

Le recueil et les préalables à l'analyse des données lors de la phase confirmatoire

Les chercheurs intéressés par la relation de l'entreprise avec son environnement, plus spécifiquement dans le cadre du champ de la stratégie, se basent sur des approches cognitives et discursives oscillant entre le déterminisme et le volontarisme (Bourgeois, 1984) et entre le fonctionnalisme et l'interprétatif (Rouleau et Séguin, 1995). Du point de vue méthodologique, la tendance ou le choix se font entre le quantitatif et le qualitatif (Duncan, 1972). Or pour notre recherche, nous avons décidé d'associer les deux méthodes pour les raisons que nous avons évoquées plus haut.

L'approche confirmatoire vise à établir des relations de cause à effet ayant trait à un phénomène social. Partant d'hypothèses, le chercheur tente de dégager à partir de questionnaires, inventaires, analyses démographiques voire d'autres sources, des données quantitatives qu'il soumet à une analyse statistique (Naguib, 2004). Les résultats de l'analyse servent alors à confirmer ou infirmer des relations supposées entre les variables.

L'examen de la littérature et les constats de la phase exploratoire nous ont permis d'avancer des hypothèses de recherche sur les relations entre les comportements stratégiques des PME et leur performance dans un environnement incertain. Ainsi, à travers

cette phase de test, notre objectif est de mieux élucider les différents éléments de notre problématique. Pour bien comprendre notre démarche, nous présentons dans ce qui suit : la construction du questionnaire et le déroulement de l'enquête, les modalités de sélection et de choix des items. Ensuite, nous spécifions les préalables à l'analyse des données lors de la phase confirmatoire.

2-1- La construction du questionnaire et le déroulement de l'enquête

Le champ de la stratégie s'est en premier lieu développé par le biais des recherches empiriques qui partent de l'étude d'un phénomène concret (les pratiques d'administration) vers la théorie (Barnard, 1938 ; Chandler, 1962 ; Crozier, 1963). Toutefois, sous l'influence de certaines disciplines notamment l'économie (Porter, 1980), les mathématiques (Ansoff, 1965) et la cybernétique (Simon, 1945), l'intérêt pour le lien entre la stratégie et la performance des entreprises s'est accentué. L'emphase est ainsi mise davantage sur le contenu que sur le processus. Ceci s'explique par le fait que le contenu se prête plus facilement à la modélisation et à la quantification (Naguib, 2004). Dans ce cadre, les analyses statistiques et la collecte des données à travers des questionnaires et éventuellement des entrevues sont les méthodes les plus utilisées.

Pour générer et rédiger notre questionnaire, le point de départ est le guide d'entretien de la phase exploratoire. Comme expliqué plus haut, ce guide aborde les principaux points soulevés dans notre recherche. Pour ce faire, nous avons retravaillé certaines questions en les adaptant aux objectifs établis à travers nos hypothèses de recherche et nous avons ajouté d'autres qui nous semblent pertinentes. C'est le cas notamment des questions liées aux variables culturelles. Nous avons également décidé d'en supprimer certaines. Nous pouvons par exemple citer la suppression de la question sur la performance financière qui selon nous n'est pas bien renseignée par les interviewés et qui a peu d'intérêt pour nos objectifs de recherche. L'élaboration du questionnaire exige une rigueur scientifique et un soin extrême. La construction de la première version s'est étalée de Décembre 2014 à Mars 2015, soit un peu plus de 3 mois. Entre le projet initial du questionnaire, et la version finale, six mois se sont écoulés et cinq moutures différentes se sont succédées.

Une fois terminée la troisième mouture, nous avons effectué un pré-test du questionnaire au près d'un échantillon réduit de 20 PME camerounaises et nous avons pu constater un certain nombre de problèmes que posait notre questionnaire. Ainsi, certaines questions étaient mal comprises, les grands thèmes du questionnaire n'étaient pas clairement présentés, la codification n'était pas claire en ce qui concernait notamment les échelles. Après

corrections de ces insuffisances, nous avons procédé à l'administration de la quatrième version du questionnaire au près des PME du secteur de la distribution du matériel informatique, qui constituent notre base de sondage au Cameroun. Après avoir obtenu 84 réponses, nous avons procédé au test du questionnaire grâce au logiciel SSPS version 20, pour apprécier la pertinence et la fidélité des échelles de mesure (alpha de Cronbach, analyse factorielle). Nous avons alors pu remarquer qu'il existait quelques valeurs manquantes du fait de la mauvaise formulation des questions liées aux comportements stratégiques, ce qui nous a permis de corriger une dernière fois le questionnaire, et de procéder à une deuxième collecte des données au près des petites et moyennes entreprises du secteur de la distribution du matériel informatique.

Le questionnaire a été administré en face à face par nos soins pour limiter les erreurs de remplissage et éventuellement apporter des éclaircissements aux répondants sur les questions mal appréhendées. Malgré le refus de certains dirigeants de participer à l'enquête, au total nous avons pu administrer 202 questionnaires dans les quatre grandes régions du Cameroun (Douala, Yaoundé, Bafoussam et Garoua). Pour y arriver, dans chaque région, une première descente sur le terrain nous a permis d'identifier les PME qui constituent notre base de sondage et de prendre rendez-vous pour l'administration du questionnaire. Nous avons pu remarquer que dans chaque région, les entreprises concernées par notre enquête étaient installées presque toutes dans un même périmètre, ce qui nous a encore facilité la tâche. En France, nous avons obtenu 101 questionnaires bien renseignés, dans les régions Hauts de France (Nord-pas-de Calais et Picardie), et Parisienne.

Notre questionnaire est bâti à partir de cinq thèmes cohérents. Il s'appuie sur un plan qui est agencé non pas selon l'ordre croissant d'émission des hypothèses, mais en respectant deux critères. Le premier est dicté par une progressivité qui implique d'aller du plus simple au plus compliqué. Le second est imposé par une succession logique dans l'ordre des questions. Au début, l'interviewé doit se familiariser avec le sujet de façon générale, avant de réfléchir et se prononcer sur des questions plus difficiles. Ainsi par exemple, nous avons délibérément laissé en dernier les questions relatives à la performance perçue et aux comportements stratégiques. L'ensemble du questionnaire est joint en annexe. Il compte 18 questions avec 63 items et a 7 pages. Nous avons tenté d'atténuer l'effet « optique » de longueur en mettant en évidence la numérotation par partie (au nombre de 5), pour éviter éventuellement de décourager les répondants :

- La première partie est consacrée aux caractéristiques internes de l'entreprise
- La deuxième partie traite de l'incertitude environnementale perçue

- La troisième partie aborde les variables culturelles
- La quatrième partie porte sur la performance perçue de l'entreprise
- La cinquième partie est destinée aux comportements stratégiques

Un message introductif portant des éléments de motivations suivants accompagne le questionnaire :

- Le cadre doctoral de l'enquête
- La garantie de la confidentialité et de l'anonymat des réponses
- Le nom de l'université et du laboratoire de recherche
- L'usage qui sera fait des résultats de la recherche

A la fin du questionnaire, on retrouve encore quelques éléments de motivations, notamment :

- Les remerciements pour la participation
- L'engagement de transmettre à ceux qui le désirent les résultats de l'enquête par courrier électronique.

L'ensemble de l'enquête par questionnaire, dans ses phases « administration », « test » et « recueil des réponses » au Cameroun, a duré 6 mois (Avril - Septembre 2015). En France, la collecte des données a duré 5 mois (Avril-Août 2016).

Par ailleurs, en l'absence d'une liste exhaustive des PME exerçant dans le secteur de la distribution du matériel informatique tant dans l'administration publique comme privée au Cameroun, nous nous sommes orientés vers une méthode d'échantillonnage non probabiliste, qui repose sur le jugement personnel du chercheur plutôt que sur le hasard pour la sélection des éléments de l'échantillon. Dans cette approche, la responsabilité incombe au chercheur de décider de manière arbitraire ou consciente des individus à interroger ; ces échantillons peuvent donner de bonnes estimations des caractéristiques de la population (Malhotra et al., 2007, p. 265). La méthode d'échantillonnage utilisée ici est donc non probabiliste.

L'utilisation de la méthode d'échantillonnage probabiliste n'est pas évidente dans les recherches en Afrique. En effet, il est difficile d'avoir accès à une base de sondage car les économies africaines sont dépourvues de listes, d'adresses et de statistiques socio-économiques fiables (Ouattara, 2003). Le Cameroun ne fait pas exception à cette règle, d'où notre choix d'opter pour un échantillon de jugement. L'échantillon de jugement se veut représentatif dans la mesure où le chercheur va interroger les individus les plus susceptibles d'éclairer et d'apporter une information pertinente sur le problème à traiter.

2-1-A- Les modalités de sélection et de choix des items

Les entretiens effectués lors de la phase exploratoire ont été très utiles pour la formulation des questions. En outre, la majorité des échelles de mesure de notre questionnaire est proposée par d'autres chercheurs. En effet, nous avons essayé, dans la mesure du possible, de nous appuyer sur des échelles et items déjà employés dans d'autres recherches ayant étudié les concepts mobilisés dans notre recherche. En effet, recourir à des instruments de mesure validés lors de recherches précédentes constitue un appui méthodologique important et assure au questionnaire un certain niveau de validité et une bonne qualité de mesure (Baumard et al., 2007). Notre revue de la littérature nous a ainsi permis de recenser les différents auteurs significatifs ayant préalablement travaillé sur nos concepts, avec le nombre d'items ainsi que leurs mesures comme le montre le tableau suivant :

Tableau 9: Les auteurs significatifs retenus pour l'élaboration du questionnaire

Construits	Nombre d'items par construit (proposé-théorique)	Auteurs significatifs	Auteurs choisis pour la mesure ; α de Cronbach
Incertitude environnementale perçue			
Turbulence perçue de l'environnement		Emery et Trist (1965) ; Lawrence et Lorsch (1967) ; Burns et Stalker (1971) ; Aldrich (1979) ; Gueguen (2001)	-Tsai et Yang (2013) ($\alpha=0.76$) - Zhou, Yim et Tse (2005) ($\alpha=0.73$) - Jaworski et Kohli, (1993) ($\alpha= 0.88$)
	6 items	Jaworski et Kohli, (1993) ($\alpha=0.68$ et 0.88)	
	2 items	Tsai et Yang (2013) ($\alpha=0.76$)	
	2 items	Zhou, Yim et Tse (2005) ($\alpha=0.73$)	
Imprévisibilité perçue de l'environnement		Knight (1921) ; Luce et Raiffa (1957) ; Lawrence et Lorsch (1967) ; Gueguen (2001)	- Sogbossi, (2009) (t=5,36 ; p = 5%)
	6 items	Duncan, (1972)	
	3 items	Sawyer, (1993) (Scores moyens=74.63 ; 68.16 ; 38.36)	
	3 items	Sogbossi, (2009) (t=5,36 ; p = 5%)	
Dynamisme perçue de		Emery et Trist (1965) ;	- Li et Liu, (2014)

l'environnement		Thomson (1967) ; Terreberry (1968) ; Miles, Snow et Pfeffer (1974); Gueguen (2001)	$(\alpha=0.74)$ - Sogbossi, (2009) ($t=22,03$; $p = 5\%$)
	6 items	Duncan, 1972 ($t= 3,453$; $p <0.01$)	
	5 items	Dess et Beard, 1984 ($\alpha=0.610$)	
	8 items	Achrol and Stern (1988) ($\alpha=0.710$)	
	8 items	Tan et Litschert, (1994) (Score moyen=5.84)	
	3 items	Zhou, Yim et Tse, (2005) ($\alpha=0.74$)	
	6 items	Gotteland, Haon, Ray et Boulé, (2008) ($\alpha=0.66$ et 0.67)	
	3 items	Sogbossi, (2009) ($t=22,03$; $p = 5\%$)	
	4 items	Li et Liu, (2014) ($\alpha=0.74$)	
Complexité perçue de l'environnement		Emery et Trist (1965) ; Thomson (1967) ; Terreberry (1968) ; Gueguen (2001)	- Sutcliffe et Huber, (1998) ($\alpha=0.60$) - Sogbossi, (2009) ($t=33,21$; $p = 5\%$) - Gotteland, Haon, Ray et Boulé, (2008) ($\alpha=0.64$ et 0.79)
	6 items	Duncan (1972) ($t= 4,388$; $p <0.001$)	
	5 items	Dess et Beard, (1984) ($\alpha=0.657$)	
	8 items	Tan et Litschert, (1994) (Score moyen=5.18)	
	5 items	Sutcliffe et Huber, (1998) ($\alpha=0.60$)	
	6 items	Gotteland, Haon, Ray et Boulé, (2008) ($\alpha=0.64$ et 0.79)	
	2 items	Sogbossi, (2009) ($t=33,21$; $p = 5\%$)	

Les auteurs significatifs pour la mesure de la culture

Construits	Nombre d'items par construit (proposé- théorique)	Auteurs signifiants	Auteurs choisis pour la mesure ; α de Cronbach
Culture			
Individualisme/Collectivisme	8 items	Hofstede, (1980, 2001)	- Srite et Kaharanna, (2006) ($\alpha=0.79$) Hofstede, (1980, 2001)
	4 items	Trompenaars, (1993)	
	10 items	Earley, (1993)	
	24 items	Singelis, (1994)	
	24 items	Triandis, (1995)	
	20 items	Gaines et al., (1997)	

	24 items	Triandis et Singelis, (1998)	
	12 items	(Yi, 2004)	
	6 items	Srite et Kaharanna, (2006) ($\alpha=0.79$)	
Masculinité/Féminité	8 items	Hofstede, (1980, 2001)	- Srite et Karahanna, (2006) ($\alpha=0.92$) Hofstede, (1980, 2001)
	9 items	Dorfman et Howell,(1988)	
	4 items	Furrer et al, (2000)	
	4 items	Avsec, (2003)	
	4 items	Vitell et al., (2003)	
	5 items	Srite et Karahanna, (2006) ($\alpha=0.92$)	
Distance hiérarchique	7 items	Hofstede, (1980, 2001)	- Srite et Kaharanna, (2006) ($\alpha=0.74$) Hofstede, (1980, 2001)
	5 items	Erez et Earley, (1987)	
	6 items	Dorfman et Howell, (1988)	
	4 items	Brockner et al., (2001)	
	7 items	Srite et Kaharanna, (2006) ($\alpha=0.74$)	
Contrôle de l'incertitude	7 items	Hofstede, (1980, 2001)	- Srite et Kaharanna, (2006) ($\alpha=0.80$) Hofstede, (1980, 2001)
	5 items	Dorfman et Howell, (1988)	
	7 items	Yoo, (1996)	
	5 items	Ang et al., (2003)	
	5 items	Vitell et al., (2003)	
	6 items	Srite et Kaharanna, (2006) ($\alpha=0.80$)	
Orientation long/court terme	4 items	Hofstede, (1980, 2001)	- Noorderhaven et Tidjani,(2001) ($\alpha=0.69$) - House, Hanges, Javidan, Dorfman, et Gupta, (2004) ($\alpha=0.80$) Hofstede, (1980, 2001)
	4 items	Sanders, (1987)	
	8 items	Hofstede et Bond, (1988)	
	1 item	Trompenaars, (1993)	
	4 items	Furrer et al., (2000)	
	5 items	Noorderhaven et Tidjani, (2001) ($\alpha=0.69$)	
	8 items	House, Hanges, Javidan, Dorfman, et Gupta, (2004) ($\alpha=0.80$)	
	6 items	Mc Guire et al., (2006)	

Les auteurs significatifs pour la mesure de la performance et du comportement stratégique

Construits	Nombre d'items par construit (proposé-théorique)	Auteurs signifiants	Auteurs choisis pour la mesure ; α de Cronbach
Performance			
Performance financière et	7 items	Ruekert, Walker, et	- Song et Parry,

non financière		Roering, (1985)	(1997) ($\alpha=0.93$)
	3 items	Ramanujam Venkatraman, (1987)	
	7 items	Irving, (1995)	
	7 items	Song et Parry, (1997) ($\alpha=0.93$)	
	7 items	Homburg, Krohmer et Workman, (1999) ($\alpha=0.73$ et 0.72)	
	3 items	Papke-Shields, Malhot et Grover, (2006)	
	3 items	Shrivastava, Mohanty, et Lakhe, (2006)	
	3 items	Rudd, Greenley Beatson, et Lings, (2008)	
Comportement stratégique			
Spécialisation	7 items	Dess et Davis (1984) (F=15,88, p=0.001)	- Wirtz, Mathieu et Schilke (2007) ($\alpha=0.74$)
	1 item	Baum et Oliver, (1996)	
	1 item	Stuart, (1998)	
	2 items	Echols et Tsai (2005)	
	4 items	Wirtz, Mathieu et Schilke (2007) ($\alpha=0.74$)	
Flexibilité		Lawrence et Lorsch (1967) ; Ansoff (1968) ; Burns et Stalker (1971) ; Drucker (1971) ; Mintzberg (1975) ; Eppink (1978) ; Jones et Ostroy (1984)	- Zhang (2006) ($\alpha=0.86$) - Tamayo et al., (2010) ($\alpha=0.84$)
	5 items	Volberda, (1998)	
	4 items	Grewal et Tansuhaj (2001) ($\alpha=0.77$)	
	5 items	Young-Ybarra et Wiersema, (1999) ($\alpha=0.81$)	
	4 items	Dröge et al., (2005)	
	5 items	Verdú et al., (2006)	
	8 items	Zhang (2006) ($\alpha=0.86$)	
	6 items	Ling-yeet et Ogunmokun (2008) ($\alpha=0.77$)	
5 items	Tamayo et al., (2010) ($\alpha=0.84$)		
Innovation	4 items	Booz Allen et Hamilton, (1982)	- Zortea-Johnston, Darroch et Matear, (2012) ($\alpha=0.71$ et 0.74)
	2 items	Cooper, (1985)	
	4 items	Eliashberg Robertson, (1988)	
	1 item	Deshpandé et al., (1993)	
	4 items	Gatignon et Xuereb, (1997)	
	2 items	Kumar et al., (2000)	
	2 items	Chandy et Tellis, (2000)	

	8 items	Zhou, Yim et Tse, (2005) ($\alpha=0.74$ et 0.73)	
	7 items	Zortea-Johnston, Darroch et Matear, (2012) ($\alpha=0.71$ et 0.74)	
	5 items	Tsai et Yang, (2013) ($\alpha=0.71$)	
Coopération	15 items	Pinto, Pinto et Prescott (1993) ($\alpha=0.82$)	- Wirtz, Mathieu et Schilke, (2007) ($\alpha=0.90$)
	6 items	Olson, Walker, et Ruekert (1995)	
	7 items	Bensaou (1997) ($\alpha=0.79$ et 0.85)	
	9 items	Rindfleisch, (2000) ($\alpha=0.79$)	
	7 items	Wirtz, Mathieu et Schilke, (2007) ($\alpha=0.90$)	

Source : Par nos soins

Ainsi, l'incertitude environnementale perçue est mesurée par la turbulence, l'imprévisibilité, le dynamisme et la complexité perçus de l'environnement.

Pour opérationnaliser la turbulence perçue de l'environnement, nous avons utilisé un item de Jaworski et Kohli (1993), un item de Zhou, Yim et Tse (2005) adapté de Jaworski et Kohli (1993), un item de Tsai et Yang (2013) adapté de Jaworski et Kohli, (1993), et un item que nous avons adapté des travaux de Gueguen (2001). Tous ces items ont été mesurés avec une échelle de likert à 5 points, allant de « Tout à fait d'accord » à « Pas du tout d'accord ».

L'imprévisibilité perçue de l'environnement a été saisie avec trois items que nous avons adaptés des travaux de Duncan (1972), Sawyerr (1993) et Sogbossi (2009), et un item adapté des travaux de Gueguen (2001). Tous ces items ont été mesurés avec une échelle de likert à 5 points, allant de « Tout à fait d'accord » à « Pas du tout d'accord ».

Le dynamisme perçue de l'environnement est mesuré avec un item de Li et Liu (2014) adapté de Dess et Beard (1984), Duncan (1972) et Tan et Litschert (1994), deux items de Sogbossi (2009) adapté de Daft et al. (1988) et Sawyerr (1993) et un item que nous avons adapté des travaux de Gueguen (2001). Tous ces items ont été mesurés avec une échelle de likert à 5 points, allant de « Tout à fait d'accord » à « Pas du tout d'accord ».

La complexité perçue de l'environnement a été mesuré avec deux items de Gotteland, et al., (2008) adapté de Tan et Litschert (1994), un item de Sogbossi (2009) adapté de Daft et al. (1988) et Sawyerr (1993), un item de Sutcliffe et Huber (1998) adapté de Dess et Beard (1984) et un item adapté des travaux de Gueguen (2001). Tous ces items ont été mesurés avec une échelle de likert à 5 points, allant de « Tout à fait d'accord » à « Pas du tout d'accord ».

Pour appréhender les variables culturelles, nous sommes partis des précédents travaux sur les orientations individualisme/collectivisme, masculinité/féminité, long terme/court terme, sur la distance hiérarchique et le contrôle de l'incertitude.

Ces variables culturelles sont mesurées à partir des items de Hofstede, (1980, 2001) issus de la Culture Survey Catalogue (Taras, 2008). Les différents scores⁹⁵ par pays sont: L'individualisme collectivisme [71 pour la France (Hofstede, 1980 ; 2001 ; Hofstede *et al.*, 2010) contre 21 pour le Cameroun (Hofstede, 1980 ; 2001 ; Hofstede *et al.*, 2010 ; Fouda, 2006)], la distance hiérarchique (68/52), le contrôle de l'incertitude (86/76) et la masculinité (43/35) sont plus élevés en France qu'au Cameroun. Ce qui signifie que nous avons à faire à deux cultures différentes.

Pour mesurer la performance perçue, nous avons utilisé quatre items adaptés de Song et Parry (1997), avec une échelle de likert à 5 points, allant de « Tout à fait d'accord » à « Pas du tout d'accord ».

Les comportements stratégiques ont été appréhendés à travers la spécialisation, la flexibilité, l'innovation et la coopération.

La stratégie de spécialisation est mesurée à travers quatre items de Wirtz, Mathieu et Schilke (2007), avec une échelle de likert à 5 points, allant de « Tout à fait d'accord » à « Pas du tout d'accord ».

La stratégie de flexibilité est saisie par quatre items que nous avons adaptés des travaux de Gueguen (2001), avec une échelle de likert à 5 points, allant de « Tout à fait d'accord » à « Pas du tout d'accord ».

La stratégie d'innovation est mesurée avec quatre items de Zortea-Johnston, Darroch et Matear (2012), avec une échelle de likert à 5 points, allant de « Tout à fait d'accord » à « Pas du tout d'accord ».

La stratégie de coopération est mesurée à travers deux items de Wirtz, Mathieu et Schilke (2007) et deux items que nous avons adaptés des travaux de Gueguen (2001). Tous ces items ont été mesurés avec une échelle de likert à 5 points, allant de « Tout à fait d'accord » à « Pas du tout d'accord ».

⁹⁵ Ces scores sont obtenus à partir du Value Survey Model (VSM94). Par exemple le score de l'index individualisme/collectivisme (IDV), est obtenu à partir de la formule suivante : $IDV = -50m(01) + 30m(02) + 20m(04) - 25m(08) + 130$; où m(01) est le score moyen obtenu pour la question 1.

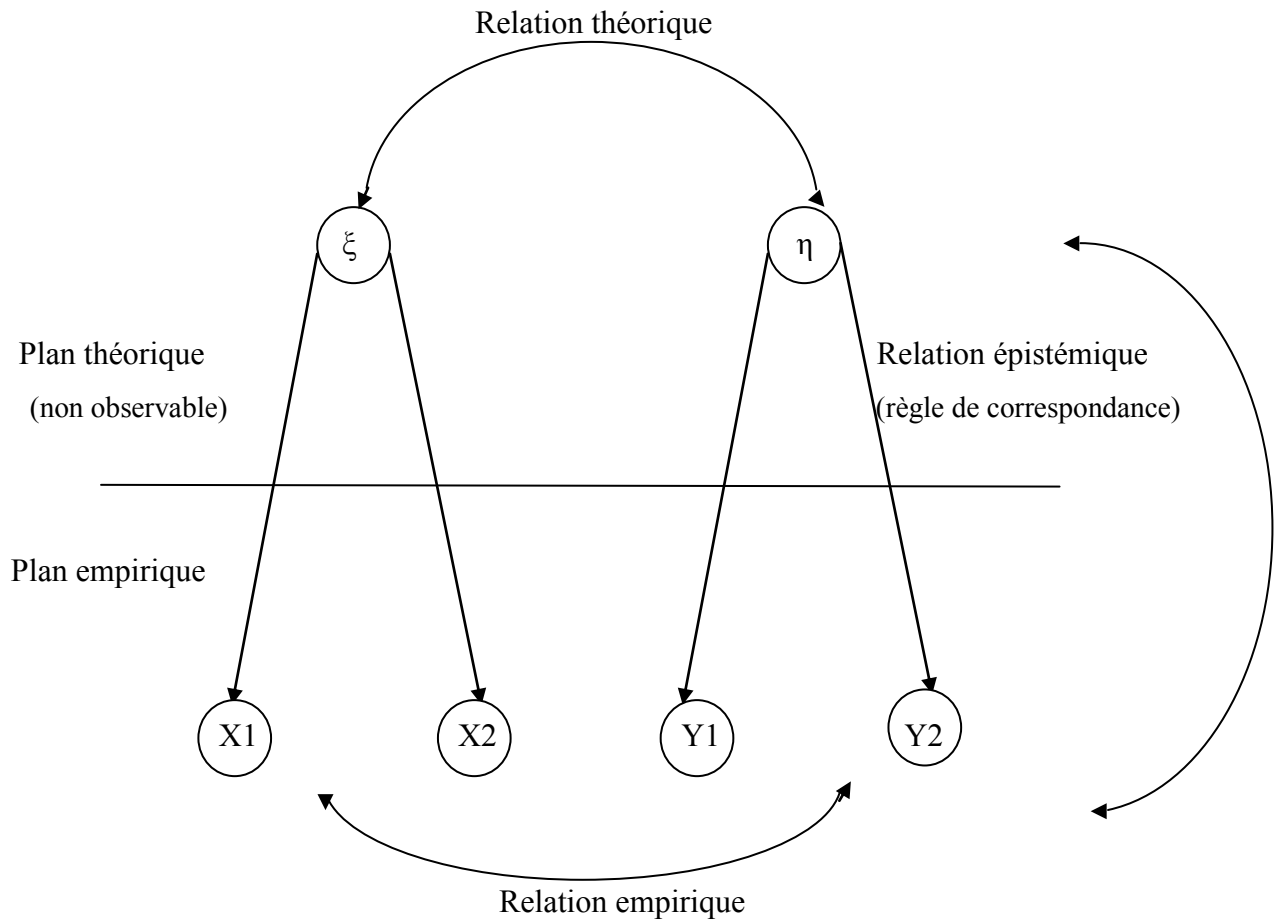
2-2- Les préalables à l'analyse des données quantitatives

Avant tout traitement statistique, nous devons procéder à la préparation des données ainsi recueillies. Celle-ci passe notamment par la vérification de la lisibilité (le questionnaire doit être lisible pour pouvoir être facilement codé), de la complétude (il s'agit de données manquantes qui déterminent la qualité de l'enquête) et de la cohérence. Une fois les données codifiées et regroupées dans un tableau, on peut dès lors procéder aux différentes mesures.

2-2-A- Le processus de mesure : définition et schéma général

Mesurer c'est établir une relation entre l'objet étudié (produit, individu, phénomène) et un symbole le représentant. De façon générale, les symboles utilisés seront le plus souvent des chiffres et la mesure peut être définie comme « les règles d'attribution de nombres à des caractéristiques des objets » (Nunnally, 1967). Il s'agit donc de mettre en correspondance l'univers réel sur lequel porte l'étude avec un système symbolique représenté par des chiffres ; les règles définissent la façon dont doit être effectuée la mesure, c'est-à-dire son opérationnalisation. La mesure établit donc une correspondance entre un niveau théorique (définition conceptuelle du phénomène étudié) et un niveau empirique (définition des indicateurs représentant ce phénomène et sur lesquels portent les opérations concrètes de mesure) (Evrard al., 2003) . Le schéma général du processus de mesure est illustré dans la figure ci-après :

Figure 17: Schéma général de la mesure



On distingue sur ce schéma trois types de relations :

-Relation théorique : il s'agit de la relation entre concepts non observables ; il faut souligner qu'il s'agit du cœur de l'étude.

-Relation empirique : il s'agit des relations entre les mesures, dont l'estimation fera l'objet du traitement des données ; la relation empirique (liaison statistique entre la réponse à deux questions) doit être considérée comme un moyen d'estimer la relation théorique entre les phénomènes non observables.

On peut déjà souligner que, si chaque concept non observable n'est opérationnalisé que par une seule variable, il y aura identité entre niveau empirique et niveau théorique, et donc pas de moyen de distinguer le phénomène étudié et son opérationnalisation, c'est-à-dire qu'on ne pourra pas cerner la qualité de la mesure ;

-Relation épistémique (ou règles de correspondance entre un phénomène et sa/ses mesures) : c'est le type de relation, qui s'établit entre le niveau non observable et le niveau empirique et qui va être au cœur de l'évaluation des mesures et de l'amélioration de leur qualité. A chaque

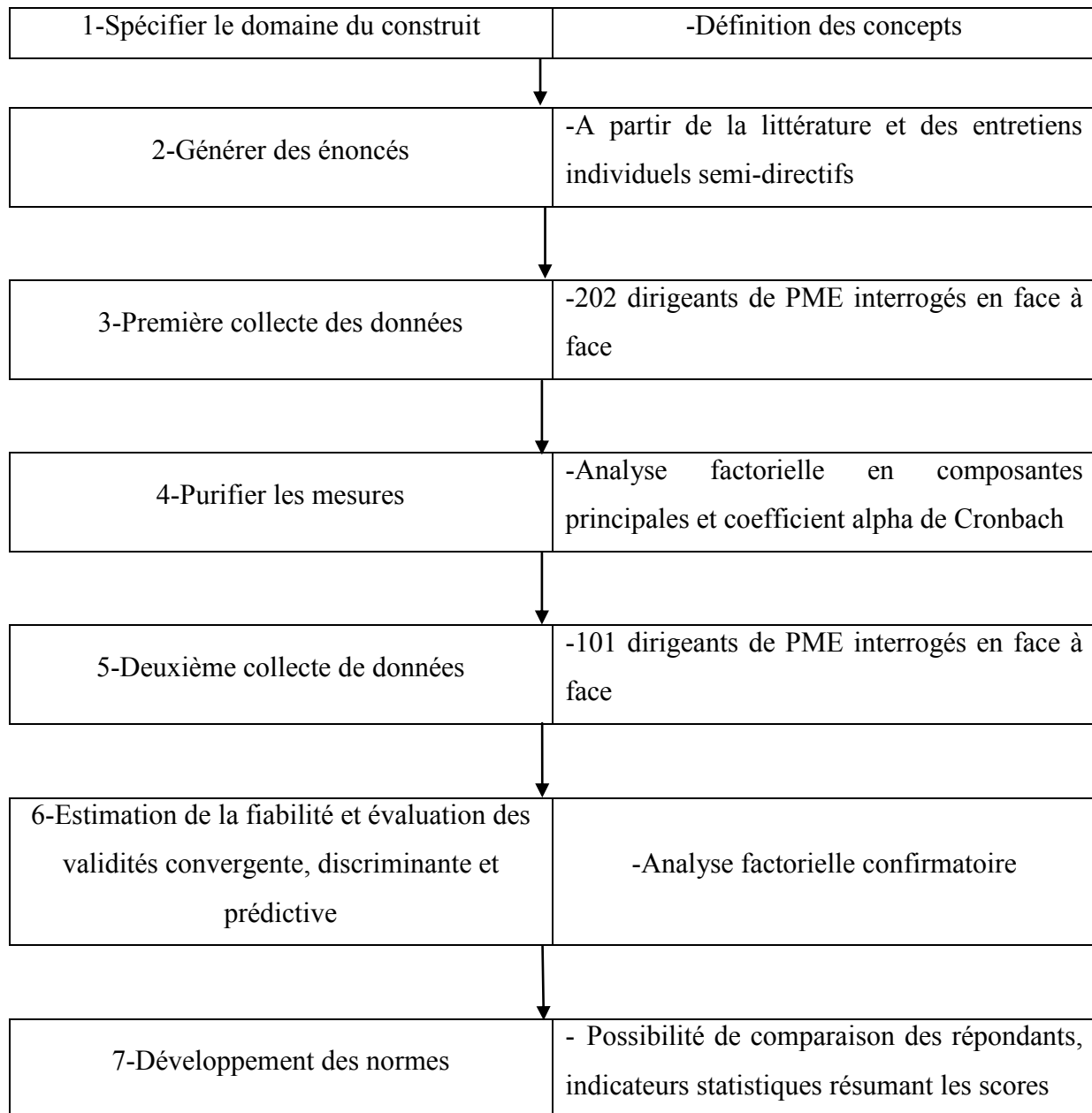
phénomène non observable étudié, on peut en effet associer plusieurs façons de le mesurer (ou indicateurs résultant d'un processus d'opérationnalisation) ; il est essentiel de percevoir que le résultat d'une mesure n'est pas le phénomène lui-même, mais un indicateur reflétant ce phénomène.

2-2-B- La procédure de construction des mesures : le paradigme de Churchill

Churchill (1979) a proposé un cadre notoire afin de développer des outils de mesure en marketing. La démarche méthodologique développée connue sous le nom de « paradigme de Churchill » vise à intégrer les connaissances concernant la théorie de la mesure ainsi que les techniques appropriées pour l'améliorer dans une procédure systématique (Evrard et al., 2003). Pour développer les mesures on peut directement tester les indicateurs ou items en utilisant le paradigme de Churchill, (1979) ou alors on peut récupérer les mesures ou items à partir de la revue de la littérature proposée par les chercheurs du domaine ou éventuellement choisir les mesures adaptées à partir des surveys lorsqu'elles existent.

S'il est vrai que la plupart de nos items sont tirés de la littérature et des surveys, il n'en demeure pas moins vrai que tout au long de notre travail nous nous sommes appuyés sur le paradigme de Churchill dont le schéma est le suivant :

Figure 18: Représentation du paradigme de Churchill (1979)



2-2-C- La théorie de la mesure : le « modèle de la vraie valeur »

Le modèle de la vraie valeur (ou True Score Model) consiste à décomposer le résultat d'une mesure en ses différents éléments : la vraie valeur (censée représenter théoriquement la mesure parfaite) et les termes d'erreur (l'erreur aléatoire et l'erreur systématique). Ce modèle peut se résumer par la formule suivante (voir Evrard, Pras et Roux, 1993 : 277-279) :

M	=	V	+	Es	+	Ea
Mesure	=	Vraie	+	Erreur	+	Erreur
Obtenue		valeur		systematique		aleatoire

-La vraie valeur : c'est la mesure « idéale », c'est-à-dire celle qui correspondrait parfaitement au phénomène étudié ; elle est, le plus souvent, impossible à atteindre directement et constitue l'« horizon » de la mesure empirique ;

-L'erreur systématique (ou biais) : ce type d'erreur provient du fait que l'instrument de mesure peut avoir un écart systématique avec le phénomène étudié ;

-L'erreur aléatoire : le phénomène mesuré par un même instrument peut être sujet à des aléas tels que les circonstances, l'humeur des personnes interrogées...

2-2-C-1- Les principes de l'analyse

L'analyse statistique a pour rôle d'aider l'utilisateur à « lire » la matrice des données recueillies. La gamme des méthodes disponibles est très vaste. On discerne cependant quelques principes généraux qui contribuent à guider le chercheur dans le choix et l'application d'une méthode. Le premier objectif de l'analyse des données est de réduire le volume des données disponibles de façon à l'amener à un format assimilable pour l'interprétation. Cette recherche est guidée par l'objectif de l'analyse : il ne s'agit pas d'extraire un petit nombre de données au hasard, mais de mettre en évidence les faits et les relations les plus importants. Le choix d'une méthode est la synthèse d'un objectif (ce qu'on cherche) et d'une contrainte (les propriétés de mesure des variables étudiées). Nous exposerons à la suite la logique du nombre de variables incluses dans le traitement, qui permet de classer les méthodes selon qu'elles concernent une variable (univariée), ou davantage (multivariée) en considérant l'analyse qui correspond aux objectifs du chercheur (description et /ou inférence) et la logique inférentielle qui se traduit dans l'application des tests statistiques.

2-2-C-1-1- L'analyse univariée

Elle consiste à examiner la distribution des modalités de réponse pour une variable, c'est-à-dire une colonne de la matrice des données (par exemple, dans le cas d'une variable nominale, il s'agit d'un tri à plat, c'est-à-dire du dénombrement des observations correspondant à chaque modalité de la variable). Deux types de problèmes sont abordés par l'analyse univariée : la description des données (abordée sous deux aspects dont la tendance

centrale et la dispersion) et l'inférence. Toutefois, les méthodes applicables dans cette analyse dépendent des propriétés de mesure de la variable étudiée. On distingue sur ce point trois niveaux : nominal, ordinal et métrique (terme générique que nous utilisons pour désigner les échelles de proportion et les échelles d'intervalle).

❖ Le cas d'une variable nominale

C'est celle qui possède le moins de propriétés mathématiques. On peut simplement compter le nombre d'observations appartenant à chaque catégorie (ou modalité) de la variable, c'est-à-dire effectuer une tabulation simple (ou tri à plat). Notons que dans le cas de l'analyse bivariée on fera plutôt des tris croisés. La tendance centrale est représentée par le mode qui est la modalité pour laquelle les observations sont les plus nombreuses. La dispersion est indiquée par les fréquences.

Une statistique inférentielle applicable aux variables nominales est le test du Chi-deux qui consiste à comparer la distribution observée des nombres d'observations par modalité à une distribution prédéterminée. Dans le cas de l'analyse bivariée le test du Chi-deux est calculé par la formule suivante :

$$\chi^2 = \sum_i \sum_j \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$$

Avec :

χ^2 : Chi-deux

n_{ij} : le nombre d'éléments de l'échantillon répondant simultanément à la modalité i de la variable X ($i = 1 \dots, r$) et à la modalité j de la variable Y ($j = 1 \dots, c$) ;

\hat{n}_{ij} : l'effectif théorique « théorique » de la case (i, j) .

On démontre que cette quantité suit une loi du χ^2 à $(r - 1)(c - 1)$ degrés de liberté ; la consultation de la table fournit la réponse. Si χ^2 observé $\geq \chi^2$ théorique, à un certain seuil de signification, on peut alors rejeter l'hypothèse nulle.

❖ Le cas d'une variable ordinale

Il s'agit ici d'une variable qualitative dont les modalités sont ordonnées. Pour une variable ordinale, la tendance centrale est représentée par la médiane, qui est la valeur qui divise la population étudiée en deux parties égales. La dispersion est indiquée par les fractiles (qui partagent la population en catégories d'effectifs égaux). Les plus usuels sont les quartiles et les déciles.

L'inférence consiste à comparer la répartition des réponses obtenue à une répartition définie a priori ; elle peut être certifiée par le test de Kolmogorov-Smirnov.

❖ Le cas d'une variable métrique

Dans ce cas, la tendance centrale est la moyenne dont la formule de calcul est la suivante :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$$

Avec :

\bar{X} = moyenne de la série ;

x_i = valeur de l'observation i ;

n = nombre d'observations de la série.

La dispersion est reflétée par plusieurs indicateurs dont les plus usuels sont la variance et l'écart-type. Leurs formules de calcul sont respectivement :

$$V(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

Avec :

$V(X)$: la variance de la variable X ;

X_i : la valeur de la variable X ;

\bar{X} : la moyenne des observations ;

n : le nombre total d'individus.

$$\sigma_X = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$$

Avec :

σ_X : l'écart-type de la variable X

X_i : la valeur de la variable X ;

\bar{X} : la moyenne des observations ;

n : le nombre total d'individus.

L'inférence dans ce cas est étudiée par le test de moyenne qui vérifie si la valeur d'une moyenne m , calculée sur un échantillon issu d'une population de variance σ^2 ,

inconnue est significativement différente d'une moyenne hypothétique μ_0 . Le choix du test dépend de trois conditions précises :

- la connaissance, a priori, de la variance du phénomène étudié dans la population,
- la taille de l'échantillon et
- la distribution de la variable dans la population.

La première condition est rarement satisfaite. Nous ne nous préoccupons donc ici que des cas où la variance de population est inconnue.

Les conditions d'application sont les suivantes :

- la population a une variance σ^2 inconnue qui doit être estimée sur l'échantillon et une moyenne μ également inconnue (posée par hypothèse égale à μ_0);
- l'échantillon est aléatoire et contient n observations indépendantes ;
- la taille n de l'échantillon est supérieure à 30 ou bien la moyenne suit dans la population une loi normale auquel cas la taille n est quelconque.

Les hypothèses :

L'hypothèse nulle à éprouver est $H_0 : \mu = \mu_0$.

L'hypothèse alternative est : $H_1 : \mu \neq \mu_0$ (pour un test bilatéral)

Ou $H_1 : \mu < \mu_0$ (pour un test unilatéral à gauche)

Ou $H_1 : \mu > \mu_0$ (pour un test unilatéral à droite).

La formule de calcul du test est :

$$t = \frac{(m - \mu_0)}{s/\sqrt{n}}$$

Sa distribution suit une loi de Student à $n - 1$ degrés de liberté. On l'appelle « test t » ou « test de Student ». Lorsque n est grand, par exemple supérieur à 30, la distribution de cette statistique suit approximativement une loi normale centrée réduite. Autrement dit,

$$t = \frac{(m - \mu_0)}{s/\sqrt{n}} = z = \frac{(m - \mu_0)}{\sigma/\sqrt{n}}$$

On peut donc prendre la décision (c'est-à-dire rejet ou non rejet de H_0) en comparant la statistique t calculée aux valeurs de la loi normale centrée réduite. Mais lorsque n est petit, par exemple inférieur à 30, il faut absolument utiliser la loi du t de Student à $n - 1$ degrés de liberté et non la loi normale z.

La variance inconnue de la population σ^2 est estimée sur l'échantillon, avec $n - 1$ degrés de liberté, par :

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m)^2$$

2-2-C-1-2- L'analyse multivariée

Le caractère combinatoire des analyses dès que le nombre de variables est assez important explique les limites des méthodes d'analyses précédentes (univariée et bivariée). C'est la raison pour laquelle il est nécessaire au chercheur de recourir à l'analyse multivariée qui lui permet le traitement simultané d'un ensemble de variables. L'objectif de notre travail n'étant pas d'évoquer de manière exhaustive toutes les méthodes d'analyse multivariée, nous nous arrêterons uniquement sur celles qui nous semblent nécessaires pour notre recherche notamment l'analyse factorielle, la régression et le modèle d'équations structurelles.

❖ l'analyse factorielle :

L'analyse factorielle est l'une des plus anciennes méthodes d'analyse des données qui a certainement fait l'objet en analyse multivariée du plus grand nombre d'applications en sciences sociales. Son origine remonte en effet aux travaux de Spearman, au début de ce siècle (Evrard et al., 2003, p. 398). Ses objectifs peuvent être appréciés de deux points de vue différents selon ces auteurs :

-d'un point de vue empirique, elle consiste à résumer l'information contenue dans un tableau de chiffres individus/variables en remplaçant les variables initiales par un nombre plus petit de variables composites ou facteurs ;

-d'un point de vue théorique, elle correspond à la démarche psychométrique de mesure des concepts non observables directement. L'analyse factorielle dans ce cas sert de révélateur à un cadre conceptuel sous-jacent masqué par « le bruit » des mesures ; les variables dans ce cas sont alors considérées comme des combinaisons d'un certain nombre de facteurs communs sous-jacents non observables appelés variables latentes.

Selon Evrard et al. (2003, p. 399), la démarche de mise en œuvre d'une analyse factorielle passe par quatre étapes : le choix d'un algorithme d'analyse et le type de données, la détermination du nombre de facteurs et l'interprétation des axes factoriels.

- Le choix d'une technique d'analyse factorielle et le type de données

On distingue deux techniques de base d'analyse factorielle : l'analyse factorielle classique ou analyse en facteurs communs spécifiques (AFCS) et l'analyse en composantes principales (ACP). A ces deux techniques, on peut ajouter l'analyse factorielle des correspondances (AFC) pour le cas des variables catégorielles. Le choix entre l'AFCS et l'ACP dépend des objectifs du chercheur. Selon Donada et Mbengue (2003), si l'objectif de recherche est simplement de résumer les données, l'ACP est le meilleur choix. Par contre, si la recherche vise plutôt à mettre en évidence une structure sous-jacente aux données (c'est-à-dire identifier les variables latentes ou des construits), alors le choix de l'AFCS s'impose. Cependant, l'analyse factorielle ne peut être menée que si les données s'avèrent adéquates à ce type d'analyse.

Mathématiquement, si l'on note X_k la variable ($k = 1, \dots, p$) et F_j le facteur, la formule de calcul de l'analyse en composantes principales (ACP) s'écrit :

$$X_k = \sum_{a_{kj}} F_j$$

Et celle de l'analyse factorielle (ou AFCS) :

$$X_k = \sum_{a_{kj}} F_j + u_k$$

Où :

u_k : est la composante spécifique (ou le terme d'erreur de mesure) de la variable X_k ;

Les a_{kj} sont les coefficients de la forme linéaire qui relie les variables et les facteurs.

Une question que le chercheur est amené à se poser à l'occasion d'une analyse factorielle est celle de savoir si les données sont « factorisables » c'est-à-dire si elles forment un ensemble suffisamment cohérent pour y rechercher des dimensions communes qui aient un sens et ne soient pas des artefacts statistiques (Evrard et al, 2003, p. 400). Les auteurs présentent deux tests qui permettent de supposer que les données peuvent faire l'objet d'une analyse factorielle : le test de sphéricité de Bartlett et le test MSA (*Measure of Sampling Adequacy*) de Kaiser, Meyer et Olkin (encore appelé test KMO).

-Le test de sphéricité de Bartlett : cette statistique, selon Galan (2003), est approximativement distribuée comme un chi-deux avec $1/2P(P-1)$ degrés de liberté. Elle teste l'hypothèse que la matrice de corrélation provient d'une population de variables qui sont

indépendantes. Le rejet de cette hypothèse signifie que les données sont appropriées pour une analyse factorielle. Autrement dit, ce test, teste si la matrice de corrélation est une matrice unité, ce qui indiquerait que le modèle factoriel est inapproprié. Il présente l'avantage de fournir, en outre, des indications sur le nombre maximum de facteurs à retenir, mais a l'inconvénient d'être pratiquement toujours satisfait sur de grands échantillons (Evrard et al, 2003, p. 400). Le test de sphéricité de Bartlett se calcule par la formule ci-après :

$$\left[(N - 1) - \left(\frac{2P + 5}{6} \right) \right] \text{Log}_e |R|$$

Où :

N est la taille de l'échantillon ;

P le nombre de variables ;

R est le déterminant de la matrice de corrélation.

Une valeur élevée de ce test sera favorable au rejet de l'hypothèse nulle. Dans le cas contraire, l'analyse factorielle s'avère inappropriée (Malhotra et al., 2007, p. 530).

-Le test MSA de Kaiser, Meyer et Olkin (ou test KMO) : ce test indique si les données sont cohérentes ensemble et si elles sont appropriées pour une analyse factorielle. Il teste si les corrélations partielles entre variables sont faibles. Le calcul du MSA se fait selon la formule ci-dessous :

$$MSA = \frac{\sum_j \sum_k r_{jk}^2}{\sum_j \sum_k r_{jk}^2 + \sum_j \sum_k q_{jk}^2}$$

Où :

r_{jk}^2 est le carré des éléments hors diagonale de la matrice de corrélation d'origine ;

q_{jk}^2 est le carré des éléments hors diagonale de la matrice de corrélation des anti-images.

Galan (2003) propose une calibration du test selon laquelle :

$MSA > 0,9 \rightarrow$ merveilleux ;

$MSA > 0,8 \rightarrow$ méritoire ;

$MSA > 0,7 \rightarrow$ moyen ;

$MSA > 0,6 \rightarrow$ médiocre ;

$MSA > 0,5 \rightarrow$ misérable ;

$MSA \leq 0,5 \rightarrow$ inacceptable.

- La détermination du nombre de facteurs

Il n'existe pas de règle générale permettant de déterminer le « bon » nombre de facteurs. Le chercheur dispose néanmoins de critères lui permettant de faire face à ce problème (Thiéart et al., 2003, p. 389), critères résultant de règles empiriques fondées sur l'expérience (Evrard et al., 2003, p. 401). Ces critères sont :

-La « spécification à priori » : dans ce cas, le chercheur sait d'avance le nombre de facteurs à retenir, souvent suggéré par la théorie, ou il peut s'agir de tester une hypothèse relative à ce nombre de facteurs. Il peut également s'agir pour le chercheur de répliquer une recherche antérieure avec le désir d'extraire exactement le même nombre de facteurs ;

-La restitution minimum : ici, le chercheur se fixe un seuil correspondant au pourcentage minimum d'information (c'est-à-dire de variance) que l'on veut restituer. On retient alors le nombre d'axes nécessaires pour atteindre ce seuil. La variance expliquée (VE) par les deux premiers facteurs F_1 et F_2 est égale à la somme des valeurs propres λ_1 et λ_2 associés à ces deux axes, divisée par la somme de l'ensemble des valeurs propres (c'est-à-dire la variance totale) :

$$VE (F_1, F_2) = \frac{\lambda_1 + \lambda_2}{\sum_{k=1}^p \lambda_k}$$

Si cette quantité est « suffisamment proche » de 1, seuls les deux premiers facteurs sont conservés, sinon le troisième facteur est introduit dans l'analyse jusqu'à ce qu'on atteigne le seuil fixé (par exemple à 80% soit 0.8) (Evrard et al., 2003, p. 401).

-La « règle de Kaiser » : le test le plus utilisé comme règle statistique pour déterminer le nombre de facteurs à extraire est le critère de racine (roots criterion) connu en contexte français sous le nom de critère de Kaiser (Galan, 2003). Si le nuage de points initial n'avait pratiquement aucune direction privilégiée (sphère par exemple), les valeurs $\lambda_1, \dots, \lambda_p$ seraient peu différentes les unes des autres et donc la variance restituée par le premier facteur serait à peu près égale à :

$$VE (F_1) = \frac{\lambda_1}{\sum_{k=1}^p \lambda_k} = \frac{1}{p}$$

ou en pourcentage : $\frac{100}{P}$

Où p est le nombre initial de variables.

En effet, si toutes les valeurs propres sont à peu près égales, on a :

$$\sum_{k=1}^p \lambda_k = p\lambda_1$$

Dans le cas où le point de départ de l'analyse est la matrice des corrélations, la règle la plus usuelle (règle de Kaiser) est de retenir les facteurs correspondant à des valeurs propres supérieures à l'unité ($\lambda > 1$).

Il faut noter que cette prescription n'est valable sans restriction que dans le cas d'une ACP effectuée sur une matrice de corrélation. Dans le cas d'une AFCS, une telle règle est trop stricte. Le chercheur peut retenir tout facteur dont la valeur propre, bien qu'inférieure à 1, est toutefois supérieure à la moyenne des communautés (c'est-à-dire des variances communes) des variables (Thiétart et al., 2003, p. 390) ;

-L'examen de la courbe des valeurs propres : l'approche ici consiste à construire un graphique (scree-test de Cattell ou « test de coude »), en mettant en abscisse le numéro des axes factoriels et en ordonnée le pourcentage d'inertie qu'ils restituent. Ensuite, on élimine les facteurs qui se trouvent après le point d'inflexion, c'est-à-dire le point où la courbe change de concavité. Une autre approche de la recherche systématisée du point d'inflexion peut être faite en recherchant le signe des écarts entre les différences de valeurs propres consécutives ; le point d'inflexion correspond à un changement de ce signe (Evrard et al., 2003, p. 406).

- L'interprétation des axes factoriels

Une fois les axes factoriels déterminés, il faut pouvoir les interpréter et le retour aux variables initiales est nécessaire. L'importance de chacune de ces variables dans la formation des dimensions latentes s'apprécie à l'aide des coefficients de corrélation entre ces variables, ce qui en facilite l'interprétation qui peut s'affiner avec la rotation des axes.

- Coefficients de corrélation entre variables initiales et facteurs

Evrard et al. (2003) affirment que si les variables initiales sont centrées et réduites, le coefficient de corrélation $r(X_k, F_j)$ est un bon indicateur pour mesurer la relation de la

variable X_k au facteur F_j (le carré de ce facteur étant le pourcentage de la variance restitué par ce facteur). Les variables initiales ayant les corrélations les plus élevées avec un facteur sont celles qui contribuent le plus à sa formation, ce qui permet son interprétation. D'après Thiétart et al., (2003), en règle générale, les corrélations supérieures à 0.3 en valeur absolue sont jugées significatives et celles supérieures à 0.5 sont très significatives. Selon Evrard et al. (2003), la règle la plus usuelle est de retenir les valeurs supérieures à 0,50.

- La rotation des axes factoriels

Une fois le nombre de facteurs déterminés il est commode d'effectuer une rotation pour faciliter l'interprétation. On distingue les rotations orthogonales des rotations obliques. Dans une rotation orthogonale, les facteurs restent orthogonaux entre eux, ce qui n'est pas le cas de la rotation oblique où les facteurs peuvent devenir corrélés les uns aux autres. Cette dernière rotation présente l'inconvénient d'être parfois plus délicate à interpréter, mais présente l'avantage de pouvoir mieux rendre compte de certaines situations (Evrard et al., 2003, p. 407). Dans tous les cas, elle donne généralement de meilleurs résultats que la rotation orthogonale (Thiétart et al., 2003, p. 392).

On distingue trois principaux types de rotation orthogonale : Varimax, Quartimax et Equamax. La première, la plus répandue, cherche à minimiser le nombre de variables fortement corrélées avec un facteur. L'algorithme Quartimax vise à faire en sorte que chacune des variables soit fortement corrélée avec un seul facteur et le moins corrélé possible avec les autres facteurs. Son inconvénient principal est que plusieurs variables peuvent être fortement corrélées avec le même facteur. Le dernier algorithme de rotation orthogonale, Equamax, semble être un compromis entre Varimax et Quartimax, son défaut étant qu'il ne donne pas des résultats très probants et est peu utilisé (Thiétart et al, 2003, p. 392).

En ce qui concerne les algorithmes de rotation oblique, on peut citer Oblimin et Promax. Ces algorithmes changent parfois de nom selon les logiciels.

Une fois que le chercheur a identifié les dimensions latentes (variables non directement observables) qui sont réputées influencer d'autres variables grâce à l'analyse factorielle, il doit s'assurer de la fiabilité et de la validité des ses mesures.

❖ Analyse de régression

La régression est sans doute parmi les méthodes explicatives, la plus connue et la plus utilisée dans le domaine des sciences économiques et particulièrement en économétrie (Evrard et al, 2003, p. 478). Pour son application, les variables utilisées doivent être quantitatives (c'est-à-dire mesurées sur des échelles de proportion ou d'intervalle), cependant il est possible d'inclure dans le modèle explicatif les variables nominales, après les avoir transformées en variables binaires (0/1) ; toutefois, l'analyse de covariance est mieux adapté à ce cas. Une autre condition d'utilisation de la régression est l'indépendance entre les variables explicatives. Une analyse factorielle en composantes principales préalable permet de réaliser la régression sur les axes factoriels obtenus, qui sont par construction indépendants.

Le modèle de régression est de la forme :

- Une variable à expliquer (notée Y) ;
- p variables explicatives (notées X_1, \dots, X_p) ;
- Une relation fonctionnelle $Y = f(X_1, \dots, X_p)$.

La relation la plus simple de la fonction f est le plus souvent choisie linéaire et est de la forme : $Y = a_0 + a_1x_1 + a_2x_2 + \dots + a_px_p$

Une fois la forme de la fonction spécifiée, il faut trouver les valeurs des paramètres (a_i) de la fonction retenue. C'est la détermination de ces paramètres qui est l'objet du calcul de la régression. Plusieurs méthodes d'estimation sont possibles, mais nous ne présenterons ici que la procédure la plus usuelle, l'estimation aux moindres carrés ordinaires (OLS, Ordinary Least Squares). Ensuite, on présentera les éléments qui permettent de déterminer la qualité du modèle, et en guide l'interprétation au niveau global par le coefficient de corrélation multiple et pour chacune des variables explicatives.

- Détermination des paramètres

Après le choix des variables explicatives et l'explicitation du modèle, il faut déterminer les paramètres de la fonction retenue. Supposons que la relation choisie soit de la forme :

$$Y = aX + b + \varepsilon$$

Où :

Y est la variable à expliquer

X est la variable explicative

a et b sont deux paramètres qu'il s'agit d'évaluer, \hat{a} et \hat{b} étant les estimations de ces paramètres.

ε est un résidu qui représente en particulier l'effet des variables non prises en compte dans le modèle.

Pour évaluer ces deux paramètres, nous disposons des mesures qui ont été effectuées sur les n individus de l'échantillon. En effet, pour chacune de ces personnes, les valeurs de Y et de X sont connues. L'application du modèle donne :

$$y_i^* = ax_i + b, i = 1, 2, \dots, n.$$

Cette relation n'est qu'approchée car la valeur y_i^* est celle prévue par le modèle (valeur théorique) qui n'est pas forcément la valeur réelle (valeur observée). La valeur observée est notée y_i . Le modèle est meilleur lorsque les différences entre y_i^* et y_i (ou résidu ε_i) sont plus faibles pour l'ensemble des individus considérés.

Il faut donc trouver les valeurs de : a_1, \dots, a_p, b , qui rendent cette différence minimum ou encore en posant :

$$\|E\|^2 = \varepsilon_1^2 + \dots + \varepsilon_n^2 = \sum_{i=1}^n \varepsilon_i^2 \text{ qui rendent les } \|E\|^2 \text{ minimum.}$$

Le choix de a et b est donc effectué pour rendre minimum les différences entre les valeurs théoriques prévues par le modèle et les valeurs observées. Dans le cas d'une régression simple (à une seule variable explicative), le modèle s'écrit :

$$y_i = ax_i + b + \varepsilon_i \text{ ou encore } \varepsilon_i = y_i - ax_i - b$$

$$\text{La forme } \|E\|^2 \text{ peut s'écrire : } \|E\|^2 = \sum_{i=1}^n (y_i - ax_i - b)^2$$

La méthode des moindres carrés consiste à choisir comme estimations des paramètres a et b les valeurs \hat{a} et \hat{b} qui minimise $\|E\|^2$. On démontre que : \hat{a} et \hat{b} annulent les dérivées partielles de $\|E\|^2$ par rapport à a et b. En résolvant le système d'équations correspondant on trouve les valeurs de \hat{a} et \hat{b} qui sont les solutions du système selon les formules suivantes :

$$\hat{a} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{b} = \bar{y} - \hat{a} \bar{x}$$

Dans le cas de la régression multiple (à plusieurs variables explicatives), le modèle se formule de manière suivante :

$$y_i = a_1 x_{1i} + a_2 x_{2i} + \dots + a_p x_{pi} + b + \varepsilon_i$$

Où : x_{1i} est la valeur de la première variable explicative mesurée sur le $i^{\text{ème}}$ individu.

La procédure de résolution est analogue à celle d'un modèle de régression simple, mais il est nécessaire d'utiliser la notation matricielle. On obtient ainsi le vecteur des paramètres en fonction des données initiales. Cette forme est alors immédiatement généralisable au cas de la régression simple.

- L'interprétation des résultats

Comme nous l'avons dit précédemment, l'interprétation des résultats d'une régression va se faire à trois niveaux : au niveau global, pour chaque variable et à l'examen des résidus.

➤ La qualité globale de la régression

La qualité de la régression effectuée est appréciée grâce aux indices suivants :

-Le coefficient de corrélation multiple (R), qui s'interprète comme un coefficient de corrélation simple ; son carré (R^2) est appelé coefficient de détermination. C'est l'indicateur usuel de la qualité de l'ajustement global. Sa formule mathématique est la suivante :

$$R^2 = \frac{\sum_{i=1}^n (y_i^* - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Il s'interprète comme le pourcentage de variance de la variable à expliquer restituée par le modèle. Plus il est proche de 1, plus les valeurs observées et les valeurs calculées par le modèle sont proches.

-Le test F de Fisher-Snedecor

Sa formule de calcul est la suivante:

$$F = \frac{n - p - 1}{n} \times \frac{R^2}{1 - R^2}$$

On cherche dans la table de Fisher-Snedecor à $(p, n - p - 1)$ degrés de liberté la valeur F théorique correspondant à un seuil α ; Si $F_{\text{obs}} > F_{\alpha}$, la régression est statistiquement significative (au seuil α).

Si le modèle est correct, le coefficient de corrélation multiple est grand, l'influence des résidus est négligeable et donc le rapport F est très grand (Evrard et al, 2003, p. 491).

-Le coefficient de variation

La racine carrée de la variance résiduelle mesure l'erreur standard de la régression. Elle s'écrit :

$$S_R = \sqrt{\frac{1}{n - p - 1} \sum (y_i - y_i^*)^2}$$

Elle donne une idée de l'imprécision de la régression exprimée dans les unités de la variable à expliquer. En la comparant à la valeur moyenne de Y, on définit le coefficient de variation :

$$C_V = \left(\frac{S_R}{\bar{y}} \right)$$

Qui est un indicateur de qualité globale de la régression.

➤ La contribution de chaque variable explicative

Le test t de Student à $(n - p - 1)$ degrés de liberté sur chaque coefficient de régression permet d'examiner si, pour chaque variable explicative, il existe une relation significative avec la variable expliquée. Cela permet d'éliminer les variables explicatives inutiles sans changer significativement la qualité de l'ajustement global.

➤ L'examen des résidus

L'hypothèse sur l'indépendance entre résidus et variables explicatives est la plus importante source de biais, mais n'est pas vérifiable (Evrard et al, 2003, p. 492). Si le modèle n'est pas adapté, il faut s'interroger sur sa spécification fondée sur les hypothèses théoriques reflétant la compréhension du phénomène étudié. Les résidus constituent un indicateur de performance de la droite de régression. Leur examen sert à estimer l'exactitude des estimations. Une analyse des résidus avec des graphiques qui fournissent des aperçus utiles

pour s'assurer que les hypothèses fondamentales et la qualité du modèle de régression ont bien été respectées est recommandée.

❖ Les modèles d'équations structurelles

C'est à la fin des années 1960 que les modèles d'équations structurelles (MES) furent développés d'abord comme modèles mathématiques sous l'impulsion des travaux de Karl Jöreskog (1973)⁹⁶, de Ward Keesling (1972) et de David Wiley (1973). Ces trois chercheurs évoluant séparément ont développé un cadre d'analyse désigné par le modèle JKW à la base des différents travaux sur les MES (Bentler, 1980). Jöreskog publie ses premiers travaux sur l'analyse factorielle au milieu des années 1960 et la première version du modèle LISREL (LInear Structural RELations) en 1970. Il cherche à améliorer la mesure des variables latentes de plus en plus utilisées tant en psychologie qu'en économétrie. La vulgarisation de ce modèle commence au milieu des années 1970 lorsque Jöreskog est rejoint par Dag Sörbom pour créer ensemble le logiciel LISREL. C'est Bagozzi qui, dans ses travaux en marketing introduisit les MES en sciences de gestion. En effet, ses travaux de 1980 (Bagozzi, 1980a et 1980b)⁹⁷, ont contribué de façon déterminante à l'essor de ces méthodes dans les disciplines de gestion. Si à la base LISREL a été le premier logiciel développé pour l'analyse des MES, de nos jours, d'autres chercheurs en statistique, en mathématique, et en psychométrie ont contribué à l'amélioration progressive des modèles et des MES. Dans cette perspective, on peut citer entre autres modèles développés : EQS, PSL, SEPATH, CALIS, AMOS (qui sera utilisé dans la présente recherche), etc.

Les modèles d'équations structurelles constituent un des développements récents les plus marquants dans les études et recherches. Ils peuvent être vus comme la conjonction de deux approches auparavant traitées séparément :

-l'analyse des variables latentes (ou non observables) développée principalement en psychologie, dans laquelle on considère que les variables directement observées (par exemple, les items dans un questionnaire) sont le reflet ou l'effet, d'autres variables non directement observables, représentant des concepts plus généraux que la formation spécifique d'un item

⁹⁶ Jöreskog K. G. (1973), "A General Method for Estimating a Linear Structural Equation System" in Goldberger A. S. and Duncan O. D. (Eds), *Structural Equation Model in Social Sciences*, New York: Seminar Press/Harcourt Brace.

⁹⁷ Bagozzi R. P. (1980a), *Causal Models in Marketing*, Wiley and Sons, New York.

Bagozzi R. P. (1980b), "Performance and Satisfaction in an Industrial Sales Force: An Examination of their Antecedents and Simultaneity", *Journal of Marketing*, Vol. 44, pp. 65-77.

particulier (par exemple, la réponse à une échelle de perception de l'incertitude environnementale est le reflet d'un phénomène mental non directement accessible : la perception que la personne interrogée a de l'incertitude environnementale étudiée). L'analyse factorielle est un exemple d'instrument pouvant être utilisé dans ce domaine.

-les modèles structurels qui visent à représenter et à estimer des relations de causalité entre les variables ; les représentations les plus utilisées prennent la forme d'un ensemble d'équations linéaires ; il s'agit des équations simultanées (en économie), ou de l'analyse des réseaux ou des cheminements, « path analysis », (en sociologie). Le cas le plus simple est celui où il n'y a qu'une seule équation, correspondant à la régression multiple. (Evrard et al, 2003, p. 560).

Les méthodes statistiques correspondantes sont fondées sur l'analyse structurelle des matrices de covariance ; une dénomination plus rigoureuse est « méthodes d'équations structurelles avec variables latentes et erreurs de mesure ». Nous allons présenter séparément les deux composantes des modèles d'équations structurelles dit de « causalité » (Evrard et al, 2003, p. 560) à savoir : l'analyse factorielle confirmatoire (modèle de mesure) et les équations simultanées (modèle de structure) ; puis nous exposerons leur intégration au sein des équations structurelles avec variables latentes et erreurs de mesure.

- L'analyse factorielle confirmatoire

Comme indiqué précédemment, on utilise le plus souvent l'analyse factorielle comme une méthode de réduction de données, permettant de résumer un ensemble de variables observées en un nombre plus réduit de facteurs ; on cherche à identifier de façon exploratoire des dimensions sous-jacentes aux variables observées. Le nombre de facteurs à retenir, puis l'interprétation et la signification des facteurs, éventuellement à l'aide d'une rotation, sont déterminés à posteriori, en fonction des règles empiriques (Evrard et al, 2003, p. 561).

L'analyse factorielle peut également être utilisée de façon confirmatoire pour tester des hypothèses définies à priori. Il s'agit de confronter aux données empiriques des hypothèses sur la structure des relations entre les variables observées et les variables latentes ou facteurs. L'analyse factorielle confirmatoire est surtout utilisée pour certifier la qualité psychométrique de résultats obtenus à l'issue d'une phase exploratoire (analyse factorielle exploratoire) (Evrard et al, 2003, p. 561). Dans un modèle factoriel restreint, lorsque les relations causales (représentées par le sens des flèches) vont des variables

latentes vers les variables observables qui les reflètent, on parle d'indicateurs « réflexifs ». Dans le cas où les sens des relations est inverse (par exemple, dans le cas où la variable latente est conçue comme un indicateur agrégé ou résumé établi à partir d'indicateurs partiels), on parle alors d'indicateurs « formatifs ». Cependant, le premier cas correspond aux applications les plus fréquentes.

Les critères d'interprétation de l'analyse factorielle confirmatoire reposent sur les éléments suivants :

-la part de variance de chaque item x_i , commune avec son facteur de rattachement. Cet indicateur permet de repérer les items faiblement liés à la variable latente qu'ils sont supposés mesurer ;

-la part de variance des facteurs composites, partagée avec l'ensemble des items x_i mesurant ce facteur ; cet indicateur peut être considéré comme un indice de fiabilité équivalent au coefficient alpha de Cronbach ;

-le pourcentage de variance « vraie », extraite de l'ensemble des items, mesurant un facteur en comparaison de la variance de l'erreur de mesure ; on considère généralement que cet indicateur doit être au moins égal à 0.5 (la variance due à l'erreur de mesure ne devant pas dépasser la variance vraie pour vérifier un niveau de convergence raisonnable) et qu'il doit être supérieur à la variance partagée par le facteur avec n'importe quel autre facteur (c'est-à-dire le carré de leur corrélation) pour vérifier la validité discriminante.

- Les modèles structurels

Ils ont pour objet de représenter des relations complexes au sein d'un ensemble de variables. Sous leur forme traditionnelle (équations simultanées, ou « path analysis »), ils permettent de résoudre les problèmes de prise en compte des interactions entre les variables et d'interprétation en terme de causalité ou tout du moins de direction de l'influence concernant la nature des relations entre variables. L'interprétation causale doit alors se situer au niveau conceptuel des variables latentes et non au niveau empirique des mesures (Evrard et al, 2003, p. 564).

On peut retenir une définition de la causalité selon Evrard et al., (2003, p. 565) comme :

-une relation asymétrique entre une cause (ou variable indépendante, ou variable exogène) et un effet (ou variable dépendante, ou endogène) : il s'agit de la causalité simple ;

-ou de façon plus générale, un réseau de relations entre des causes et des effets : il s'agit de la causalité complexe ;

-dans certains cas, la relation entre deux variables peut être réciproque : on parle alors de causalité circulaire.

Une autre interrogation fréquente du chercheur concerne l'effet d'une variable (modératrice ou médiatrice) sur la relation entre une variable explicative et une variable à expliquer. Les méthodes d'équations structurelles contribuent à répondre à ces questions.

Les équations structurelles représentent un réseau de relations sous forme d'un ensemble d'équations linéaires reliant des variables endogènes aux variables exogènes. Ces méthodes peuvent être considérées comme des extensions de la régression, elles contribuent à résoudre un système d'équations simultanées représentant le réseau des relations entre variables. En économie, l'accent est plutôt mis sur la prédiction des variables endogènes, dans le contexte d'applications à des fins prévisionnelles ; alors qu'en sociologie, l'accent est mis davantage sur le calcul des coefficients reliant les variables à des fins de compréhension du réseau de relation.

La construction d'un modèle d'équations structurelles comprend quatre étapes :

- la spécification : c'est la formulation des hypothèses causales à tester ;
 - l'identification : il s'agit de préciser les paramètres à calculer, et de comparer leur nombre aux données disponibles ;
 - l'estimation : c'est le calcul des paramètres ; différentes méthodes d'estimation sont disponibles selon la forme particulière du modèle (par exemple, selon qu'il y'a ou non causalité circulaire : modèles non récursifs ou récursifs) ;
 - le test : c'est la vérification de l'adéquation du modèle aux données, au niveau général (l'ensemble du modèle), ou analytique (chaque paramètre).
- Les équations structurelles avec variables latentes : le modèle général

Le modèle général utilisé pour l'analyse de causalité va résulter de la conjonction :

- d'un modèle d'équations structurelles linéaire reliant deux ensembles de variables latentes : des variables exogènes (ξ) et des variables endogènes (η), dont le système d'équation est représenté comme suit :

$$\eta = \beta\eta + \Gamma\xi + \zeta$$

- de deux modèles de mesure reliant les variables latentes aux variables observées qui leur correspondent respectivement (soit X correspondant à ξ et Y à η) dont les systèmes d'équation sont :

$$X = \Lambda_x\xi + \varepsilon$$

$$Y = \Lambda_y\eta + \delta$$

Où ζ , ε et δ sont les résidus respectifs des variables latentes endogènes (η), des variables observées exogènes (X) et endogènes (Y).

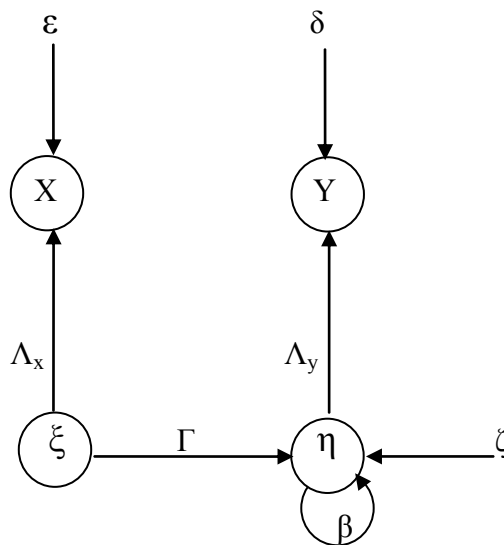
Avec :

- **Y** est le vecteur des variables endogènes (dépendantes) observées ;
- **X** est le vecteur des variables exogènes (indépendantes) observées ;
- **η** (*ETA*) est le vecteur des variables endogènes latentes ;
- **ξ** (*KSI*) est le vecteur des variables exogènes latentes ;
- **Λ_y** (*LAMDA Y*) est la matrice des coefficients reliant les variables endogènes latentes aux variables endogènes observées ;
- **Λ_x** (*LAMDA X*) est la matrice des coefficients reliant les variables exogènes latentes aux variables exogènes observées ;
- **ε** (*EPSILON*) est un vecteur de résidus pour les variables endogènes observées ;
- **δ** (*DELTA*) est un vecteur de résidus pour les variables exogènes observées ;
- **ζ** (*ZETA*) est le vecteur des résidus pour les variables endogènes latentes ;
- **Γ** (*GAMMA*) est la matrice des relations causales entre variables latentes exogènes et endogènes ;

β (*BETA*) est la matrice des relations causales entre variables latentes endogènes.

Schématiquement le modèle général se présente ainsi :

Figure 19: Le modèle général des équations structurelles avec variables latentes



Dans le modèle LISREL, l'estimation des paramètres est fondée comme pour l'analyse factorielle confirmatoire sur la minimisation d'une fonction du maximum de vraisemblance, rapprochant la matrice de covariance S calculée sur les données et la matrice théorique Σ (fondée sur le modèle testé).

- L'interprétation des résultats

L'évaluation des résultats issus d'une analyse par le modèle d'équations structurelles est fondée sur des indicateurs de la qualité de l'ajustement entre les données empiriques et le modèle testé ; ces indicateurs se situent à trois niveaux :

-Au niveau global (c'est-à-dire pour l'ensemble du modèle) : ici on peut utiliser un critère (« overall goodness of fit ») mesurant l'écart entre les matrices S et Σ . On peut également compléter ce critère global par d'autres critères formulés en terme de variance expliquée (GFI « Goodness of Fit Index », et AGFI « Adjusted Goodness of Fit Index »). Enfin, un autre critère consiste à étudier les résidus (c'est-à-dire les écarts entre résultats calculés et données) ; l'indice global correspondant est le RMR (« Root Mean square Residual »).

-Au niveau des parties du modèle : pour chacune des trois parties principales du modèle (les deux modèles de mesure et la structure des équations), un indicateur peut également être calculé. Il s'agit du coefficient de détermination. Il varie de 0 à 1 et l'ajustement est d'autant meilleur que le coefficient est proche de 1.

-Au niveau de chaque paramètre : chaque paramètre peut enfin faire l'objet d'un test partiel de signification. Un niveau d'exigence élevé consiste à n'accepter le modèle comme valide que si tous les paramètres sont significatifs et du signe attendu.

2-2-D- Fiabilité et validité de la mesure

Le problème de fiabilité est concerné par l'erreur aléatoire. En effet, un instrument est fiable si son utilisation répétitive donne les mêmes résultats. Un instrument est d'autant plus fiable que l'erreur aléatoire est minimale.

Le problème de la validité dépend de l'ensemble des termes d'erreur, c'est-à-dire de l'écart entre la mesure obtenue et la vraie valeur (inconnue).

Ainsi, un instrument de mesure doit satisfaire aux critères suivants :

- La fiabilité (ou fidélité) : si on mesure un phénomène plusieurs fois avec le même instrument, on doit obtenir le même résultat.
- La validité : les instruments de mesure choisis doivent permettre d'appréhender le mieux possible le phénomène que l'on cherche à mesurer.
- La sensibilité : il s'agit de se doter d'un instrument capable d'enregistrer des variations assez fines du phénomène mesuré.

Les deux premiers critères sont les plus généralement retenus pour juger de la qualité d'une mesure et sont intégrés dans le modèle dit « de la vraie valeur » qui constitue le fondement de la théorie de la mesure.

2-2-D-1- Les techniques d'amélioration de la fiabilité

Carmines et Zeller (1990)⁹⁸ présentent quatre méthodes qui permettent d'estimer la fiabilité :

- La méthode du « test-retest » : elle consiste à effectuer le même test sur les mêmes individus à des périodes différentes. On calcule ensuite un coefficient de corrélation entre les deux tests successifs. Plus la valeur de ce coefficient est proche de 1, plus les mesures sont proches et plus fiable est l'outil utilisé ;

⁹⁸ Carmines E. et Zeller R. (1990): *Reliability and Validity Assessment*, Sage, London, cités par Drucker-Godard C., Ehlinger S. et Grenier C. (2003), « Validité et fiabilité de la recherche », in Thietart R.-A. (sous la direction de), *Méthodes de recherche en management*, 2^{ème} édition, Dunod, Paris, pp. 257-287.

-La méthode des formes alternatives : elle consiste à interroger les mêmes sujets avec des outils différents au même moment. Par exemple dans un questionnaire, plusieurs questions permettant de mesurer le même phénomène sont introduites, mais en le formulant de façon différente ;

- La méthode du « split half » (ou des deux moitiés) : selon Evrard et al. (2003, p. 304), elle consiste à utiliser le même instrument de mesure au même moment avec des sujets différents, puis on scinde ensuite l'échantillon en deux moitiés et on compare les résultats de ces deux moitiés. Drucker-Godard et al. (2003) proposent plutôt d'utiliser le même outil au même moment sur les mêmes individus mais en séparant l'ensemble des items d'une échelle en deux moitiés, chaque moitié devant être représentative de l'attitude que l'on cherche à mesurer et contenir un nombre suffisamment important d'items pour être significative. Ensuite, on calcule un coefficient de corrélation, notamment celui de Spearman-Brown, sur les réponses obtenues dans chaque moitié ;

-La méthode de la cohérence interne : des méthodes ont été développées pour pallier à la principale limite de la méthode précédente. Elles visent à estimer des coefficients de fiabilité qui mesurent la cohérence interne de l'échelle. Le plus connu et le plus utilisé de ces coefficients est l'alpha de Cronbach. Cependant, le Rhô de Ksi ou rho de Jöreskog créé par Jöreskog (1973) s'avère être plus efficace voire plus pertinent pour la mesure de la fiabilité. Pour cette raison, de nos jours dans la pratique des recherches, cet indicateur est de plus en plus préféré au traditionnel alpha de Cronbach. Ces deux coefficients seront utilisés dans la présente recherche pour apprécier la cohérence interne des échelles.

2-2-D-1-1- L'alpha de Cronbach

La mesure d'un phénomène ou d'une variable latente se fait souvent à l'aide d'une échelle multiple, c'est-à-dire d'un certain nombre de questions ou d'items. Une fois les dimensions de la variable identifiées, il est important de savoir si les items retenus pour mesurer la variable ou chacune de ses dimensions sont cohérents c'est-à-dire reflètent le même phénomène ou ne mesurent pas des facettes différentes de celui-ci. L'alpha de Cronbach permet de juger de cette cohérence. Le coefficient alpha (α) s'exprime par la formule suivante :

$$\alpha = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum_i \sigma_i^2}{\sum_i \sigma_i^2 + 2 \sum_{i,j} \sigma_{ij}} \right)$$

Où :

k est le nombre de questions (ou items) ;

σ_i^2 est la variance de l'item i (erreur aléatoire) ;

$\sigma_{i,j}$ est la covariance entre l'item i et l'item j ;

Si les items sont censés mesurer le même phénomène, ils doivent être corrélés ; la covariance $\sigma_{i,j}$ doit être élevée ;

-si $\sigma_{i,j}$ est élevé, α est proche de 1,

-si $\sigma_{i,j}$ est faible, α est proche de 0.

Ainsi, quand les questions ont un α qui se rapproche de 1, l'échelle (c'est-à-dire l'ensemble des items) a une bonne cohérence interne ; les questions censées mesurer la même chose mesurent effectivement la même chose. On pourra donc prendre la somme des scores des items comme mesure synthétique. Quand les questions mesurent des phénomènes différents, l' α se rapproche de 0 et la cohérence interne de l'échelle est faible.

Il n'existe pas de distribution statistique connue permettant de conclure si l' α est acceptable ou non. En revanche, des seuils empiriques, donnés par l'expérience des études en psychométrie, peuvent servir de référence. On considère ainsi que, pour une étude exploratoire, l' α est acceptable s'il est compris entre 0,6 et 0,8 ; pour une étude confirmatoire, une valeur supérieure à 0,8 est recommandée (Evrard et al, 2003, p. 350).

En dépit de ces préconisations, il est à noter que la valeur du coefficient Alpha est influencée par le nombre d'items et la nature du construit. Denis Darpy (2002) (cité par Evrard et al, 2003, p. 309) affirme qu'une méta-analyse du coefficient alpha de Cronbach sur 30 années dans la littérature marketing montre que les échelles publiées à 2 ou 3 items ont un alpha moyen de 0,73 alors que les échelles composées de plus de trois items ont un alpha moyen de 0,77. Cette limite justifie la préférence du Rhô de Jöreskog dans certains cas, notamment dans les analyses confirmatoires.

2-2-D-1-2- Le Rhô de Jöreskog

On vient de le voir, l'évaluation de la fiabilité de cohérence interne peut se faire à l'aide du traditionnel Alpha de Cronbach. Mais avec la méthode des équations structurelles, le rhô de validité interne ou rhô de Jöreskog (1973) est l'outil le plus adapté puisqu'il intègre de manière très explicite les termes de l'erreur (Roussel et al, 2002, p. 55). Créé par Jöreskog (1973), d'où son nom « Rhô de Jöreskog » (on l'appelle aussi « Rhô de Ksi »), ce coefficient de fiabilité de cohérence interne s'appuie sur une mesure des contributions factorielles ou

loadings des items ou indicateurs. Il a été conçu pour compléter les analyses factorielles confirmatoires. Selon Roussel et al. (2002, p. 78), le Rhô de Jöreskog serait, contrairement au coefficient Alpha, moins sensible au nombre d'items et devrait être privilégié dans les phases confirmatoires de développement des échelles si l'on a recours à l'AFC. Il n'existe pas de règles précises pour interpréter ce coefficient. On considère que la fiabilité du construit est bonne si la valeur du Rhô est supérieure à 0,7 ou 0,8 (Fornell et Larker, 1981). Son objectif est de vérifier que les indicateurs spécifiés dans le modèle représentent suffisamment bien le construit. Il se calcule grâce à la formule suivante :

$$\rho_{(A)} = \frac{\left(\sum_{i=1}^n \lambda_i\right)^2}{\left(\sum_{i=1}^n \lambda_i\right)^2 + \sum_{i=1}^n \text{var}(\varepsilon_i)}$$

Avec :

$\rho_{(A)}$: rhô de Joreskög, qui est l'indice de fiabilité de cohérence interne de la mesure de la variable théorique A ;

λ_i : contribution factorielle (*loading*) de l'indicateur i sur la variable théorique issue d'une AFC ;

$\text{var}(\varepsilon_i) = 1 - \lambda_i^2$, c'est la variance du résidu de l'indicateur i.

2-2-D-2- Les techniques d'amélioration de la validité

La validité s'exprime sur le plan interne et externe. La validité d'un instrument de mesure renvoie à l'idée que celui-ci mesure à la perfection ce qu'il est censé mesurer et non autre chose. Il convient de noter que l'on peut appréhender la validité pas seulement au niveau de l'instrument de mesure mais aussi au niveau du modèle de la recherche (on parle de validité de la recherche). Evrard et al. (2003) distinguent trois types de validité :

-La validité « faciale » (ou validité de contenu) : elle est fondée sur le jugement du chercheur, appuyé par les pairs de sa communauté scientifique au sein duquel l'instrument utilisé fait l'objet d'un consensus. La validité « faciale » rejoint ainsi la notion « d'intersubjectivité ». Il s'agit en particulier de savoir si la mesure opératoire capture les différents aspects (ou « facettes ») du phénomène étudié.

-La validité de trait (ou validité de construit) : ce type de validité nous intéresse plus particulièrement ; Il s'agit de savoir si les indicateurs construits sont une bonne représentation du phénomène à étudier. Deux formes permettent d'en vérifier la pertinence : il s'agit de la validité convergente et de la validité discriminante. En effet, lorsqu'on mesure la validité interne, il faut vérifier si les indicateurs qui sont supposés mesurer le même phénomène sont corrélés (validité convergente). En revanche, si les indicateurs sont supposés mesurer des phénomènes différents, ils doivent être faiblement corrélés entre eux car ils doivent permettre de discriminer les phénomènes entre eux (validité discriminante). Il faut préciser que cette démarche est une démarche de réfutation. On ne parvient jamais à démontrer la validité absolue, mais seulement la non validité des résultats. La validité de trait consiste à vérifier d'une part si la mesure est spécifique et d'autre part si elle mesure bien le trait qu'elle est censée mesurer. Cette validité se vérifie à l'aide de la méthode multi traits multi méthodes (MTMM).

Cette méthode est très lourde à mettre en œuvre et ne donne pas toujours des résultats satisfaisants. Fornell et Larcker, (1981) proposent une méthode alternative au MTMM qui consiste à comparer le pourcentage de variance qu'une variable latente (facteur) partage avec ses mesures (items) au pourcentage de variance qu'elle partage avec d'autres variables latentes. La validité convergente est assurée si la variable latente partage au moins 50 % de sa variance avec chacune de ses mesures. Elle est obtenue par la formule suivante :

$$\rho_{vc(\eta)} = \frac{\sum_{i=1}^n (\lambda_i^2)}{\sum_{i=1}^n (\lambda_i^2) + \sum_{i=1}^n \text{var}(\varepsilon_i)}$$

Où :

$\rho_{vc(\eta)}$ est le critère de validité convergente de la variable latente η ou variance partagée avec les mesures

λ_i est le « corrélateur » de l'item i avec le facteur ; $i = 1$ à p items

$\text{var}(\varepsilon_i)$ est la variance des termes d'erreur ($\text{var}(\varepsilon_i) = 1 - \lambda_i^2$).

Si la valeur de $\rho_{vc(\eta)}$ est inférieure à 0,50, cela signifie que la variance due aux erreurs de mesure est supérieure à la variance capturée par le construit, ce qui revient à dire que la validité des indicateurs individuels (y_i) ainsi que celle du construit η est remise en question.

Le coefficient de validité discriminante est obtenu en élevant au carré, les corrélations observées entre les variables latentes. Selon Fornell et Larcker (1981), le $\rho_{vc(\eta)}$ peut être utilisé pour évaluer la validité discriminante. Pour satisfaire la validité discriminante, les deux inéquations suivantes doivent être vérifiées :

$$\rho_{vc(\eta)} > \gamma^2$$

et

$$\rho_{vc(\xi)} > \gamma^2$$

Cela signifie que deux variables latentes (η et ξ), différentes théoriquement doivent également être distinctes statistiquement. Pour que cette condition soit satisfaite, il faut s'assurer que la variance moyenne extraite (ρ_{vc}) de chaque construit soit supérieure au carré des corrélations que celui-ci partage avec les autres construits (γ^2).

-La validité nomologique (ou validité prédictive) : elle concerne la liaison entre les concepts (et les « construits » qui les opérationnalisent). Il s'agit de savoir si les relations entre les mesures d'un concept et celles d'autres concepts sont ou non en conformité avec les prédictions issues de la théorie fondée sur les recherches précédentes. La validité nomologique et la validité « faciale » sont souvent en interaction.

2-2-E- L'évaluation de l'ajustement global du modèle d'équations structurelles

Il existe un grand nombre d'indices pour évaluer l'ajustement global du modèle de mesure (comme du modèle structurel). Dans les analyses, les indices les plus utilisés sont : le GFI (*Goodness of Fit Index*), l'AGFI (*Adjusted Goodness of Fit Index*), et le NFI (*Normed Fixed Index*). Galan (2003) souligne que le problème de ces indices est qu'ils sont dépendants de la taille de l'échantillon, ce qui n'est pas le cas des indices les plus récents que sont le *Tucker-Lewis Index* (TLI ou NNFI) et le *Comparative Fit Index* (CFI). A ceux-ci, il faut ajouter le *Root Mean Square Error* (RMSEA). De façon générale on distingue : les indices absolus, les indices relatifs et les indices de parcimonie.

2-2-E-1- Les indices absolus

Ils permettent d'évaluer dans quelle mesure le modèle théorique reproduit les données observées. Plusieurs indices sont disponibles : certains testent la qualité de l'ajustement (indice dits de « Goodness of Fit », d'autres testent au contraire le mauvais ajustement (indices

dits de « non-centralité »). Certains indices présentent non seulement une valeur centrale, mais aussi un intervalle de confiance.

2-2-E-1-1- indices de Goodness of Fit

-L'indice du Khi deux: Le principe est de comparer la matrice de covariances (ou de corrélations) prédite par le modèle avec celle que donnerait un modèle « saturé » qui prédirait parfaitement les données observées. On obtient un indice de χ^2 (en fait un pseudo- χ^2) qui devrait idéalement être non significatif (contrairement aux analyses classiques). En pratique il l'est souvent, surtout avec les échantillons de plus de 200 sujets car la taille du χ^2 augmente avec la taille de l'échantillon. Cet indice est fonction de la taille de l'échantillon et est lié à la multi normalité. Le Khi deux est calculé avec une valeur seuil : $(N - 1) * F(V)$ degrés de liberté.

Où : $F(V)$ correspond à la fonction de vraisemblance (< 0) et N est la taille de l'échantillon

-L'indice GFI (Goodness of Fit Index): il est l'œuvre de Joreskog et Sörbom (1993)⁹⁹, dans le cas d'une estimation selon le maximum de vraisemblance. C'est la part relative de la variance/covariance expliquée par le modèle. Peu sensible à la taille de l'échantillon, il est cependant très sensible à la complexité du modèle. Il varie entre 0 et 1. Lorsqu'il est proche de 1, le modèle est de bonne qualité et lorsqu'il est proche de 0, le modèle n'est pas acceptable. Selon Croutsche (2002), il faut que cet indice soit supérieur à 0,9 pour que le modèle soit de bonne qualité.

-L'indice AGFI (Adjusted Goodness of Fit Index): nous le devons aux deux auteurs précédents (Jöreskog et Sörbom). C'est la part de variance/covariance expliquée par le modèle ajusté par le nombre de variables par rapport au nombre de degrés de liberté. L'AGFI est peu sensible à la taille de l'échantillon mais très sensible à la complexité du modèle. Cet indice s'interprète exactement comme le précédent. La différence est que l'AGFI prend en compte le nombre de paramètres estimés. Plus ceux-ci sont nombreux, plus on a de chance que le modèle soit de qualité. Il est à noter que AGFI est toujours inférieur à l'indice GFI. Sa formulation est :

$$AGFI = 1 - \left[\frac{k(k + 1)}{2ddl} \right] (1 - GFI)$$

⁹⁹ Jöreskog K. et Sörbom D. (1993): LISREL 8: User's reference guide. Chicago: Scientific Software International, cités par Evrard Y. Pras B. et Roux E. (2003) en collaboration avec Desmet P., Dussaix A-M et Lilien G: Market : études et recherches en marketing, 3^{ème} édition, Dunod, 684 p.

-L'indice RMR (Root Mean square Residual): Il consiste à étudier les résidus (c'est-à-dire les écarts entre résultats calculés et données). Il se calcule en prenant la racine carrée du carré moyen des résidus. Les résidus peuvent être interprétés comme reflétant l'influence des facteurs non pris en compte dans le modèle. Plus la valeur de cet indice est proche de zéro, plus l'ajustement est de qualité. Dans tous les cas, RMR ne doit pas être $> 0,10$. Le modèle est estimé de bonne qualité si $RMR < 0,05$.

2-2-E-1-1- indices de non-centralité (sorte de Badness of Fit)

-L'indice RMSEA (Root Mean Square Error of Approximation): de Steiger (1989)¹⁰⁰, Il représente la différence moyenne par degré de liberté, attendue dans la population totale et non dans l'échantillon. Il est indépendant de la taille de l'échantillon et de la complexité du modèle et dispose d'un intervalle de confiance associé (90 %). Cependant, son seuil d'acceptation varie selon les auteurs, inférieur à 0.05 ou 0.08 (Evrard et al., 2003). Sa formule est :

$$RMSEA = \sqrt{\frac{2F_{min}}{[p(q + 1)]}}$$

Où:

F_{min} est la fonction de vraisemblance à son minimum

p le nombre de variables manifestes ;

q le nombre de paramètres à estimer.

- **PNI** : paramètre de non centralité de la population, très sensible à la complexité du modèle mais peu à la taille de l'échantillon.

- **PNNI** de McDonald (1996): Il s'agit du PNI normé.

-**Gamma 1 et 2** : variantes des GFI/AGFI calculés selon la matrice de variance/covariance induite par la population totale et non l'échantillon.

¹⁰⁰ Steiger J. H. (1989): *EzPATH: Causal modeling*. Evanston, IL: SYSTAT, cité par Evrard Y. Pras B. et Roux E. (2003) en collaboration avec Desmet P., Dussaix A-M et Lilien G: *Market : études et recherches en marketing*, 3^{ème} édition, Dunod, 684 p.

2-2-E-2- Les indices incrémentaux ou relatifs

Ces indices visent à comparer deux modèles ; ils sont généralement utilisés pour comparer le modèle testé au modèle le plus contraint qui est le « modèle d'indépendance » où toutes les corrélations entre variables observées seraient nulles. Les principaux indices de cette famille sont :

-L'indice NFI (Normal Fit Index) : Cet indice représente la proportion de la variance totale entre les variables expliquées par le modèle testé lorsque le modèle nul est pris comme référence. Pour être acceptable, sa valeur doit être supérieure à 0,9. Sa formule mathématique est :

$$NFI = \frac{T_0 - T_m}{T_0}$$

Où :

$T_0 = \chi^2$ du modèle de base (modèle nul)

$T_m = \chi^2$ du modèle à tester.

-L'indice NNFI (Non-Normal Fit Index) ou TLI (Tucker Lewis Index) : Cet indice compare le manque d'ajustement du modèle à tester à celui du modèle de base. Sa valeur permet d'estimer l'amélioration relative par degré de liberté, du modèle à tester par rapport au modèle de base. Le TLI ou NNFI n'est pas recommandé pour les petits échantillons ($N < 150$). Sa formule mathématique est :

$$TLI = NNFI = \frac{\left[\left(\frac{T_0}{ddl_0} \right) - \left(\frac{T_m}{ddl_m} \right) \right]}{\left[\frac{T_0}{ddl_0} - 1 \right]}$$

Où :

ddl_0 est le nombre de degrés de liberté du modèle de base ;

ddl_m est le nombre de degrés de liberté du modèle à tester.

-L'indice IFI (Incremental Fit Index) : Cet indice compare le manque d'ajustement du modèle à tester à celui du modèle de base. Comme tous les indices de cette nature, sa valeur est jugée acceptable au seuil de 0,9. Et est moins sensible que le TLI/NNFI pour les petits échantillons.

Sa formulation mathématique est :

$$IFI = \frac{T_0 - T_m}{T_0 - ddl_m}$$

Où ddl_m est le nombre de degrés de liberté du modèle à tester.

-L'indice CFI (Comparative Fit Index) : Cet indice mesure la diminution relative du manque d'ajustement. Cette diminution est estimée suivant la distribution non centrée du χ^2 du modèle à tester par rapport au modèle de base. Il corrige une lacune du NFI en remplaçant la distribution centrée du χ^2 par la distribution non centrée du χ^2 (Roussel et al, 2002, p. 67). Le CFI est identique au RNI (*Relative Noncentrality Index*) de McDonald et Marsh (1990)¹⁰¹.

2-2-E-3- Les indices de parcimonie

Ils visent à tenir compte du fait qu'il est autant plus facile d'estimer un modèle que le nombre de paramètres à estimer est plus élevé ; pour cela ils complètent les indicateurs relatifs en tenant compte du nombre de degré de liberté du modèle testé et du modèle servant de base à la comparaison.

-Le Chi-deux normé : C'est le Chi-deux divisé par le nombre de degré de liberté. Il permet de déceler les modèles « sur-ajustés » et « sous-ajustés ». Il permet également de distinguer parmi plusieurs modèles alternatifs, lequel est le plus parcimonieux. Sa valeur doit être la plus faible possible et ne doit pas dépasser 3, certains auteurs autorisent 5 (Galan, 2003, p. 181).

Sa formule de calcul est la suivante : $\chi^2_{normé} = \frac{\chi^2}{ddl_m}$

-L'indice PNFI (Parsimony Normed Fit Index) : Le PNFI résulte de l'ajustement du NFI par rapport aux degrés de liberté du modèle nul et du modèle testé. Il possède les mêmes caractéristiques que le NFI et ne peut être utilisé que pour la comparaison de différents modèles alternatifs.

Sa formulation mathématique est : $PNFI = \left(\frac{ddl_m}{ddl_0} \right) NFI$

¹⁰¹ Mc Donald R.P. et Marsh H.W. (1990): "Choosing a multivariate model: Noncentrality and goodness of fit", Psychological Bulletin, Vol.107, pp. 247-255.

-L'indices AIC (Akaike Information Criterion) Cet indice issu de la théorie de l'information vise à pénaliser les modèles complexes. Il peut dans une certaine mesure être considéré comme un indice de parcimonie puisqu'il corrige la statistique globale d'ajustement pour tenir compte du nombre de paramètres à calculer. Il est particulièrement utilisé pour la comparaison des modèles. Sa formulation mathématique est la suivante :

$$AIC_M = \chi^2 - 2ddl_m$$

Tableau 10: Synthèse des valeurs souhaitées généralement admises des principaux indices d'ajustement

	Indices	Valeurs souhaitées
Indices absolus	χ^2 et scaled χ^2	+ petite possible (voir p associé)
	GFI, AGFI, Gamma 1 et 2	> 0.90
	PNI	+ petite possible
	PNNI	> 0.95
	RMR et SRMR	Le plus petit possible, au gré du chercheur
	RMSEA	< .08 ou mieux < .05
Indices incrémentaux	NFI et ρ de Bollen	> .90
	NNFI et δ de Bollen	> .90
	CFI et RNI	> .90
Indices de parcimonie	χ^2 normé	+ petite possible, entre 1, 2-3, voire 5
	AIC	+ petite possible (comparaison)
	PNFI	+ forte possible (comparaison)

Source : Adapté de Roussel et al. (2002) et Galan (2003).

En définitive, ce chapitre nous a permis de présenter les choix méthodologiques de notre recherche. Nous avons dans la première section exposé et décrit le positionnement épistémologique et l'approche méthodologique adoptée. Ces différents choix et la nature de notre problématique nous ont conduits à poursuivre deux étapes de recherche pour mener notre étude empirique, à savoir une étape exploratoire et une étape confirmatoire. En effet, pour mettre en place nos deux phases de recherche, une réflexion sur les choix méthodologiques est alors posée pour comprendre la différence entre une approche qualitative et une approche quantitative. Pour bien circonscrire notre problématique, nous avons enfin exposé le processus général de notre recherche.

Dans la deuxième section, nous avons présenté les phases de recueil et de préparation d'analyse des données. Dans un premier temps nous avons décrit le recueil des données qualitatives lors de la phase exploratoire, en détaillant la conduite des entretiens au près des dirigeants de PME et le guide d'entretien utilisé ; puis nous avons exposé les préalables à l'analyse des données de la phase exploratoire. Dans un deuxième temps, nous avons décrit le recueil des données quantitatives en exposant la construction du questionnaire et les étapes de déroulement de l'enquête ; puis nous avons présenté les préalables à l'analyse des données lors de la phase confirmatoire.

Comme le suggère de plus en plus la recherche en sciences de gestion, nous avons adopté une approche méthodologique mixte (qualitative et quantitative) qui assure la multiplication des angles de vue et permet, par conséquent, de produire des résultats plus riches et de meilleure qualité. Les résultats de la phase exploratoire et de la phase confirmatoire sont présentés dans le Chapitre 4.