



HAL
open science

Comment limiter les biais liés au choix des échelles de mesure dans les études marketing ?

Antoine Legardinier

► **To cite this version:**

Antoine Legardinier. Comment limiter les biais liés au choix des échelles de mesure dans les études marketing ?. Gestion et management. 2013. dumas-00933905

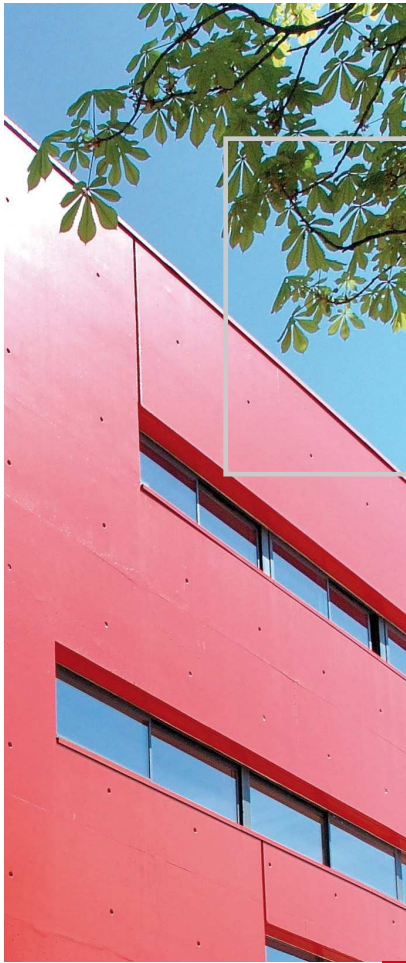
HAL Id: dumas-00933905

<https://dumas.ccsd.cnrs.fr/dumas-00933905v1>

Submitted on 14 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Mémoire de stage / recherche

Comment limiter les biais liés au choix des échelles de mesure dans les études marketing ?



Présenté par : Antoine LEGARDINIER

Nom de l'entreprise : IFOP

Tuteur entreprise : Julien BELIN

Tuteur universitaire : Jean-Luc GIANNINI

**Master DEG 1^{ère} année mention Marketing – FI
Master Marketing
2012 - 2013**



Avertissement :

L'IAE de Grenoble, au sein de l'Université Pierre-Mendès-France, n'entend donner aucune approbation ni improbation aux opinions émises dans les mémoires des candidats aux masters en alternance : ces opinions doivent être considérées comme propres à leur auteur.

Tenant compte de la confidentialité des informations ayant trait à telle ou telle entreprise, une éventuelle diffusion relève de la seule responsabilité de l'auteur et ne peut être faite sans son accord.

Mémoire de stage / recherche

Comment limiter les biais liés au choix des échelles de mesure dans les études marketing ?



Présenté par : Antoine LEGARDINIER

Nom de l'entreprise : IFOP

Tuteur entreprise : Julien BELIN

Tuteur universitaire : Jean-Luc GIANNINI ONI

REMERCIEMENTS

Dans un premier temps, je souhaite remercier **Julien BELIN**, directeur du pôle services de l'IFOP pour la confiance qu'il m'a accordé, durant ces cinq mois de stage.

Je tiens également à remercier tous les membres de ce service qui m'ont chaleureusement accueilli, ce qui m'a permis de m'intégrer très rapidement. Je pense tout particulièrement à **Florence PUIS**, **Marion LAMBOLEY**, **Ornella JEUDANE**, **Simon ALLARD** & **Colas CASTIGNOLES** qui se sont montrés très disponibles tout au long de ce stage pour répondre à mes questions sur le fonctionnement du service. Vous n'avez pas hésité à m'aider dans mes tâches quotidiennes, et à répondre à mes questions que ce soit sur les enquêtes réalisées ou sur le métier. Merci de m'avoir apporté du temps pour découvrir votre métier : celui de chargé d'études. Merci aussi à **Céline FERRET**, **Jérôme BENOIT** et **Rudy BOURLES** pour vos conseils très précieux fournis durant cette période de stage. Grâce à vous, j'ai une vision très complète sur le métier de chargé d'études en institut de sondage.

Merci également à **Elodie BANQ**, **Christophe JOURDAIN**, **Mathieu PRIGENT** et **Emilie FERNANDES** de m'avoir permis de mettre en place l'étude présentée dans ce mémoire.

Enfin, je souhaite remercier les enseignants de l'IAE de Grenoble, et tout particulièrement **Jean-Luc GIANNELLONI**, tuteur de mon stage pour sa disponibilité et ses conseils durant cette période de stage de fin d'année.

Résumé

Ce mémoire a pour objectif de mettre en évidence les biais liés à l'utilisation des échelles de mesure dans les questionnaires. Après avoir présenté les différentes échelles de mesure qui existent et qui sont fréquemment utilisées dans les études de marché, nous montrerons comment en construire une via les différentes étapes du paradigme de Churchill (1979).

Au terme d'une étude menée avec l'omnibus *on line* de l'IFOP auprès de 1000 internautes Français, nous tenterons de mettre en avant les principales différences d'interprétation des résultats lorsqu'on intègre une modalité neutre, et le « ne sait pas » dans l'échelle de mesure.

Mots clés

Echelles de mesure ; biais ; modalité neutre ; NSP

REMERCIEMENTS.....	4
RESUME.....	5
SOMMAIRE.....	6
AVANT-PROPOS	7
1) L'INSTITUT FRANÇAIS D'OPINION PUBLIQUE – IFOP : PIONNER DES ETUDES	7
2) CHIFFRES IMPORTANTS.....	8
3) L'ORGANISATION ET LE FONCTIONNEMENT EN BREF.....	8
4) LES FORCES DE L'IFOP.....	9
INTRODUCTION.....	10
PARTIE 1 : QU'EST-CE QU'UNE ECHELLE DE MESURE ?	12
1) LES ECHELLES DE MESURE EXISTANTES.....	12
A) INVENTAIRE DES ECHELLES DE MESURE EXISTANTES.....	12
B) APPLICATIONS STATISTIQUES	15
2) LES ECHELLES DE MESURE INEXISTANTES.....	18
A) CONSTRUIRE UNE ECHELLE DE MESURE	18
B) LE MODELE DE CHURCHILL	18
C) PERTINENCE D'UNE ECHELLE DE MESURE.....	19
PARTIE 2 : PRINCIPAUX BIAIS DE MESURE.....	21
1) LIMITES & BIAIS DE MESURE DES ECHELLES	21
A) BARRIERES DE LA LANGUE.....	21
B) NOMBRE DE MODALITES	22
C) LES NON REponses : NSP.....	24
D) LES EFFETS DE HALO.....	25
2) MINIMISER LES BIAIS DE MESURE.....	27
A) CONTOURNER LES BIAIS DE MESURE	27
B) FAIRE APPEL A DES INSTRUMENTS DE MESURE PLUS ADAPTES.....	30
PARTIE 3 : ETUDE D'UN BIAIS LIE AU CHOIX DES ECHELLES.....	32
1) METHODOLOGIE	32
A) RAPPEL DU CONTEXTE	32
B) MISE EN PLACE DE L'ETUDE	32
C) LES ECHELLES A TESTER.....	33
D) LE CHOIX DES QUESTIONS	34
E) PROFILS SOCIODEMOGRAPHIQUES.....	35
2) PRINCIPAUX RESULTATS	36
A) LE « NI D'ACCORD, NI PAS D'ACCORD » COMME MODALITE REFUGE	36
B) LE « VOUS NE SAVEZ PAS, VOUS N'AVEZ PAS D'OPINION » COMME MODALITE REFUGE DANS UNE QUESTION PEU IMPLIQUANTE. 38	38
C) « NI D'ACCORD, NI PAS D'ACCORD » <i>VERSUS</i> « VOUS N'AVEZ PAS D'OPINION, VOUS NE SAVEZ PAS »	39
D) LES FEMMES AVOUENT PLUS FACILEMENT QU'ELLES NE SAVENT PAS	41
3) APPORTS DE L'ETUDE.....	42
4) LIMITES DE L'ETUDE	43
CONCLUSION	45
BIBLIOGRAPHIE.....	46
ANNEXES.....	48

AVANT-PROPOS

La présente introduction a pour objectif de présenter l'IFOP, l'un des principaux instituts d'opinion et d'étude de marché en France, dans lequel j'ai effectué mon stage.

Avant d'accomplir cette courte introduction, il est important de revenir sur ce qui a déterminé mon choix. En effet, je souhaitais compléter mes précédents stages chez BVA en tant qu'assistant chargé de traitement, et chez LH2 en tant qu'assistant chargé d'études dans un nouvel institut d'études de marché. Mon but était donc d'approfondir mes connaissances et compétences pour devenir chargé d'études.

L'IFOP, par son histoire (pionnier en France) et sa renommée (6^{ème} institut français) en font un acteur incontournable et majeur sur son marché mais aussi dans la société. Effectivement, cet institut apporte – au-delà des nombreuses études ad hoc – des réflexions de fond sur des sujets sociaux, économiques et politiques. C'est donc tout naturellement que j'ai postulé pour un stage de chargé d'études à l'IFOP.

1) L'Institut Français d'Opinion Publique – IFOP : Pionner des études

Jean STOETZEL (professeur de psychologie sociale à la Sorbonne) rencontre le sociologue et statisticien américain Georges GALLUP qui lui fait comprendre « la nécessité pour les entreprises et les institutions administratives et politiques de mieux connaître les besoins et les motivations des différents acteurs de la vie économique et sociale ». C'est dans cette optique, qu'il fonde en 1938 l'institut français d'opinion publique – plus connu sous le sigle IFOP – pionner sur le secteur des études de marché et de sondages en France.

L'institut gagnera sa légitimité avec l'annonce de la mise en ballottage du Général de Gaulle en 1965, alors que tous étaient convaincus de sa victoire. Cette déclaration permis dans le même temps d'asseoir le bien fondé des chiffres comme représentation de la réalité, ce qui explique en partie la raison pour laquelle aujourd'hui encore une grande majorité des études sont quantitatives (72% chez l'IFOP). L'IFOP a également développé plusieurs techniques d'enquête - telles que les estimations aux sorties de votes ou les écoutes radio – qui sont aujourd'hui utilisées par la plupart des instituts.

L'IFOP verra ses concurrents arriver bien plus tard sur le marché (TNS en 1963, GFK et BVA en 1970, Ipsos en 1975, Médiamétrie en 1985 et bien d'autres ont suivi) ce qui lui a permis d'acquérir un savoir-faire et une expérience reconnue, notamment dans les études politiques. Malgré son avance, l'IFOP doit sans cesse se moderniser et s'adapter aux évolutions du marché pour rester parmi les leaders face à une concurrence accrue.

2) Chiffres importants

Date de création	1938
Chiffre d'affaires (2012)	39 M€
Effectifs total du groupe (2012)	210 (Dont plus de la moitié à Paris)
Répartition de l'activité (2012)	69% de marketing, 19% d'études opinion, 12% de Panels
Classement (2011)	6 ^{ème} institut français
Présence mondiale	4 zones géographiques : <ul style="list-style-type: none"> ✓ l'Europe (Paris) ✓ l'Amérique du Nord (Toronto) ✓ l'Amérique du Sud (Buenos Air) ✓ l'Asie (Shanghai, Hong Kong) Réalisation d'études dans plus de 50 pays

3) L'organisation et le fonctionnement en bref

Laurence PARISOT (à la tête du MEDEF jusqu'au 3 juillet 2013, et actionnaire principale de l'IFOP) prend la direction de l'IFOP en 1990 jusqu'en 2008 où elle a été remplacée par Stéphane TRUCHI (ancien PDG d'IPSOS France) et nommée vice-présidente. Cette date marque également un grand changement en termes d'organisation interne puisque l'IFOP va se structurer autour de 6 principaux pôles 'métier' qui permettent ainsi de mieux répondre aux attentes clients en développant des expertises sur certains secteurs d'activité à savoir :

Services

3 secteurs principaux : les finances & l'assurance, l'énergie et le transport

Opinion & stratégie d'entreprises

Études d'opinion, d'image institutionnelle et corporate, sondages électoraux

Santé

Études sur l'ensemble des problématiques santé

Consumer

Études marketing autour de 3 domaines : Grande Consommation, Bien être

Média & Numérique

Études d'audiences, de développement des médias notamment d'Internet

Luxe

Études sur l'ensemble des activités du secteur du luxe (parfums, mode, loisirs...)

A cela s'ajoute de nombreux postes transversaux (ex: service juridique, les ressources humaines...) ainsi que trois départements : Opération (études Omnibus, Grandes enquêtes, Panel), Quali et Planning Stratégique (cahier des tendances, box IFOP trends10, benchmarking...)

L'IFOP adapte son organisation interne aux évolutions du marché en restructurant ses équipes ou en créant de nouveaux pôles (ex : le pôle Beauty en 2012). La structuration en « pôle d'activité » est similaire à celle de nombreux instituts (Ipsos, TNS...) ce qui peut s'expliquer par la demande des clients qui cherchent une simplification de l'offre (un repérage facile) et une expertise secteur (valeur ajoutée).

4) Les forces de l'IFOP

L'IFOP peut aujourd'hui répondre à un large panel de problématiques clients quel que soit leur nature (privé/public, PME, firmes multinationales, institutions gouvernementales...). Sa mission ? Aider ses clients à prendre des décisions stratégiques en apportant une valeur ajoutée sur les études en termes de connaissances et en anticipant les tendances de demain. Pour cela, l'entreprise utilise :

- **Sa légitimité historique** (vision plus large de la société et expérience incontournable)
- **Son esprit pionnier** (recherche constante des tendances à venir notamment par l'observation des pratiques sur les réseaux ou des pays émergents)
- **Son internationalisation et sa volonté d'avoir une approche globale** (études multi pays)

INTRODUCTION

La conduite d'une étude à l'IFOP est classique pour un institut d'études marketing. Tel que j'ai pu le constater lors de mes différents stages, la mise en place de l'étude fait généralement suite à un cahier des charges transcrit par le client qui souhaite répondre à une problématique clairement définie. Les objectifs assignés à cette étude sont présentés dans ce cahier des charges. Le prestataire, à savoir l'institut d'études doit alors répondre via la proposition commerciale avec la méthodologie qui sera employée pour atteindre les objectifs souhaités (échantillon, quotas, méthode d'enquête, trame du questionnaire, déroulement du terrain, analyse des résultats, relations avec le commanditaire), les moyens humains et techniques mobilisés pour l'étude, ainsi qu'un planning détaillé de la démarche. Nous nous intéresserons ici qu'à la partie « études quantitatives » des sociétés d'études.

Lorsque le client valide le projet auprès de son prestataire, l'étude peut alors démarrer. Une réunion de lancement avec le client et généralement organisé afin de redéfinir clairement les objectifs. La méthode de collecte des données a préalablement été définie dans la proposition commerciale, il s'agit donc de construire un questionnaire adapté à ce mode de collecte, et qui permettra de répondre aux objectifs souhaités par l'entreprise.

L'objectif de ce mémoire s'inscrit dans la lignée de la phase de création d'un questionnaire en institut d'études marketing. En effet, le questionnaire est un outil essentiel dans la collecte des données. C'est grâce à cet outil que le recueil de l'information est possible. L'analyse des résultats en dépend fortement. Il est donc important de construire un questionnaire de qualité : la qualité des données est effectivement directement liée au questionnaire.

Lorsque l'institut de sondage lance une étude, plusieurs biais peuvent cependant intervenir. Le choix de la méthode de collecte peut être un premier biais. En optant pour un terrain online (CAWI¹), téléphonique (CATI²) ou face-à-face (CAPI³), des biais peuvent effectivement intervenir. Une collecte des données online ne nous permettra pas d'interroger les individus qui n'ont pas d'accès à Internet. Par téléphone, nous ne pourrions pas utiliser de visuels dans les enquêtes. Il est par ailleurs compliqué d'effectuer des questionnaires téléphoniques dont la durée excède les quinze minutes : la qualité des données en dépend fortement. La formulation des questions peut également faire intervenir de nombreux biais : les questions à poser ne doivent pas influencer les répondants dans leur formulation. Elles doivent également être posées dans le langage du répondant (et non celui de l'entreprise). Il faut donc choisir la méthode de collecte la plus adaptée à l'étude, mais surtout minimiser les biais qui peuvent intervenir dans la formation d'une enquête par questionnaire. Toutefois,

¹ CAWI : Computer-Assisted Web Interviewing

² CATI : Computer-Assisted Telephone Interviewing

³ CAPI : Computer-Assisted Personal Interviewing

nous ne présenterons pas les biais liés aux méthodes de collecte, ni les biais liés à la construction d'un questionnaire dans ce mémoire.

Dans ce dernier, nous nous concentrerons exclusivement sur les biais dans la mesure. Les instruments de mesure, et particulièrement les échelles sont des outils utilisés pour quantifier les données, pour les catégoriser. Elles permettent de donner une valeur à un attribut que l'on souhaite mesurer.

En marketing, les échelles de mesure sont généralement utilisées afin d'évaluer une attitude, une opinion, un comportement. Ces informations à recueillir sont donc nombreuses. Il peut s'agir de mesurer la satisfaction : « *Dans quelle mesure les clients de ma société sont satisfaits des produits proposés ?* » On peut également mesurer l'accord : « *Dans quelle mesure les français sont-ils d'accord avec les affirmations caractérisant mon produit ?* » Ou encore mesurer l'intention : « *Est-ce que les clients de ma société ont l'intention de refaire appel à mes services ?* » Etc.

Les échelles de mesure permettent donc de donner des éléments de réponse à une société en quête d'informations. Cependant, le choix d'une échelle de mesure est important puisque les résultats vont en dépendre.

Pour mener à bien ces mesures, de nombreuses échelles sont disponibles. Le choix de l'une ou de l'autre va dépendre de ce que l'on souhaite mesurer. Nous présenterons dans un premier temps les échelles de mesure qui existent et qui sont fréquemment utilisées en institut d'études, mais également comment en construire une, lorsqu'il n'existe pas d'échelle adaptée à ce que l'on souhaite mesurer.

Nous dégagerons ensuite les principaux biais liés au choix des échelles de mesure, mais aussi comment ces derniers peuvent être contournés.

Enfin, nous profiterons de ce stage en société d'étude pour tenter de répondre à une problématique liée au choix des échelles de mesure, avec le cas de la modalité centrale, également appelée modalité « neutre ».

PARTIE 1

Qu'est-ce qu'une échelle de mesure ?

Lorsqu'une entreprise, ou la société d'études à laquelle elle a fait appel, souhaite mener une étude de marché pour répondre à une problématique, plusieurs possibilités s'offrent à elle : elle peut mener une étude qualitative pour recueillir des éléments qualitatifs et mieux comprendre les clients. Ou alors, elle peut choisir d'effectuer une étude quantitative si elle souhaite obtenir des résultats chiffrés.

Lorsque l'entreprise choisit de mettre en place une étude quantitative, plusieurs types de question peuvent s'intégrer au questionnaire : les questions ouvertes, les questions fermées. Ces dernières peuvent être posées sous forme d'échelles. Nous nous intéresserons dans la suite de ce mémoire exclusivement aux questions posées sous formes d'échelle. Les échelles sont très utiles dans une analyse quantitative puisqu'elles permettent effectivement de quantifier la donnée. Cependant, ces échelles sont nombreuses. Selon Eric Vermette (1991), on recense dans les manuels spécialisés de recherche marketing « plus d'une vingtaine de sortes d'échelles utilisables dans un questionnaire. Le choix d'une échelle ou d'une autre va dépendre de ce que l'on souhaite mesurer. Pour Vermette, l'efficacité de l'échelle retenue dépend de trois principaux axes de qualité : la qualité méthodologique (Comment l'instrument de mesure a été développé ?), opérationnelle (Rapidité et simplicité de mise en œuvre ?) et la qualité décisionnelle (L'échelle développée permet-elle de trancher aisément ?).

Nous verrons dans cette première partie quelles sont les différents types d'échelle de mesure qui existent et, celles qui sont utilisées actuellement. Dans une seconde partie, nous verrons comment construire une échelle de mesure lorsqu'il n'existe pas d'outil disponible pour mesurer ce que l'on cherche.

1) Les échelles de mesure existantes

a) Inventaire des échelles de mesure existantes

Les échelles de mesures existantes sont les échelles qui sont les plus utilisées en société d'études. Ce sont généralement des échelles qui **ont fait leur preuve** dans le temps. Le choix d'une échelle de mesure plutôt qu'une autre dépend des critères que l'on souhaite mesurer. Zaichkowsky, (1984), dans une étude tentant de différencier deux concepts étroits (l'implication et le risque perçu) montre à quel point il est difficile de choisir l'échelle de mesure.

Les échelles de mesure à catégorie spécifique permettent de mesurer l'intention, le degré d'accord, etc. Ces échelles de mesure sont également appelées « **échelles d'attitudes** ». Cette notion est difficile à déterminer. Dans le MERCATOR, l'attitude est définie comme étant formée de « croyances, de sentiments et d'intentions d'agir ». C'est pourquoi, les mesures d'attitude reprennent les composantes de l'attitude : éléments cognitifs, affectifs et conatifs.

En marketing, il existe deux grandes catégories d'échelles de mesure :

- **Les échelles métriques** : (échelles continues) Ces échelles sont destinées à catégoriser des données quantitatives. On compte parmi elles :
 - **Les échelles d'intervalles** : (ou échelle relative) Ce sont des échelles numériques ou l'intervalle entre chaque unité a été arbitrairement défini. Toutefois, l'intervalle défini est constant entre deux unités. En ce sens, le zéro ne signifie pas une absence de phénomène mais est un point de repère. (C'est le cas du 0°C dans la mesure de la température.) Cette échelle est très peu utilisée en marketing et est davantage utilisée à des fins scientifiques.
 - **Les échelles de ratio** : (ou de rapport, absolue) Ces échelles présentent également un intervalle régulier entre chaque unité (comme les échelles d'intervalles). Cependant, le zéro n'est pas arbitrairement défini. Il décrit une absence de phénomène : le zéro correspond à une « existence naturelle » (Giannelloni et Vernet, 2001)
 - **Les échelles numériques** : Dans une étude de marché, ces échelles sont généralement utilisées pour attribuer une note. Elles ont l'avantage de pouvoir mettre en exergue un certain nombre de tests statistiques (Cf. « *Applications statistiques* » p14). Les échelles numériques sont connues pour leurs utilisations dans le système scolaire de notation.
- **Les échelles non-métriques** : Ces échelles sont destinées à catégoriser des données qualitatives. On compte parmi elles :
 - **Les échelles nominales** : Ces échelles permettent de classier les données qualitatives. Les chiffres ou nombres, lorsqu'ils sont présents sur l'étiquette « jouent uniquement le rôle d'étiquette permettant d'identifier des objets, des propriétés ou des événements » (Green, Tull et Albaum, 1988). Dans le cadre d'une enquête sur le revenu des français par exemple, cette échelle permettrait de classier notre échantillon selon leur sexe :

Veuillez indiquer votre sexe :

- Homme (1)
- Femme (2)

- **Les échelles ordinales** : Ces échelles permettent également de classifier les données qualitatives mais l'ordre des modalités proposées a un sens : la relation entre les données observées est transitive. Toujours dans le cadre d'une enquête sur le revenu des français, cette échelle pourrait être utilisée pour classifier les individus enquêtés selon leur niveau de diplôme :

Veillez indiquer votre niveau de diplôme :

- < BAC (1)
- BAC (2)
- BAC+1 / BAC+2 (3)
- BAC+3 / BAC+4 (4)
- > BAC+5 (5)

L'échelle de **Likert** fait partie de la catégorie des échelles ordinales. Cette échelle est très populaire et est souvent utilisée dans les études marketing. Elle permet en effet d'avoir des nuances dans les réponses des interrogés. Contrairement aux échelles dichotomiques (oui/non), elle permet en effet d'avoir différents degrés de jugement dans l'analyse des résultats. Selon Likert, les échelles de mesure les plus fiables sont celles qui sont en sept modalités. Pourtant, les échelles sont majoritairement utilisées en cinq modalités dans les enquêtes menées par les sociétés d'études. Nous évoquerons ces différences dans les principaux biais liés au choix des échelles de mesure (Cf. *Partie 2 : Principaux biais de mesure*, p19).

En complément de ces deux grandes catégories d'échelle, on retrouve également une différence entre les échelles de type **numérique** et les échelles de type **sémantique**. Nous verrons dans la partie sur les applications statistiques ci-après que le choix de l'une ou de l'autre peut avoir son importance dans la réalisation d'analyses statistiques.

L'échelle **numérique**, telle que son nom l'indique, est davantage réservée à des attributions de notes (généralement entre 1 et 10). Ces échelles numériques sont également appelées échelles mathématiques. La note donnée permet alors au répondant de nuancer son jugement : le 1 étant défini comme la plus mauvaise note et le 10 comme la meilleure note que l'on peut attribuer.

Les échelles **sémantiques** sont quant à elles construites sur la base de mots. Ces échelles permettent donc de mesurer autrement l'information. Ces dernières ont de plus l'avantage d'être comprises facilement.

Contrairement aux échelles numériques qui permettent de nuancer un jugement, dans le cas des échelles sémantiques, chaque proposition permet de mettre en avant un aspect différent du jugement grâce à des mots/phrases courtes. Si la construction de ces échelles est compliquée et fait intervenir de nombreux biais (Cf. *Partie 2 : Principaux biais de mesure*, p19), les échelles sémantiques sont très utilisées en marketing. En effet, nous verrons par la suite que ces échelles permettent de construire des graphiques intéressants grâce à la réalisation d'analyses statistiques simples. On peut alors comparer des **profils** : d'où son utilisation fréquente en marketing. Parmi les échelles sémantiques les plus célèbres, et les plus utilisées dans les études marketing, on retrouve l'échelle sémantique

différentielle d'Osgood. Cette échelle est basée sur l'opposition entre deux items de part et d'autre de l'échelle. Le répondant doit alors se placer dans les cases intermédiaires situées entre les deux antonymes. Cette échelle est présentée dans la figure 1 ci-dessous.

Comment jugeriez-vous l'envoi du colis sur les critères suivants ?

Lent	___ / ___ / ___ / ___ / ___ / ___ / ___	Rapide
Cher	___ / ___ / ___ / ___ / ___ / ___ / ___	Peu cher

Figure 1 : L'échelle sémantique différentielle d'Osgood

L'échelle de **Stapel** présentée dans la figure 2 est construite en dix niveaux. Cette échelle peut être décrite comme « une échelle verbale en 10 points unipolaires allant de -5 à +5 qui mesure la direction et l'intensité simultanément » selon Hawkins, Albaum, et Best (1974). Cette échelle est construite verticalement : les notes positives en haut et les notes négatives en bas. La sélection du « +5 » correspond à un accord total avec le terme employé dans l'échelle, alors que le « -5 » correspond à un désaccord total.

Comment jugeriez-vous l'envoi du colis sur les critères suivants ?

+5		+5
+4		+4
+3		+3
+2		+2
+1		+1
Rapidité		Prix
-1		-1
-2		-2
-3		-3
-4		-4
-5		-5

Figure 2 : L'échelle de Stapel

Chaque échelle de mesure possède ainsi ses propres caractéristiques morphologiques. Nous verrons dans la partie qui suit que le choix de l'une ou de l'autre est étroitement lié au traitement statistique que l'on souhaite en tirer.

b) Applications statistiques

Le choix d'une échelle ou d'une autre dépend de la nature même de la variable que l'on souhaite étudier. La distinction entre les différents types d'échelle est primordiale puisque l'analyse statistique en dépend. En effet, le traitement statistique apporté aux différentes échelles dépend des caractéristiques morphologiques de chaque échelle de mesure.

Prenons le cas des échelles dichotomiques (Cf. Annexe1). Ces échelles permettent de répondre aux questions soit par « oui », soit par « non ». Cette échelle permet donc de scinder en

deux notre échantillon et par la suite effectuer des tests de **comparaisons de proportions**. A partir de ces échelles, nous pouvons donc effectuer des groupes plus ou moins homogènes qui vont nous permettre d'affiner notre lecture des résultats de l'étude.

Selon le type d'analyse que la société d'études souhaite resituer à son client, l'échelle va pouvoir varier. Par exemple, dans le cas d'une étude de satisfaction, les échelles de notes permettent de construire des matrices d'importance-satisfaction. Cette matrice, initialement présentée par Martilla et James (1977) est très riche pour le client, et permet d'axer les points d'amélioration prioritaires à apporter. Pour ce type de matrice, on demande généralement aux répondants d'attribuer une note de 1 à 10 sur une batterie de critères correspondant à ce que l'on souhaite évaluer dans la satisfaction, puis on leur demande au début ou à la fin du questionnaire d'attribuer une note correspondant à la satisfaction globale du client vis-à-vis du mesuré (De nombreuses interrogations surgissent quant au placement de la question globale *à priori* ou *à postériori*. Ganassali, (2008) étudie « l'impact de la disposition de l'évaluation globale ». Nous n'interviendrons pas dans ce débat.). Le fait d'attribuer des notes sur 10 via les échelles de mesure est fondamental dans ce type de d'analyse, puisque ce sont ces notes qui vont nous permettre de **calculer les corrélations** des différents critères à la satisfaction globale. Cela nous permet ainsi de mesurer la liaison des différents critères à la satisfaction générale, et donc de construire la matrice. Lorsque ces échelles sont ordinales, on pourra d'ailleurs utiliser le « V » de Cramer qui est plus adapté à ce type d'échelle. Les échelles numériques ont donc l'avantage de permettre la réalisation d'analyses statistiques plus complexes telles que des ACP ou mapping.

Pour les analyses telles que présentées ci-dessus, les échelles numériques sont donc nécessaires. Cependant, pour les échelles de type Likert, lorsqu'elles sont considérées comme étant **métriques**, c'est-à-dire lorsque l'intervalle entre chaque modalité est considéré comme constant, on peut utiliser les techniques disponibles pour les variables quantitatives. En effet, en attribuant un score à chaque modalité, on peut **transformer l'échelle de Likert en échelle numérique**, et par conséquent, effectuer les analyses statistiques qui en découlent. Par ailleurs, en appliquant cette méthode qui consiste à pondérer chaque modalité, cela va également nous permettre d'obtenir la moyenne de classement. La moyenne obtenue est généralement donnée entre deux modalités, ce qui permet de situer la tendance centrale des répondants sur l'échelle. Toutefois, il est important de faire attention au **codage** de ces échelles. Celui-ci a son importance dans l'analyse des résultats. Si l'affirmation utilisée dans l'échelle est en opposition avec ce que l'on souhaite mesurer, alors il ne faut pas oublier d'inverser l'échelle de mesure au moment du traitement. Le codage des items « négatifs » et « positifs » doit alors être inversé au moment du traitement. (Giannelloni et Vernet, 2001). Ainsi, le « 1 » devient « 7 », et le « 7 » devient « 1 ». L'exemple suivant en figure 3 permet d'illustrer ces propos. Dans celui-ci, nous demandons aux répondants de juger la qualité d'une serviette. L'affirmation C, (Serviette fine) ne va pas dans le même sens que les autres affirmations, à savoir, une serviette de bonne qualité. Lors du traitement des données, le codage de cette affirmation doit alors être inversé.

Concernant les caractéristiques suivantes sur la serviette, êtes-vous d'accord avec les affirmations proposées ?

	Pas du tout d'accord	Ni d'accord, ni pas d'accord					Tout-à-fait d'accord	
A : La couleur est agréable	1	2	3	4	5	6	7	
B : La serviette est douce	1	2	3	4	5	6	7	
C : La serviette est fine	1	2	3	4	5	6	7	



Après traitement :

A :	1	2	3	4	5	6	7
B :	1	2	3	4	5	6	7
C :	7	6	5	4	3	2	1

Figure 3 : Traitement statistique des échelles de Likert

Les échelles sémantiques différentielles peuvent aussi utiliser cette méthode de **pondération des modalités** pour effectuer certaines analyses statistiques. Cependant, il est préférable d'utiliser les échelles de notes si l'intention de l'étude est d'effectuer des analyses statistiques fines. Ces échelles permettent en effet d'avoir un modèle bien plus robuste dans le traitement de nos données : avec une échelle numérique, la note attribuée par le répondant correspondra effectivement à la note utilisée dans les analyses statistiques. Or, avec une échelle ordinale où l'on attribue une note par modalité (sous l'hypothèse que celle-ci est métrique), il est probable que la modalité choisie par le répondant ne corresponde pas à la note qui lui est affecté après le traitement statistique.

Par ailleurs, les différentiels sémantiques sont très utilisés en marketing. Outre leur simplicité de compréhension de la part des répondants, ces échelles permettent d'obtenir des graphiques très intéressants en marketing, « **sans traitements statistiques sophistiqués** » (Giannelloni et Vernet, 2001). En effet, grâce à de simples calculs de moyenne, nous pouvons obtenir des profils sémantiques qui permettent de comparer facilement des marques ou produits, sur une batterie de

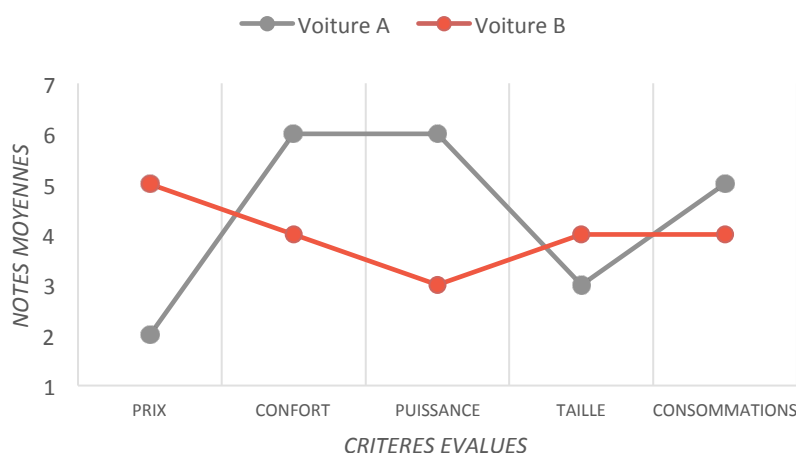


Figure 4 : Exemple de profils comparés critères mesurés dans l'étude. C'est le cas de la figure 4 présentée ci-dessous.

2) Les échelles de mesure inexistantes

a) Construire une échelle de mesure

Malgré le nombre important d'échelles de mesure disponibles pour effectuer des études via questionnaire, **il n'existe pas toujours d'échelle adaptée pour mesurer ce que l'on recherche**. Dans ce cas, il est possible de créer sa propre échelle de mesure. De nombreux articles de recherche montrent comment mettre en place un instrument de mesure. Christian Derbaix (1997) a mis au point une échelle pour mesurer « *l'implication de l'enfant* », soulignant que les échelles de mesures utilisées de manière générale servent à analyser le comportement des adultes. L'échelle pour mesurer le « *butinage* » par Cindy Lombart (2004) est un deuxième exemple qui montre que la construction d'une échelle de mesure a déjà fait l'objet de plusieurs recherches. La création de ces échelles permet de mesurer une attitude, un comportement, un sentiment dont on ne dispose pas des outils nécessaires pour les mesurer.

b) Le modèle de Churchill

Pour mettre en place un nouvel instrument de mesure, il existe plusieurs paradigmes. Le plus connu et le plus célèbre d'entre eux est le **paradigme de Churchill** (1979), complété en 1988 par Gerbing, et en 1993 par Roehrich. Ce paradigme est majoritairement utilisé par l'ensemble des chercheurs qui souhaitent développer leurs propres échelles de mesure plutôt que d'utiliser les instruments qui existent déjà. De ce fait, bien que l'application du paradigme de Churchill suive une procédure stricte, la méthodologie reste commune à tous les chercheurs qui développent leurs échelles. Le paradigme de Churchill est constitué de huit étapes résumées dans la figure 5 ci-après.

La première étape de cette démarche consiste à définir la « *nature du construit* », c'est-à-dire définir implicitement **ce que l'on cherche à mesurer**. Cette première étape est importante puisqu'elle permet de « *délimiter ce qui est et ce qui n'est pas à inclure dans la mesure* » (Lombart, 2004). Pour effectuer cette première étape, une étude qualitative peut être lancée. Elle permettra par ailleurs de définir la liste des items. L'étude qualitative, via ses entretiens, permet ainsi d'obtenir des informations supplémentaires que la revue de littérature ne peut fournir sur la nature du construit.

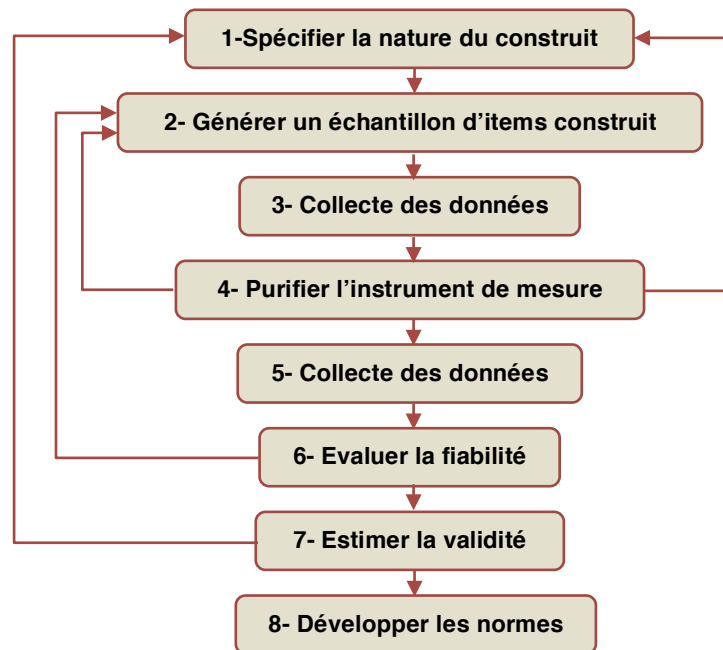


Figure 5 : La démarche de Churchill (1979)

La deuxième étape consiste à **créer un ensemble d'items** qui permettent de mesurer le construit. Elle peut s'effectuer à partir des entretiens menés lors de la phase qualitative en étape 1, mais également grâce à la revue de la littérature. Cette étape du paradigme de Churchill apparaît être la plus difficile pour le chercheur. La liste des items proposée doit être associée aux dimensions définies dans la nature du construit. Lorsque cette liste d'items a été validée par des experts, une première collecte des données peut avoir lieu. La purification des données suite à cette collecte, va alors nous indiquer si les items choisis permettent de mesurer précisément ce que l'on cherche, grâce à l'indicateur de fiabilité que nous développerons dans la partie suivante. Cette étape du paradigme de Churchill va alors permettre de résumer les données, les réduire et par la suite modifier voir supprimer des items. En fonction de ces résultats, des retours aux étapes 2 et 3 sont à prévoir. Si les items sont conservés, une collecte finale des données peut alors être entreprise.

L'estimation de la **fiabilité**, puis de la **validité** de l'échelle de mesure doit alors être effectuée. Dans le cas où ces deux conditions ne sont pas satisfaites, il faut revoir la liste des items. Ces deux notions fondamentales pour une échelle de mesure vont être présentées dans la partie suivante : elles permettent en effet d'affirmer qu'une échelle de mesure est pertinente ou non, pour ensuite être utilisée.

c) Pertinence d'une échelle de mesure

Pour les échelles de mesure existantes, comme pour les échelles de mesure inexistantes : deux conditions essentielles doivent être remplies pour pouvoir les utiliser. Si nous n'avons pas présenté ces deux conditions dans la partie sur les échelles de mesure existantes, c'est que l'ensemble de des échelles de mesure qui y sont présentées remplissent déjà ces conditions : c'est d'ailleurs ce qui

explique pourquoi ces échelles de mesure sont si **connues** et si **utilisées**. Pour Vernet (1991), « *la fiabilité de la plupart des instruments de mesure utilisés aujourd'hui dans les études marketing quantitatives a été prouvée* ». Toutefois, il est important de revenir sur ce qui définit la fiabilité et la validité d'une échelle de mesure :

Fiabilité : Selon Jolibert et Jourdan (2003), « *une échelle est dite fiable lorsqu'elle donne les mêmes résultats lors de mesures répétées quelles que soient les personnes qui l'utilisent et quel que soit le moment ou le test est effectué* ». Pour vérifier si l'échelle de mesure donne des résultats identiques lors de mesures répétées, deux méthodes principales s'offrent à nous ; le **test-retest** (un même questionnaire est administré deux fois à un même échantillon, sur deux périodes distinctes) et le **split-half** (cette méthode étudie la corrélation entre deux groupes d'un même échantillon, sur deux ensembles d'items). Nuannaly (1967) indique que la fiabilité d'un instrument de mesure est déterminée pour un *alpha de Cronbach* compris entre 0,60 et 1,00. Toutefois, cet *alpha de Cronbach* nécessite d'être plus élevé selon l'objectif de la recherche (0,80 pour une étude appliquée).

Validité : Avant de valider une échelle de mesure, cette dernière doit impérativement vérifier la première condition de fiabilité (Churchill, 1979) : une échelle valide est forcément fiable. A noter que la réciproque n'est pas vraie. Il existe différents niveaux de validation d'une échelle. Pour Messick (1989), la validité d'une échelle de mesure est définie de la façon suivante : « *Un instrument est valide lorsqu'il mesure réellement ce qu'il est censé mesurer* ». Il existe différents niveaux de validation selon Penta, Arnould et Decruynaere (2005). La validation de contenu consiste à savoir si l'échelle mesure bien ce qui est étudié. La validité convergente (ou discriminante) permet d'évaluer si l'échelle mesure uniquement le construit en question. La validité monologique permet de vérifier si le construit mesuré est corrélé avec d'autres mesures relativement proches.

D'autres conditions s'ajoutent également à la fiabilité et à la validité d'une échelle de mesure. Il est en effet pertinent de prendre en compte certains critères comme la facilité de réponse de la part de l'interviewé, à savoir si celui-ci est capable de répondre clairement et spontanément aux échelles de mesure qui lui sont proposées.

PARTIE 2

Principaux biais de mesure

Nous avons vu en première partie qu'une échelle de mesure ne peut être utilisée que si elle respecte des conditions strictes : la fiabilité et la validité des échelles sont effectivement deux qualités nécessaires pour être utilisées.

Toutefois, même lorsque ces conditions sont validées, les biais dans la mesure peuvent surgir. Si ces biais sont parfois inévitables, le rôle de l'utilisateur des échelles de mesure (cas du chargé d'études en institut d'études marketing) est de les **minimiser**. En effet, lorsque le chargé d'études choisit d'utiliser une échelle de mesure, son choix est souvent justifié par des contraintes extérieures : le coût de la mise en œuvre permet-il d'utiliser l'instrument de mesure le plus adapté ? La durée de l'enquête est-elle suffisante pour utiliser les échelles de mesure les plus pertinentes ? Les contraintes liées au mode de recueil sélectionné permettent-elles l'utilisation de l'instrument ? Les problèmes dans la mesure interviennent alors soit dans l'instrument de mesure lui-même, soit dans son utilisation. Nous verrons dans cette partie quels sont les biais de mesure les plus courants, et comment ils sont parfois traités.

1) Limites & biais de mesure des échelles

a) Barrières de la langue

Lorsqu'un chercheur met au point une échelle de mesure, la langue est un frein à l'utilisation de cette dernière dans un pays de langue étrangère à celle où l'échelle a été développée. Les travaux de Jean-Louis Chandon et Boris Bartikowski (2004) montrent à quel point la **traduction d'une échelle** est un exercice difficile pour le chercheur : la traduction d'un mot peut effectivement avoir une signification différente d'un pays à autre. Ainsi, un adjectif comme « adéquat » est généralement significatif de manque de performance chez les français, alors que pour les anglais, cet adjectif aura pour valeur l'indifférence. Une simple traduction n'a donc aucun intérêt. Le questionnaire doit être adapté. Cependant, la traduction n'est pas toujours possible ce qui compromet l'utilisation de l'échelle de mesure. Nous retiendrons qu'une échelle validée dans un pays ne peut pas être utilisée dans un autre pays sans être **testée** et **validée** (Chandon, Bartikowski, et Gierl, 2006).

Par ailleurs, tous les pays n'ont pas la même réaction aux regards des différentes échelles qui leurs sont proposées. Certains pays auront plus de mal à utiliser les **différentiels sémantiques** que d'autres, ce qui peut avoir des conséquences négatives sur les statistiques de traitement de ces échelles de mesure. La traduction d'une échelle de mesure est donc un obstacle majeur pour les chercheurs qui essaient d'adapter une échelle de mesure créée par des confrères étrangers.

La méthode de Thurstone sur laquelle se sont basés ces travaux de Chandon et Bartikowski révèle par ailleurs que la valeur accordée aux adjectifs employés dans une échelle de mesure peut varier entre deux individus qui parlent la même langue, en fonction de la **classe sociale** de ces individus. Une échelle à caractère sémantique peut alors introduire des biais importants si celle-ci est destinée à des individus de classes sociales différentes.

D'autres chercheurs ont également mis en avant ces problèmes liés à la langue dans l'utilisation des échelles de mesure. Ruth Menahem, dans « *Le différenciateur sémantique* » (Menahem, 1968) met en avant un autre type de biais lié aux langues, celui de la « **bipolarité** ». Les échelles sémantiques différentielles (cas des échelles d'**Osgood**) nécessitent en effet que les deux termes sélectionnés soient antagonistes. La distance entre les deux antonymes par rapport à la modalité centrale doit alors être exactement la même. S'il est difficile de choisir deux **antonymes**, Menahem montre tout de même qu'il **n'existe pas de règle** précise pour les définir, ce qui constitue donc une limite à la création de ce type d'échelle. Il démontre par ailleurs que le choix des qualificatifs dépend fortement du concept que l'on souhaite décrire : ainsi, les qualificatifs d'une échelle ne peuvent pas être utilisés pour n'importe quel type de mesure. La langue devient donc une barrière à l'utilisation des échelles de mesure.

b) Nombre de modalités

Dans cette partie, nous nous intéresserons à la difficulté liée au **nombre de modalités** à proposer dans une échelle de mesure. De nombreuses interrogations font effectivement leurs apparitions quant au **nombre d'items à proposer**. Pour Likert, les échelles de mesure les plus fiables sont celles qui comportent **sept modalités**. Cependant, de nombreuses adaptations de cette échelle ne présentent que cinq, voire quatre modalités. Quel est donc le nombre de modalités à adopter pour que l'instrument sélectionné mesure idéalement le construit ? L'échelle de Likert en quatre est souvent retenue dans les études marketing. Pourtant, certaines échelles telles que les échelles à deux niveaux et les échelles psychologiques présentent une bonne validité. Les recherches à ce sujet sont peu nombreuses, c'est pourquoi cette interrogation fera l'objet de notre étude présentée dans la troisième partie.

Il va de soi que plus une échelle de mesure est composée de modalités, plus elle est **précise**. Dans les mesures de l'attitude, on utilise généralement 7 ± 2 modalités (Peter, 1979). D'ailleurs, le marketing a conscience que les échelles avec de fortes intensités sont celles qui donnent des résultats plus précis. Pour des raisons de praticité et de mesure des résultats, le marketing privilégie généralement les échelles qui ont un nombre de modalités peu élevé. En effet, la taille d'un échantillon est parfois limitée dans les études marketing. Pour permettre d'analyser les résultats, et effectuer des tests de significativité, il faut s'assurer d'avoir des **bases solides**. Lorsque les échelles

de mesure comportent un nombre de modalités important, la taille de l'échantillon doit pareillement être conséquente, au risque de devenir un biais pour l'échelle.

Si pour Likert, une échelle de mesure idéale doit comporter **sept options**, d'autres études vont également dans ce sens. En effet, la recherche a pu démontrer qu'au-delà de sept modalités, les personnes interrogées ont du mal à effectuer un choix parmi la liste proposée. De ce fait, les réponses données aux questions où l'échelle présente un nombre d'échelons important risque d'inciter les répondants à fournir des réponses au hasard.

Quel que soit le nombre de modalités retenues dans l'échelle de mesure, une question subsiste. Faut-il opter pour une échelle à nombre de modalités **pair** ou **impair** ? Cette question consiste finalement à déterminer si l'on choisit de faire intervenir la modalité neutre dans les échelles de mesure ou non. En effet, les résultats obtenus lorsque la modalité neutre est proposée sont-ils identiques lorsqu'elle ne l'est pas ? Nous tenterons de vérifier l'hypothèse que les résultats sont différents en la présence / absence de cette modalité dans l'étude expérimentale présentée en partie 3.

A cette question qui consiste à opter pour une échelle de mesure à nombre de modalités pair ou impair, deux possibilités s'offrent donc aux chargés d'études qui souhaitent mener leurs études :

- **Echelles de mesure à modalités impaires** : Dans ce cas de figure, cela signifie généralement que l'on fait intervenir une modalité centrale dans l'échelle de mesure. Sur une échelle de Likert mesurant le degré de satisfaction d'un bien ou d'un service, la modalité centrale se traduirait par « Ni satisfait, ni insatisfait ». Lorsque l'on fait intervenir cette modalité centrale, c'est que l'indifférence des répondants vis-à-vis du sujet mesuré est un résultat intéressant qui donne une information importante pour celui qui la mesure. Le taux de répondants inclut dans la modalité centrale doit alors être présenté comme un résultat à part entière.
- **Echelle de mesure à modalités paires** : lorsqu'on emploie ce type d'échelle, les individus qui y sont soumis n'ont pas la possibilité de donner une réponse qui exprime une indifférence face à l'affirmation proposée. Cette méthode oblige donc les répondants à ce placer d'un côté ou d'un autre de l'échelle. Il s'agit donc d'une échelle à choix forcé puisque le répondant est contraint de se positionner. On appelle cette méthode mesure ipsative.

Le choix d'opter pour une échelle à nombre de modalités paire ou impaire dépend de plusieurs facteurs. Néanmoins, ce choix est important puisque l'échelle utilisée peut avoir des conséquences importantes sur l'interprétation des résultats qui s'en suit. Lorsqu'on opte pour une échelle **impaire**, la modalité centrale s'inscrit parfois comme une **modalité « refuge »**. Utiliser cette échelle peut donc avoir des effets néfastes sur la qualité de l'étude et par conséquent, sur l'interprétation des résultats. En effet, lorsqu'une modalité centrale est proposée, un individu qui n'est

pas impliqué dans son questionnaire va se placer plus facilement dans cette modalité « refuge » pour passer plus rapidement la question. C'est le cas lorsque les questionnaires sont trop longs. Or, lorsqu'un nombre important de répondants se positionnent sur cette modalité, les résultats de la question ne peuvent pas toujours être exploités puisqu'il n'y a pas assez de répondants de part et d'autre de l'échelle. Les sociétés d'études évitent donc d'utiliser ces échelles puisqu'elles ne permettent pas de fournir des éléments de réponse. D'ailleurs, Presser et Schuman (1980) ont montré que **ce point médian** est souvent choisi s'il est « explicitement proposé ». Cet effet est d'avantage marqué lorsque la modalité neutre est placée au point médian de l'échelle que lorsqu'elle est placée à la fin. (Brignier, 1991).

D'autres biais liés au nombre de modalités interviennent également, notamment lorsqu'il s'agit d'échelles **paire**s. Bien que les répondants soumis à ce type d'échelle sont plus impliqués puisqu'ils doivent obligatoirement se placer d'un côté ou de l'autre de l'échelle, des biais subsistent. En effet, en l'absence de modalité centrale, les individus interrogés sur ces échelles vont souvent se placer sur les **modalités à niveau d'implication plus faibles**. Dans le cas de notre échelle de Likert adapté à la satisfaction, si le bien ou service évalué n'a que peu d'incidence pour les personnes interrogées, la proportion d'individus « plutôt satisfait » sera plus conséquente. Malgré que le bien ou service évalué n'évoque pas d'insatisfaction pour la personne interrogée, ce bien ou service ne génère pas spécialement une « bonne satisfaction ». Cette remarque fait intervenir le concept de la **zone d'indifférence** mis en avant par de nombreux auteurs et repris par Bery et Parsuraman (1991) : « *un consommateur est satisfait s'il perçoit une performance supérieure à la borne supérieure de sa zone d'indifférence* ». Ce concept de la zone d'indifférence, est subtil et est parfois difficile à définir. Ainsi, lorsque Jean-Louis Chandon tente de créer une échelle de mesure pour classer les répondants en « satisfaits », « indifférents » et « insatisfaits », il note que l'énoncé « normal » est compris comme une évaluation neutre de performance. L'étude mise en place dans la partie 3 va nous permettre d'étudier l'impact lié à la présence / absence de la modalité centrale sur le reste des échelons.

c) Les non réponses : NSP

Faut-il laisser le choix aux répondants de ne pas répondre à la question ? Une échelle de mesure doit-elle intégrer la proposition « Vous n'avez pas d'opinion, vous ne savez pas » ? Tel que le nombre de modalités à intégrer dans une échelle de mesure, ces interrogations méritent réflexion. En effet, permettre aux répondants de « passer » la question sous prétexte que ce dernier ne sait pas ou n'as pas d'opinion fait intervenir des biais, dont certains sont similaires à ceux liés à l'intégration de la modalité neutre évoquée ci-dessus.

Lorsqu'on propose aux interrogés de **ne pas répondre**, on prend effectivement le risque de se retrouver avec un taux élevé de non-réponse (modalité refuge). Non ne reviendrons pas sur ce type

de biais qui réduit l'analyse des questions lorsque certaines modalités ont des bases trop faibles (Cf. *Nombre de modalités*, p20).

Cependant, lorsqu'on fait le choix justifié d'intégrer la proposition « NSP » dans la mesure, le taux de NSP devient un **résultat à part entière** pour l'étude. Or, pour les mêmes raisons que la modalité médiane, les sociétés d'études ont tendance à privilégier les **mesures ipsatives** pour avoir des réponses concrètes à fournir aux clients. Payne avait dès 1930 testé le rôle du NSP. Le fait de demander aux individus s'ils n'avaient « pas d'opinion » permettait d'enregistrer plus de huit points de réponse par rapport à l'absence du NSP. Ainsi, quand on autorise les enquêtés à dire « qu'ils n'ont pas d'avis » cela augmente fortement le nombre de NSP. Un sondeur/institut va avoir tendance à minimiser les NSP car il souhaite avant tout que les individus interrogés se positionnent ; là où un sociologue nous dira qu'il préfère garder le NSP pour avoir uniquement des « vraies réponses ».

Permettre aux interrogés de répondre NSP aux différentes questions qui leur sont posées fait intervenir le sujet de la **qualité des données** : lorsque les répondants n'ont absolument aucune idée, et ne savent pas comment se prononcer à la question qui leur est soumise, que vont-ils répondre ? On supposera que lorsqu'ils répondent, la réponse fournie se situe soit au niveau de la **modalité neutre** (lorsque cette dernière est proposée), soit au niveau de la modalité **la moins impliquante** (lorsqu'il n'y a pas de modalité centrale). Nous vérifierions cette hypothèse dans l'étude mise en place en partie 3.

Cette interrogation sur la qualité des données surgit de façon plus importante dans le cas des questionnaires en ligne : en effet, les questionnaires sont programmés pour que chaque individu interviewé soit en mesure de fournir une réponse. Si l'individu ne sait pas quoi cocher à une question et que le champ « Vous n'avez pas d'opinion, vous ne savez pas » ne lui est pas proposé, il doit tout de même fournir une réponse pour pouvoir continuer le questionnaire. Dans ce cas, deux possibilités s'offrent à lui. Soit il répond au hasard à la question parmi la liste des propositions qui lui sont fournies, au risque de fausser les résultats et de biaiser les interprétations qui s'en suivent. Soit il choisit de quitter l'enquête, ce qui affaiblit le taux de réponse de l'étude. Dans le cas des questionnaires en ligne, et plus particulièrement lorsqu'il s'agit de panels, la gratification offerte pour avoir terminé un questionnaire suggère que les enquêtés fournissent des réponses sélectionnées au hasard : la qualité des données est alors fortement remise en question.

d) Les effets de Halo

La qualité des données est une nouvelle fois remise en question via les nombreux effets de Halo que les chercheurs en psychologie sociale ont pu mettre en évidence. Ces effets de Halo, bien qu'indépendants de l'échelle de mesure employée, sont toutefois étroitement liés à ces types de mesure. En effet, si le type d'échelle utilisé dans le questionnaire n'est pas directement remis en cause, il s'agit ici de la méthode de mesure via les échelles qui est elle-même remise en cause.

Le **halo affectif**, tel que le définit Aurifeille est « l'influence de l'affectif sur les perceptions » (Aurifeille J.M, 1991). Le halo affectif, dans une étude sur la satisfaction consisterait à surévaluer les différents critères proposés dans l'étude, dû à la perception positive du consommateur vis-à-vis du bien ou du service évalué. Cet effet de halo est cependant difficile à mesurer. Un consommateur addictif à un produit ou une marque aurait donc tendance à surévaluer tous les critères proposés pour défendre la marque, ce qui peut biaiser l'interprétation des résultats. Ayant conscience de l'existence du halo affectif, on peut se demander si l'échelle de mesure est l'instrument le plus adapté.

Pourtant, même si l'effet de halo affectif est un élément essentiel en marketing, un autre effet de halo, plus proche du sujet étudié dans ce mémoire mérite d'être davantage développé : le **halo cognitif**. Le halo cognitif est à l'origine d'un certain nombre de biais dont il est important de connaître l'existence. Ce type de halo intervient généralement à cause d'un questionnaire mal construit, favorable à l'apparition de cet effet. Selon Stéphane Ganassali (2008), l'effet de Halo peut s'exprimer sous deux formes :

- « **Effet de cohérence** » : Lorsque les répondants sont soumis à une batterie d'items à évaluer, les derniers critères à évaluer dépendent des réponses données aux premiers critères. Cela se traduit par le fait que les individus évitent de donner des réponses en opposition avec ce qui a pu être donné précédemment. Ainsi, les réponses fournies ne sont pas incohérentes l'une par rapport à l'autre. Par exemple, et dans le cadre d'une enquête de satisfaction, si l'individu attribue la note de 5 sur 10 à un critère qu'il juge moyen, et qu'il doit évaluer un second critère qui lui paraît moins performant que le premier, alors, et par soucis de cohérence, la note qu'il va attribuer sera normalement inférieure au 5 attribué précédemment. D'où le nom « d'effet de cohérence » donné par Ganassali (2008). Si on avait demandé au répondant d'évaluer le second critère en premier, il est légitime de penser que la note attribuée aurait été supérieure. En cours d'enquête, les réponses données ont donc une influence sur les réponses qui vont être données. Cependant, on note ici que la hiérarchie dans la notation des critères est respectée, par soucis de cohérence.

Ainsi, on peut dire que globalement les critères obtiennent des notes plus élevées si le premier critère évalué est important, et inversement. (Brignier, 1991). Toutefois, l'ordre des critères n'a pas d'importance, et les critères évalués n'influencent pas la notation des critères suivants lorsqu'il s'agit d'échelles qualitatives, composées d'un faible nombre de libellés verbaux (Schwarz et Wyer, 1985).

- « **Effet de fatigue** » : Cet effet est de plus en plus marqué lorsque la liste des critères à évaluer est assez longue. « L'effet de fatigue » tel que le définit Ganassali peut intervenir quelque-soit le mode d'administration du questionnaire (Téléphone, Internet, face-à-face, papier). En effet, en cours de questionnaire, l'attention du répondant peut s'affaiblir pour les importantes listes de critères à évaluer : au fur et à mesure de l'évaluation, le répondant subit une « baisse d'attention ». La réflexion apportée par le répondant est alors plus faible. En fin

de questionnaire, et lorsque les questions enchaînées se ressemblent, l'implication des répondants est de moins en moins forte : c'est l'effet de « halo cognitif » présenté par Ganassali. La lassitude des répondants se traduit par une baisse de l'attention qui entraîne « une diminution de la discrimination dans les réponses ». Ainsi, lorsque la liste des critères à évaluer est importante, les réponses attribuées aux échelles placées l'une sous l'autre auront tendance à être similaires (Par téléphone, cet effet de fatigue peut se manifester par la redondance des notes attribuées par le répondant sur un ton identique d'une question à une autre. En CAWI, sur une liste d'items à évaluer ou les échelles sont présentées l'une au-dessous de l'autre, l'interviewé aura tendance à « cliquer » en descendant la souris toujours sur la même note).

2) Minimiser les biais de mesure

a) Contourner les biais de mesure

Pour limiter les biais de mesure, il est parfois possible de les contourner via les différents modes de présentation de ces échelles. De nombreuses techniques adaptent effectivement les échelles de mesure telles que présentées ci-dessus, de façon à réduire au maximum les biais. Nous verrons dans cette partie quelles sont ces techniques.

« Vous ne savez pas ? »

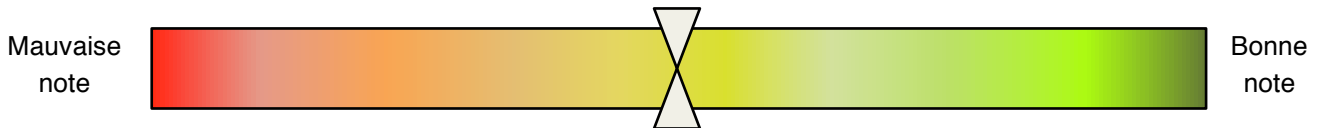
Précédemment, nous avons vu quel était l'enjeu d'intégrer la modalité « Vous n'avez pas d'opinion, vous ne savez pas » dans une étude marketing. Si l'intégration de cette option paraît inévitable dans le cas d'un questionnaire en ligne, ce biais peut être contourné dans le cas d'un questionnaire administré par téléphone ou en face-à-face. En effet, lorsque l'échelle de mesure sera prononcée par l'enquêteur, les chargés d'études peuvent demander aux enquêteurs de ne pas citer la modalité NSP. Ainsi, le taux de NSP ne correspondra qu'aux individus qui ne savent véritablement pas se prononcer. Cette option permet en effet de dissuader les répondants qui souhaiteraient passer plus facilement la question, et les impliquent ainsi d'avantage dans cette dernière.

Les échelles continues à curseur

Avec l'émergence d'Internet ces dernières années, les enquêtes en ligne sont de plus en plus nombreuses. Si ce mode de recueil ne s'adapte pas à toutes les cibles, il a l'avantage d'être beaucoup moins cher, et bien plus rapide que les autres modes de recueil jusque-là utilisés (Giannelloni et Vernet, 2001). Mais cet avantage n'est pas le seul, les enquêtes *on line* ont permis de développer de nouvelles méthodes qui limitent les biais de mesure. En effet, sur Internet, l'échelle de mesure peut tout simplement être remplacée par un système de curseur (Cf. Figure 6). Cette technique permet

d'obtenir des scores **bien plus précis** qu'une échelle de note, ou une échelle de Likert ce qui permet d'affiner les analyses statistiques. En effet, lorsque le répondant va déplacer le curseur sur la zone, la note qui va être récupérée par cette technique permet d'obtenir plusieurs chiffres après la virgule. Cela permet ainsi de donner plus de liberté à un interviewé qui souhaite nuancer son jugement ; dans une échelle de notes de 1 à 10, les répondants peuvent rarement attribuer la note de 4,5. Ils doivent opter pour soit pour le 4, soit pour le 5. Cette limite disparaît avec les échelles à curseur.

Déplacer le curseur sur la zone qui permet d'exprimer votre jugement :



Les échelles d'icônes (*smiling faces*)

Pour faire face aux barrières de la langue, certaines échelles de mesure telle que celle de Likert (Adaptée à la satisfaction, au degré d'accord, ou à l'intention par exemple) peuvent être reprises sur la base d'une échelle avec des images : elles se nomment « échelles graphiques ». Si la compréhension des mots d'un pays à un autre s'arrête généralement aux frontières qui les séparent, les images ont l'avantage de faire passer un message similaire quel que soit le pays. Les couleurs sont également un bon moyen de franchir la barrière de la langue (Cf. figure 6). Il faut cependant être vigilant au choix des couleurs qui n'ont pas toujours les mêmes significations d'un pays à un autre. Avec une échelle à icônes, les smileys traduisent des sentiments positifs ou négatifs via l'expression des sourires (Cf. figure 7). L'utilisation des échelles d'icônes permettent alors de contourner la barrière de la langue. Ces échelles de mesure sont **ludiques**, et d'une grande **facilité de compréhension**. Cependant, l'échelle à icônes est considérée comme moins « cognitive » que l'échelle de Likert, et fait plutôt appel à « l'affectif » (Derbaix, 1997).

De plus, cette échelle est très intéressante puisque des études ont pu montrer que son **taux de retour est plus important**, notamment parce qu'elle permet d'interroger des populations illettrées. Elle permet également d'interroger une cible compliquée à atteindre, mais importante pour certaines entreprises dans leur stratégie marketing : les enfants. Les échelles d'icônes permettent effectivement de capter plus facilement l'attention et sont très bien perçues auprès des jeunes. (Jolibert et Baumgartner, 1981).

En utilisant ce type d'échelle, on peut ensuite mettre en correspondance ces icônes sur des supports sémantiques. C'est ce qu'a tenté Vernet (1992), dans une étude sur les propriétés des échelles d'icônes : ainsi, trois points d'ancrage ont pu être définis comme repère pour les différentes expressions des *smiley* d'une échelle. Ces points d'ancrage servent de repère pour les traitements statistiques. L'échelle à icônes présente alors les mêmes propriétés qu'une échelle d'intervalle.

Cette échelle permet ainsi de contourner les barrières de la langue, mais il faut être vigilant avec son utilisation. En effet, ces échelles sont **limitées en terme graphique** : le placement des yeux,

l'inclinaison, et l'intensité du sourire peuvent conclure à des résultats parfois différents (Vernette, 1992). Il est donc nécessaire de vérifier les propriétés de chaque nouvelle figure proposée dans les échelles à icônes.



Figure 7 : Un exemple d'échelle à icônes

Réduire les effets de Halo

Pour limiter certains effets de Halo comme ceux présentés en début de partie, différentes techniques peuvent être utilisées selon le mode d'administration choisi par les chercheurs ou les chargés d'études.

L'effet de fatigue mis en évidence par Ganassali peut effectivement être atténué dans les enquêtes en ligne, en proposant des **menus déroulants** plutôt que d'afficher les échelles l'une sous l'autre. En effet, par effet de fatigue, les répondants peuvent donner des réponses identiques (de manière consciente ou inconsciente) au fur et à mesure que les questions sont posées. En proposant des menus déroulant plutôt que des puces à cocher, ces derniers sont alors obligés d'aller chercher la réponse dans la liste fournie par le menu. L'effort est alors obligatoire de la part du répondant et les réponses données ne subissent donc pas l'effet de fatigue qui consiste à toujours donner la même réponse (du fait que les échelles soient placées identiquement l'une sous l'autre). Toutefois, les menus déroulants sont **coûteux en termes de temps** et **assez fastidieux** pour le répondant : cette technique n'est pas aussi pratique qu'une coche dans les puces prévues à cet effet. L'utilisation de cette pratique doit alors être limitée au risque de faire apparaître de nouveaux biais dans la mesure.

D'autre part, nous avons vu avec **l'effet de cohérence**, que, lorsqu'il s'agit d'attribuer des notes à une liste d'affirmations, les personnes interrogées utilisent les réponses qui ont déjà été données pour attribuer la suivante. Ainsi, l'ordre des affirmations tel qu'il est proposé aux enquêtés peut avoir une influence sur l'ensemble des notes attribuées. S'il existe une cohérence entre les réponses données, une question se pose : qu'est-ce qu'une **bonne note** pour l'individu qui vient de les fournir ? La définition d'une « bonne note » peut varier en fonction de l'ordre d'apparition des items. De même, peut-on dire que la définition d'une « bonne note » est la même pour tous les répondants ? Pour permettre de répondre à ces interrogations, on peut choisir d'utiliser une question ouverte en plus des notes qui ont été attribuées afin de les justifier (*Pourquoi avez-vous attribué cette note ?*). Ces questions ouvertes sont très riches et sont un bon complément d'information pour ceux qui analysent les résultats. De surcroît, en demandant aux répondants « ce qu'est une bonne note pour eux », cela permet d'affiner les comparaisons entre les différents individus. Si ces techniques sont très enrichissantes, elles sont néanmoins très complexes à étudier. Le traitement de ces questions demande alors beaucoup plus de temps.

b) Faire appel à des instruments de mesure plus adaptés

Dans certains cas, il est nécessaire dans le cadre d'une étude marketing de faire appel à la **mémoire** ou à la **fréquence d'utilisation**. La mesure peut alors faire appel aux échelles. Pourtant, les échelles de mesure employées sont parfois très peu adaptées à ce que l'on souhaite mesurer réellement, et la mesure en est faussée. En effet, la **perception du temps est très subjective** ce qui complexifie sa mesure (Giannelloni et Vernet, 2001). Les échelles de mesure doivent alors permettre de répondre facilement, ce qui nécessite l'usage d'un vocabulaire clair et concis. C'est pourquoi, il est préférable d'utiliser des « ancrages temporels objectifs » plutôt que des adverbes de temps. Cela permet ainsi d'éviter que deux adverbes se chevauchent pour au final exprimer la même chose.

Toutefois, le vocabulaire utilisé dans les échelles de mesure n'est pas le sujet de cette partie. En effet, pour ce type de mesure, c'est l'instrument lui-même qui est directement remis en cause plutôt que le contenu de l'échelle de mesure. Nous verrons dans cette dernière partie quelles sont les méthodes qui existent pour mesurer l'attitude qui est souvent étudiée en marketing, autrement que via les échelles de mesure, notamment pour les notions compliquées à définir telle que la perception du temps, ou la fréquence d'utilisation.

En effet, pour savoir à quelle fréquence la cible qu'on étudie écoute la radio / regarde une chaîne de télévision, on peut décider de se baser sur des éléments déclaratifs via les questionnaires. Cependant, il est parfois judicieux et plus intéressant d'utiliser de nouveaux « **outils de mesure** ». Dans un questionnaire administré, afin d'obtenir l'information souhaitée, il serait nécessaire de demander aux interviewés une question sous la forme suivante : « *A quelle fréquence regardez-vous la chaîne de télévision XYZ ?* » ou « *Au cours des 6 derniers mois...* ». Or, ces questions sont très compliquées pour obtenir une information fiable. Il est effectivement difficile pour le répondant de juger le temps réellement passé devant la chaîne de télévision, ou de se souvenir d'un évènement passé il y a quelques mois. C'est pourquoi, une question ouverte permet parfois de « mieux mesurer directement le temps total » (Vernet, 1991). Par ailleurs, **des mesures bien plus précises** existent et permettent d'obtenir cette information. C'est le cas de la société d'études Médiamétrie qui est spécialisée dans la mesure d'audience et qui fait appel à un panel de répondant. L'individu indique alors en temps réel, le temps qu'il passe devant la chaîne. Un boîtier développé par la société d'études permet alors de savoir très précisément combien de temps l'individu (dont on connaît ses caractéristiques) a passé devant la chaîne.

D'autres techniques permettent également d'obtenir des mesures de manière nettement plus précise qu'un questionnaire. L'*Eye tracking* permet par exemple d'obtenir des informations très fines concernant les endroits où le regard s'est déposé : cette information n'est pas possible à retenir dans un questionnaire où l'on demande aux répondants de se rappeler de ce qui a été vu dans un supermarché par exemple.

Ces techniques ne sont pas les seules. Toutefois, si elles montrent les limites liées aux échelles de mesure dans les questionnaires, elles n'entrent pas dans le cadre de notre mémoire. Nous étudierons la problématique liée au choix d'une échelle ou d'une autre dans la partie qui suit.

PARTIE 3

Etude d'un biais lié au choix des échelles

1) Méthodologie

a) Rappel du contexte

Tels que nous les avons vus précédemment, les biais liés au choix des échelles de mesure sont nombreux et déterminants dans l'analyse qu'il s'en suit. Les biais dans la mesure sont cependant inévitables. Supprimer un biais peut effectivement se faire au détriment d'un autre. Nous avons pu le voir avec l'échelle de Stapel : cette échelle permet d'avoir des éléments précis sur ce que l'on souhaite mesurer. Cette échelle demande bien plus d'implication de la part du répondant. Pourtant, lorsqu'elle est trop utilisée dans une même étude, l'échelle peut perdre l'attention de l'interviewé. Il faut donc apprendre à minimiser ces biais tout en maximisant la qualité de l'information que l'on souhaite mesurer.

Parmi les biais présentés dans la partie 2, il y en a un qui a retenu mon attention. Il n'existe en effet pas d'étude concrète sur les conséquences liées **au nombre de modalités** dans une échelle de mesure. De plus, certaines échelles de mesure excluent les modalités neutres. Pourtant, il est légitime de penser qu'en présence d'une modalité neutre dans l'échelle de mesure, les résultats peuvent différer de ceux menés avec une échelle où cette modalité n'existe pas.

b) Mise en place de l'étude

L'expérience présentée ci-après, et menée avec l'aide de l'IFOP aura pour but de démontrer qu'il existe effectivement **une différence significative liée au choix de l'échelle de mesure** pour une même question. Pour mener à bien cette expérience et vérifier cette hypothèse, nous avons donc choisi de poser deux questions à des échantillons de même taille. Pour que les comparaisons entre les deux échantillons soient possibles, il est nécessaire d'avoir un nombre de répondants suffisants. Afin de pouvoir respecter cette dernière condition, nous avons choisi d'intégrer les deux questions dans l'omnibus Omcawi⁴ de l'IFOP. Cet omnibus s'effectue sur un échantillon de plus de 1 000 répondants. Les conditions de mise en œuvre (les délais de réalisation, et le coût lié à l'intégration des deux questions dans l'omnibus IFOP) nous contraignent à scinder l'échantillon de 1 000 répondants en quatre échantillons appariés. Après appariement, les quatre échantillons scindés seront alors

⁴ 1000 personnes de 18 ans et plus, interrogées sur Internet toutes les semaines. Terrain du mardi au vendredi – résultats : le mardi - Réalisable en Régions, en France, en Europe.

comparables deux-à-deux. Les deux questions vont apparaître dans une vague de l'omnibus de manière à pouvoir tester plusieurs échelles de mesure sur les mêmes questions.

c) Les échelles à tester

Afin de tester l'impact de la présence *versus* absence de la modalité neutre sur les réponses, quatre échelles différentes vont être testées. L'expérience se déroulera sur une seule vague de l'omnibus. Pour cette vague, l'échantillon sera splitté en quatre : les quatre échelles de mesure testées seront donc soumises à des échantillons de 250 individus. Nous avons donc choisi de tester quatre échelles de mesure basées sur celle de Likert en intégrant ou en retirant la modalité neutre.

Un premier échantillon de répondants sera donc soumis aux deux questions à laquelle il devra répondre à l'aide de l'échelle suivante, appelée dans notre étude « échelle A » :

Echelle A :

- Tout-à-fait d'accord
- Plutôt d'accord
- Plutôt pas d'accord
- Pas du tout d'accord

Pour le deuxième échantillon, une modalité neutre va être intégrée à l'échelle de mesure. La deuxième partie de l'échantillon de la première vague de l'omnibus sera donc soumise à l'échelle suivante :

Echelle B :

- Tout-à-fait d'accord
- Plutôt d'accord
- Ni d'accord, ni pas d'accord
- Plutôt pas d'accord
- Pas du tout d'accord

Comme nous avons pu le voir dans la partie 2, cette deuxième échelle intègre la modalité neutre qui peut également être interprétée comme étant la modalité « refuge ». En effet, lorsque l'échelle en quatre modalités est employée, cela oblige le répondant à se positionner d'un côté ou d'un autre de l'échelle. Ce qui n'est pas le cas avec la modalité neutre où le répondant peut choisir de ne pas « concrètement » se positionner. Afin de déterminer si cette modalité est liée à un manque d'opinion, nous avons jugé intéressant de mettre en perspective ce qu'il se passe quand on intègre la modalité « Vous n'avez pas d'opinion ». Les deux autres échelles vont nous permettre de le mettre en exergue. Nous pourrons voir dans quelle mesure la modalité « Ni d'accord, ni pas d'accord » équivaut à un manque d'intérêt vis-à-vis du sujet en question.

Notre troisième échantillon aura alors la même échelle que le premier, mais la modalité « Vous n'avez pas d'opinion, vous ne savez pas » y sera intégrée :

Echelle C :

- Tout-à-fait d'accord
- Plutôt d'accord
- Plutôt pas d'accord
- Pas du tout d'accord
- Vous n'avez pas d'opinion, vous ne savez pas

Enfin, pour le quatrième échantillon nous laisserons la possibilité au répondant de sélectionner sa réponse parmi les modalités de base dites « signifiantes », mais il aura également le choix entre les modalités « Ni d'accord, ni pas d'accord » et « Vous n'avez pas d'opinion, vous ne savez pas ». La distinction entre ces deux dernières modalités nous permettra de distinguer si la part de « Ni d'accord, ni pas d'accord » de l'échelle précédemment testée correspond davantage à un manque d'information de la part du répondant, ou davantage à un avis neutre sur le sujet en question. Pour ce dernier échantillon, l'échelle se présentera donc sous la forme suivante :

Echelle D :

- Tout-à-fait d'accord
- Plutôt d'accord
- Ni d'accord, ni pas d'accord
- Plutôt pas d'accord
- Pas du tout d'accord
- Vous n'avez pas d'opinion, vous ne savez pas

d) Le choix des questions

Pour tester les différentes échelles présentées ci-dessus, deux questions vont être posées à l'ensemble des échantillons soumis aux quatre échelles. Ces deux questions seront identiques dans nos quatre enquêtes afin de nous concentrer sur le sujet de l'étude : les échelles de mesure.

Pour identifier les éventuels biais liés au choix des échelles, nous avons donc choisi de poser deux questions. Toutefois, les deux questions qui seront posées auront des niveaux d'implication différents.

Le niveau d'implication des répondants peut effectivement jouer sur les réponses données. Lorsque que l'interviewé est questionné sur un sujet qu'il maîtrise, il donnera plus facilement son opinion que s'il ne connaît pas du tout le sujet. L'**implication du répondant** est donc déterminante dans les réponses fournies.

C'est pourquoi, nous avons souhaité poser deux questions à niveau d'implication différent pour étudier nos échelles de mesure. Elles vont nous permettre d'identifier l'impact de la présence du « Ni d'accord, ni pas d'accord » dans l'échelle de mesure, l'impact du « Vous n'avez pas d'opinion » mais également l'impact de ces deux modalités lorsqu'elles sont confrontées l'une à l'autre.

La première question qui sera posée aux quatre échantillons sera donc une question où les opinions semblent être ancrées, alors que la seconde question sera davantage tournée sur un sujet où les opinions ne sont pas encore formées.

Pour pouvoir intégrer les deux questions à l'omnibus IFOP, nous devons choisir une thématique en lien avec le secteur d'activité qui souhaite poser les questions : le pôle services⁵.

Pour la question où l'on suppose que le niveau d'implication du sujet est plus important, et que les répondants ont probablement un avis, nous avons choisi un sujet lié au prix. En effet, lorsqu'il s'agit d'argent, on suppose que les répondants ont plus facilement un avis à fournir. Nous avons donc choisi de demander aux répondants ce qu'ils pensent du rôle de l'Etat dans la fixation du prix de l'électricité en France.

En revanche, pour la question concernant un plus faible niveau d'implication de la part des interviewés, et *à fortiori* nécessite que le répondant ait un intérêt plus conséquent pour le sujet, la thématique du gaz de schiste qui fait actuellement débat en France a été choisie. Une question portant sur le gaz de schiste va nous permettre d'établir un lien ou non entre la sélection des modalités neutres et le niveau d'implication des répondants. Cette question va effectivement nous permettre de voir dans quelle mesure le répondant choisit une réponse alors qu'il ne dispose pas spécialement d'information à l'égard de ce sujet. Par ailleurs, cette question permettra également à l'IFOP de disposer d'une information que la société ne possède pas encore : l'opinion des Français sur l'exploitation du gaz de schiste.

Les questions posées à l'omnibus IFOP seront donc les suivantes (Cf. Annexe 2) :

- 1) *Dans quelle mesure êtes-vous d'accord avec l'affirmation suivante ?*
L'état français doit intervenir dans la fixation du prix de l'électricité.

[Intégration de l'échelle de mesure : Echelle A, B, C ou D]

- 2) *Dans quelle mesure êtes-vous d'accord avec l'affirmation suivante ?*
L'exploitation du gaz de schiste doit être autorisée en France.

[Intégration de l'échelle de mesure : Echelle A, B, C ou D]

e) Profils sociodémographiques

L'omnibus CAWI de l'IFOP permet aux clients ayant choisi d'y intégrer leur(s) question(s), d'obtenir des résultats selon les dimensions sociodémographiques des répondants. L'omnibus est un bon moyen pour ces clients d'effectuer de courtes analyses statistiques à partir de ces quelques questions, pour un moindre coût.

⁵ Le pôle services de l'IFOP est constitué de 3 secteurs principaux : finances & l'assurance, l'énergie et le transport.

Lorsque l'omnibus CAWI de l'IFOP sera terminé, nous obtiendrons les données correspondant aux questions que nous avons choisi de poser, pour les quatre échelles de mesure. Parallèlement à ces données, nous recevrons les profils sociodémographiques des répondants. Cela va nous permettre de croiser les données issues de nos questions par les variables de type âge, sexe, PCS ou encore niveau de diplôme.

Le fait de pouvoir lier les résultats que nous allons obtenir avec nos différentes échelles de mesure va nous permettre d'**établir des profils selon les caractéristiques des répondants**. Il s'agira notamment de démontrer l'hypothèse que les femmes ont plus tendance à répondre « ne pas savoir » que les hommes.

2) Principaux résultats

Le terrain de cette étude s'est déroulé du mercredi 17 au vendredi 19 juillet. Les principaux résultats qui vont nous permettre de valider ou non nos hypothèses seront présentés dans cette partie.

Afin d'analyser les résultats, les quatre échantillons ont été redressés à partir de nos données de cadrage (celles qui nous ont permis de constituer l'échantillon de 1000 répondants), de manière à obtenir quatre échantillons appariés. Cela nous permet ainsi de comparer les quatre échantillons deux-à-deux. Ces analyses ont été effectuées à partir du logiciel Sphinx.

Les résultats tels qu'ils sont affichés dans la suite de ce rapport, présentent les base brutes (colonne « Nb »), et les pourcentages redressés (colonne « %cit. »).

a) Le « Ni d'accord, ni pas d'accord » comme modalité refuge

Q1 - L'état français doit intervenir dans la fixation du prix de l'électricité.

Echelle A	Nb	% cit.	Echelle B	Nb	% cit.
Tout à fait d'accord	94	37%	Tout à fait d'accord	103	41%
Plutôt d'accord	134	↗ 54%	Plutôt d'accord	86	↘ 35%
Plutôt pas d'accord	13	5%	Ni d'accord, ni pas d'accord	42	17%
Pas du tout d'accord	9	4%	Plutôt pas d'accord	10	4%
Total	250	100%	Pas du tout d'accord	9	4%
			Total	250	100%

Q2 - L'exploitation du gaz de schiste doit être autorisée en France.

Echelle A	Nb	% cit.
Tout à fait d'accord	33	12%
Plutôt d'accord	76	30%
Plutôt pas d'accord	61	25%
Pas du tout d'accord	80	32%
Total	250	100%

Echelle B	Nb	% cit.
Tout à fait d'accord	35	14%
Plutôt d'accord	39	16%
Ni d'accord, ni pas d'accord	79	31%
Plutôt pas d'accord	33	13%
Pas du tout d'accord	64	26%
Total	250	100%

- La proportion d'individus sur cette modalité est significativement *supérieure* sur cette échelle que celle qui est comparée
- La proportion d'individus sur cette modalité est significativement *inférieure* sur cette échelle que celle qui est comparée

Au regard des résultats que peut fournir cette étude, un premier constat apparaît entre les deux premières échelles (échelle A et échelle B) : quel que soit le degré d'implication des questions auxquelles les répondants ont été soumis, on remarque en effet qu'une différence significative est présente entre les modalités des échelles A et B. En effet, lorsque les répondants sont soumis à une échelle en quatre modalités (Tout à fait d'accord, plutôt d'accord, plutôt pas d'accord, pas du tout d'accord), le taux de « Plutôt d'accord » est surreprésenté. A contrario, en présence de la modalité centrale « Ni d'accord, ni pas d'accord » (échelle B), ce taux diminue significativement au profit de la nouvelle modalité.

Cette remarque met donc en évidence que la modalité centrale « Ni d'accord, ni pas d'accord » est une **modalité refuge** pour les répondants. N'ayant pas de voies de sortie sur l'échelle A, les personnes interrogées se placent au niveau de l'échelon « Plutôt d'accord ». Dans le cas de la première question où le niveau d'implication des répondants est *à priori* plus important, les résultats obtenus après l'utilisation de l'échelle A permettent d'affirmer que les répondants sont globalement d'accord avec le fait que l'Etat doit intervenir dans la fixation du prix de l'électricité : le taux de répondants « plutôt d'accord » (54%) est cependant bien plus élevé que le taux de répondants « tout à fait d'accord » (37%). Avec l'échelle B, même si globalement les résultats démontrent que les français sont d'accord avec le fait que l'Etat doit intervenir dans la fixation du prix de l'électricité, le taux d'individus « tout à fait d'accord » est identique à l'échelle A. Cependant, le taux de « plutôt d'accord » est bien plus faible dans le cas de l'échelle B. Le taux d'individus « tout à fait d'accord » (41%) devient alors supérieur au taux d'individus « plutôt d'accord » (35%).

Les sociétés d'études qui analysent ces résultats ont alors fortement intérêt à **ne retenir que le taux de « tout à fait d'accord »** qui est beaucoup plus « pur » lorsque les échelles proposées ne sont qu'en quatre modalités. Dans le cas de la seconde question où l'implication des répondants semble être moins importante, puisqu'il s'agit d'un sujet que les Français maîtrisent moins, on retrouve cette répartition, affectée au taux de « Plutôt d'accord » et « Plutôt pas d'accord » : le sujet étant plus délicat à traiter, les avis divergent. En l'absence de modalité neutre, lorsque les répondants sont ni d'accord, ni pas d'accord ou lorsqu'ils ne savent pas, ces personnes interrogées se placent dans les échelons centraux.

b) Le « Vous ne savez pas, vous n'avez pas d'opinion » comme modalité refuge dans une question peu impliquante

Q1 - L'état français doit intervenir dans la fixation du prix de l'électricité.

Echelle A	Nb	% cit.
Tout à fait d'accord	94	37%
Plutôt d'accord	134	54%
Plutôt pas d'accord	13	5%
Pas du tout d'accord	9	4%
Total	250	100%

Echelle C	Nb	% cit.
Tout à fait d'accord	91	36%
Plutôt d'accord	120	48%
Plutôt pas d'accord	22	9%
Pas du tout d'accord	6	2%
Vous n'avez pas d'opinion	15	6%
Total	254	100%

Echelle A	Nb	% cit.
Tout à fait d'accord	33	12%
Plutôt d'accord	76	↗ 30%
Plutôt pas d'accord	61	25%
Pas du tout d'accord	80	32%
Total	250	100%

Echelle C	Nb	% cit.
Tout à fait d'accord	25	9%
Plutôt d'accord	54	↘ 21%
Plutôt pas d'accord	44	19%
Pas du tout d'accord	79	30%
Vous n'avez pas d'opinion	52	20%
Total	254	100%

Q2 - L'exploitation du gaz de schiste doit être autorisée en France.

Afin de vérifier que les résultats obtenus sont identiques dans le cas où la modalité « Vous n'avez pas d'opinion, vous ne savez pas » est proposée, et dans le cas où elle ne l'est pas, nous avons comparé les résultats de la même manière que pour la modalité « Ni d'accord, ni pas d'accord ».

Dans le cas où la question est impliquante (où l'on suggère que les répondants ont quoi qu'il arrive un avis à fournir), nous détectons aucune différence significative entre les quatre modalités de l'échelle A, et ces quatre mêmes modalités de l'échelle C. La modalité « Vous n'avez pas d'opinion » n'apparaît donc pas comme une voie de sortie puisque 6% seulement des individus s'y placent : les répondants ont un avis et ils se placent d'un côté ou de l'autre de l'échelle de mesure.

En revanche, dans le cas où la question demande d'être davantage renseigné sur le sujet pour pouvoir y répondre (comme c'est le cas avec notre question portant sur le gaz de schiste), on remarque que le taux de répondant qui se place sur l'échelon « Vous n'avez pas d'opinion, vous ne savez pas » est élevé (20%). Par ailleurs, en comparant les quatre modalités de l'échelle A avec celles de l'échelle C, on constate que toutes ces modalités sont équivalentes d'une échelle à un autre, sauf pour la modalité « Plutôt d'accord » qui est significativement sous représentée (30% contre 21%). De la même façon que pour la modalité neutre de notre échelle B avec le « Ni d'accord, ni pas d'accord », on peut mettre en évidence que les répondants ne savent pas quoi répondre parce qu'ils ne sont ni d'accord, ni pas d'accord, ou parce qu'ils n'ont pas d'avis, sont significativement plus

présents dans la modalité « Plutôt d'accord ». Cette dernière s'inscrit alors comme **modalité refuge** lorsque la question est peu **impliquante**.

c) « Ni d'accord, ni pas d'accord » versus « Vous n'avez pas d'opinion, vous ne savez pas »

Nous avons vu précédemment que les modalités « Ni d'accord, ni pas d'accord » et « Vous n'avez pas d'opinion, vous ne savez pas » peuvent devenir des modalités refuges lorsqu'elles sont proposées. Dans le cas où elles ne le sont pas, on retrouve ces répondants majoritairement dans les échelons en « plutôt ». Toutefois, nous souhaitons vérifier si ces deux modalités, considérées comme neutres jouent un rôle similaire lorsqu'elles sont proposées et lorsqu'elles ne le sont pas.

A première vue, en comparant deux-à-deux les échelles A et B, et les échelles A et C (vues précédemment), et dans le cadre de notre première question plus impliquante, on constate une première différence : le taux de « Plutôt d'accord » diminue significativement en présence de la modalité « Ni d'accord, ni pas d'accord » (au profit de cette dernière) de l'échelle B. Cette remarque n'est pas valable en présence de la modalité « Vous ne savez pas, vous n'avez pas d'opinion » de l'échelle C, où aucune différence significative entre les modalités est à noter. Ce premier constat montre donc que ces deux modalités n'ont pas le même rôle de « refuge » lorsqu'elles sont proposées. Ainsi, on peut mettre en évidence le fait que les individus se placent plus facilement dans une modalité en « Ni d'accord, ni pas d'accord » que dans une modalité en « Vous n'avez pas d'opinion, vous ne savez pas ».

Par ailleurs, lorsque les questions sont moins impliquantes et demandent de connaître un minimum le sujet pour pouvoir y répondre, le basculement des modalités en « plutôt » vers la modalité neutre existe quel que soit cette modalité neutre (« Ni d'accord, ni pas d'accord » ou « Vous n'avez pas d'opinion, vous ne savez pas »). Toutefois, ce basculement est plus important dans le cas où l'échelon « Ni d'accord, ni pas d'accord » est proposé : on perd plus d'individus sur l'ensemble des modalités en « plutôt » quand « Ni d'accord, ni pas d'accord » est proposé que lorsque « Vous n'avez pas d'opinion, vous ne savez pas » est proposé.

Cette différence se confirme lorsque les deux échelons sont directement comparés l'un par rapport à l'autre (entre les deux échelles B et C) :

Q1 - L'état français doit intervenir dans la fixation du prix de l'électricité.

Echelle B	Nb	% cit.
Tout à fait d'accord	103	41%
Plutôt d'accord	86	↓ 35%
Ni d'accord, ni pas d'accord	42	↑ 17%
Plutôt pas d'accord	10	↓ 4%
Pas du tout d'accord	9	4%
Total	250	100%

Echelle C	Nb	% cit.
Tout à fait d'accord	91	36%
Plutôt d'accord	120	↑ 47%
Plutôt pas d'accord	22	↑ 9%
Pas du tout d'accord	6	2%
Vous n'avez pas d'opinion	15	↓ 6%
Total	254	100%

Q2 - L'exploitation du gaz de schiste doit être autorisée en France.

Echelle B	Nb	% cit.
Tout à fait d'accord	35	14%
Plutôt d'accord	39	16%
Ni d'accord, ni pas d'accord	79	31%
Plutôt pas d'accord	33	13%
Pas du tout d'accord	64	26%
Total	250	100%

Echelle C	Nb	% cit.
Tout à fait d'accord	25	9%
Plutôt d'accord	54	21%
Plutôt pas d'accord	44	19%
Pas du tout d'accord	79	30%
Vous n'avez pas d'opinion	52	20%
Total	254	100%

Ainsi, on peut mettre en évidence que les modalités « Ni d'accord, ni pas d'accord » ou « Vous n'avez pas d'opinion » sont significativement différentes : 17% contre 6% dans le cas de la première question, et 31% contre 20% pour la seconde question. Cela confirme ainsi que les répondants se placent plus facilement dans la modalité « Ni d'accord, ni pas d'accord » que dans la modalité « Vous n'avez pas d'opinion, vous ne savez pas » : on constate donc que les individus ont du mal à affirmer qu'ils n'ont pas d'opinion. Il semble donc plus facile pour eux de déclarer avoir un avis neutre.

Cette différence se confirme également en confrontant les deux modalités à l'intérieur d'une même échelle (échelle D) :

Q1 - L'état français doit intervenir dans la fixation du prix de l'électricité.

Echelle D	Nb	% cit.
Tout à fait d'accord	100	40%
Plutôt d'accord	86	34%
Ni d'accord, ni pas d'accord	41	16%
Plutôt pas d'accord	13	5%
Pas du tout d'accord	4	2%
Vous n'avez pas d'opinion	9	3%
Total	253	100%

Q2 - L'exploitation du gaz de schiste doit être autorisée en France.

Echelle D	Nb	% cit.
Tout à fait d'accord	19	7%
Plutôt d'accord	37	16%
Ni d'accord, ni pas d'accord	65	25%
Plutôt pas d'accord	30	14%
Pas du tout d'accord	76	29%
Vous n'avez pas d'opinion	26	9%
Total	253	100%

En calculant l'intervalle de confiance pour les différentes modalités on obtient les résultats suivants pour la question 1 :

$$\begin{aligned} \text{IC}_{\text{Ni d'accord, ni pas d'accord} - \text{Q1}} &= 0,16 \pm 1,96^6 \times \text{RACINE} ((0,16 \times (1 - 0,16)) / 252) \\ &= [11\% ; 21\%] \end{aligned}$$

$$\text{IC}_{\text{Vous n'avez pas d'opinion, vous ne savez pas} - \text{Q1}} = [2\% ; 6\%]$$

Ainsi, on constate via ces analyses qu'il existe une différence significative entre les deux échelons puisque ces deux intervalles de confiance ne se chevauchent pas. Les deux modalités jouent donc un rôle différent dans l'échelle de mesure, ce qui confirme les résultats présentés plus haut.

Concernant la seconde question, on obtient les intervalles de confiance suivants :

$$\text{IC}_{\text{Ni d'accord, ni pas d'accord} - \text{Q2}} = [20\% ; 30\%]$$

$$\text{IC}_{\text{Vous n'avez pas d'opinion, vous ne savez pas} - \text{Q2}} = [5\% ; 13\%]$$

Les résultats pour la seconde question démontrent aussi qu'il existe une différence significative entre le taux de répondant de la modalité « Ni d'accord, ni pas d'accord » et « Vous n'avez pas d'opinion, vous ne savez pas ». L'implication de la question n'influence donc pas le choix d'une modalité ou d'une autre, puisque la modalité centrale « **Ni d'accord, ni pas d'accord** » est **davantage sélectionnée** par les répondants. Ces résultats démontrent ainsi qu'il y a une préférence de la part des répondants à **donner un avis neutre**, plutôt que **d'avouer ouvertement ne pas avoir d'opinion** ou **ne pas savoir**.

d) Les femmes avouent plus facilement qu'elles ne savent pas

Grâce au logiciel Sphinx, nous avons pu croiser les deux questions posées à l'omnibus CAWI de l'IFOP avec les critères sociodémographiques des répondants. Les différents tableaux croisés obtenus vont nous permettre de détecter des différences significatives entre les individus, selon ces critères.

Afin de détecter un comportement différent selon l'âge, le sexe, la profession, ou encore le niveau de diplôme, la totalité des échelles de mesure pour les deux questions a été croisée avec les variables qui correspondent à ces critères.

Toutefois, **aucune différence significative** n'a pu être mise en évidence concernant **l'âge, la profession** et le **niveau de diplôme**. Seules des disparités liées à la région habitée, et entre hommes et femmes ont pu surgir : ces disparités n'apparaissent cependant que sur la question n°2 qui est moins impliquante (sur l'exploitation du gaz de schiste en France).

Au niveau des régions (Cf. Annexe 3-c), on détecte des différences significatives sur l'échelle D : la **région Parisienne** est significativement sous-représentée au niveau de la modalité « Ni

⁶ Pour une confiance de 95% la borne supérieure de $Z_{0,975} = 1,96$

d'accord, ni pas d'accord » alors que la modalité « Tout à fait d'accord » est significativement surreprésentée. Les Parisiens semblent être en mesure de se positionner plus facilement que le reste des Français. Cette prise de position plus facile pour les habitants de la région Parisienne mérite réflexion. Une piste serait que les habitants de cette région sont davantage informés de l'actualité grâce aux nombreux journaux quotidiens gratuits, ce qui leur permet d'avoir une opinion plus formée sur ces questions qui traitent de l'actualité.

Pour ce qui est des différences entre **les hommes et les femmes**, il n'existe pas de lien selon le sexe du répondant lorsque la question est impliquante. Néanmoins, lorsque la question nécessite d'être un minimum informée pour pouvoir y donner son avis, plusieurs constats apparaissent : les hommes sont significativement plus présents dans la modalité extrême de l'échelle « Tout à fait d'accord » (Cf. Annexe 3-a et 3-b). Ces derniers ont donc plus de facilité à donner un avis tranchant à ce type de question (20% de « Tout à fait d'accord » pour les hommes contre 8% pour les femmes de l'échelle B et 11% contre 4% sur l'échelle D).

Mais cette remarque n'est pas le seul constat qui montre que **les femmes ont davantage de mal à se positionner que les hommes sur ce sujet**. En effet, dans les deux cas où la modalité « Vous ne savez pas, vous n'avez pas d'opinion » est proposée (échelle C et échelle D), les femmes sont significativement surreprésentées : 29% contre 11% sur l'échelle C, et 16% contre 5% sur l'échelle D. Si ces différences sont notables, il est important de préciser qu'aucune différence significative n'existe entre les hommes et les femmes sur la modalité « Ni d'accord, ni pas d'accord ». Il ne s'agit donc pas pour elles de se placer dans la modalité neutre de l'échelle. Les femmes avouent plus aisément que les hommes ne pas avoir d'opinion, ce qui confirme notre hypothèse de départ que les femmes ont plus tendance à répondre « ne pas savoir » que les hommes.

3) Apports de l'étude

Dans le cadre de cette étude qui a pu être menée avec l'aide de l'IFOP, plusieurs éléments intéressants se dégagent. En effet, nous avons pu mettre en avant que les résultats d'une question posée sous la forme d'une échelle de mesure sont variables du nombre et du type de modalités qui y sont proposées. Ainsi, lorsqu'une échelle de mesure n'intègre pas la modalité centrale « Ni d'accord, ni pas d'accord », l'étude met en évidence qu'un individu incapable de se placer de part et d'autre de l'échelle de mesure va sélectionner « par défaut » l'échelon « Plutôt d'accord ». Si les sociétés d'études ne souhaitent pas intégrer cette modalité dans leurs mesures, notamment parce qu'elles souhaitent (elles ou leurs clients) que les répondants se positionnent d'un côté ou de l'autre de l'échelle, alors elles doivent tenir compte d'un fort taux de « plutôt d'accord ». Dans ce cas, les sociétés d'études ont intérêt à se concentrer sur le taux de répondants « Tout à fait d'accord » qui est beaucoup plus magistral. Cette remarque est valable quel que soit le niveau d'implication de la question.

Toutefois, si l'étude a pu mettre en évidence ces différences, elle montre tout de même que les échelons situés aux extrémités sont équivalents quelle que soit l'échelle utilisée. Les modalités en « tout à fait » et en « pas du tout » ne sont pas influencées par la présence d'une modalité neutre. Cette remarque est importante puisque ces échelons sont finalement ceux qui comptent le plus dans une étude puisqu'ils sont les plus tranchants.

L'analyse des résultats de l'étude démontre par ailleurs que la modalité « Ni d'accord, ni pas d'accord » est différente de « Vous n'avez pas d'opinion, vous ne savez pas ». Ces deux modalités ont effectivement deux rôles distincts : la modalité « Ni d'accord, ni pas d'accord » est significativement plus choisie que « Vous n'avez pas d'opinion, vous ne savez pas », ce qui fait d'elle une modalité refuge. Intégrer « Vous n'avez pas d'opinion, vous ne savez pas » permet ainsi de recueillir l'opinion des individus sans prendre le risque que ces derniers la sélectionne plus facilement : seuls les individus interrogés qui n'ont vraiment pas d'avis vont la choisir. Lorsque cette dernière est proposée sur une échelle avec les quatre modalités, le « ne sait pas » permet tout de même d'obliger les répondants à se positionner : les répondants semblent vouloir se positionner plutôt qu'affirmer ne pas avoir d'avis. Toutefois, lorsque la question est moins impliquante (parce qu'elle traite d'un sujet où peu de répondants ont un avis), le taux de répondant qui sélectionne « Vous n'avez pas d'opinion, vous ne savez pas » augmente, au profit du « plutôt d'accord ». Dans ce cas, les deux modalités neutres ont le même rôle de porte de sortie. Même si cela peut être dérangent à première vue, il semble être plus cohérent de proposer le « Vous n'avez pas d'opinion, vous ne savez pas » compte tenu de la nature même de la question qui nécessite de disposer d'informations pour pouvoir y répondre. Le taux de « ne sait pas » est en effet plus significatif que le « ni d'accord, ni pas d'accord » dans une question impliquante.

4) Limites de l'étude

Dans cette étude, nous avons cherché à identifier les conséquences liées à l'intégration, et/ou à l'absence des modalités dites « neutres ». Plusieurs limites apparaissent dans cette étude. Le mode d'administration des questions via l'omnibus CAWI en constitue la première : nos questions ont été placées en fin d'enquête. Elles ont alors pu être soumises à l'effet de fatigue tel que nous l'avons mis en évidence dans la partie 2. Le nombre de questions présent dans l'omnibus étant important.

Par ailleurs, il serait utile de tester ces résultats sur davantage de questions : nous avons choisi une question impliquante et une question non impliquante mais il aurait pu être intéressant de tester ces échelles sur d'autres sujets de questionnement. En effet, les différences obtenues entre les hommes et les femmes concernant la question portant sur le gaz de schiste peuvent être dues au fait que les femmes sont davantage respectueuses de l'environnement (ce qui expliquerait donc les taux de « tout à fait d'accord » plus élevés chez les hommes), ou tout simplement parce qu'elles s'intéressent/connassent moins le sujet. Il aurait également fallu tester ces échelles de mesure sur

d'autres indicateurs tels que l'intention ou la satisfaction afin d'en tirer des conclusions plus générales sur les échelles de mesure.

Enfin, les deux questions ont toujours été posées l'une après l'autre dans le même sens : d'abord la question portant sur l'intervention de l'Etat dans la fixation du prix, puis la question sur l'exploitation du gaz de schiste en second lieu. Il aurait fallu mettre en place un aléa pour éviter de faire intervenir un biais lié à l'ordre des questions (cette hypothèse a été envisagée mais compte tenu de l'aléa déjà utilisé pour les échelles de mesure, cette option n'a pas pu être programmée).

CONCLUSION

A l'heure actuelle, les sociétés d'études continuent d'utiliser les échelles de mesure avec quatre modalités dans leurs études de marché. Il suffit de regarder les nombreux sondages qui existent pour s'en rendre compte.

L'objectif de ce mémoire était de montrer que ces intentions de la part des sociétés d'études, mais également de la part du commanditaire lui-même peuvent conduire à des biais dans l'interprétation des résultats. En continuant d'utiliser des échelles sans modalité centrale dans leurs études de satisfaction par exemple, le taux de répondants « satisfait » est souvent faussé puisque surévalué. Les sociétés d'études ont conscience de l'existence de ce biais. Elles encouragent d'ailleurs leurs clients à intégrer cette modalité centrale. Pourtant, les rapports d'études font souvent abstraction de ce biais. En effet, les commanditaires y voient un intérêt à utiliser des échelles de mesure en quatre : le sous-total satisfait est bien plus important.

Si les résultats de notre étude ont permis de démontrer que les individus n'ayant pas d'avis sur une question se positionnent généralement sur la première modalité positive de l'échelle, l'étude montre également que les résultats sur les échelons les plus extrêmes sont identiques quel que soit l'échelle et le nombre de modalités utilisé. Toutefois, l'étude a pu montrer que les résultats varient en fonction de l'implication de la question. En présence d'une question non impliquante, la modalité « Vous n'avez pas d'opinion, vous ne savez pas » sera davantage sélectionnée. En revanche, quel que soit le niveau d'implication de la question, la modalité « Ni d'accord, ni pas d'accord » sera choisie, au détriment du premier échelon positif de l'échelle.

Par ailleurs, cette étude est très intéressante pour l'IFOP et moi-même puisqu'elle m'a permis d'apporter des informations capitales sur un sujet récurrent de la vie quotidienne du chargé d'études : le type d'échelle à employer lors de la construction d'un questionnaire.

Du point de vue de l'IFOP, et plus particulièrement du Pôle Services où j'ai effectué mon stage, le contenu même de l'étude permet de donner des éléments de réponse intéressants sur l'opinion des Français : Si les Français semblent être majoritairement d'accord avec l'intervention de l'Etat dans la fixation du prix de l'électricité, ils sont beaucoup plus partagés et semblent même être défavorables à l'exploitation du gaz de schiste en France.

BIBLIOGRAPHIE

Ouvrages

- Giannelloni JC, Vernet E. (2001), Etudes de marché 2^{ème} édition, Vuibert, 258-272
- Lendrevie J, Lévy J, Lindon D. (2009), Mercator : théorie et pratique du marketing 9^{ème} édition, Dunod, 66-67, 135
- Menahem R. (1968), Le différenciateur sémantique. In: L'année psychologique, vol. 68, 451-465
- Penta M, Arnould C, Decruynaere C. (2005), Développer et interpréter une échelle de mesure – Application du modèle de Rash, Mardaga, 116-119

Périodiques

- Aurifeille J-M. (1991), Proposition d'une méthode de mesure du halo affectif en marketing, *Recherche et Applications en Marketing*, vol.4, 59-77
- Bartikowski B, Chandon JL et Gierl H. (2006), Calibration internationale des échelles sémantiques, *Décisions Marketing n°43-44*, 207-219
- Brignier JM. (1991), L'influence des échelles de mesures sur les réponses collectées dans les enquêtes, *Recherche et Applications en marketing*, vol.4, 1-26
- Chandon JL. (2004), Une échelle ordinale permettant de classer les répondants en « satisfait », « indifférent » et « insatisfait », *Recherche et Applications en marketing*, vol.19, 39-53
- Cox BP. (1980), *The optimal Number of Response Alternatives for a Scale: A Review*, *Journal of Marketing Research*, 407-422
- Derbaix C. (1997), L'implication et l'enfant : proposition d'une échelle de mesure, *Recherche et Applications en marketing*, vol.12, 45-68
- Ganassali S. (2008), La Halo Cognitif : Un biais potentiel dans les méthodes de catégorisation des attributs de la satisfaction, *Décisions Marketing n°50*, 39-47
- Goueron J. (2011), L'intégration d'une mesure de l'attachement à la marque dans les études de satisfaction, *La Revue des Sciences de Gestion, Direction et Gestion n°252 – Marketing et comportement*, 109-118
- Hawkins DI, Albaum G, Best R, (1974), *Stapel Scale or Semantic Differential in Marketing Research?*, *Journal of Marketing Research*
- Lombart C. (2004), Le butinage : proposition d'une échelle de mesure, *Recherche et Applications en marketing*, vol.19, 1-30

Schwarz N. et Wyer R S. (1985), Effects of rank ordering stimuli on magnitude ratings of these and other stimuli, *Journal of Experimental Social Psychology*, vol 21, 30-46

Vernette E. (1991), L'efficacité des instruments d'études : évaluation des échelles de mesure, *Recherche et Applications en marketing*, vol.6, 43-65

Vernette E. (1992,) Evaluation des propriétés d'intervalles des échelles à icônes, Acte du congrès, 447-461

Webographie

Site Internet de l'ifop, consulté en juin à l'adresse <http://www.ifop.com>

Les 100 premiers instituts d'études marketing et opinion 2011 - Marketing Magazine n°158 consulté le 22 juin à l'adresse <http://www.e-marketing.fr/Marketing-Magazine/Article/Les-100-premiers-instituts-d-etudes-marketing-et-opinion-2011-43219-1.htm>

Mesurer la satisfaction du client, consulté le 29 juin 2013 à l'adresse <http://www.avantage-rh.com/article-la-satisfaction-du-client-un-enjeu-majeur-pour-le-medef-64615427.html>

Qualité de la mesure, consulté le 6 juillet 2013 à l'adresse <http://www.irdp.ch/edumetrie/lexique/echelle.htm>

Echelles multiples, consulté le 07 juillet 2013 à l'adresse <http://www.jybaudot.fr/Sondages/echeloncomp.html>

Caractéristiques d'une échelle de mesure, consulté le 07 juillet 2013 à l'adresse <http://www.market-audit.com/Echelles.html>

Les différentiels sémantiques, consulté le 12 juillet 2013 à l'adresse <http://www.mercator-publicitor.fr/lexique-marketing-definition-differentiel-semantique-echelle-osgood>

Les échelles de mesure, consulté le 13 juillet 2013 à l'adresse <http://www.marketing-strategique.com/Echelles.htm>

Avantages et inconvénients des échelles de mesure, consulté le 15 juillet 2013 à l'adresse <http://www.init-marketing.fr/chelle/introduction-46/>

Avantages et inconvénients des échelles de mesure, consulté le 15 juillet 2013 à l'adresse http://ressources.auneg.fr/nuxeo/site/esupversions/83e876d5-3c45-45cb-a888-2af03045ca8e/MarFonPda/co/L2_4_2_redaction_questionnaire.html

Le paradigme de Churchill, consulté le 17 juillet 2013 à l'adresse http://theses.univ-lyon3.fr/documents/getpart.php?id=lyon3.2011.othmane_j&part=308843

Calculs statistiques en fonction des échelles de mesure, consulté le 19 juillet à l'adresse <http://help.surveymonkey.com/>

Adapter les échelles en fonction des types de recueil, consulté le 19 juillet à l'adresse <http://www.init-marketing.fr/chelle/chelle-semantique/>

Nouvelles pistes de mesure, consulté le 19 juillet 2013 à l'adresse

ANNEXES

Annexe n°1

Principales échelles de mesure

Echelle dichotomique :

La livraison du colis vous a-t-elle semblé rapide ?

- Oui**
 Non

Echelle sémantique différentielle (Echelle d'Osgood) :

Comment jugeriez-vous l'envoi du colis sur les critères suivants ?

Lent ___ / ___ / ___ / ___ / ___ / ___ / ___ **Rapide**
Cher ___ / ___ / ___ / ___ / ___ / ___ / ___ **Peu cher**

Echelle à support sémantique :

Comment jugeriez-vous la durée de réception du colis ?

- Très lente** **Plutôt lente** **Moyenne** **Rapide** **Très rapide**

Echelle de Stapel :

Comment jugeriez-vous l'envoi du colis sur les critères suivants ?

+5	+5
+4	+4
+3	+3
+2	+2
+1	+1
Rapidité	Prix
-1	-1
-2	-2
-3	-3
-4	-4
-5	-5

Echelle de Likert :

Indiquez votre degré d'accord ou de désaccord avec le jugement suivant ?

	Pas du tout d'accord		Ni d'accord, ni pas d'accord			Tout à fait d'accord	
A : La livraison a été rapide	1	2	3	4	5	6	7
B : Le prix est abordable	1	2	3	4	5	6	7

Echelle à icônes :

La livraison du colis vous a-t-elle semblé rapide ?



Annexe n°2

Aperçu des questions telle qu'elles ont été posées en ligne dans l'omnibus IFOP

Changeons de sujet.
Dans quelle mesure êtes-vous d'accord avec l'affirmation suivante ?

Une réponse par ligne

	Tout à fait d'accord	Plutôt d'accord	Plutôt pas d'accord	Pas du tout d'accord
L'état français doit intervenir dans la fixation du prix de l'électricité	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

0% 100%

Vos réponses sont enregistrées au fur et à mesure que vous remplissez le questionnaire. Merci de ne pas utiliser le bouton "précédent" de votre navigateur.

Si vous rencontrez un problème technique, merci de contacter : [Support](#)

©Ifop 2013

Aperçu de la question 1 sur une échelle en quatre modalités.

Changeons de sujet.
Dans quelle mesure êtes-vous d'accord avec l'affirmation suivante ?

Une réponse par ligne

	Tout à fait d'accord	Plutôt d'accord	Plutôt pas d'accord	Pas du tout d'accord	Vous n'avez pas d'opinion, vous ne savez pas
L'exploitation du gaz de schiste doit être autorisée en France	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

0% 100%

Vos réponses sont enregistrées au fur et à mesure que vous remplissez le questionnaire. Merci de ne pas utiliser le bouton "précédent" de votre navigateur.

Si vous rencontrez un problème technique, merci de contacter : [Support](#)

©Ifop 2013

Aperçu de la question 2 sur une échelle en cinq modalités.

Annexe n°3-a

Résultats de l'étude selon le sexe – Q1

Q1 - L'état français doit intervenir dans la fixation du prix de l'électricité.

	Homme		Femme		Total	
	N	% cit.	N	% cit.	N	% cit.
Tout à fait d'accord	52	40%	42	35%	94	38%
Plutôt d'accord	64	50%	70	58%	134	54%
Plutôt pas d'accord	7	5%	6	5%	13	5%
Pas du tout d'accord	6	5%	3	2%	9	4%
Total	129	100%	121	100%	250	100%

	Homme		Femme		Total	
	N	% cit.	N	% cit.	N	% cit.
Tout à fait d'accord	55	46%	48	37%	103	41%
Plutôt d'accord	33	28%	53	41%	86	34%
Ni d'accord, ni pas d'accord	20	17%	22	17%	42	17%
Plutôt pas d'accord	5	4%	5	4%	10	4%
Pas du tout d'accord	7	6%	2	2%	9	4%
Total	120	100%	130	100%	250	100%

	Homme		Femme		Total	
	N	% cit.	N	% cit.	N	% cit.
Tout à fait d'accord	49	40%	42	32%	91	36%
Plutôt d'accord	54	45%	66	50%	120	47%
Plutôt pas d'accord	6	5%	16	12%	22	9%
Pas du tout d'accord	3	2%	3	2%	6	2%
Vous n'avez pas d'opinion	9	7%	6	5%	15	6%
Total	121	100%	133	100%	254	100%

	Homme		Femme		Total	
	N	% cit.	N	% cit.	N	% cit.
Tout à fait d'accord	54	43%	46	36%	100	40%
Plutôt d'accord	42	34%	44	34%	86	34%
Ni d'accord, ni pas d'accord	19	15%	22	17%	41	16%
Plutôt pas d'accord	4	3%	9	7%	13	5%
Pas du tout d'accord	3	2%	1	<1%	4	2%
Vous n'avez pas d'opinion	3	2%	6	5%	9	4%
Total	125	100%	128	100%	253	100%

Annexe n°3-b*Résultats de l'étude selon le sexe – Q2***Q2 - L'exploitation du gaz de schiste doit être autorisée en France.**

	Homme		Femme		Total	
	N	% cit.	N	% cit.	N	% cit.
Tout à fait d'accord	21	16%	12	10%	33	13%
Plutôt d'accord	43	33%	33	27%	76	30%
Plutôt pas d'accord	23	18%	38	31%	61	24%
Pas du tout d'accord	42	33%	38	31%	80	32%
Total	129	100%	121	100%	250	100%

	Homme		Femme		Total	
	N	% cit.	N	% cit.	N	% cit.
Tout à fait d'accord	24	↑ 20%	11	↓ 8%	35	14%
Plutôt d'accord	23	↑ 19%	16	↓ 12%	39	16%
Ni d'accord, ni pas d'accord	33	28%	46	35%	79	32%
Plutôt pas d'accord	12	10%	21	16%	33	13%
Pas du tout d'accord	28	23%	36	28%	64	26%
Total	120	100%	130	100%	250	100%

	Homme		Femme		Total	
	N	% cit.	N	% cit.	N	% cit.
Tout à fait d'accord	17	14%	8	6%	25	10%
Plutôt d'accord	29	24%	25	19%	54	21%
Plutôt pas d'accord	20	17%	24	18%	44	17%
Pas du tout d'accord	42	35%	37	28%	79	31%
Vous n'avez pas d'opinion	13	↓ 11%	39	↑ 29%	52	20%
Total	121	100%	133	100%	254	100%

	Homme		Femme		Total	
	N	% cit.	N	% cit.	N	% cit.
Tout à fait d'accord	14	↑ 11%	5	↓ 4%	19	8%
Plutôt d'accord	21	17%	16	13%	37	15%
Ni d'accord, ni pas d'accord	27	22%	38	30%	65	26%
Plutôt pas d'accord	13	10%	17	13%	30	12%
Pas du tout d'accord	44	35%	32	25%	76	30%
Vous n'avez pas d'opinion	6	↓ 5%	20	↑ 16%	26	10%
Total	125	100%	128	100%	253	100%