chapitre précédent). Néanmoins, les résultats du test de Lawshe invalident l'instrument.

Ces résultats décevants ont forcé une étude approfondie du processus d'élaboration, un questionnement sur la valeur de l'échantillon, un questionnement par rapport aux instructions qui leur ont été données et une remise en question de la valeur du procédé de validation de Lawshe dans le contexte.

Ce constat ramène au problème abordé par Lawshe (1975), soit la difficulté de démontrer objectivement la validité d'un instrument de sélection (basé sur la validité de contenu). Ce manque de moyen quant à la démonstration de la validité de contenu d'instruments de sélection demeure un domaine de recherche à investiguer, d'autant plus que de tels types d'instruments sont très utilisés.

Il serait intéressant d'appliquer les recommandations suggérées suite à cette recherche et vérifier ensuite la validité de contenu de l'instrument ainsi

conçu. Chaque étape étant en quelque sorte validée séparément, les résultats devraient être plus simples à analyser. Plusieurs hypothèses pourraient être mises à l'épreuve. Ces questionnements toujours en suspend sont déstabilisants en fin de processus mais n'est-ce pas le propre de la recherche que de susciter autant de questions que de réponses?

RÉFÉRENCES

- ADAMS, J.R., BARNDT, S.E. & MARTIN, M.D.** (1979). *Managing by project management*, États-Unis : Universal Technology Corporation.
- ANASTASI, A.** (1997). *Psychological Testing*, 7^e Edition, New York, Macmillan Publishing Company.
- BASS, B.M.** (1981). *Handbook of Leadership : Revised and expanded edition*. New York: Free Press.
- BENNIS, W.** (1989). *Why Leaders Can't Lead*, San Francisco ; Jossey Bass.
- BENTSON, C.** (1987). *Everything you always wanted to know about the In Basket test but didn't know who to ask*. Document présenté au 16^e « International Congress on the Assessment Center Method », Tampa, Floride.
- BORMAN, W.C.** (1982). *Validity of behavioral assessment for predicting recruiter performance*. *Journal of Applied Psychology*, 67, 1, 3-9.
- BOWERS, D.G., & SEASHORE, S.E.** (1966). *Predicting organizational effectiveness with a four-factor theory of leadership*. *Administrative Science Quarterly*, 11, 238-263.
- BRAY, D.W., GRANT, D.L.** (1966). *The assessment Center in the Measurement of Potential for Business Management*, *Psychological Monographs*, Vol. 80, No 17.
- BUCK, L.S.** (1975). *The use of a criterion-referenced tests in personnel selection: A summary status report*, (Technical Memorandum 75-6), Washington, DC: Civil Service Commission.
- BYHAM, W.C.** (1979). *Applying assessment center technology to a personnel system*. New Orleans.
- BYHAM, W.C., THORNTON, G.C.** (1970). *Assessment Centers: A New Aid in Management Selection*, *Studies in Personnel Psychology*, Vol. 2, No 2, pp 21-35.
- CARROLL, S.J.Jr. & GUILLEN, D.J.** (1987). *Are the classical management functions useful in describing managerial work?* *Academy of Management Review*, 12, 38-51.

- CLELAND, D. I. & KING, W. R.** (1988). *Project management handbook*. New York: Van Nostrand Reinhold. 997 pages.
- COOK, M.** (1988). *Personnel selection and productivity*, Chichester, Angleterre: John Wiley.
- CRONBACH, L.J.** (1970). *Essentials of psychological testing*, New York: Harper & Row. 650 p.
- DECLERC, R.P., EYMERY, P. & CRENER, M.A.** (1980). *Le management stratégique des projets*, Paris, France: Éditions Hommes et Techniques. 198
- DEVELLIS, R.F.** (1991). *Scale development*. Newbury Park, CA: Sage. 121 p.
- DUHAMEL, C.** (1992). *Le choix des tests de recrutement, guide du professionnel*, Ed. d'Organisation, Paris. 179 p.
- DUNCAN, W. R. & al.** (1996). *A Guide to the Project Management Body of Knowledge*. Project Management Institute, Upper Darby, United States.
- DUNETTE, M.** (1994). *Handbook of industrial and organizational psychology*. Palo Alto, Ca : Consulting Psychologist Press.
- FAYOL, H.** (1949). *General and industrial management*. London: Pitman.
- FLANAGAN, J.C.** (1954). *The critical incidents technique*, Psychological Bulletin, 51, 327-358.
- FREDERIKSEN, N., JENSEN, O. & BEATON, A.E.** (1972). *Prediction of Organizational Behavior*. Pergamon General Psychology Series, New York : Pergamon Press Inc.
- GATEWOOD, R.D. & FEILD, H.S.** (1994). *Human Ressource Selection*. The Dryden Press, Third Edition.
- GAUGLER, B.B., ROSENTHAL, D.B., THORNTON, G.C.III, & BENTSON, C.** (1987). *Meta-analysis of assessment center validity*. Journal of Applied Psychology, 72, 493-511.
- GOLDSTEIN, G. & HERSEN, M.** (1990). *Handbook of Psychological Assessment*, Second Edition. Pergamon General Psychology Series, New York: Pergamon Press Inc.
- GUION, R.M.** (1965). *Personnel testing*, New York: McGraw Hill.

- HEMPHILL, J.K., D.E. GRIFFITHS, FREDERIKSEN, N.** (1962). *Administrative performance and Personality*. New York: Teachers College Bureau of Publication, Columbia University.
- HOUSE, R.J., & MITCHELL, T.R.** (1974). *Path-goal theory of leadership*. *Contemporary Business*, 3, 81-98.
- HUCK, J.R.** (1977). The research base. In J.L. Moses and W.C. Byham (Eds.), *Applying the assessment center method*. New York: Pergamon.
- HUNTER, J.E. & HUNTER, R.F.** (1984). *Validity and utility of alternative predictors of job performance*. *Psychological Bulletin*, 96, 72-98.
- IRWIG, H.G.** (1979). *Project and Functional Managers in Construction: A Comparison Using a Critical Functions Approach*. Proceeding of the Eleventh Project Management Institute Seminar/Symposium, Atlanta, Georgia.
- ISENBERG, D.J.** (1986). *Thinking and managing: A verbal protocol analysis of managerial problem solving*. *Academy of Management Journal*, 29, 775-788
- KERLINGER, F.N.** (1973). *Foundations of behavioral research* (2 éd. rev.). New York: Holt, Rinehart and Winston.
- KLEIMAN, L.S. & FALEY, R.H.** (1985). *The implications of professional and legal guidelines for court decisions involving criterion-related validity: A review and analysis*. *Personnel Psychology*, 38, 803-833.
- KOMAKI, J.L.** (1986). *Toward effective supervision: An operant analysis and comparison of managers at work*. *Journal of Applied Psychology*, 71, 270-278.
- LAWSHE, C.H.** (1975). *A quantitative approach to content validity*, *Personnel Psychology*, 28, 563-575.
- LOPEZ, F.M.** (1966). *Evaluation executive decision making, the « in-basket » technique*, New York: American Management Association inc.
- LUTHANS, F. & LOCKWOOD, D.L.** (1984). *Toward an observation system for measuring leader behavior in natural settings*. In J.G. Hunt, D. Hosking, C.A. Schriesheim, & R. Stewart (Eds.), *Leaders and managers: International perspectives on managerial behavior and leadership*. New York: Pergamon Press.

- MAHONEY, T.A., JERDEE, T.H., & CARROLL, S.J.Jr.** (1965). *The jobs of management*. Industrial Relations, 4, 97-110.
- MÉNARD, P.** (1995). *La gestion de projet*, Ordre des Comptables Généraux licenciés du Québec. Montréal.
- MINTZBERG, H.** (1973). *The nature of managerial work*. New York: Harper & Row.
- MOREAU, A.** (1983). *Relation entre la performance à un test « In Basket » et le rendement au cours d'étudiants en administration scolaire*, Université du Québec à Trois-Rivières. Mémoire de maîtrise en éducation.
- MORSE, J.J., & WAGNER, F.R.** (1978). *Measuring the process of managerial effectiveness indicators*. Academy of Management Journal, 21, 23-35.
- MOSES, J.L. & BYHAM, W.C.** (1977). *Applying the Assessment Center Method*. Pergamon Press. New York. 310 pages.
- MUELLER, R.** (1979). *L'entreprise et la gestion de l'innovation*. Paris, Publi-Union. 266 pages.
- MUSSIO, S. J. & SMITH, M.K.** (1979). *Content validity : A procedural Manual*. Minneapolis, Civil Service Commission. 101 p.
- O'BRIEN, J. A.** (1975). *Les systèmes d'information de gestion, la perspective du gestionnaire utilisateur*. St-Laurent, Québec. Editions du Renouveau Pédagogique. 768 pages.
- PETERS, T. J.** (1993). *L'entreprise libérée : libération management*. Dunod, Paris.
- PETERS, T. J., & WATERMAN, R.H.Jr.** (1982). *In search of excellence : Lessons from America's best-run company*. New York : Harper & Row.
- PETTERSEN, N.** (1997). *Éléments de psychométries appliqués à la gestion des ressources humaines*. Département d'administration et d'économique, U.Q.T.R, non publié.
- PETTERSEN, N.** (1991). *Utilisation des mises en situation à des fins de sélection*. Université du Québec à Trois-Rivières, École Supérieure de Gestion, Casablanca, Maroc. 37 p.
- PETTERSEN, N.** (1991). *Selecting project managers: An integrated list of predictors*. Project Management Journal, Vol 22, no 2, pp 21-26.

- PETTERSEN, N.** (1991). *What do we know about the effective project manager?* International Journal of Project Management, Vol 9, no 2, pp 99-104.
- PETTERSEN, N.** (1989). *Que font les gestionnaires de projet? Tour d'horizon*, Gestion 2000 Management et prospective, no 5, octobre-novembre, pp 87-108.
- ROBERTSON, I.T. & ILES, P.A.** (1988). *Approches to managerial selection*, In C.I. Cooper & I. Robertson (éd.). International Review of Industrial and Organizational Psychology. Chichester, Angleterre: John Wiley.
- RODGER, D.** (1988). *Activity tests*. Elizabeth Sidney (éd.), Managing Recrutement (4e éd. rev.), pp. 172-187. Aldershot, Angleterre: Gower.
- SACKETT, P.R., & DREHER, G.F.** (1982). *Constructs and assessment center dimensions: Some troubling empirical findings*. Journal of Applied Psychology, 67, 401-410.
- SACKETT, P.R., & WILSON, M.A.** (1982). *Factors affecting consensus judgment process in managerial assessment centers*. Journal of Applied Psychology, No 67, Vol 1, pp 10-17
- SERIEYX, H.** (1993). *Le big bang des organisations : quand l'entreprise, l'État, les régions entrent en mutation*. Paris : Calmann-Lévy. 342 p.
- SCHMITT, N.** (1977). *Interrater agreement and dimensionality and combination of assessment center judgments*. Journal of Applied Psychology, 62, 171-176.
- SCHMITT, N., NOE, R.A., MERITT, R. & FITZGERALD, M.P.** (1984). *Validity of assessment center ratings for prediction of performance ratings and school climate of school administrators*. Journal of Applied Psychology, 69, 207-213.
- SNYDER, N. & GLUECK, W.F.** (1980). *How managers plan: The analysis of manager's activities*. Long Range Planning, 13, 70-76.
- SOCIETY FOR INDUSTRIAL AND ORGANIZATIONAL PSYCHOLOGY (SIOP).** (1987). *Principles for validation and use of personnel selection procedures* (3 éd. rév.). College Park, MD.
- SPITZ, C.J.** (1982). *The project leader: a study of task requirements, management skills and personal style*. Thèse de doctorat. Western University.

- STOGDILL, R.M.** (1963). *Manual for the Leader Behavior Description Questionnaire - Form XII*. Columbus: Ohio State University, Business Research.
- STOGDILL, R.M.** (1974). *Handbook of leadership: A survey of the literature*. New York: Free Press.
- TAPERNOUX, F.** (1984). *Les Centres d'évaluation, Assessment Center*. Ed. Payot, Lausanne. 470 p.
- THORNTON G.C., & BYHAM W.C.** (1982). *Assessment centers and managerial performance*. NY : Academic Press.
- THORNTON, G.C.** (1992). *Assessment Centers in Human Resource Management*. Addison-Wesley Publishing Company.
- TOFLER, A.** (1991). *Les nouveaux pouvoirs : savoir, richesse et violence à la veille du XXI^e siècle*. Paris : Fayard. 658 pages.
- TIFFIN, J. & McCORMICK, E. J.** (1967). *Psychologie Industrielle*. Paris : Presses Universelles de France.
- TOFLER, A.** (1988). *Le chaos management*. Paris : InterEditions. 610 pages.
- TURNAGE, J.L. & MUCHINSKY, P.M.** (1984). *A comparison of predictive validity of assessment evaluation versus traditional measures in forecasting supervisory job performance: Interpretive implications criterion distortion for the assessment paradigm*. Journal of Applied Psychology, 69, 595-602.
- TZINER, A.** (1990). *Organization staffing and work adjustment*. New York: Praeger.
- TZINER, A., & DOLAN, S.** (1982). *Validity of an assessment center for identifying future female officers in the military*. Journal of Applied Psychology, 67, 728-736.
- TZINER, A., JEANRIE, C. & CUSSON, S.** (1993). *La sélection du personnel : concepts et applications*. Laval : Agence d'Arc. 227 pages.
- URWICK, L.F.** (1952). *Notes on the theory of organization*. New York: American Management Association.
- WERTHER, W., DAVIS, K., LEE-GOSSELIN, H.** (1990). *La gestion des ressources humaines*, Ed. McGraw Hill, Montréal, p. 232.

- WEXLEY, K.N.** (1984). *Personnel training*. In M. R. Rosenweig and L.W. Porter (Eds.), Annual review of psychology. Palo Alto, CA: Annual Reviews.
- YUKL, G.A.** (1987). *A new taxonomy for integrating diverse perspectives on managerial behavior*. Working Paper, SUNYA.
- YUKL, G.A.** (1987). *Development of a new measure of managerial behavior*. Paper presented at the Eastern Academy of Management Meetings, Boston.
- YUKL, G.A.** (1983). *Development of a taxonomy of managerial behavior by factor analysis*. Paper presented at Society for Organizational Behavior Annual Meeting, Minneapolis, Minnesota.
- YUKL, G.A.** (1981). *Leadership in organizations*, Englewood Cliffs, NJ: Prentice-Hall.
- YUKL, G.A. & CLEMENCE, J.** (1984). *A test of path-goal theory of leadership using questionnaire and diary measures of behavior*. Proceedings of the twenty-first annual meeting of the Eastern Academy of Management.
- YUKL, G.A. & NEMEROFF, W.** (1979). *Identification and measurement of specific categories of leadership behavior: A progress report*. In J.G. Hunt and L.L. Larson (Eds.), Crosscurrents in Leadership. Carbondale: Southern Illinois University Press.
- YUKL, G.A. & VAN FLEET, D.** (1981). *Cross-situational, Multi-method research on military leader effectiveness*. Organizational Behavior and Human Performance. 30, 87-108.

APPENDICE A

Définitions des dimensions (Yuki, 1987)

APPENDICE A

Définitions des dimensions (Yukl, 1987)

Information et communication :

Capacité de divulguer les informations pertinentes à ses subordonnés et à ses collègues. Le gestionnaire saura faire circuler les informations nécessaires à son équipe, afin que celle-ci puisse effectuer adéquatement son travail. Cette capacité de communiquer peut prendre plusieurs formes : répondre clairement à une question, téléphoner afin de divulguer de l'information, convoquer une réunion afin de mettre l'équipe à jour au sujet des nouveaux développements, envoyer des messages électroniques, faire circuler des lettres ou mémos, etc.

Résolution de problèmes et gestion de crises:

La résolution de problème, tout comme la planification et l'organisation, implique de prendre connaissance de la problématique, d'analyser les causes et de décider d'un plan d'action (Isenberg, 1986). Cependant, il y a une différence importante entre ces deux dimensions, en ce sens que la résolution de problème fait référence à une réponse à une perturbation du travail normal, telles un bris de matériel, un manque de fourniture, une plainte d'un client, un accident, etc., tandis que la planification sous-entend une anticipation dans le temps. Ainsi, la résolution de problème implique une décision à court terme afin de permettre au fonctionnement normal de reprendre son cours. Par contre, tel qu'il est expliqué

ci-dessous, la planification sous-entend une action à long terme et une anticipation des problèmes à venir.

Planification et organisation :

Cette catégorie de comportements inclut des décisions administratives au sujet des objectifs, des priorités, des stratégies, de la structure, de la distribution des ressources et des responsabilités et de l'horaire des activités. En d'autres mots, elle signifie décider quoi faire, comment le faire, qui doit le faire et quand cela doit être fait. Le but de la planification et de l'organisation est donc d'assurer l'efficience et l'efficacité du travail de l'unité ou du département (Yukl, 1987).

Comme il a été mentionné précédemment, il est important de distinguer l'implication à court terme de la résolution de problèmes, de l'implication à plus long terme, plus visionnaire de la planification et de l'organisation. Par contre, cette distinction est plus théorique que pratique. Elle permet de mieux comprendre et mesurer la performance administrative. Mais, dans la réalité, ces deux catégories de comportements se situent sur un continuum, puisqu'il y a plusieurs similitudes dans le processus cognitif et décisionnel qui est impliqué (Yukl, 1987).

Il y a plusieurs variétés de planification incluses dans les responsabilités du gestionnaire. Notons entre autres la planification stratégique (politiques et objectifs), la planification opérationnelle ou planification d'actions (développement

d'étapes ou d'horaires afin d'implanter un changement). La dimension planification et organisation comprend évidemment le pôle organisation, qui implique la distribution des ressources selon les diverses activités en cours, et en fonction de leur importance respective.

Finalement, la dimension planification et organisation suppose le développement de procédures afin de répondre aux problèmes potentiels ou crises à venir, d'où l'aspect visionnaire qui permet de prévoir les problèmes à venir plutôt que de réagir cas par cas.

Consultation et délégation :

Cette catégorie implique que le gestionnaire encourage la participation de ses subordonnés dans la prise de décision. Plusieurs degrés de participation sont possibles: vérifier une opinion auprès de ses subordonnés, consulter avant de prendre une décision, demander à un individu ou à un groupe de prendre une décision selon certaines limites spécifiques, etc. Une caractéristique de cette capacité de la part du gestionnaire est de stimuler l'intérêt et le défi chez les subordonnés afin de favoriser leur développement. De plus, ce type de gestion peut dégager le gestionnaire, en partie, du poids de ses responsabilités et ainsi, le rendre plus efficace.

Clarification des rôles et objectifs :

Cette catégorie implique une compréhension claire des responsabilités reliées aux différents postes, des tâches de chacun et de la performance attendue et elle sous-entend la communication claire de ces attentes. L'objectif de ce comportement est de coordonner les activités des différents membres de l'équipe, sous la responsabilité du gestionnaire. En contraste avec l'aspect motivation, qui sous-entend d'énergiser un subordonné, l'aspect de la clarification des rôles est un comportement d'orientation et de guide. Cette dimension de la performance du gestionnaire peut prendre différents aspects, de la simple directive donnée, à la rencontre plus formelle afin de formuler les attentes en terme de performance et de responsabilités.

Comme la dimension information, ce comportement implique la communication auprès des subordonnés. Cependant, la clarification des rôles implique une directive claire à propos des attentes, tandis que l'aspect information ferait plus référence à la communication d'informations pertinentes afin de faciliter le travail du subordonné. Cette distinction n'implique pas que ces deux comportements ne soient pas reliés. Quelquefois, un incident pourra mesurer un seul aspect (une dimension), tandis qu'un autre problème pourra impliquer les deux comportements. L'importance de distinguer ces deux comportements dépend du chevauchement induit par les problèmes présentés. Lors de la cueillette des incidents critiques de Yukl et Van Fleet (1982), seulement huit pour-cent (8%) des

incidents impliquaient une redondance dans les deux catégories. Ces résultats suggèrent donc l'attrait de les considérer séparément dans la mesure. Par contre, il est important de considérer cet aspect dans le choix des items, afin d'assurer l'indépendance de ces deux dimensions dans la mesure.

Suivi et contrôle :

Ce comportement implique d'être informé des activités de chacun, des progrès et de la qualité du travail. Le suivi peut prendre différentes formes telles la surveillance, la rédaction de rapports de rendement, les rencontres avec les subordonnés afin de vérifier les progrès, l'inspection de la qualité du travail, l'évaluation de la progression du projet auprès de clients ou sous-traitants, etc. Le premier objectif du suivi et contrôle est de maintenir la stabilité des opérations et de faciliter les ajustements en fonction du déroulement du projet. Cette dualité entre la surveillance et l'ajustement est souvent présentée dans la littérature sous l'appellation simple de contrôle. Par contre, cette dualité est plus présente sous l'expression suivi et contrôle.

Motivation :

Ce comportement implique l'utilisation de diverses techniques d'influence auprès des subordonnés afin de générer de l'enthousiasme au travail, une bonne coopération vis-à-vis des ordres et requêtes, ainsi que l'accord face aux objectifs des diverses tâches. La première conception de cette dimension était le

« leadership ». Ce concept a fourni multiples preuves de son importance dans l'efficacité administrative, comme l'ont mentionné Kipnis (1976) et Kotter (1985) (Yukl, 1981). Dans des recherches plus récentes, on inclut au concept de leadership tout l'aspect charismatique de la motivation (Bass, 1985 ; House, Schein, 1985 ; Sashkin, 1986 ; Tichy & Devanna, 1986 ; tous cités dans Yukl, 1987).

Support :

Le support inclut une vaste gamme de comportements par lesquels le gestionnaire démontre sa considération, son acceptation et sa préoccupation des besoins et sentiments des autres. Les comportements de support ont pour but de construire et de maintenir de bonnes relations interpersonnelles. Un gestionnaire amical a plus de chances de susciter la coopération, l'amitié et la loyauté. Un second objectif du comportement de support est d'induire une plus grande satisfaction au travail pour les subordonnés et les collègues.

Reconnaissance et récompenses :

Cette catégorie inclut toute récompense directe, telle une augmentation de salaire, ainsi que les récompenses moins tangibles, comme la reconnaissance en public et les compliments. Cette dimension sous-tend une double intention : moduler le comportement des subordonnés et induire une plus grande satisfaction pour ceux-ci. Les renforcements positifs ont fait l'objet de recherches extensives

depuis plusieurs décennies en psychologie. En effet, la littérature concernant le leadership a révélé un intérêt grandissant pour ce type de renforcements (Greene, 1976 ; Oldham, 1976 ; Podsakoff & Todor, 1985 ; Sims & Szilagyi, 1975 ; Yukl & Kanuk, 1978 ; tous cités dans Yukl, 1987). De récentes recherches sur l'efficacité de certaines organisations américaines ont fourni la preuve de l'importance de la reconnaissance et des récompenses appropriées afin de stimuler le travail des employés (Peters & Waterman, 1982).

Harmonie et travail d'équipe :

Cette catégorie inclut une large variété de comportements, comme le développement du travail d'équipe, l'identification de l'employé par rapport à son unité de travail, ainsi que la résolution constructive des conflits. Le premier objectif de ce comportement est de bâtir et de maintenir des liens au sein de l'équipe de travail, autant entre les pairs que vis-à-vis les supérieurs. Le deuxième but consiste à établir, favoriser une cohésion dans le travail de l'équipe de même qu'une identification comme membre important au sein de cette dite équipe.

Développement de carrière :

Cette dimension de comportement du gestionnaire prépare les divers éléments de son équipe à développer des traits performants chez eux, afin de viser d'éventuelles possibilités d'avancement dans l'entreprise. Cela peut prendre différentes formes. Par exemple, cela peut inclure de critiquer de façon

constructive des déficiences au niveau de la performance attendue, d'encourager l'individu à développer certains traits positifs, de proposer un entraînement, une formation ou une spécialisation, d'encourager les ambitions personnelles, de servir de mentor, de déléguer des responsabilités favorisant le développement, etc.

Cette dimension pourrait avoir des consonances communes avec la catégorie clarification des rôles. Le principal facteur qui différencie ces deux catégories concerne la perspective de temps. Le développement de carrière fait référence au long terme, à l'encouragement vers des buts qui s'inscrivent dans un continuum, tandis que la clarification des rôles a pour but une efficacité à plus court terme. Cependant, ces deux catégories ne doivent pas être perçues comme dichotomiques mais plutôt comme faisant partie du même type de comportement variant sur un continuum de temps. La distinction entre ces deux comportements est donc bien claire aux deux extrémités du continuum mais devient plus ambiguë dans une perspective plus intermédiaire. Par exemple, l'encouragement et l'encadrement de certains traits utiles au travail présent peuvent aussi soutenir des perspectives à plus long terme, reliées à d'éventuelles possibilités d'avancement. Il est donc très important, afin d'éviter un chevauchement dans la mesure, que les items soient soigneusement choisis de façon à maintenir la distinction entre ces deux dimensions.

Représentation :

Cette catégorie inclut plusieurs comportements visant à promouvoir et à défendre les intérêts du département ou de l'organisation auprès d'autres entreprises ou d'autres unités administratives. Par exemple, la représentation inclut des actions comme : être le représentant pour le département, négocier avec des concurrents ou administrateurs d'autres unités, être en relation avec des administrateurs plus élevés dans l'échelon hiérarchique afin de susciter du support ou des ressources financières, recruter et interviewer de nouveaux candidats pour l'équipe, être l'ambassadeur de l'entreprise auprès de visiteurs, etc.

Inter-relations départementales :

Comme la représentation, la catégorie inter-relations départementales implique des relations avec l'extérieur de l'équipe immédiate. De plus, cette catégorie inclut, elle aussi, le mandat de promouvoir et de défendre les intérêts du département. Cependant, le gestionnaire est ici préoccupé par l'intégration du travail de son équipe à l'intérieur de l'organisation entière. Il s'agit donc de relations à l'intérieur de son entreprise, mais en lien avec les autres départements. Par exemple, il peut s'agir de développer des relations de coopération avec d'autres gestionnaires de l'entreprise, de développer des sources d'informations à propos d'événements externes, d'organiser des réunions avec des clients afin d'améliorer leur satisfaction, de réunir des gestionnaires d'autres départements afin de mieux coordonner le travail d'ensemble, de rencontrer les dirigeants afin de

mieux concilier les buts de l'unité et ceux de l'organisation dans son ensemble, etc.

Certains de ces comportements se retrouvent dans des catégories comme: l'information, la consultation, le suivi et le contrôle, l'harmonie, la planification et la résolution de problèmes. Par contre, même si l'on peut retrouver certaines similitudes, l'objectif qui sous-tend le comportement est différent.

APPENDICE B

Les catégories de comportements du *PMBOK*

APPENDICE B

Les catégories de comportements du *PMBOK*

(1) La planification du projet

S'assurer que les divers éléments du projet sont bien coordonnés, planifier le développement du projet ainsi que son exécution en détaillant les différentes activités, coordonner les changements à l'intérieur des différentes étapes du projet.

(2) La définition en séquences des différentes étapes du projet

Cette dimension de la gestion de projet implique la capacité de subdiviser le projet en petites étapes plus faciles à gérer. Cette étape permet de mesurer l'étendue et les limites du projet.

(3) Les délais

Ce domaine implique des ajustements dans le temps et les effectifs afin de faire face aux impondérables qui pourraient venir interférer dans les délais établis.

(4) Le budget

Cela implique une planification des ressources nécessaires à la réalisation du projet (personnel, équipement, etc.) ainsi que tous les réajustements budgétaires en cours de projet.

(5) Le contrôle de qualité

Il importe d'identifier les standards de qualité adéquats pour la réalisation du projet. Cette dimension implique de vérifier les résultats spécifiques afin de déterminer s'ils sont à la hauteur des standards de qualité préalablement établis et d'identifier les voies d'ajustement afin d'éliminer les causes de performances insuffisantes.

(6) Les ressources humaines

Toute intervention dans des domaines tels : assigner les rôles, les responsabilités et la hiérarchie pour le personnel nécessaire à la réalisation du projet, favoriser le développement des attitudes personnelles et collectives qui permettent de maximiser le rendement.

(7) La communication

Déterminer le type d'information nécessaire à chaque personne ressource à l'intérieur du projet : qui a besoin de telle information? Quand en a-t-il besoin? Comment doit-elle lui être communiquée? Cette dimension inclut la capacité de rédiger des rapports faisant état du stade d'avancement du projet, de la mesure des progrès et des prévisions. De plus, il importe de distribuer l'information adéquate à chaque phase du projet. Cet aspect peut comporter la capacité de communiquer avec les médias afin de fournir des informations à l'externe et de faire de la publicité.

(8) La gestion du risque et des impondérables

Capacité d'identifier quels risques peuvent affecter le projet et être apte à se procurer l'information nécessaire afin d'y faire face. Évaluer le risque et les interactions afin de mesurer le niveau possible d'imprévus à considérer dans le projet. Répondre aux changements et imprévus à l'intérieur du processus d'avancement du projet.

(9) Les fournitures - matériel

Déterminer les fournitures nécessaires à la réalisation de chacune des étapes du projet. Se documenter sur les produits nécessaires et identifier les fournisseurs potentiels. Obtenir des cotations, des soumissions des fournisseurs éventuels, établir des relations adéquates avec le fournisseur qui aura été retenu. Enfin, s'assurer que les livraisons sont effectuées dans les délais et adapter en fonction des difficultés éventuelles.

APPENDICE F

Le protocole de cueillette des incidents critiques

APPENDICE F

Le protocole de cueillette des incidents critiques

Tel que présenté aux gestionnaires de projet lors de la cueillette des incidents

(Pettersen et Grandmont, 1987)

- A. Introduction (inspirée de Flanagan, 1954, p.340-1)
 - 1. Présentation des chercheurs et du cadre de la recherche
 - a) Commandités par le département d'Administration et d'économique de l'U.Q.T.R., ainsi que par la commission PRETAGEC, nous sommes actuellement à faire une recherche sur le gestionnaire de projets au Québec afin d'alimenter la mise à jour de notre programme de formation en gestion de projets. En effet, le comité réseau se questionne quant au profil que devrait présenter le candidat à la fin de sa maîtrise en gestion de projets.
 - b) Cette étude effectuée par Yves Grandmont et Normand Pettersen, professeurs au Département d'Administration et d'économique de l'U.Q.T.R., avec la participation de Hubert Wallot, directeur de la maîtrise en gestion de projets à l'UQAC.
 - 2. Objectif de la recherche
 - a) L'objectif de cette étude est d'analyser les activités du gestionnaire de projets afin d'identifier les exigences que comporte la gestion de projets en terme d'habiletés, de connaissances et de qualités personnelles. Nous voulons donc

apprendre en détails comment agissent les gestionnaire de projets afin de pouvoir dire ce qu'ils font d'efficace ou d'inefficace.

b) L'étude servira à court terme à revoir et à améliorer s'il y a lieu l'ensemble des cours de Maîtrise en gestion de projets. À long terme, les résultats pourront être utilisés à la construction de tests situationnels de type « in-basket », servant à évaluer les compétences en gestion de projets.

3. Échantillon des observateurs

a) Les personnes retenues pour l'étude ont été choisies pour (1) leur niveau de connaissance et d'expérience comme en témoigne leur poste occupé, (2) ou parce qu'ils sont actuellement inscrits à la Maîtrise en gestion de projets (UQTR, UQAC et UQAM).

b) Ainsi, chaque répondant choisi devra être en mesure de fournir des informations sur les pratiques et l'efficacité en gestion de projets.

4. Confidentialité des réponses

a) Il va de soi que l'on s'engage à traiter vos réponses avec confidentialité et à sauvegarder en toutes circonstances l'anonymat des répondants et des personnes décrites dans les incidents critiques.

b) Afin de nous aider à respecter cette entente, nous vous prions de bien vouloir suivre les règles suivantes :

- Ne pas indiquer les noms réels des personnes en cause dans les incidents.

- Placer l'ensemble de vos réponses dans les enveloppes prévues à cette fin, et remettre l'enveloppe cachetée à la personne désignée.

B. Présentation de la méthode des « Incidents critiques »

1. Description

a) La méthode des « incidents critiques » a été introduite par John, C. Flanagan en 1954. Malgré son âge, c'est une des meilleures méthodes lorsqu'il s'agit d'analyser les tâches importantes d'une fonction et d'en identifier les exigences. D'un fonctionnement simple, cette méthode est recommandée par plusieurs spécialistes parce qu'elle se base sur des comportements observés plutôt que sur des jugements (Campbell et al., 1970; Carroll & Schneider, 1982; Dunnette, 1976; Latham et al., 1979; Schwab et al., 1975). En effet, la méthode consiste simplement à demander à des observateurs de relater des exemples réels de comportements de personnes occupant une fonction donnée qui ont été particulièrement efficaces ou particulièrement inefficaces par rapport à la réalisation de la mission et des objectifs de cette fonction.

b) On entend par incident un événement observable suffisamment complet en lui-même pour permettre des inférences au sujet des comportements de la personne observée et de leurs conséquences. Un incident pour être retenu doit rencontrer les critères suivants :

- (1) L'observateur doit connaître la mission et les objectifs de la fonction étudiée.

- (2) L'incident doit avant tout porter sur les comportements observables d'une personne occupant la fonction à étudier.
 - (3) L'incident doit survenir dans une situation où il est possible de saisir les événements et le contexte qui ont conduit aux comportements observés.
 - (4) L'intention de la personne observée et ses comportements doit être très claire pour l'observateur.
 - (5) Les conséquences de l'incident doivent être suffisamment évidentes de sorte qu'il n'y ait pas de doute quant à leurs effets sur la mission et les objectifs de la fonction concernée.
- c) On entend par critique, un incident qui a un effet important sur la réalisation des objectifs de la fonction étudiée. Un incident pour être critique doit rencontrer les critères suivants :
- (6) Il doit être pertinent dans le sens où l'incident a un impact direct ou indirect sur la réalisation de la mission et des objectifs de la fonction observée. Cet impact peut être à court, moyen ou long terme.
 - (7) L'impact doit être suffisamment important pour avoir un effet significatif sur la réalisation de la mission et des objectifs de la fonction. L'impact peut être positif ou négatif, en autant que son importance soit significative.

2. Exemples d'incidents critiques

- a) Incident positif recueilli auprès d'officiers des Forces armées américaines (Flanagan, 1951, p.29).

« Une directive reçue du quartier général demandait certaines informations détaillées qui apparaissaient à prime abord impossible à recueillir dans un laps de temps aussi court. Après discussion, divers points de vue furent émis comme par exemple faire rapport au quartier général que certaines de ces informations sont inconnues, demander un délai plus long, etc. Un officier dit alors : « C'est un ordre! Nous devons obtenir ces informations au lieu de discuter. Confiez-moi ce travail, et je m'en chargerai. » tel que demandé, on lui a assigné la tâche et immédiatement il passa à l'action. Les informations avaient à être rassemblées pour plus d'une centaine d'endroits répartis sur l'ensemble du territoire américain. Il a formé une équipe qui a recueilli les informations, et une autre pour les compiler selon les exigences de la directive. L'échéancier a été respecté grâce à l'initiative et au dynamisme de cet officier qui n'ont pas été freinés par l'impossibilité apparente du problème. (...) »

b) Incident négatif (Flanagan, 1951, p.29).

« Cet officier servait comme adjudant dans ma section. Mon adjudant précédent était très énergique et s'acquittait de ses tâches administratives de façon autonome et avec très peu de supervision. Ce nouvel officier a donc fait ses débuts avec moi sous ce même type de supervision. L'administration commençant à ralentir et les rapports étant en retard, j'ai dû vérifier. J'ai trouvé que cet officier était très minutieux, à un point tel que la moindre opération devait attendre sa signature. Un après-midi, j'ai surpris des employés qui avaient fini un travail depuis midi et qui, trois heures plus tard, attendaient toujours la signature de l'adjudant

sur leur passe sans pouvoir quitter la base. Je suis allé vérifié avec l'adjudant; j'ai appris qu'à midi il était trop occupé pour signer leur passe et qu'il s'en occuperait aussitôt qu'il aurait fini un travail urgent. J'ai alors signé moi-même les passes afin qu'ils puissent s'en aller. Après quelques vérifications additionnelles, j'ai découvert que ces retards étaient devenus fréquents. J'ai décidé que cet homme était trop lent pour mon unité, et je m'en suis débarrassé. »

3. Exercice

Pour chacun des cas suivants, dites s'il s'agit d'incident critiques; justifiez votre réponse à l'aide des critères ci-haut.

a) Incident 1. « Un rapport de fin de mois pour un lot de travail important indiquait un écart de 25% supérieur au budget, ce qui à première vue annonce un dépassement de plus de 15% pour l'ensemble du projet. Le gérant de projet fait une étude approfondie des données, réévalue ce qui reste à réaliser et fait des propositions à ses supérieurs visant à minimiser l'impact de cet écart, sans affecter de façon significative l'échéancier des travaux. Il propose de changer la séquence des travaux et de redistribuer le travail aux équipes les plus efficaces. Il obtient le feu vert de son supérieur et procède aussitôt. Finalement, le projet se termine avec un dépassement de 5%. Son supérieur conscient du caractère risqué de ce projet, est très heureux du dénouement et compte recommander cet individu pour la gérance d'un nouveau projet d'envergure que la firme est sur le point de se voir accorder. »

b) Incident 2. « Le ministère de l'environnement entreprend des démarches pour faire arrêter les travaux de démolition d'une ancienne bâtisse parce que celle-ci abriterait selon leurs dires des nids d'une espèce très rare d'hirondelles. Le représentant du Ministère fait des représentations auprès du gérant de projet l'invitant à suspendre temporairement les travaux jusqu'à ce qu'on ait trouvé une solution à ce problème. Le gérant de projet trouve ça ridicule et indique qu'il n'est pas question d'arrêter un projet de 5M\$ pour une histoire d'hirondelles. Cinq jours plus tard, un huissier vient sur le chantier et dépose une injonction ordonnant l'arrêt immédiat des travaux jusqu'à nouvel ordre. Le gérant de projet fait une colère et refuse de cesser les travaux. Les policiers sont alors dépêchés sur les lieux et font arrêter les travaux. Quelques jours plus tard, le gérant de projet est sommé de comparaître en cours. »

c) Incident 3. « Pierre est gérant de projet sur un chantier de construction pour une bâtisse industrielle de plusieurs M\$. Pierre se trouve soudainement au centre d'un conflit entre deux de ses contremaîtres, soit Richard (finition intérieure) et Jacques (électricité). Lundi matin dernier, Richard entre subitement dans le bureau de Pierre et éclate : « Si tu ne t'occupes pas de Jacques avant 5h.00 ce soir, demain, je serai malade pour longtemps, c'est clair? Ça fait plus d'un mois que je te chante la même chose : je ne pourrai pas encore rencontrer mon échéancier parce que je dois toujours attendre après lui, car il est incapable de livrer ses travaux à temps. Je te le répète, ça s'annonce pour être le même scénario que pour le chantier précédant. » Vers 3h.00 de l'après-midi, Pierre convoque ces deux

individus à son bureau et leur dit ce qui suit : « Ce n'est pas la première fois que cela se produit. Veuillez une fois pour toute vous entendre et régler votre problème de coordination et me ficher la paix. J'ai autre chose à faire que de régler vos chicanes. Au travail maintenant. » Mardi matin, Richard téléphone à la secrétaire de Pierre et lui demande d'avertir son patron qu'il est très malade, qu'il ne peut se présenter au chantier et qu'il ne sait pas quand il pourra reprendre son travail. »

d) Incident 4. « Un gérant de projet, dont les coûts augmentent beaucoup plus vite que prévus, s'entête à ne pas aviser ses supérieurs; il espérait ne dépasser les coûts que par un faible pourcentage, et ainsi ne pas avoir à avouer l'erreur d'avoir mal estimé les coûts. Ne réussissant pas à comprimer les coûts, il finit par avouer son problème à ses supérieurs, en leur faisant une nouvelle demande de fonds. Les fonds sont accordés, non sans friction, et le projet continue. Le même scénario se reproduit, mais avec un dépassement moindre. Le projet est réalisé avec des coûts beaucoup plus élevés que prévus. Il faut dire ici que lors de l'analyse du projet au début, un estimé des coûts avait été fait pas un tiers, et n'avait pas été accepté par la direction parce que trop élevé. Le gérant avait alors fait reprendre l'estimé par un de ses subordonnés, et c'est cet estimé, moins élevé, qui avait été approuvé et servait de budget pour le projet. »

C. Personnages étudiés

1. Définition des principaux types de gestionnaires de projets.

a) Gérant (ou directeur) de projets : celui à qui la haute direction d'une

entreprise confie l'entière responsabilité de la réalisation d'un produit donné tout en maximisant l'utilisation des ressources nécessaires humaines et matérielles (ex.: directeur de chantier de construction).

b) Chargé de projets : celui qui a la responsabilité de gérer le cheminement général du projet à travers ses différentes phases, à savoir les autorisations, les études de faisabilité, les appels d'offre, etc. (ex. : ingénieur municipal).

c) Ingénieur de projets : celui à qui incombe la responsabilité technique, à savoir la réalisation des plans et devis ainsi que la surveillance lors de l'exécution des travaux sur le terrain (ex. : ingénieur travaillant au département d'ingénierie dans une usine de papier).

d) Coordonnateur de projets : adjoint au gérant ou parfois au chargé de projet.

Note : Le gérant, le chargé et l'ingénieur de projets sont habituellement tous considérés comme des gestionnaires de projet « fonctionnels » (« line »), alors que le Coordonnateur se retrouve plutôt dans une fonction d'adjoint (« staff »). Le gérant, le chargé et l'ingénieur de projets ont tous la responsabilité de gérer des projets ou une de ses parties; leur travail se distingue toutefois d'abord et surtout par l'envergure et la complexité grandissante des projets en cause à mesure que l'on passe de l'ingénieur au chargé, puis au gérant de projets.

2. Type de gestionnaire retenus pour l'étude

Les sujets étudiés dans la présente recherche regroupent tous les types de gestionnaires définis ci-haut, à l'exception des ingénieurs de projets considérés

plus comme des professionnels techniques que comme des gestionnaires. cela veut donc dire que le gestionnaire observé et mis en cause dans l'incident doit être nécessairement un gérant, un chargé ou un Coordonnateur de projets. Il n'est pas nécessaire cependant que ce soit toujours le même personnage d'un incident à l'autre.

3. Personnes observées.

a) Le personnage mis en cause dans un instrument peut être :

- (1) soit le répondant lui-même
- (2) soit son supérieur hiérarchique propre
- (3) ou un de ses subordonnés
- (4) si aucun de ces cas n'est possible, le répondant peut décrire le(s) gestionnaire(s) de projets le(s) plus près de lui.

Note : rappelez-vous qu'il n'est pas nécessaire que ce soit toujours le même personnage d'un incident à l'autre.

b) Localisation et secteur d'activité :

- secteur privé, public ou para-public;
- projets durs ou mous;
- sous une structure matricielle, fonctionnelle ou par projet;
- oeuvrant dans des entreprises québécoises ou dans des entreprises avec succursale au Québec;
- faisant affaire ou non à l'étranger.

4. Comportement observé

a) Type d'activités en général;

- conception, planification, organisation et contrôle;
- coordination des différents intervenants;
- direction des subordonnés;
- relations avec les supérieurs;
- relations externes;
- etc.

b) Exemples de comportements spécifiques

- maintient le déroulement du projet à l'intérieur des délais tout en respectant les spécifications initiales;
- rencontre un sous-traitant afin de voir un nouveau procédé;
- supporte et encourage les membres de l'équipe;
- assiste à un colloque afin de garder le contact avec ses principaux compétiteurs;
- etc.

D. Cueillette des incidents critiques

1. Rappel des principaux critères

a) Avant de passer à la cueillette des incidents critiques, rappelez-vous qu'un incident pour qu'il soit retenu comme « critique » doit rencontrer les critères suivants.

- Il doit être pertinent dans le sens que tout incident décrit doit avoir un impact direct ou indirect sur la réalisation de l'objectif général de la fonction de gestionnaire de projets. L'impact peut être à court , moyen ou long terme.

- Cet impact doit être suffisamment important dans le sens que l'incident décrit à un effet significatif sur la réalisation de l'objectif général de la fonction de gestionnaire de projets. L'impact peut être positif ou négatif.

b) Rappelez-vous également qu'un incident doit être rédigé en respectant le plus possible les règles suivantes.

- L'incident doit avant tout porter sur les comportements observables

- L'incident doit raconter les événements et le contexte qui ont conduit aux comportements observés.

- L'intention de la personne observée et ses comportements doivent être décrits très clairement.

- Les conséquences de l'incident doivent être relatées de façon évidente de sorte qu'il n'y ait pas de doute quant à leurs effets sur la mission et les objectifs de la fonction concernée.

2. Directives

a) Voyons maintenant ce que chacun d'entre vous peut relater comme incident critique. Dans un premier temps, chacun va essayer de rédiger un incident à impact négatif. Une fois terminé, nous regarderons ensemble ces incidents.

b) Nous allons reprendre l'exercice, sauf que cette fois vous allez tenter de relater un incident positif.

c) Dans l'étude, on s'attend à ce que chacun d'entre vous puissiez recueillir en tout dix incidents critiques, soit cinq positifs et cinq négatifs. Nous vous demandons donc votre collaboration pour rédiger huit autres incidents d'ici la prochaine rencontre. Profitant de cette période d'observation additionnelle, vous pourrez certainement être à l'affût de nouveaux incidents qui pourraient survenir dans votre organisation.

3. Spécifications des incidents et des observateurs

a) Nous vous demandons de ne rédiger qu'un incident pas page ; si un incident prend plus d'une page, commencez une nouvelle page pour l'incident suivant.

b) Pour chacun des incidents rédigés, il est important d'indiquer clairement (de préférence au début de la page) les éléments suivants:

- Est-ce un incident positif ou négatif?
- Est-ce vous ou un autre qui est le personnage principal?
- Si ce n'est pas vous, quelle est la fonction du personnage observé?

APPENDICE H

Consignes aux experts

APPENDICE H

Consignes aux experts

Adressée à chacun des experts,

Suite à notre conversation téléphonique, je vous fais parvenir mes items et matrice tel que convenu. Je tiens tout d'abord à vous remercier de votre collaboration. Je suis consciente que votre horaire est chargé et j'apprécie énormément votre contribution, votre expertise me sera très utile.

Afin de clarifier mes attentes, je vous résume en quelques mots ma démarche. J'ai élaboré un « in-basket » pour la sélection de gestionnaires de projet. Pour ce faire, une recension de différentes recherches m'a permis de construire une matrice comprenant, en abscisse, des dimensions (Yukl, 1987) et en ordonné, des domaines de la gestion de projet (*PMBOK*). Vous remarquerez que des domaines ont été éliminés. Soit parce qu'ils étaient difficiles à mesurer dans le contexte (mon test se situe en phase de réalisation de projet, ainsi tout domaine incluant une planification préalable a dû être éliminé), ou encore parce que ces dits domaines étaient redondants avec les dimensions en abscisse. Suite à l'élaboration, une étape de pré-test était nécessaire, c'est à ce moment que je sollicite votre contribution...

En effet, afin de s'assurer de la représentativité des items pour la gestion de projet, il importe de consulter des spécialistes du domaine. Voici la démarche que je vous propose:

1 - Lecture des documents de support. Ceux-ci doivent placer le candidat dans une situation suffisamment claire pour lui permettre de répondre aux items. De plus, elle doit être représentative de la gestion de projet. Pourriez-vous SVP m'indiquer vos commentaires et/ou suggestions.

2 - Afin de s'assurer de couvrir l'ensemble du domaine de contenu (gestion de projet), classement des items à l'intérieur de la matrice (ex: item 1 - cellule suivi et contrôle x délais). À noter qu'un item peut se retrouver dans plus d'une cellule.

3 - Les items sont-ils représentatifs? Suffisamment ambigus? Comportent-ils suffisamment d'informations? Le degré de difficulté est-il adéquat? Commentaires et/ou suggestions pour en améliorer la qualité.

Encore merci de votre collaboration,

bien à vous,

Nathaly Jutras

(Les consignes telles que présentées à l'Appendice F, le contexte de la simulation et les documents de support (Appendice D et E), la matrice ainsi que les items (Appendice G) ont été remis avec la lettre)

APPENDICE I

Consignes aux candidats

Test de Lawshe

APPENDICE I

Consignes aux candidats

Test de Lawshe

Madame, Monsieur,

Dans le cadre d'un mémoire de maîtrise, j'ai élaboré un instrument de sélection de personnel, un « in-basket » adapté à la gestion de projet. Un « in-basket » présente des situations critiques (qui demandent une intervention de la part du gestionnaire) prises dans le « panier de gestion », celles-ci se présentent sous différentes formes (notes à l'agenda, mémos, lettres, téléphones, etc.). Le candidat doit y répondre en un temps donné et il est évalué à partir des stratégies qu'il mettra de l'avant afin de solutionner les problèmes qui lui sont présentés. Ce test comporte des items ainsi que des documents de support comprenant des informations sur l'entreprise fictive (budget, organigramme, mandat, etc.).

Afin d'assurer la validité de contenu du test, il importe de consulter des spécialistes du domaine à mesurer, en l'occurrence la gestion de projet. Pour ce faire, j'utilise un test a posteriori proposé par Lawshe (1975). Cette démarche consiste à coter la pertinence de chacun des items proposés dans le test.

C'est à cette étape de ma recherche que je sollicite votre collaboration. Afin de réunir les informations nécessaires pour mes statistiques, je vous demanderais de répondre à un court formulaire comprenant des informations pertinentes pour mieux décrire l'échantillon des répondants. Ensuite, veuillez SVP lire chacun des items et répondre en fonction de votre opinion, à savoir:

Est-ce que l'habileté (ou connaissance) mesurée par cet item est

- Essentielle
- Utile mais non essentielle, ou
- Non nécessaire

à la performance à l'emploi.

Traduction libre de: « Is the skill (or knowledge) measured by this item, Essential, useful but not essential, or not necessary to the performance of the job » (Lawshe, 1975).

Je n'ai pas jugé utile de vous fournir les documents de support afin de ne pas alourdir votre lecture inutilement mais vous pouvez prendre pour acquis que le candidat aura toutes les informations nécessaires afin de répondre aux items lors d'une passation éventuelle.

Merci infiniment de votre collaboration,

Bien à vous,

Nathaly Jutras

INFORMATIONS CONCERNANT LES RÉPONDANTS

Nom: _____

Âge: _____

Sexe: _____

Nombre d'années d'expérience en gestion: _____

Nombre d'années d'expérience en gestion de projet: _____

Domaine de spécialisation (construction, ingénierie, international, etc.): _____

FEUILLE RÉPONSE

SVP Entourez la réponse appropriée

Question:

Est-ce que l'habileté (ou connaissance) mesurée par cet item est

- Essentielle
- Utile mais non essentielle, ou
- Non nécessaire

à la performance à l'emploi.

Item 1: Agrandissement non conforme aux plans et devis

Essentielle Utile mais non essentielle Non nécessaire

Item 2: Incapacité à manipuler matériel d'examen adéquatement

Essentielle Utile mais non essentielle Non nécessaire

Item 3: Demandes de transfert pour votre service

Essentielle Utile mais non essentielle Non nécessaire

Item 4: Rédaction de documents sur les buts et objectifs de l'Hôpital de Jour

Essentielle Utile mais non essentielle Non nécessaire

Question:

Est-ce que l'habileté (ou connaissance) mesurée par cet item est

- Essentielle
- Utile mais non essentielle, ou
- Non nécessaire

à la performance à l'emploi.

Item 5:	Problèmes avec l'implantation d'un logiciel		
	Essentielle	Utile mais non essentielle	Non nécessaire
Item 6:	Délai de livraison (appareils physiothérapie)		
	Essentielle	Utile mais non essentielle	Non nécessaire
Item 7:	Michel surpris à dormir sur son chiffre		
	Essentielle	Utile mais non essentielle	Non nécessaire
Item 8:	Informations aux CLSC et médecins des services offerts		
	Essentielle	Utile mais non essentielle	Non nécessaire
Item 9:	Conflit de personnalité Jean Racine versus Nicole Dumais		
	Essentielle	Utile mais non essentielle	Non nécessaire

Question:

Est-ce que l'habileté (ou connaissance) mesurée par cet item est

- Essentielle
- Utile mais non essentielle, ou
- Non nécessaire

à la performance à l'emploi.

Item 10: Infirmiers et infirmières dénigrent le travail des préposés aux bénéficiaires

Essentielle Utile mais non essentielle Non nécessaire

Item 11: Comportements déplacés de la part des ouvriers envers des patients

Essentielle Utile mais non essentielle Non nécessaire

Item 12: Dépassement du budget au niveau des salaires

Essentielle Utile mais non essentielle Non nécessaire

Item 13: Livraison retardée de 4 semaines et plus...

Essentielle Utile mais non essentielle Non nécessaire

Item 14: Partage du temps d'Yvan Cournoyer, médecin

Essentielle Utile mais non essentielle Non nécessaire

Question:

Est-ce que l'habileté (ou connaissance) mesurée par cet item est

- Essentielle
- Utile mais non essentielle, ou
- Non nécessaire

à la performance à l'emploi.

Item 15: Taux d'absentéisme élevé

Essentielle Utile mais non essentielle Non nécessaire

Item 16: Contremaître remplacé, retard de construction

Essentielle Utile mais non essentielle Non nécessaire

Item 17: C.S.S.T. exige l'engagement d'un inspecteur de sécurité

Essentielle Utile mais non essentielle Non nécessaire

Item 18: Comptes à payer en retard

Essentielle Utile mais non essentielle Non nécessaire

Item 19: Directives contradictoires

Essentielle Utile mais non essentielle Non nécessaire

Question:

Est-ce que l'habileté (ou connaissance) mesurée par cet item est

- Essentielle
- Utile mais non essentielle, ou
- Non nécessaire

à la performance à l'emploi.

Item 20: Désordre lors du changement de chiffre

Essentielle Utile mais non essentielle Non nécessaire

Item 21: Ancien gestionnaire en poste faisait peu état des règles de fonctionnement

Essentielle Utile mais non essentielle Non nécessaire

Merci encore,

Nathaly Jutras

(Les 21 items tels que présentés à l'Appendice G ont été fournis aux participants.)

APPENDICE J

Résultats des coefficients de validité de contenu
(sujets regroupés par domaine de spécialisation)

Coefficient de validité - sujets regroupés par domaine**Domaine d'expertise : informatique**

Item 1 :	0,00	Item 12 :	0,00
Item 2 :	0,00	Item 13 :	1,00
Item 3 :	-1,00	Item 14 :	-1,00
Item 4 :	1,00	Item 15 :	0,00
Item 5 :	-0,50	Item 16 :	0,00
Item 6 :	0,00	Item 17 :	1,00
Item 7 :	-1,00	Item 18 :	-0,50
Item 8 :	0,00	Item 19 :	-0,50
Item 9 :	0,00	Item 20 :	-1,00
Item 10 :	0,00	Item 21 :	0,50
Item 11 :	0,00		

Nombre de sujets : 4

Coefficient de validité - sujets regroupés par domaine**Domaine d'expertise : construction**

Item 1 :	0,00	Item 12 :	0,33
Item 2 :	0,00	Item 13 :	0,33
Item 3 :	-0,67	Item 14 :	-0,67
Item 4 :	0,33	Item 15 :	-0,33
Item 5 :	-0,33	Item 16 :	-1,00
Item 6 :	0,33	Item 17 :	0,33
Item 7 :	-1,00	Item 18 :	-1,00
Item 8 :	1,00	Item 19 :	-1,00
Item 9 :	-0,33	Item 20 :	-1,00
Item 10 :	0,33	Item 21 :	-0,33
Item 11 :	-0,33		

Nombre de sujets : 6

Coefficient de validité - sujets regroupés par domaine**Domaine d'expertise : ingénierie**

Item 1 :	0,60	Item 12 :	0,47
Item 2 :	-0,60	Item 13 :	0,47
Item 3 :	-0,33	Item 14 :	0,07
Item 4 :	0,43	Item 15 :	-0,07
Item 5 :	0,33	Item 16 :	-0,20
Item 6 :	0,20	Item 17 :	0,07
Item 7 :	-0,33	Item 18 :	-0,60
Item 8 :	-0,60	Item 19 :	-0,20
Item 9 :	-0,07	Item 20 :	-0,20
Item 10 :	-0,47	Item 21 :	-0,33
Item 11 :	-0,33		

Nombre de sujets : 15

Coefficient de validité - sujets regroupés par domaine**Domaine d'expertise : services**

Item 1 :	0,33	Item 12 :	-1,00
Item 2 :	-1,00	Item 13 :	0,33
Item 3 :	0,33	Item 14 :	-0,33
Item 4 :	0,33	Item 15 :	0,33
Item 5 :	0,33	Item 16 :	-0,33
Item 6 :	-0,33	Item 17 :	0,33
Item 7 :	-0,33	Item 18 :	-0,33
Item 8 :	-0,33	Item 19 :	-1,00
Item 9 :	0,33	Item 20 :	-0,33
Item 10 :	-0,33	Item 21 :	0,33
Item 11 :	-0,33		

Nombre de sujets : 3

Coefficient de validité - sujets regroupés par domaine**Domaine d'expertise : divers**

Item 1 :	0,40	Item 12 :	0,20
Item 2 :	-0,60	Item 13 :	-0,20
Item 3 :	-0,60	Item 14 :	-0,60
Item 4 :	0,20	Item 15 :	0,00
Item 5 :	-0,20	Item 16 :	-0,40
Item 6 :	-0,40	Item 17 :	0,00
Item 7 :	-0,80	Item 18 :	-0,56
Item 8 :	-0,80	Item 19 :	-0,20
Item 9 :	-0,20	Item 20 :	-0,80
Item 10 :	-0,40	Item 21 :	-0,20
Item 11 :	-0,60		

Nombre de sujets : 10

Coefficient de validité - sujets regroupés par domaine**Domaine d'expertise : non spécifié**

Item 1 :	-0,20	Item 12 :	0,60
Item 2 :	0,20	Item 13 :	0,20
Item 3 :	-0,60	Item 14 :	-0,20
Item 4 :	0,60	Item 15 :	0,20
Item 5 :	0,20	Item 16 :	0,20
Item 6 :	-0,20	Item 17 :	0,60
Item 7 :	-0,60	Item 18 :	-1,00
Item 8 :	0,20	Item 19 :	-0,20
Item 9 :	0,20	Item 20 :	-0,60
Item 10 :	-0,20	Item 21 :	0,60
Item 11 :	-0,20		

Nombre de sujets : 5