



CENTRE DE RECHERCHE DMSP

**Comment mesurer l'apprentissage de groupe ?
Construction d'une échelle de mesure bi-dimensionnelle**

Valentine Galtier

Cahier n°348

Octobre 2005

Valentine Galtier

Docteur en sciences de gestion
Centre de recherche DMSP DRM (CNRS UMR 7088)
Université Paris IX Dauphine
Place Maréchal de Lattre de Tassigny
75775 Paris Cedex 16

valentine.galtier@free.fr
valentine.galtier@dauphine.fr

COMMENT MESURER L'APPRENTISSAGE DE GROUPE ? CONSTRUCTION D'UNE ECHELLE DE MESURE BI-DIMENSIONNELLE

Résumé

Dans le contexte du jeu de simulation Markstrat, une échelle de mesure de l'apprentissage de groupe est élaborée sur la base des préconisations du paradigme de Churchill (1979). La démarche méthodologique exposée met en évidence la possibilité de mesurer quantitativement le résultat de l'apprentissage de groupe de manière fiable et valide. La mesure construite comporte deux dimensions : une dimension dite « substantive » et une dimension dite « procédurale ».

Mots Clefs : apprentissage de groupe, apprentissage organisationnel, échelle de mesure

Abstract

In the context of Markstrat simulation game, a group learning scale has been developed based on Churchill's work (1979). Our methodology reveals that it is possible to measure group learning by developing a specific scale. This scale appears as reliable and valid. The scale falls into two dimensions we called « knowledge » dimension and « procedural ».dimension.

Key Words : group learning, organizational learning, measurement scale

MESURER L'APPRENTISSAGE DE GROUPE : UNE APPLICATION DU PARADIGME DE CHURCHILL

Introduction

Le but de cet article est de proposer un nouvel outil méthodologique : une mesure quantitative de l'apprentissage de groupe. L'apprentissage de groupe est défini sur la base de l'apprentissage organisationnel comme « un phénomène collectif d'acquisition et d'élaboration de compétences, qui plus ou moins profondément, plus ou moins durablement modifie la gestion des situations et les situations elles-mêmes ». (Koenig, 1994, page 78). Le groupe est considéré comme un cas particulier d'organisation. Plus précisément nous nous intéressons à des groupes dits primaires c'est à dire de taille réduite et à la structure organisationnelle non formalisée (Anzieu et Martin, 1968). En outre, nous verrons que l'apprentissage est à la fois un processus et un résultat et nous avons choisi de nous intéresser à l'aspect résultat de l'apprentissage de groupe, résultat que nous définissons comme l'ensemble des connaissances partagées par le groupe. La connaissance est définie de manière large comme une représentation partagée et englobe la notion de compétence définie comme un savoir agir reconnu.

Dans le contexte du jeu de simulation Markstrat, une échelle de mesure de l'apprentissage de groupe est élaborée sur la base des préconisations du paradigme de Churchill (1979). La démarche méthodologique exposée met en évidence la possibilité de mesurer quantitativement le résultat de l'apprentissage de groupe de manière fiable et valide.

Dans un premier temps, nous nous attarderons sur le niveau d'analyse du groupe puis nous discuterons des mesures existantes de l'apprentissage. Nous présenterons ensuite la démarche méthodologique suivie pour mettre au point notre échelle de mesure de l'apprentissage de groupe et nous en présenterons les résultats.

1. Remarque préalable : le groupe comme unité d'analyse

Nous avons choisi de considérer le groupe comme unité d'analyse de notre recherche. Le groupe est en effet utilisé pour étudier l'apprentissage organisationnel comme le montrent par exemples les travaux de Cangelosi et Dill dès 1965 ou ceux de Simon en 1991. Pour définir l'apprentissage de groupe et en développer une mesure, nous nous sommes basés sur les travaux menés sur l'apprentissage organisationnel.

Le groupe est défini comme un ensemble de personnes interdépendantes dans la poursuite d'un objectif. En ce sens, le groupe est considéré comme un cas particulier d'organisation, comme un modèle d'organisation simple : c'est à dire de taille réduite, et à la structure organisationnelle non formalisée c'est à dire qu'il n'y a pas par exemple de hiérarchie entre les membres du groupe. On parle ainsi de groupe restreint ou groupe primaire pour le distinguer du groupe secondaire qu'est l'organisation (Anzieu et Martin, 1968).

Le groupe est donc une unité intermédiaire entre l'individu et l'organisation, considérée comme un groupe secondaire. Cette unité permet de rendre compte des tensions entre les aspects individuels et collectifs de l'apprentissage.

Le choix de se situer au niveau du groupe conduit à préciser la définition de l'apprentissage de groupe par rapport à celle de l'apprentissage organisationnel. Pour définir l'apprentissage de groupe, la définition de l'apprentissage organisationnel est appliquée au niveau du groupe. Ainsi, nous définissons l'apprentissage de groupe de la même manière que l'apprentissage organisationnel c'est à dire comme « un phénomène collectif d'acquisition et d'élaboration de compétences, qui plus ou moins profondément, plus ou moins durablement modifie la gestion

des situations et les situations elles-mêmes ». (Koenig, 1994, page 78). En effet, le groupe, en tant que modèle d'organisation simple, est comme nous l'avons expliqué un moyen d'appréhender certains processus de la vie organisationnelle comme l'apprentissage.

Par ailleurs, l'apprentissage est à la fois un processus et un résultat. Nous avons choisi de nous intéresser à l'aspect résultat de l'apprentissage défini comme la connaissance partagée par les membres du groupe. Cette définition large du résultat de l'apprentissage permet d'englober la notion de compétence définie comme un savoir agir reconnu.

La distinction entre apprentissage de groupe et apprentissage organisationnel tient à la définition du groupe précédemment choisie. Ainsi, dans le cadre du groupe, modèle d'organisation simple, un certain nombre de questions liées à l'apprentissage organisationnel sont écartées telles que :

- les obstacles liés à la taille, à la dispersion géographique des individus,
- ou encore la prise en compte d'une structure formalisée (telle qu'une hiérarchie).

La création d'une nouvelle échelle se justifie par rapport au contexte ou à l'insuffisance des échelles existantes pour saisir le domaine du construit. La revue des mesures d'apprentissage existantes permet de rendre compte du besoin de développer une nouvelle mesure de l'apprentissage de groupe.

2. Le besoin d'une échelle de mesure : limites des mesures actuelles

La tendance générale des travaux sur l'apprentissage organisationnel est l'utilisation d'études qualitatives. Il s'agit de repérer des traces d'apprentissage en suivant l'évolution des pratiques et des visions existant dans l'organisation. L'objectif de ce type de recherche souvent abductive est de contribuer à une compréhension en profondeur de l'apprentissage. Le champ s'est donc considérablement développé grâce à ses travaux. Il nous semble donc intéressant aujourd'hui de travailler dans une optique plus déductive et de mettre en place des méthodologies différentes comme par exemple les méthodes quantitatives.

Dans le souci de nous inscrire parmi les travaux existants, nous souhaitons contribuer à proposer une mesure opérationnelle quantitative du concept d'apprentissage organisationnel et plus précisément, dans notre cas, d'apprentissage de groupe. En effet, dans une optique de méthodologie quantitative, les questions qui se posent alors touchent à l'opérationnalisation du construit d'apprentissage collectif. Comment mener des travaux plus quantitatifs en mobilisant ce concept d'apprentissage ou encore comment comparer en terme d'apprentissage différentes organisations ou différents groupes effectuant une même tâche ? Il est possible de trouver des éléments de réponse dans la littérature mais ceux-ci nous semblent insuffisants pour permettre l'opérationnalisation du concept.

Une première solution consisterait à mesurer la performance du groupe ou de l'organisation (Carley, 1992). Cette assimilation entre apprentissage et performance s'appuie sur les travaux portant sur les courbes d'apprentissage. Cette mesure offre l'avantage de se situer au niveau du groupe ou de l'organisation, ce qui évite les problèmes d'agrégation pour passer du niveau individuel au niveau collectif ou organisationnel. Elle est facile à obtenir et permet de comparer différentes organisations. Cependant, il est clair qu'amélioration de la performance n'est pas synonyme d'apprentissage organisationnel. La performance peut être une des manifestations de l'apprentissage organisationnel, mais la performance n'est pas toujours le proxy d'un apprentissage. Ainsi, mesurer la performance peut refléter un apprentissage organisationnel ou apprentissage de groupe, mais pas toujours... C'est pourquoi, nous pensons qu'il est

indispensable de combiner la performance avec d'autres indicateurs d'apprentissage ou de construire une mesure spécifique de l'apprentissage.

Si la performance est insuffisante en raison principalement du fait qu'elle n'est pas forcément une manifestation de l'apprentissage, il faut donc se tourner vers la mesure d'autres éléments. Les travaux récents de Bourgeon et Tarondeau (2000) offrent une autre solution. Les auteurs proposent de mesurer l'apprentissage organisationnel ou en l'occurrence l'apprentissage de groupe, en examinant la présence ou l'absence de conditions favorables à l'apprentissage sur la base des travaux de Shrivastava (1983). Il s'agit de mesurer l'existence d'un système d'apprentissage favorable à l'apprentissage dans une équipe de projet et ce par questionnaires administrés auprès des chefs de projets. Là encore, une limite importante apparaît : mesurer les conditions favorables à l'apparition d'un phénomène et le phénomène lui-même sont deux choses distinctes. En fait, les auteurs mesurent une probabilité d'apparition de l'apprentissage.

Enfin, si l'on se réfère aux travaux sur les échelles sociométriques, il n'existe pas à ce jour de mesure de l'apprentissage organisationnel, ni de l'apprentissage de groupe. En particulier, la large sélection de mesures dites institutionnelles ou organisationnelles proposée par Miller (1991) n'offre pas une telle mesure.

Le *Tableau 1* ci-après présente les principales mesures de l'apprentissage organisationnel utilisées dans la littérature. Comme nous l'avons vu, la mesure de l'apprentissage organisationnel peut être quantitative avec la mesure de la performance ou de la présence de conditions favorables à l'apprentissage. Nous avons vu les limites de ces mesures. En outre comme nous l'avons évoqué, les recherches sur l'apprentissage sont généralement qualitatives : la mesure de l'apprentissage sur la base des réflexions théoriques de Cyert et March (1963), Argyris et Schön (1978) ou encore de Weick (1991), consiste à suivre les modifications des pratiques, comportements et/ou des connaissances/représentations, etc. partagées. Ce dernier type de mesure est propre à chaque cas.

Tableau 1 Les principales mesures de l'apprentissage organisationnel

Indicateurs utilisés	Phénomènes Décrits	Base théorique (Auteurs)	
Performance	Courbes d'apprentissages Adaptation Correction d'erreurs	Wright (1953) Ascher (1954) Carley (1992)	
Modifications des théories pratiquées/ des réponses : Modification de la réponse donnée dans une situation de stimulation identique.	Apprentissage comportemental	Cyert et March (1963), Weick (1991)	Argyris et Schön (1978),
Modification des théories affichées : utilisation de proxys liés à chaque cas.	Apprentissage cognitif Niveaux d'apprentissage		
Mesure de la présence de conditions favorables à l'apprentissage	Présence d'un système d'apprentissage	Shrivastava (1983) Bourgeon et Tarondeau (2000)	

Dans le cadre d'une méthodologie quantitative, il apparaît donc utile de développer un nouvel outil pour mesurer l'apprentissage.

3. Vers une nouvelle mesure de l'apprentissage : intérêt d'une échelle multi-items

Au vu des avantages et des limites des possibilités existantes de mesure de l'apprentissage organisationnel, développer une échelle de mesure, dans un premier temps, de l'apprentissage de groupe est un point fondamental. Cet outil de mesure, tout en restant synthétique, devrait permettre, au travers du développement de plusieurs items, de saisir les différentes facettes du phénomène.

En effet, il ressort de la revue de la littérature que l'apprentissage organisationnel et donc de groupe est un phénomène aux manifestations variées. L'apprentissage peut donc être considéré comme une variable latente non directement observable, ou plus précisément nécessitant plusieurs "points de vue" d'observation. D'où la nécessité d'une échelle multi-items. Comme l'explique Churchill (1979) un seul item, tend à catégoriser les phénomènes en un nombre limité de catégories. Par exemple un item unique, " votre organisation a appris telle ou telle chose", avec un format de réponse en 5 points, ne permet de distinguer que 5 types différents.

L'échelle multi-items permet, grâce à l'utilisation combinée de plusieurs items de répondre à plusieurs difficultés :

- saisir les différentes facettes d'un phénomène,
- pourvoir faire des distinctions plus précises entre les unités analysées,
- accroître la fiabilité de la mesure et réduire les erreurs de mesure.

En outre, comme le montrent les travaux en sociométrie, il est effectivement possible de mesurer des phénomènes sociaux (Miller, 1991). Développer une mesure de l'apprentissage de groupe est donc envisageable.

L'intérêt d'une telle mesure multi-items et graduée de l'apprentissage de groupe est :

- sur le plan théorique, une meilleure compréhension du résultat d'apprentissage ;
- sur le plan empirique, avoir un indicateur plus précis, simple et relatif de l'apprentissage permettant de comparer l'apprentissage de différents groupes en observant leur score global d'apprentissage et la structure de ce score et de mesurer l'intensité du phénomène;

Pour établir cet indicateur synthétique, nous nous référons aux travaux sur la construction d'échelle. Dans la partie suivante, nous présentons le contexte de la recherche. Nous présenterons ensuite notre démarche et les choix effectués pour construire notre mesure de l'apprentissage de groupe.

4. Le contexte de la recherche : Markstrat, jeu de simulation

Le contexte de la recherche est constitué par le jeu de simulation Markstrat. Plus précisément, Markstrat est un jeu de simulation de gestion marketing d'une entreprise dans un environnement compétitif et dynamique. C'est ainsi une mise en situation de décision de gestion qui se rapproche du réel : l'information est complexe, abondante et imparfaite.

Dans notre cas, nous avons utilisé le jeu Markstrat sur différentes promotions de MSG de Dauphine. Ainsi,

- chaque groupe de TD représente une industrie : à l'intérieur de chaque industrie, 6 firmes sont en concurrence sur un ou deux marchés.
- Les 6 firmes ont des situations au départ équivalentes.

- Le but du jeu est de faire augmenter le cours de l'action par une bonne gestion de son entreprise.
- Le jeu est divisé en périodes de jeu : une période correspondant virtuellement à un an et à la prise d'un ensemble de décisions (prix de vente, distribution, gestion des vendeurs, des stocks, recherche et développement etc.).
- Le modèle prend en compte les décisions des différentes firmes et l'état du marché pour rendre de nouvelles données en début de chaque nouvelle période.

Dans ce contexte, la méthodologie mise en place pour construire l'échelle d'apprentissage de groupe suit les recommandations du paradigme de Churchill et est présentée dans la partie suivante.

5. Le cadre méthodologique : le processus de construction de l'échelle

Pour développer notre mesure de l'apprentissage, nous avons adapté le paradigme de Churchill à la réalité de notre processus de recherche. Le paradigme de Churchill (1979) est une démarche méthodologique visant à la construction d'échelle multiple a posteriori. Ainsi, l'échelle est affinée sur les données d'enquête. Le paradigme de Churchill (1979) est aujourd'hui contesté (Rossiter, 2002), mais il n'en reste pas moins d'actualité car il offre des règles précises et simples pour construire des échelles de mesure fiables. Les différentes étapes de cette démarche et les techniques méthodologiques utilisées sont décrites dans le *Tableau 2*.

Tableau 2 La construction d'échelles (paradigme de Churchill) et les méthodes utilisées

Etapes	Description	Techniques recommandées par Churchill (1979)	Techniques méthodologiques complémentaires
Etape 1 : Définition du domaine conceptuel	Définir précisément ce que l'on souhaite mesurer.	Revue de la littérature Réflexion théorique. Conserver les définitions existantes.	
	Générer un échantillon d'items	Se baser sur la littérature Entretiens qualitatifs ou tables rondes Intuition	Etudes exploratoires
Etape 2 : La phase exploratoire	Collecte de données		
	Purifier la mesure : Mesure de la consistance interne (α de Cronbach) Réflexion sur les dimensions sous-jacentes (AFE)	Coefficient α de Cronbach Analyses factorielles exploratoires (AFE)	
	Collecte de données		
Etape 3 : La phase de validation	Vérifier la fiabilité et la validité de l'instrument de mesure	Sur de nouvelles données,	
	Fiabilité	Analyse Factorielle (AFE) Coefficient α de Cronbach	Analyse Factorielle Confirmatoire (AFC) Coefficient ρ de Joreskog
	Validité de construit	Validité convergente/discriminante : corrélations bivariées Validité nomologique	Analyse Factorielle Confirmatoire (AFC) Coefficients d'ajustement du modèle

Nous avons adapté cette démarche à notre recherche et à l'élaboration de la mesure de l'apprentissage de groupe. Notre processus de recherche n'étant pas linéaire, l'étude des dimensions et des items a bénéficié tout au long de ce travail de la compréhension graduelle du construit d'apprentissage de groupe.

5.1.Choix méthodologiques préliminaires

Le principe de la mesure développée est de mesurer une variable dite "latente" (l'apprentissage) par des éléments réels périphériques (les items de l'échelle).

Avant d'aller plus avant dans le développement de ces items, il convient de poser le format que nous souhaitons donner à cette échelle afin d'en assurer une certaine objectivité. Ainsi, nous nous sommes attachés à répondre à deux questions :

- Quels aspects du phénomène souhaitons-nous mesurer : ses manifestations, ses conséquences ou ses antécédents ?
- Quel va être le format de l'échelle de mesure et de l'échelle de notation ?

En outre, l'apprentissage de groupe étant un concept se situant au niveau du groupe, sa mesure pose la question du niveau de mesure et du niveau d'analyse ainsi que la question de l'agrégation des données.

5.1.1. Manifestations / Conséquences / Antécédents

Les indicateurs utilisés pour construire une échelle peuvent être des manifestations, des antécédents ou des conséquences du phénomène que l'on souhaite mesurer. Le paradigme de Churchill (1979) ne préconise rien à ce sujet.

Pour le développement de la mesure de l'apprentissage de groupe, nous avons choisi d'utiliser les manifestations du phénomène. Choisir les antécédents ou les conséquences nous semblent moins fiable. En effet, au vu de la complexité du concept, il est difficile d'établir un lien sûr entre l'apprentissage et ses antécédents ou ses conséquences.

Avant de se lancer dans l'élaboration d'items et la construction d'échelle, il nous faut préciser la nature de notre échelle et donc opérer deux grands types de choix concernant l'échelle de mesure et l'échelle de notation. Ces deux choix sont respectivement l'objet de chacune des sous-parties suivantes.

5.1.2. L'échelle de mesure : une échelle d'intervalle

Le choix de l'échelle de mesure est crucial : il va spécifier le rapport des chiffres à la réalité. Ces chiffres vont-ils représenter des quantités, des "étiquettes", des rangs,...? En outre, le choix du type d'échelle de mesure conditionne les traitements statistiques ultérieurs possibles.

Il existe quatre grands types d'échelle de mesure (Evrard et al, 2000) : les échelles de proportion, d'intervalle, ordinales et nominales. Pour choisir parmi ces catégories, nous nous sommes basés sur la démarche proposée par Evrard et al (2000). Cette démarche nous conduit à choisir une échelle d'intervalle, c'est-à-dire une échelle dont les catégories ou échelons sont ordonnées, dont les intervalles sont égaux et dont le zéro est fixé arbitrairement. Ce type d'échelle permet ainsi de mesurer l'intensité du phénomène et d'utiliser les méthodes d'analyse de données classiques et nécessaires au développement de l'échelle (par exemple l'analyse factorielle).

5.1.3. L'échelle de notation

L'échelle de notation concerne le format de réponse qui sera proposé au répondant. Il existe de très nombreuses formulations d'échelles. Nous avons ainsi choisi de retenir une notation avec des ancrages sémantiques à chaque extrémité et une visualisation graphique des catégories et intervalles. Le format de réponse proposé au répondant est le suivant (cf. *Figure 1*) :

Figure 1 Format du support de l'échelle

	← Pas d'accord	D'accord →
1. Nous avons des représentations partagées du jeu qui guident nos comportements	/ _____ /	/ _____ /

Nous avons choisi ce type d'échelle car elle offre plusieurs avantages : facilité d'administration, mise en évidence de l'égalité des intervalles, répartition des réponses.

Churchill (1979) met en effet en évidence la facilité de mise en œuvre de ce support et souligne également que ce support permet une interprétation plus facile des différentes catégories. Par

ailleurs, ce type de support renforce la perception d'équidistance des intervalles (Wildt et Mazis, 1978) Elle donne en effet l'impression d'une "règle continue". Enfin, pour parer le point faible de ce type de format graphique, à savoir la concentration des réponses sur la case centrale (Churchill, 1979), nous avons opté pour des ancrages sémantiques modérés aux extrémités de l'échelle : « pas d'accord » et « d'accord ».

Un dernier choix méthodologique préliminaire concerne la question du niveau de mesure et du niveau d'analyse.

5.2. Niveau de mesure et niveau d'analyse : la question de l'agrégation

Nous avons choisi de faire mesurer l'apprentissage de groupe par chacun des membres du groupe. En effet, si nous avons demandé une réponse de groupe, cela aurait pu conduire à une surestimation de l'apprentissage. Dans le cas d'une mesure individuelle, nous collectons l'avis de chaque membre du groupe sur l'apprentissage de son groupe ce qui permet d'avoir une vision plus nuancée et sans doute plus réelle de l'apprentissage. La question cruciale est alors celle de l'agrégation des données collectées pour obtenir un score d'apprentissage de groupe.

Plusieurs précautions sont à prendre lorsque l'on souhaite agréger des données de manière à placer la mesure à un niveau collectif (Goodman et al, 1987 ; Roberts, Hulin et Rousseau, 1978 ; Van de Ven et Ferry, 1980).

Reprenons chacune des recommandations et confrontons-les à notre mesure de l'apprentissage de groupe.

1. Une « théorie de la composition » (Rousseau, 1985) doit justifier l'agrégation. Comme nous l'avons expliqué, il existe deux principales justifications possibles : soit la signification de la variable est la même au niveau individuel et au niveau du groupe (Van de Ven et Ferry, 1980), soit la variable au niveau du groupe est considérée comme une « macro-perception » ou une vision partagée par le groupe (James, 1982, page 219). Dans le cas de l'apprentissage de groupe, la revue de la littérature insiste sur le fait que ce concept est différent de l'apprentissage individuel : le concept d'apprentissage n'a donc pas le même sens aux deux niveaux. En outre, l'apprentissage de groupe ne peut se réduire à la simple somme des apprentissages individuels. Ainsi, dans le cadre de notre mesure, nous nous plaçons dans la deuxième perspective, celle de la macro-perception, qui est cohérente avec la définition que nous avons retenue de l'apprentissage de groupe.
2. Une deuxième recommandation est que les items de mesure fassent référence au niveau auquel se positionne la recherche (Van de Ven et Ferry, 1980). Dans le développement de notre mesure et des items, nous prêterons donc attention à ce qu'ils fassent référence à ce niveau d'analyse. Ainsi, nous emploierons les termes « nous », « mon équipe », « notre équipe »¹ etc. Cette démarche permet de placer les répondants dans le cadre de leur groupe.

Une fois ces deux éléments précisés, il nous faut choisir la méthode correspondante pour agréger les données. Il existe en effet trois méthodes principales à notre disposition : la moyenne, la variance, le minimum/maximum.

¹ Nous employons le terme « équipe » pour le questionnaire, car le terme « groupe » se réfère pour les étudiants à leur groupe de TD et créait donc pour eux des confusions.

Pour choisir notre méthode d'agrégation, nous nous sommes basés sur l'idée que l'apprentissage de groupe est plus que la somme des apprentissages individuels ou du moins est autre chose. Nos items suivant la théorie de la composition choisie, vont mesurer la perception de l'apprentissage de groupe par chaque membre du groupe. Etant donné que nous avons défini l'apprentissage de groupe comme l'ensemble des connaissances partagées, nous avons retenu la méthode d'agrégation par le minimum. En effet, si l'on retient la perception de l'apprentissage de groupe la plus petite, on est sûr que les membres du groupe sont tous d'accord sur le fait que ce niveau d'apprentissage est acquis par leur équipe c'est-à-dire que cet apprentissage est effectivement partagé par les différents membres du groupe. Ensuite, certains membres peuvent penser que leur groupe a appris plus, mais en tout cas, le groupe a atteint en principe au moins le niveau le plus petit déclaré.

6. La construction de l'échelle de l'apprentissage de groupe

La première étape de la construction de la mesure est la définition du domaine du construit. Cette étape de définition est essentielle car elle conditionne la validité de contenu de l'échelle.

6.1. Définition du domaine du construit

Selon Churchill (1979), la spécification du domaine du construit se base sur la revue de la littérature. En outre, au vu du caractère nécessairement cumulatif de la recherche, il est préconisé de retenir une définition existante sauf si elle nécessite réellement d'être modifiée ou complétée.

Comme nous l'avons expliqué, nous calquons notre définition de l'apprentissage de groupe sur celle de l'apprentissage organisationnel. La définition théorique de l'apprentissage organisationnel retenue est celle de Koenig (1994). Cette définition englobe un certain nombre de phénomènes différents (acquisition, élaboration, etc.) et recouvre à la fois un processus et un résultat. L'opérationnalisation de l'apprentissage implique un certain nombre de choix et de précisions.

Tout d'abord, nous avons choisi de nous situer au niveau d'analyse du groupe. Ensuite, notre choix est de nous intéresser à l'aspect RESULTAT de l'apprentissage de groupe. Pour définir de manière opérationnelle ce résultat de l'apprentissage et avoir la matière et le vocabulaire nécessaires pour développer les items, nous avons repris les travaux fondateurs sur l'apprentissage et les travaux plus récents afin de savoir comment est qualifié le résultat de l'apprentissage organisationnel. Ainsi, le résultat de l'apprentissage consiste non seulement en de nouvelles compétences, mais également en une "théorie de l'action organisationnelle" (Argyris et Schön, 1978), ou encore une "connaissance organisationnelle" nouvelle (Duncan et Weiss, 1979; Lyles et Schwenk, 1992). Si ces conceptualisations semblent recouvrir des domaines différents, elles ont en réalité deux points communs mis à jour par Klimecki et Lassleben (1998) :

1. Elles s'appuient sur la notion traditionnelle de la connaissance, à savoir que la connaissance est une représentation de la réalité,
2. Elles se réfèrent à la connaissance organisationnelle comme une connaissance qui est partagée par les membres de l'organisation.

La notion de partage est un élément caractéristique de l'apprentissage organisationnel et de l'apprentissage de groupe.

Par ailleurs, concernant la distinction entre approche comportementale et cognitive, nous considérons, que ces deux approches sont complémentaires. Nous choisissons de dépasser cette dichotomie en optant pour une approche sociale de l'apprentissage organisationnel (Nicolini et Mezner, 1995) : les facettes comportementales et cognitives de l'apprentissage sont liées et l'apprentissage naît des interactions sociales dans un contexte précis. Ainsi, comme nous le verrons plus loin (partie 6.2.1.), notre contexte d'étude a également une influence sur l'opérationnalisation du concept et est un des objets de l'apprentissage.

En outre, pour préciser la nature de l'apprentissage organisationnel, nous soulignons qu'il existe différents niveaux d'apprentissage : l'apprentissage organisationnel ou de groupe peut être d'ampleur différente (simple boucle, double boucle et triple boucle). La connaissance partagée créée peut prendre plusieurs formes suivant les niveaux d'apprentissage (comme le montrent les travaux de Snell et Chak, 1998). En outre, nous considérons que ces différents niveaux peuvent coexister et forment un continuum (Snell et Chak, 1998)

6.2.La génération des items

L'étape de la génération des items se base sur l'étape de définition du domaine et également sur une phase qualitative permettant de saisir la réalité opérationnelle du construit que l'on veut mesurer.

6.2.1. La nécessaire prise en compte du contexte

Pour la génération de nos items, nous avons accordé une place importante à la réflexion sur le contexte de la mesure. En effet, l'apprentissage est un phénomène situé, c'est-à-dire relié à une tâche précise et à un cadre précis. Dans le cas de la construction d'une échelle de mesure de l'apprentissage de groupe, la prise en compte du contexte de l'étude apparaît cruciale. En effet, le contenu ou le résultat de l'apprentissage est relié au contexte de l'apprentissage. L'apprentissage résulte en des connaissances, des compréhensions et des cadres d'action ad hoc, c'est-à-dire ancrés dans l'activité authentique du travail (Brown et Duguid, 1991 et 2001).

Pour prendre en compte ce contexte, nous avons donc mené plusieurs études exploratoires qui, mises en perspective avec la définition du domaine, nous ont fourni les éléments opérationnels et le vocabulaire nécessaire pour générer les items. En outre, les études exploratoires permettent également de définir le niveau d'apprentissage de départ, que l'on peut fixer à zéro : les individus sont mis en groupe et cette « mise en groupe » au début du jeu définit le début de l'apprentissage du groupe.

6.2.2. Le développement des items

Comme il est recommandé par le paradigme de Churchill, nous avons généré un échantillon d'items en nous basant sur notre expérience du contexte (études exploratoires) et sur notre définition du domaine (revue de la littérature).

Nous avons ainsi formulé 73 items. Le nombre important d'items ainsi générés permet en principe d'atteindre une bonne fiabilité interne (Zaichkowsky, 1985; DeVellis, 1991).

Dans un second temps, nous avons effectué un retour sur la littérature et mené un pré-test visant à vérifier la « clarté » des items pour les répondants. Nous avons ainsi éliminé un certain

nombre d'items pour leur non pertinence par rapport à la définition retenue, nous avons reformuler d'autres items et formuler des items supplémentaires plus spécifiques au contexte.

Nous avons ensuite réalisé un pré-test afin d'apprécier la compréhension des items par les sujets soumis à la recherche. Pour ce pré-test, nous avons soumis le questionnaire à 6 équipes d'étudiants de l'Université du Maine participant au jeu Markstrat (ce premier pré-test nous a également servi aux pré-tests de nos autres échelles). Ce pré-test nous a conduit à modifier les items dont la formulation était incorrecte : à double signification, ou difficilement compréhensibles pour les répondants.

Finalement, nous avons conservé et formulé 43 items, ce qui est supérieur au nombre préconisé par Nunnally (1978). Les 43 items ont été soumis aux experts.

6.2.3. La consultation d'experts : l'adéquation des items à la définition du domaine

Afin de nous assurer de l'adéquation de nos items à la définition du construit retenue, nous avons mis en œuvre une consultation d'experts. Cette méthode est fréquemment utilisée (Zaichkowsky, 1985).

Suivant la procédure classiquement utilisée (Zaichkowsky, 1985), nous avons donc invité deux experts² à juger de la pertinence des énoncés par rapport à la définition de l'apprentissage organisationnel retenue. Les experts nous ont fourni une aide précieuse en jugeant les énoncés par rapport à la définition retenue selon les modalités suivantes :

Clairement représentatif du concept d'apprentissage de groupe	A peu près représentatif du concept d'apprentissage de groupe	Absolument pas représentatif du concept d'apprentissage de groupe
---	---	---

Nous avons ainsi pu nous assurer de la qualité de nos items : sur 43 items, 20 items sont clairement représentatifs de l'apprentissage organisationnel, 9 sont à peu près représentatifs, 13 ne sont absolument pas représentatifs.

La première étape (définition du construit) et la deuxième étape (génération des items) étant finalisées³, nous sommes passés à l'étape d'épuration conseillée par le paradigme de Churchill (1979).

6.3.L'épuration de la mesure

La méthode d'analyse employée pour épurer la mesure est l'analyse factorielle. L'analyse factorielle est une méthode statistique qui a pour objectif de résumer un grand nombre de variables en un petit nombre de facteurs dans le but de faciliter l'interprétation d'un nuage de points. Ainsi l'analyse factorielle permet d'analyser les dimensions sous-jacentes à un construit.

Les objectifs de cette première étude sont :

- dégager la structure factorielle de la mesure de l'apprentissage de groupe,
- épurer et réfléchir sur la dimensionnalité de la mesure : repérer (et éliminer) les dimensions non représentatives du construit,
- épurer et réviser l'échantillon d'items pour la deuxième étude exploratoire.

² Nous les remercions pour leurs précieuses réponses.

³ Remarque : la consultation d'experts a été menée en parallèle de l'étape d'épuration et de collecte des données

6.3.1. L'échantillon (étude 1)

L'étude a été menée sur 270 sujets répartis en 60 équipes. Les participants sont des étudiants en 1^{ère} année de Maîtrise de Sciences de Gestion. Deux équipes ont été exclues des traitements statistiques du fait de nombreuses variables manquantes lors de la saisie des questionnaires. L'étude a donc été menée sur 262 participants, soit 58 équipes.

Présentons maintenant les résultats de cette démarche d'épuration.

6.3.2. Résultats finals de l'épuration

Afin de faciliter la présentation et la lecture des résultats, seule la solution factorielle finale est indiquée.

Selon la dernière analyse factorielle avec rotation Varimax, après épuration et suivant la méthode complémentaire de l'alpha de Cronbach, nous mettons en évidence 3 composantes ou facteurs. Ces facteurs représentent les dimensions c'est-à-dire les manifestations distinctes de l'apprentissage prises en compte par la mesure. Les indices KMO et le test de Bartlett qui indiquent que les données peuvent être soumises à la méthode de l'analyse factorielle sont corrects (KMO=0,665 ; Bartlett=127,558, $p=0,000 < 0,05$). Le pourcentage de variance expliquée par les trois composantes ou dimensions mises en évidence par les épurations successives est satisfaisant (72,35%).

L'analyse des composantes ou des dimensions se fait par rapport à la contribution des items à cette composante. Les items donnent du « sens » à la dimension en question. Le *Tableau 3* de la matrice des composantes après rotation indique la contribution de chaque item à chacune des composantes, facteurs ou dimensions identifiées par l'analyse factorielle.

Tableau 3 Structure factorielle de l'échelle d'apprentissage de groupe

Variance Totale Expliquée = 72,358%	Composante		
	1	2	3
Nous sommes capables d'anticiper la réaction de l'environnement à nos choix stratégiques.	,874		
Nous savons comment atteindre nos objectifs.	,785		,265
Toutes les compétences clefs sont en place.	,761	,350	
Les membres de l'équipe ont une compréhension partagée des situations (du jeu, de la stratégie...)	-,248	,836	
Mon équipe a acquis des règles de fonctionnement partagées/admises par chacun de ses membres.	,186	,775	
Nous avons développé des cadres communs d'action.	,285	,729	,138
A ce stade du jeu, nous maîtrisons parfaitement le choix du circuit de distribution.			,895
A ce stade du jeu, nous maîtrisons parfaitement la fixation des objectifs publicitaires.	,116	,200	,851
Pourcentage de variance expliquée	34,672%	20,779%	16,907%
Alpha de Cronbach du facteur	0,8404	0,8069	0,6844
Significativité du test de Bartlett	0,000		
Indice KMO	0,665		

Après analyse des items, les trois dimensions mises en évidence sont présentées dans le *Tableau 4* :

Tableau 4 Dimensions de l'apprentissage de groupe suite à l'analyse factorielle

Dimension 1	Dimension « substantive »	Connaissances, compétences et savoirs qui permettent au groupe de réaliser sa tâche
Dimension 2	Dimension « procédurale »	Capacité à interpréter les informations au niveau du groupe afin de favoriser de nouveaux apprentissages
Dimension 3	Dimension « contextuelle »	Connaissances spécifiques au contexte et à la tâche réalisée.

Il convient maintenant de s'intéresser à la fiabilité de la mesure pour définir la structure finale de la mesure.

6.3.3. Fiabilité et structure factorielle finale de la mesure

L'alpha de Cronbach donne une estimation de la fiabilité en calculant la corrélation de l'ensemble des items du test à un autre échantillon d'items similaires provenant du même univers d'items. Plus ce coefficient est élevé, plus la fiabilité est grande et plus l'erreur de mesure est petite. La règle de décision est d'avoir un alpha qui soit compris entre 0,7 et 0,9 (Kline, 1993). Un coefficient faible indique que l'échantillon d'items ne capture pas correctement le construit. Les résultats finaux pour notre mesure sont les suivants (*Tableau 5*) :

Tableau 5 Structure finale de la mesure et fiabilité de la mesure de l'apprentissage de groupe

<i>Dimensions</i>	<i>Nombre d'items</i>	<i>Alpha de Cronbach</i>
Dimension 1 : Dimension substantive	3	0,8404
Dimension 2 : Dimension procédurale	3	0,8069
Dimension 3 : Dimension contextuelle	2	0,6844

Les alphas des dimensions 1 et 2 sont satisfaisants. L'alpha de la dimension 3 révèle une certaine faiblesse. Cependant, un alpha de Cronbach supérieur à 0,6 peut être acceptable dans le cas du développement d'échelle (Evrard, Roux et Pras, 2000 ; Nunnally, 1967).

Nous avons ainsi finalisé les deux premières étapes de la construction de l'échelle. Il nous reste maintenant à en estimer les qualités et à nous intéresser à sa validité et fiabilité afin d'en affiner la structure.

6.4. La validation de l'échelle

6.4.1. L'analyse factorielle confirmatoire

L'analyse factorielle confirmatoire permet de tester le bon ajustement de l'échelle d'apprentissage de groupe. Plus exactement, l'ajustement se réfère à la capacité d'un modèle ou d'une échelle à reproduire les données. Dans ce but, le logiciel AMOS 4.00 a été utilisé pour mener les différentes analyses factorielles confirmatoires. Il s'agit de confronter aux données collectées des hypothèses sur la structure des relations entre les variables mesurées ou observées (les items) et les variables latentes ou facteurs, c'est-à-dire les dimensions de l'apprentissage de groupe.

Pour mettre en place notre analyse factorielle confirmatoire nous avons suivi les cinq phases recommandées par Bollen et Long (1993) : spécification du modèle, identification du modèle, estimation du modèle, test de son ajustement, re-spécification du modèle.

6.4.2. L'échantillon (étude 2)

Une collecte de données a été menée pour procéder à l'analyse confirmatoire de la mesure de l'apprentissage de groupe. Cette étude 2 a été menée sur 237 sujets répartis en 60 équipes. Les participants sont des étudiants en 1^{ère} année de Maîtrise de Sciences de Gestion de l'Université Paris IX Dauphine.

Les conditions de passation du questionnaire sont similaires à l'étude 1. Le libellé des items entre cette étude et l'étude 1 est inchangé. Ainsi, nous avons agrégé les observations de ces deux études afin d'avoir suffisamment de données pour mener l'analyse factorielle confirmatoire. L'analyse a donc été menée sur 118 équipes.

6.4.3. Les résultats

L'ESTIMATION ET L'ADÉQUATION DU MODÈLE

Les paramètres estimés par le logiciel AMOS ainsi que les indices permettant de juger de la qualité de l'ajustement du modèle. Certains indicateurs sont significatifs et suggèrent un bon ajustement du modèle (GFI, RMR, CAIC). Cependant, un certain nombre d'indicateurs restent limites par rapport aux seuils de décision (AGFI, IFI, NFI) et surtout certains indicateurs révèlent un problème d'ajustement du modèle (RMSEA). L'indicateur RMSEA indique un problème de présence de résidus. Ce résultat et les contradictions entre les indices pour juger de l'ajustement, laisse penser qu'une respécification du modèle pourrait être nécessaire.

LA RESPÉCIFICATION ET L'ADÉQUATION DU MODÈLE FINAL

L'étude des résidus standardisés laisse apparaître le problème. En effet, la valeur prise par le résidu entre aoco22_1 et aores_12 est supérieure à $|2,54|$, seuil recommandé par Steenkamp et Van Trijp (1991). Cette valeur indique donc le besoin de respécifier le modèle au niveau de ces variables. A partir de ces différents indices et des indices de modifications, nous choisissons de respécifier le modèle en éliminant l'item aoco22_1 (« Toutes les compétences sont en place »). D'un point de vue qualitatif, le caractère très affirmatif et radical de l'item explique peut-être qu'il soit la source d'une forte erreur aléatoire et ne lui permet donc pas de saisir toute la variance possible du construit.

Tous les indices d'ajustement de l'analyse factorielle confirmatoire menée le modèle respécifié c'est-à-dire sur les 7 items restants sont améliorés et témoignent d'un bon ajustement du modèle autant de manière absolue que de manière relative par rapport au modèle de 8 items.

6.4.4. La nouvelle structure de la mesure de l'apprentissage de groupe

Une nouvelle analyse factorielle est alors conduite sous SPSS sur les 7 items retenus suite à l'analyse confirmatoire. Cette nouvelle analyse conduit à fusionner les dimensions contextuelle et substantive. Cela est qualitativement justifié : en effet les items de la dimension « contextuelle », sont relatifs à la connaissance acquise lors du jeu (par exemple : «A ce stade du jeu, nous maîtrisons parfaitement la fixation des objectifs publicitaires »), au même titre que les items de la dimension « substantive » (par exemple : « Nous savons comment atteindre nos objectifs »).

Nous optons donc pour la structure finale de la mesure comprenant deux dimensions : une dimension « procédurale » et une dimension « substantive ». La structure issue de l'analyse factorielle sur 7 items est la suivante (Tableau 6) :

Tableau 6 Structure finale de l'échelle d'apprentissage de groupe

(Matrice des composantes après rotation)

Variance Totale Expliquée = 56,850%	Composante	
	1	2
A ce stade du jeu, nous maîtrisons parfaitement la fixation des objectifs publicitaires.	,747	,116
A ce stade du jeu, nous maîtrisons parfaitement le choix du circuit de distribution.	,739	
Nous sommes capables d'anticiper la réaction de l'environnement à nos choix stratégiques.	,712	
Nous savons comment atteindre nos objectifs.	,706	,282
Les membres de l'équipe ont une compréhension partagée des situations (du jeu, de la stratégie,...)		,839
Mon équipe a acquis des règles de fonctionnement partagées/admises par chacun de ses membres.		,731
Nous avons développé des cadres communs d'action.	,225	,696
Pourcentage de variance expliquée	34,689%	22,161%
Alpha de Cronbach du facteur	,7131	,6433
Significativité du test de Bartlett	,000	
Indice KMO	,674	

Les indicateurs restent satisfaisants. L'alpha de la dimension 2 est cependant assez faible. Mais d'un point de vue qualitatif, cette structure nous semble plus justifiée et plus pertinente. C'est pourquoi, nous privilégions les critères qualitatifs aux critères statistiques et nous conservons cette structure bidimensionnelle. Finalement, l'échelle d'apprentissage de groupe retenue est la suivante (Encadré 1) :

Encadré 1 L'échelle finale d'apprentissage de groupe

<p>Dimension A : Dimension « substantive »</p> <p>1A. Nous sommes capables d'anticiper la réaction de l'environnement à nos choix stratégiques. (aoeten4)</p> <p>2A. Nous savons comment atteindre nos objectifs. (aoeten2)</p> <p>3A. A ce stade du jeu, mon équipe maîtrise parfaitement la fixation des objectifs publicitaires. (aoco15-5)</p> <p>4A. A ce stade du jeu, mon équipe maîtrise parfaitement le choix du circuit de distribution. (aoco15-8)</p> <p>Dimension B : Dimension « procédurale »</p> <p>1B. Les membres de l'équipe ont une compréhension partagée des situations (du jeu, de la stratégie,...). (aoesu2)</p> <p>2B. Mon équipe a acquis des règles de fonctionnement partagées/admises par chacun de ses membres. (aoesu3)</p> <p>3B. Nous avons développé des cadres communs d'actions. (aoesu12)</p>

Nous nous intéressons maintenant aux qualités de cette échelle de mesure (fiabilité et validité).

7. Les qualités de l'échelle de mesure

Un instrument de mesure doit remplir deux critères essentiels : la fiabilité et la validité. Un instrument est dit « fiable » si, suite à des utilisations répétées de l'instrument de mesure, on obtient le même résultat. Un instrument de mesure est valide s'il mesure bien ce que l'on cherche à mesurer. Nous allons maintenant présenter les outils disponibles pour juger de la fiabilité et de la validité de notre échelle de mesure, ainsi que les résultats obtenus.

7.1. La fiabilité de l'échelle

Pour juger de la fiabilité de l'échelle de mesure construite, nous avons retenu la méthode la plus utilisée et recommandée par Churchill (1979), à savoir le calcul de l'alpha de Cronbach, la corrélation inter-items et le rho de Joreskog.

Nous avons présenté précédemment les premiers résultats concernant la fiabilité au travers du calcul des alphas de Cronbach. Concernant la corrélation inter-items, Bearden et al (1993) suggèrent que des corrélations inter-items supérieures à 0,3 sont très satisfaisantes. L'analyse factorielle confirmatoire nous permet d'estimer la fiabilité par le rho de Joreskog. L'intérêt de cet indicateur est qu'il est indépendant du nombre d'items présents dans l'échelle contrairement à l'alpha de Cronbach.

Les résultats obtenus sont satisfaisants et témoignent que la mesure est fiable (*Tableau 7*). Comme l'explique Vernet (1991), des coefficients proches de 0,7 sont tout à fait convenables pour une création d'échelle. De plus, Nunnally (1967) ainsi qu'Evrard et al (2000) estiment qu'un alpha supérieur à 0,6 est acceptable pour une recherche exploratoire.

Tableau 7 *Fiabilité de l'échelle d'apprentissage de groupe*

	<i>Dimension 1</i>	<i>Dimension 2</i>
Alpha de Cronbach	0,7131	0,6433
Corrélations inter-items	0,3832	0,3760
Rho de Joreskog	0,7131	0,7708

7.2. La validité de la mesure

La validité s'intéresse à l'écart entre la mesure obtenue et la vraie valeur de la mesure, ce qui revient à répondre à la question : « L'échelle de mesure mesure-t-elle ce qu'elle est supposée mesurer ? ». Répondre à cette question est difficile puisque ce que l'on mesure n'est pas directement observable. C'est pour cette raison que l'on décompose la validité en différentes formes qui constitueront autant de sous-questions. Le processus de validation consiste alors à accumuler des preuves.

Classiquement on distingue 3 types de validité (Peter, 1981 ; Evrard et al, 2000) : la validité « faciale » ou de contenu, la validité de trait ou de construit et la validité nomologique.

7.2.1. La validité « faciale » ou de contenu

Elle consiste à estimer dans quelle mesure les différents aspects de la mesure sont représentatifs du concept étudié. Cette validité est fondée sur le jugement du chercheur et de ses pairs. La qualité de la validation de contenu va dépendre fortement de la précision de définition du concept mesuré et de l'accord des experts avec l'opérationnalisation de cette définition (étape de la construction des items). Compte tenu de la procédure suivie pour le développement des items : étude exploratoire, consultation des personnes de la population étudiée, consultation d'experts, nous pensons avoir garanti la validité de contenu.

7.2.2. La validité de trait ou de construit

Il s'agit ici de s'assurer que la mesure mesure bien ce qu'elle est censée mesurer en s'assurant que l'on mesure bien un phénomène et rien que ce phénomène. Pour cela, on décompose la validité de trait en validité convergente et validité discriminante (Zaltman, Pinson et Angelmar, 1973) :

- la validité convergente : les indicateurs censés mesurer le même phénomène doivent être corrélés⁴,
- la validité discriminante : les indicateurs censés mesurer des phénomènes différents ou des facettes distinctes du même phénomène doivent être faiblement corrélés (Evrard et al, 2000).

Aucune échelle d'apprentissage de groupe n'existant, nous avons pensé tester la validité de trait en observant les corrélations entre les scores obtenus grâce à l'échelle développée et le score obtenu à un test de connaissance sur le jeu Markstrat. Suivant notre définition l'apprentissage de groupe est un concept distinct de la « somme des apprentissages individuels ». Ainsi, tester les corrélations entre les scores sur les dimensions substantive et procédurale avec le score de bonnes réponses, illustrerait la validité discriminante de l'échelle d'apprentissage de groupe. Les corrélations obtenues sont effectivement extrêmement faibles mais ne sont pas significatives ce qui ne nous permet pas de conclure.

En outre, toujours concernant la validité discriminante, les faibles corrélations entre les dimensions obtenues par l'analyse factorielle confirmatoire témoignent d'une bonne validité à ce niveau.

7.2.3. La validité nomologique

La validité nomologique concerne les liens entre concepts. Elle consiste à savoir si les relations entre la mesure du concept étudié et la mesure d'autres concepts sont en conformité avec les prédictions issues de la théorie. Cette forme de validité peut être établie au travers du test d'un modèle conceptuel.

⁴ Lorsque cette corrélation est supérieure à 0,7, on peut dire que la mesure a une bonne validité convergente. (Kline, 1993)

CONCLUSION

Dans cet article, nous avons présenté la mesure de l'apprentissage de groupe que nous avons construite dans le cadre du jeu de simulation Markstrat. La validité externe de l'échelle proposée est faible étant donné que la mesure est contextualisée. Cependant, la démarche méthodologique exposée propose une solution de mesure de l'apprentissage de groupe et constitue une alternative aux mesures existantes telles que la mesure par la performance par exemple. En outre la mesure obtenue met en évidence une structure claire à deux dimensions (substantive et procédurale) qui aidera sans doute à la construction future d'échelles de mesure de l'apprentissage de groupe ou de l'apprentissage organisationnel.

Références

- Anzieu, D. and J.-Y. Martin (1968), *La dynamique des groupes restreints*, Paris: Presses universitaires de France.
- Argyris, C. and D.A. Schön (1978), *Organizational learning: a theory of action perspective*, Reading: Addison-Wesley.
- Bearden, W.O., Netemeyer, R.G. and M.F. Mobley (1993), *Handbook of marketing scales: Multi item measures for marketing and consumer behavior research*, Newbury Park: Sage.
- Bollen, K. A. and J.S. Long (1993) *Testing structural equation models*, Newbury Park: Sage.
- Bourgeon, L. and J.-C. Tarondeau (2000), L'apprentissage dans les organisations transversales : le cas de l'organisation des projets de R&D, In B. Quélin and J.-L. Arrègle (Eds.), *Le management stratégique des compétences*, Paris: Ellipse, 287-346.
- Brown, J. and P. Duguid (2001), Knowledge and Organization: A Social-Practice Perspective, *Organization Science*, 12:2, 198-213.
- Brown, J. and P. Duguid (1991), Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation, *Organization Science*, 2:1, 40-57.
- Cangelosi, V. and W.R. Dill (1965), Organizational Learning: Observations toward a Theory, *Administrative Science Quarterly*, 10:2, 175-203.
- Carley, K. (1992), Organizational learning and personnel turnover, *Organization Science*, 3:1, 20-46.
- Churchill, G. A. (1979), A paradigm for developing better measures of marketing constructs, *Journal of Marketing Research*, 16:1, 64-73.
- Cyert, R. M., and J.G. March (1963), *A behavioral theory of the firm*, Englewood Cliffs: Prentice-Hall.
- DeVellis, R. F. 1991. *Scale development: theory and applications*. Newbury Park: Sage.
- Duncan, R. B. and A. Weiss (1979), Organizational Learning: Implications for organizational design, In B. Staw, & L. L. Cummings (Eds.), *Research in organizational behavior*, Greenwich: JAI Press, Vol. 1: 75-123.
- Evrard, Y., Pras, B and E. Roux (2000), *Market : Etudes et recherches en marketing, Fondements, Méthodes* (2e ed.), Paris: Dunod.
- Goodman, P. S., Ravlin, E. C. and L. Argote (1987), Understanding Groups in Organizations, In B. M. Staw and L. L. Cummings (Eds.), *Research in organizational behavior: an annual series of analytical essays and critical reviews*, Greenwich: JAI Press, Vol. 9: 1-71.
- James, L. R. (1982), Aggregation bias in estimates of perceptual agreement, *Journal of Applied Psychology*, 67:2, 219-229.
- Klimecki, R. and H. Lassleben (1998), Modes of Organizational Learning, *Management Learning*, 29:4, 105-130.
- Kline, P. (1993), *Personality: the psychometric view*, London: Routledge.
- Koenig, G. (1994), L'apprentissage organisationnel : repérages des lieux, *Revue Française de Gestion*, 97:Janvier-Février, 76-83.
- Miller, D.C (1991), *Handbook of research design and social measurement* (5th ed.), Newbury Park: Sage.
- Nicolini, D. and M.B. Mezner (1995), The Social Construction or Organizational Learning: Conceptual and Practical Issues in the Field, *Human Relations*, 48:7, 727-746.
- Nunnally, J. C. (1967), *Psychometric theory*, New York: McGraw-Hill.
- Nunnally, J. C. (1978), *Psychometric theory* (2d ed.), New York: McGraw-Hill.
- Peter, J. P. (1981), Construct validity: A review of basic issues and marketing practices, *Journal of Marketing Research*, 18:2, 133-145.

- Roberts, K. H., Hulin, C. L. and D.M. Rousseau (1978), *Developing an interdisciplinary science of organizations*, San Francisco: Jossey-Bass.
- Rossiter, J. R. (2002), The COARSE procedure for scale development in marketing, *International Journal of Research in Marketing*, 19:4, 305-335.
- Rousseau, D. M. (1985), Issues of level in organizational research: Multi-level analysis and cross-level perspectives, In B. M. Staw, & L. L. Cummings (Eds.), *Research in organizational behavior: an annual series of analytical essays and critical reviews*, Greenwich: JAI Press, Vol. 7: 1-37.
- Shrivastava, P. (1983), A Typology of Organizational Learning Systems, *Journal of Management Studies*, 20:1, 7-29.
- Snell, R. and A.M.-K. Chak (1998), The Learning Organization: Learning and Empowerment of Whom ?, *Management Learning*, 29:3, 337-364.
- Steeckamp, J. B. and M. Van Trijp (1991), Omission and commission in judgement and choice, *Journal of Experimental Social Psychology*, 27:1, 76-105.
- Van de Ven, A. H. and D.L. Ferry (1980), *Measuring and assessing organizations*, New York: Wiley.
- Weick, K. E. (1991), The Nontraditional Quality of Organizational Learning, *Organization Science*, 2:1, 116-124.
- Wildt, A. R. and M.B. Mazis (1978), Determinants of scale response: Label versus position, *Journal of Marketing Research*, 15:2, 261-267.
- Zaichkoswsky, J. L. (1985), Measuring the involvement inventory, *Journal of Consumer Research*, 12:4, 341-352.
- Zaltman, G., Pinson, C., and R. Angelmar (1973), *Metatheory and consumer research*, New York: Holt Rinehart and Winston.