



UNIVERSITE DE LISALA

**CENTRE INTERUNIVERSITAIRE DE RECHERCHE  
PLURIDISCIPLINAIRE (CIREP)  
STATUT : UNIVERSITE PUBLIQUE  
Web : [www.cirep.ac.cd](http://www.cirep.ac.cd)  
Email : [info@cirep.ac.cd](mailto:info@cirep.ac.cd)**

# **NOTES DE COURS DE METHODOLOGIE DE RECHERCHE SCIENTIFIQUE EN PSYCHOLOGIE**

## **OBJECTIF GÉNÉRAL**

L'objectif général de ce cours est de familiariser les étudiants avec les différentes étapes du processus de recherche en psychologie, de la formulation d'une question de recherche à la communication des résultats.

## **OBJECTIFS SPÉCIFIQUES**

- ✓ Comprendre les principes fondamentaux de la recherche scientifique en psychologie.
- ✓ Apprendre à formuler une question de recherche pertinente et précise.
- ✓ Maîtriser les différentes méthodes de collecte et d'analyse des données en psychologie.
- ✓ Savoir interpréter les résultats et les présenter de manière claire et rigoureuse.
- ✓ Acquérir des compétences en rédaction scientifique et en communication orale.
- ✓ Développer un esprit critique vis-à-vis des recherches existantes en psychologie.

## **OBJECTIFS OPÉRATIONNELS**

- Réaliser une revue de la littérature sur un sujet donné en psychologie.
- Concevoir et mettre en œuvre une étude empirique en respectant les normes éthiques.
- Analyser les données collectées à l'aide de logiciels statistiques appropriés.
- Rédiger un rapport de recherche structuré et argumenté.
- Présenter les résultats de manière claire et convaincante lors d'une soutenance.

## **INTRODUCTION**

La recherche scientifique est généralement un processus qui demande beaucoup de lucidité d'esprit afin de parvenir à l'objectif recherché. « La recherche scientifique est un processus systématiquement et intentionnellement orienté et ajusté en vue d'innover ou d'augmenter la connaissance dans un domaine donné ». Mais lorsqu'il s'agit de la recherche scientifique dans les sciences sociales, le chercheur s'implique, vit et ressent la situation, il devient une partie de l'action : « faire émerger un projet de recherche n'est pas chose aisée lorsqu'on est soi-même partie prenante de l'action. Quelle perte de temps et combien d'hésitations et de découragements lorsqu'il s'agit de concevoir un projet, c'est-à-dire d'identifier précisément ce que l'on cherche ».

La recherche scientifique permet au chercheur dans le domaine de la psychologie de découvrir les relations entre les facteurs ainsi que les relations entre les attitudes, les opinions et les comportements qui les sous-tendent. Elle doit obéir à une démarche strictement méthodique. Cette démarche est l'articulation d'un ensemble d'activités successives organisées et structurées dans une pensée logique répondant à des règles objectives et précises. La recherche scientifique impose au chercheur une certaine objectivité de façon à pouvoir contrôler ses représentations, ses opinions et ses affects de manière à rester neutre. Elle est rigoureusement appelée à apporter des réponses vérifiables et palpables en vue de contribuer à résoudre un problème ou à enrichir un domaine d'un savoir donné. Dans cet esprit, l'outil permettant la collecte des données pour parvenir à cette fin est capital. Cet ouvrage n'est pas un manuel de méthodologie générale, il est consacré essentiellement à la construction des outils de recherche en psychologie. En étant illustré par des exemples de divers domaines de la psychologie, rédigé dans un style simple et renforcé par des exemples simplifiés loin de toute complication, cet ouvrage répond aux besoins des étudiants de tous les niveaux.

## **La pertinence du choix des outils**

### **1.1 – Pourquoi des outils ?**

Les outils de recherche sont indispensables à la collecte des informations de manière objective et systématique. La psychologie n'est une science que grâce à la fiabilité de ses instruments de mesure.

### **1.2 – Choix des outils**

Les outils de recherche ne sont pas utilisés de manière aléatoire, car le choix doit répondre exactement à ce que l'on veut mesurer. Avant la rédaction d'un questionnaire, la préparation d'un entretien, d'une observation ou d'une expérience, on doit au préalable définir les hypothèses générales et les objectifs afin de pouvoir cerner les différents aspects du problème que nous cherchons à comprendre ou à résoudre. L'utilisation d'un outil ou d'un autre est fonction de plusieurs facteurs à savoir :

#### ***1.2.1 – La méthode***

Le choix de l'outil en psychologie dépend en grande partie de la méthode utilisée, car la méthode qualitative et la méthode quantitative n'utilisent pas forcément le même outil.

#### ***1.2.2 – la nature de l'étude***

La nature de l'étude impose nécessairement le choix de l'outil propice.

#### ***1.2.3 – l'objet de l'étude***

L'objet de l'étude « est une délimitation plus précise à partir de l'idée, et une délimitation de son "champs" avec un maximum de clarté ». (Mucchielli. R) Si on prenait l'exemple de l'absentéisme des cadres d'une entreprise et l'échec scolaire dans une école, nous allons constater que sans doute chaque objet va faire appel à un outil qui lui est convenable.

#### ***1.2.4 – Les hypothèses de travail***

Elles tendent à énoncer une relation entre des faits significatifs sous l'aspect d'une loi plus ou moins générale et aide à classer les faits observés dont l'interprétation permet la déduction d'hypothèses qui, une fois vérifiées, constituent un élément de la théorie. Les hypothèses tendent à trouver des solutions à différentes sortes de questions. Elles prennent naissance à partir d'une observation ou de faits opérés au cours d'une recherche. Elles peuvent être aussi le résultat d'une élaboration purement théorique à partir d'un

ensemble de connaissances.

Les hypothèses varient en fonction de leur étendue, elles peuvent parvenir à une généralisation après avoir fait l'objet d'un test ou d'une expérience sur un objectif donné. Il est donc tout à fait lucide de déclarer que l'hypothèse aide à préciser l'outil de recherche.

En sciences sociales, on distingue trois types d'hypothèses suivant leur niveau d'abstraction : Des hypothèses qui supposent une existence entre deux variables. On les utilise pour quantifier des comportements : A titre d'exemple, chercher un rapport entre le taux de divorce et le revenu des ménages. Des hypothèses prévoyant l'existence de liens raisonnables entre deux variables : Tel est le cas de comportements dus aux aspects culturels.

Des hypothèses concernant des relations entre variables analytiques. Elles stipulent la formulation de relations entre variables complexes ou entre variables dépendantes, indépendantes et intermédiaires. Prenons l'exemple d'une étude sur l'influence de la classe sociale du niveau de croyance sur le taux de fécondité.

Pour être valables, les hypothèses doivent être utilisées sous certaines conditions :

L'hypothèse doit être vérifiable. Elle doit mettre en œuvre des faits réels et ne pas comporter de jugements de valeur.

Enfin, elle doit se rattacher à une théorie existante et être en harmonie avec l'actualité scientifique. Toute hypothèse non explicite ne peut que rendre l'objectif incorrect et n'aboutira qu'à des fins scientifiques erronées.

#### **1.2.5 – La recherche bibliographique**

La recherche bibliographique apporte une aide précieuse au chercheur pour dégager des pistes de tâtonnements afin d'arrêter le choix des outils.

#### **1.2.6 – La population-cible**

La population de l'étude a un rôle primordial dans le choix des méthodes et des outils. Si on est appelé à étudier le même comportement chez des enfants et des adultes, on peut recourir à l'observation chez les enfants et à l'entretien chez les adultes.

#### **Recherche en sciences sociales ?**

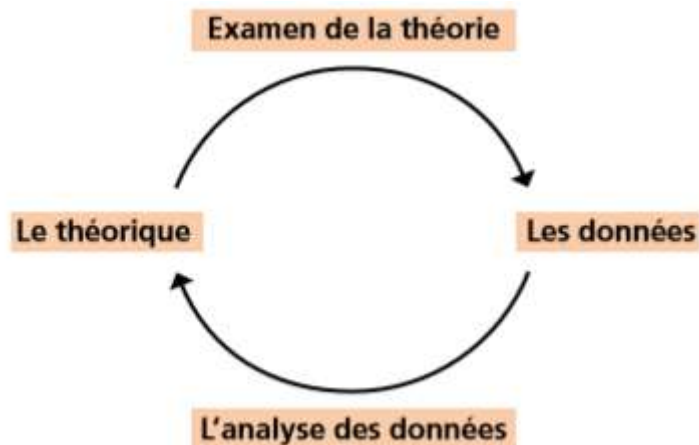
La recherche scientifique académique repose sur le recours

systematique à des méthodes et procédures spécifiques pour obtenir des informations ou pour révéler les relations entre les variables de la société. La recherche scientifique vise à mettre en lumière de nouvelles informations ou à en vérifier d'anciennes afin d'augmenter ou de vérifier les connaissances. Par conséquent, la recherche scientifique repose sur l'examen d'hypothèses afin de comprendre ou d'analyser un phénomène donné au sein de la société.

La connaissance scientifique est une connaissance organisée, soumise à des règles et des principes méthodologiques, à laquelle nous ne pouvons accéder sans suivre ces principes et y adhérer. Par conséquent, nous constatons que les connaissances scientifiques diffèrent dans de nombreux cas des analyses et des idées qui prévalent dans la société car les travaux de recherche s'efforcent d'approfondir le sujet, de collecter des informations objectives et de les analyser en dehors des idées préconçues, des analyses étroites et des stéréotypes.

### **Qu'est-ce que la connaissance scientifique ?**

Les scientifiques disent souvent qu'ils ne parlent pas sans fondement mais s'appuient sur des faits certains. Ils entendent dire par là que les connaissances scientifiques qu'ils avancent ne viennent pas du néant, mais sont plutôt bâties sur la base de théories existantes et d'études antérieures. Les études scientifiques sélectionnent les théories en rassemblant et en analysant des données et des preuves, puis en reformulant les théories en fonction des résultats de l'analyse des nouvelles informations et données. Ainsi les connaissances s'accumulent et la science progresse. Certaines hypothèses sont donc exclues, tandis que d'autres sont examinées pour s'assurer de leur capacité à expliquer certains phénomènes de société.



En ce sens, la connaissance scientifique est un savoir fondé sur l'interprétation de la société qui s'appuie sur des preuves et des données matérielles et objectives dans un cadre théorique spécifique. Les études théoriques qui traitent des idées et des opinions sans les étayer par des preuves et des données ne sont pas considérées comme « scientifiques ».

Les études qui collectent des statistiques et des preuves sans les interpréter et analyser systématiquement les données ne sont pas non plus considérées comme des études « scientifiques » parce que l'interprétation et l'analyse scientifiques font défaut. On peut donc dire que la connaissance scientifique se situe au point de jonction de la théorie et de la pratique.

### **Fonctions et objectifs de la recherche scientifique**

La recherche scientifique peut comporter six fonctions ou objectifs principaux :

- **Diagnostic** : Le diagnostic est l'une des fonctions les plus importantes de toute recherche scientifique car la caractérisation de la population étudiée et le diagnostic du phénomène sur lequel se concentre la recherche est la première étape, voire la plus importante, de la recherche scientifique.
- **Exploration** : L'une des fonctions de la recherche scientifique est également de rechercher des informations, d'explorer des faits et de recueillir des preuves et des données.
- **Interprétation** : Basée sur le diagnostic et l'exploration, l'interprétation est le troisième objectif de la recherche scientifique. Déterminer un phénomène ou un modèle dans la société et rechercher

des informations suffisantes pour le comprendre amènent le chercheur/la chercheuse à fournir une interprétation ou une analyse précise du phénomène étudié.

- **Prévision** : La prédiction ou l'extrapolation est la cible de nombreuses études scientifiques qui surveillent l'évolution temporelle des phénomènes sociaux ou celles qui étudient les relations entre différents facteurs de la société et l'étendue de leur influence les uns sur les autres. Dans ces cas, la prévision scientifique est basée sur l'étude des modèles sociaux et l'observation stricte des phénomènes et des comportements dans la société.
- **Contrôle** : En raison de la nature du travail de recherche et de sa capacité à détecter des modèles sociaux ainsi qu'à prédire sur la base des données et des informations du groupe d'une manière scientifique précise, le contrôle, la maîtrise et la planification deviennent des fonctions de base de la recherche scientifique.
- **Archivage** : Enfin, la sixième fonction de la recherche scientifique est de créer une banque de données et des archives de données dont d'autres chercheurs peuvent bénéficier.

### **Caractéristiques des connaissances scientifiques**

Il existe six caractéristiques de la connaissance scientifique, à savoir :

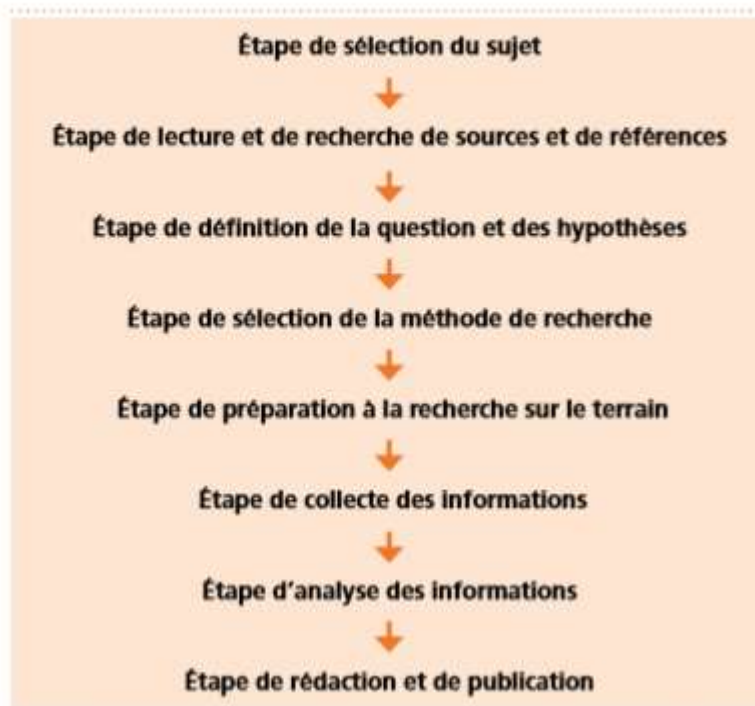
- **L'accumulation** : La recherche scientifique ne part pas de zéro mais bénéficie plutôt de ce qui a été précédemment publié dans son domaine. Elle apporte une alternative ou une nouveauté, ou encore prouve des connaissances antérieures. C'est ainsi que les connaissances scientifiques augmentent et s'accumulent avec chaque nouvelle recherche ajoutée à la littérature scientifique.
- **L'organisation** : La connaissance scientifique est une connaissance organisée et systématique qui peut être évaluée avec des preuves qui soient claires et spécifiques.
- **La causalité** : La causalité est une question complexe en sciences sociales, et il convient de vérifier qu'il existe une relation causale et pas seulement une corrélation entre deux variables.



- La **précision** : Précision dans le choix de la méthodologie de la recherche scientifique et dans l'utilisation des termes et concepts.
- **L'objectivité** : Le chercheur/la chercheuse doit être neutre, faire abstraction autant que possible de lui/elle-même, et étudier les faits et les données tels qu'ils sont en réalité.
- La **généralisation** : La généralisation n'est possible que si l'échantillon est représentatif. Ainsi, les résultats d'une recherche qualitative ne peuvent pas être généralisés car l'échantillon est souvent de petite taille et non représentatif.
- **Chacune** de ces caractéristiques sera examinée plus largement dans les sections suivantes.

### Étapes de la recherche scientifique

La recherche scientifique comprend huit étapes :



- Étape de sélection du sujet : La recherche scientifique commence par la sélection d'un sujet. Habituellement, ce choix résulte d'un intérêt personnel pour un sujet spécifique, ou d'un changement de société, ou encore de l'intérêt du bailleur de fonds pour un sujet spécifique, ou enfin de la disponibilité de nouvelles informations (comme Wikileaks). La motivation pour la recherche peut avoir une raison politique.

Comme le sujet de recherche est d'habitude général et vaste, il est nécessaire de définir un angle spécifique pour la recherche et de passer du sujet à la question de la recherche.

- Étape de définition de la question et des hypothèses : La question et les hypothèses de recherche sont déterminées sur la base de lectures préalables et d'une réflexion. La question de recherche doit être claire, précise et il doit être possible d'y répondre. Les hypothèses doivent être spécifiques, vérifiables et réversibles. Il existe quatre types de questions de recherche : la question descriptive, la question comparative, la question évolutive et la question théorique. Plus d'une question peut être incluse dans une recherche.

Types de questions de recherche		
Question descriptive	Qu'est ce qui s'est passé ?	Les résultats des filles aux examens en pays X sont-ils meilleurs que ceux des garçons ?
Question comparative	Est-ce que cela s'est passé ailleurs ?	Cela s'est-il produit dans d'autres pays ?
Question évolutive	Est-ce que cela s'est passé à travers le temps ?	Cela s'est-il produit avant ?
Question théorique	Pourquoi cela s'est-il passé ?	Pourquoi les résultats des filles aux examens sont meilleurs que ceux des garçons ?

- Étape de sélection de la méthode de recherche : Cette étape comprend le choix de la méthodologie qualitative ou quantitative en fonction de la question posée ainsi que la répartition des thèmes principaux et sous-thèmes sur des fondements et des critères clairs. Une structure est construite à ce stade également pour la recherche et la classification des informations (assignation de titres principaux, sous-titres et titres partiels : parties, sections, chapitres, branches, etc.).
- Étape de préparation à la recherche de terrain : Cette étape comprend la sélection de l'échantillon et la rédaction du questionnaire (recherche quantitative) ou la préparation des questions de recherche (recherche qualitative). C'est à ce moment qu'il convient de réfléchir à des plans alternatifs, travailler en réseau et contacter des personnes

qui peuvent faciliter la recherche et aider à atteindre l'échantillon souhaité.

- Étape de collecte d'informations : La recherche sur le terrain commence alors ; le chercheur/la chercheuse collecte des informations selon la méthodologie utilisée et ce dans le cadre de l'éthique de la recherche scientifique convenue.
- Étape d'analyse des informations : A l'issue de sa recherche sur le terrain, le chercheur/la chercheuse se retrouve au milieu d'une pléthore d'informations. Alors que doit-il/elle faire ? La première étape consiste à passer au crible les informations obtenues, en donnant la priorité aux sources originales, en scrutant les informations fiables de plus d'une source et en se concentrant sur les références les plus récentes, que ce soit pour leurs statistiques, leurs chiffres, ou leur documentation. Ce faisant il/elle devra écarter les informations qui ne sont pas directement liées au sujet de recherche afin de ne pas prendre de direction erronée et de gagner du temps et des efforts. Certains programmes peuvent être utilisés pour l'analyse scientifique. Ainsi :
  - Pour la recherche quantitative, on peut utiliser Excel, SPSS, STATA, Matlab
  - Pour la recherche qualitative, on peut utiliser Nvivo, MAXQDA, Atlas
- Étape de rédaction et de publication : Il existe certaines méthodes de formulation et d'édition des résultats de l'étude et certains moyens de diffusion des travaux de recherche scientifique.

### **Objectivité et subjectivité**

La subjectivité et l'objectivité ont été parmi les sujets qui ont suscité de vifs débats parmi les chercheurs/chercheuses au tournant du XXe siècle. Malgré le désir d'impartialité et d'objectivité totale dans la recherche scientifique, les sociologues admettent qu'il est très difficile d'éliminer complètement la subjectivité ; De ce fait, une distinction est faite entre l'objectivité scientifique et la recherche subjective.

L'objectivité scientifique consiste à restreindre le chercheur/la

chercheuse à la méthodologie scientifique de la recherche. Il doit collecter des données et les traiter telles quelles, et sans modification, en fonction des opinions ou penchants. En ce sens, l'objectivité consiste à ne pas inclure les opinions et les points de vue du chercheur/de la chercheuse dans le processus de recherche qui doit suivre les règles qui lui sont spécifiques. Quant à la subjectivité, c'est le facteur humain dans l'étude, car le chercheur/la chercheuse est un être humain et a des tendances personnelles, des croyances et des opinions qui peuvent être une motivation première pour choisir un sujet de recherche ou pour choisir un cadre théorique d'analyse et d'interprétation. Par conséquent, les chercheurs/chercheuses s'efforcent de réduire leur subjectivité et de faire preuve du plus d'objectivité possible pour aborder la question de recherche mais il reste toujours une marge de subjectivité dans le choix du sujet et la rédaction de l'étude.

Par conséquent, une distinction peut être faite entre l'objectivité de la méthode de recherche et l'objectivité de la rédaction de la recherche. L'objectivité de la méthode scientifique est l'une des caractéristiques les plus importantes et l'une des constantes fondamentales de tout travail de recherche sérieux qui aspire à présenter de nouvelles connaissances scientifiques. Quant à la subjectivité de l'écriture, elle se limite au style de l'écrivain.e, aux sujets sur lesquels il/elle choisit de se concentrer et à la manière d'interpréter les indices et les données. La subjectivité peut être contrôlée et réduite en clarifiant les limites et en en discutant dans le cadre de la méthodologie. La possibilité de subjectivité dans la rédaction de l'étude n'exclut pas que des limites et des méthodes spécifiques soient requises pour la rédaction de la recherche scientifique qui nécessite l'utilisation d'une terminologie précise et d'analyses systématiques, en écartant toute pensée et opinion qui ne sont basées sur des données et analyses scientifiques.

La subjectivité peut être réduite par la présence de plus d'un chercheur dans le projet de recherche ou en sollicitant l'aide de chercheurs/chercheuses et de collègues afin qu'ils lisent la recherche et fassent des observations sur la méthode de traitement du sujet et de rédaction de l'étude. Il est également important que le chercheur ou la

chercheuse se forme à toujours se poser la question inverse : pourquoi pas ? Les efforts fournis par le chercheur/la chercheuse pour aborder la question sous différents points de vue et de l'extraire du domaine privé, qui est affecté par ses opinions et ses penchants, peut grandement aider à réduire la subjectivité et à renforcer l'argumentation et l'analyse du chercheur/de la chercheuse.

### **Éthique de la recherche scientifique**

Le premier objectif de tout.e chercheur/chercheuse est d'obtenir des informations et des données. Or toutes les méthodes d'obtention d'informations ne sont pas légales et éthiques. L'éthique de la recherche scientifique exige le respect de la vie privée des participant.es à la recherche, la préservation de leurs droits et le respect de leurs opinions, ainsi que la préservation de la sécurité des participant.es et du chercheur/de la chercheuse, et ce à n'importe quel prix ! Il est vrai que l'éthique de la recherche scientifique limite souvent l'accès à l'information. Cependant, la recherche scientifique s'accorde aujourd'hui à donner la priorité au respect de l'éthique de la recherche, même au détriment de l'accès à l'information.

Cet aspect a émergé à la suite de l'étude publiée par Laud Humphreys (Humphreys 1970)<sup>1</sup>. Dans cette étude, Humphreys a pu étudier des aspects importants du comportement des hommes dans les toilettes publiques et la relation entre la sexualité et le genre aux Etats-Unis en observant les usagers des toilettes publiques en Louisiane et en collectant des informations importantes et nouvelles qu'il aurait été impossible de collecter par d'autres moyens. Cette étude a toutefois déclenché un grand débat dans les cercles académiques car le chercheur observait les gens à leur insu et sans leur consentement ; ce qui était considéré comme une violation claire et flagrante de leur vie privée.

Il y a débat concernant la dégradation de la qualité de la recherche lorsque les participant.es savent qu'ils sont surveillé.es car la plupart des comportements humains changent lorsqu'ils/ elles se rendent compte qu'ils/elles sont observé.es. Les informations que le chercheur/la chercheuse obtient ou « arrache » lorsque les participant.es ne savent pas qu'il/elle les observe sont beaucoup plus riches que les informations

recueillies « de manière éthique ». Cependant, il y a un coût moral considérable à ne pas informer et obtenir le consentement des personnes surveillées car il y a atteinte à leur intimité et à leur vie privée, ce qui peut leur nuire lorsque l'étude est publiée.

Par conséquent, afin de préserver le respect des droits des individus et groupes qui participent aux études scientifiques, nous ne pouvons aujourd'hui faire aucun travail de recherche qui ne respecte pas les règles de l'éthique de la recherche mentionnée ci-dessous :

- **Honnêteté** : Le chercheur/la chercheuse doit être honnête et clair.e avec les participant.es à la recherche. Il/elle doit leur expliquer la recherche et son objectif et leur donner accès aux informations de base qui constituent le fondement de sa recherche. Le chercheur/la chercheuse doit également être honnête avec ses lecteurs/ lectrices en transférant des informations de manière honnête et véridique sans falsifier aucune information ou compléter des informations partielles basées sur des théories antérieures ou ses opinions personnelles.
- **Anonymat** : L'une des conditions de base pour respecter l'éthique de la recherche scientifique est de protéger l'identité des participant.es à la recherche en ne donnant pas leur vrai nom ou en n'utilisant pas d'indices qui pourraient conduire à révéler leur véritable identité.
- **Confidentialité** : La question de la vie privée concerne la protection des données que le chercheur/la chercheuse a collectées pendant la période de recherche. Les données contiennent de nombreuses informations privées et précises. Le chercheur/la chercheuse doit garantir la confidentialité des informations et les conserver dans un endroit sûr de sorte que personne ne puisse y accéder ou y consulter les données. Habituellement, les informations sont détruites une fois la recherche terminée, surtout si elles contiennent des informations confidentielles et sensibles.
- **Confiance** : Le chercheur/la chercheuse doit essayer de construire une relation de confiance avec les participants à la recherche afin d'obtenir une plus grande coopération et des résultats plus précis et crédibles. Ainsi, lorsque le/la participant.e fait confiance au

chercheur/à la chercheuse, il/elle est généralement plus généreux/généreuse, franc/ franche et précis.e dans ses réponses et ses informations.

- **Consentement** : Le chercheur/la chercheuse doit toujours s'assurer d'obtenir le consentement des participant.es avant de commencer tout travail de recherche sur le terrain. Le consentement est généralement écrit en demandant au participant/à la participante de signer une déclaration de consentement pour participer à la recherche. Cette déclaration comprend une explication claire du but de l'étude et de ce qu'elle exigera du/de la participant.e. De plus, cette déclaration doit expliquer au/à la participant.e quels sont ses droits pendant et après l'étude. Le chercheur/la chercheuse ne doit, à aucun stade de la recherche, utiliser la méthode de l'intimidation pour obtenir des informations ou pour faire pression sur le/la participant.e pour qu'il/elle ne se retire pas de la recherche.
- **Retrait** : Le retrait est considéré comme l'un des droits les plus importants des participant. es à toute recherche scientifique. Le temps que le/la participant.e consacre à la recherche relève de sa décision personnelle, d'autant plus que la plupart des participant.es à l'étude sont généralement des volontaires. Le consentement à participer à la recherche ne garantit pas que le/la participant.e restera jusqu'à la fin et le retrait est un droit que le chercheur/la chercheuse doit respecter. Par conséquent, il est toujours conseillé au chercheur/à la chercheuse d'essayer d'atteindre le plus grand nombre possible d'individus dans l'échantillon en s'assurant que son étude couvre un groupe d'individus suffisant car il est attendu que certains des participant.es se retireront pendant la recherche.
- **Enregistrement audio ou photographie**: Le chercheur/la chercheuse n'a pas le droit d'enregistrer des voix ou de prendre des photos ou des vidéos sans que les participant.es ne soient informés et consentants. Il n'est pas correct que le chercheur/ la chercheuse demande le consentement du/de la participant.e après avoir terminé

l'enregistrement ou pris la photographie, car l'approbation doit toujours avoir lieu avant de commencer la recherche.

- **Tromperie ou faux espoirs:** Dans de nombreux cas, les participant.es à la recherche pensent que leur participation leur permettra d'améliorer leurs conditions de vie. Par

## **MÉTHODES DE RECHERCHE QUANTITATIVES, MÉTHODES DE RECHERCHE QUALITATIVES**

Les méthodes qualitatives comme les méthodes quantitatives s'inscrivent dans la longue histoire des sciences humaines et sociales dont les origines remontent à l'Antiquité. Au xiie siècle, on retrouve déjà la « querelle des universaux » qui impliquait fondamentalement un questionnement sur la réalité ontologique des concepts généraux et sur la différence entre expérience singulière et lois générales. En psychologie, l'ouvrage de Parot et Richelle retrace ces « querelles » et analyse leurs implications actuelles, en particulier, dans les différenciations entre méthodes qualitatives, que l'on peut rapprocher d'une recherche du singulier, et les méthodes quantitatives, davantage rattachées à la recherche de lois universelles. Durant le xixe siècle, la psychologie naissante se confronte aux questions de méthodologie et aux grandes orientations en vigueur: quantitative, expérimentale et qualitative ; ces dernières coexisteront jusqu'au début du xxe siècle. Cependant, en souhaitant légitimer la scientificité de la discipline, un certain nombre de travaux contribuent peu à peu à une rigidification des postures méthodologiques, de pair avec la progressive prédominance des méthodes quantitatives, s'affirmant comme les seules scientifiques. On assiste alors à un déclin apparent des usages des méthodes qualitatives, surtout à partir de la fin des années 1930. Avec la montée des recherches comportementalistes et, par la suite, cognitivistes, les méthodes reconnues comme « scientifiques » se réduisent pendant une certaine période aux méthodes expérimentales et quantitatives, en vigueur dans les sciences de la vie.



## Quelle est la méthodologie de la recherche scientifique?

### Déterminer la méthodologie de recherche

Il existe deux types de méthodes de recherche utilisées en sciences sociales : l'approche quantitative et l'approche qualitative. Le choix de la méthodologie dépend du type de question de la recherche et des moyens disponibles compte-tenu de la situation sécuritaire, de l'accès à l'échantillon ou de la présence d'une équipe de chercheurs/chercheuses préparé.es, etc. Le financement joue également un rôle fondamental dans la détermination de la méthodologie de recherche car le coût de la recherche varie selon la méthodologie utilisée et, généralement, le coût de la recherche quantitative est plus élevé que celui de la recherche qualitative.

### La recherche quantitative comparée à la recherche qualitative

La recherche qualitative comprend différents types de méthodologies, telles que l'enquête sociale ou statistique ou l'étude des réseaux sociaux. Quant à la recherche qualitative, elle peut inclure l'approche ethnographique (observation et entretiens), l'approche expérimentale ou la méthode d'étude de cas. Le tableau ci-dessous montre une comparaison entre la recherche quantitative et qualitative.

Recherche quantitative	Recherche qualitative
S'appuie sur la méthode d'observation, d'entretien et de références, loin des méthodes statistiques	Des normes et des tests ou des listes d'évaluation sont utilisés pour collecter des données et des informations avec des méthodes statistiques
Des informations riches et approfondies La collecte et la classification des informations prennent du temps	Informations générales, pas toujours exactes La collecte et la transcription des informations sont plus rapides et plus faciles

Le chercheur/la chercheuse ne contrôle aucune des variables	Le chercheur/la chercheuse exerce un contrôle complet sur toutes les variables liées à la variable dépendante
Biais : Biais du chercheur/de la chercheuse Biais de l'observateur/de l'observatrice	Biais : Biais de non-réponse Effet de la formulation
Un petit échantillon non représentatif Il ne peut pas être généralisé	Échantillon représentatif Peut être généralisé
Les rapports sont rédigés dans un style analytique et narratif	Les données sont utilisées dans la rédaction de rapports, de tableaux et de graphiques

### **Triangulation**

La triangulation dans la recherche scientifique signifie l'utilisation de plus d'une méthodologie (ou source ou chercheur/chercheuse) au cours du processus de recherche. La triangulation comprend toutes les informations que le chercheur/la chercheuse a collectées à partir de différentes sources ou méthodologies pour déterminer la cohérence des preuves entre les sources d'informations et des données.

Dans de nombreux cas, les données de la méthode quantitative sont insuffisantes à elles seules pour comprendre certains phénomènes sociaux ou certains de leurs aspects, tels que les attitudes, les opinions et les valeurs sociales. En effet, ces données n'offrent pas une compréhension approfondie mais alertent plutôt certain.es chercheurs/chercheuses sur la nécessité d'utiliser la méthode qualitative en plus de la méthode quantitative étant donné qu'elle est plus complète et plus approfondie que la vision globale ; ce qui contribue à la précision de l'analyse. Par conséquent, la combinaison de méthodes quantitatives et qualitatives est très utile pour une étude plus complète.

### **L'évaluation du comportement comparée à celle des attitudes**

Il existe une grande différence entre l'étude des situations et l'étude du comportement. Les attitudes et les opinions sont étudiées au moyen d'entretiens, de questionnaires et d'enquêtes (questions directes), tandis que le comportement est étudié par l'observation et la collecte de documents (étude indirecte). Par exemple, les schémas de manifestations dans la société peuvent être étudiés en observant des manifestations ou en compilant des

articles, des photos ou des vidéos prises lors de manifestations et en étudiant les comportements. L'étude du comportement évite au chercheur/à la chercheuse de se heurter à de nombreux problèmes méthodologiques liés à la validité des réponses dans les études mesurant les attitudes et les opinions. L'évaluation du comportement peut être une approche scientifique plus appropriée aux circonstances de guerre ou de conflit car elle n'inclut pas de relation directe entre le chercheur/la chercheuse et les membres de la communauté à étudier ; ce qui réduit les obstacles éthiques à la recherche dans le contexte des guerres et des conflits.

### **Utilisation de sources secondaires**

L'utilisation de sources secondaires peut être très importante, en particulier dans les études axées sur l'évaluation du comportement. Les sources secondaires comprennent les archives de journaux et de magazines, des documentaires ou des programmes d'information, des photos, des déclarations officielles, des rapports d'organisations locales ou internationales, etc. Les sources secondaires de recherche peuvent constituer une partie essentielle de la recherche en termes d'informations, ou bien elles peuvent être utilisées pour vérifier et valider les informations recueillies par d'autres moyens.

### **La démarche comparative**

Cette approche de recherche se concentre sur la comparaison des similitudes et des différences entre les phénomènes sociaux afin de découvrir les facteurs ou les conditions qui accompagnent l'émergence d'un phénomène social ou d'un modèle de comportement spécifique. La comparaison peut être:

- **Longitudinale:** On compare un phénomène dans plus d'une communauté/région au cours d'une même période.
- **Transversale:** On compare un phénomène dans une seule population sur une longue période de temps afin d'étudier le développement du phénomène et son évolution dans le temps.

L'analyse comparative se fait selon quatre cas de comparaison qui sont les suivants:

- Comparer une variable dans des sociétés similaires, comme étudier le niveau de scolarité de la classe ouvrière dans deux sociétés industrielles.
- Comparer plusieurs variables dans des sociétés similaires, comme l'étude de la situation économique (taux de revenus, taux de chômage, prix des matières premières, etc.) dans les pays qui ont connu les « révolutions arabes ».
- Etudier la relation entre plusieurs variables dans une société, comme l'étude de la relation entre taux de fécondité, classe sociale et zone géographique (urbaine et rurale) dans la société algérienne.
- Etudier la relation entre plusieurs variables dans des sociétés différentes, comme l'étude de la relation entre le genre et le revenu par habitant dans les sociétés du Moyen-Orient et d'Afrique du Nord en comparaison avec les sociétés scandinaves.

La recherche comparative se distingue des autres sur plusieurs aspects. Premièrement, la recherche comparative aide à accroître la capacité du chercheur/de la chercheuse à fournir des explications plus solides du phénomène étudié, car ces explications sont basées sur des preuves recueillies auprès de plusieurs sociétés ou sur une longue période ce qui réduit l'influence des facteurs de hasard et des biais culturels. Deuxièmement, la recherche comparative renforce la capacité du chercheur/de la chercheuse à élargir la gamme des variables étudiées qui sont incluses dans la conception de la recherche. Il/elle utilise à cet effet divers indicateurs tirés de plus d'une société, tels que les indicateurs qui sont utilisés pour mesurer le statut social et qui comprennent le revenu et la profession. Dans certains pays, ils incluent également le lieu de résidence et la lignée familiale. Troisièmement, la recherche comparative permet d'utiliser des facteurs et des aspects culturels et sociaux de chaque société ; ce qui renforce également le pouvoir des interprétations et augmente leur résilience face à la critique.

### **Étude de cas**

Contrairement à l'approche comparative, l'étude de cas constitue un type de recherche scientifique qui se concentre sur une communauté ou un

cas. Cette approche fournit au chercheur/à la chercheuse des données quantitatives et qualitatives sur plusieurs facteurs liés à des individus, des institutions ou des groupes sociaux dans des situations spécifiques. Ces données comprennent à la fois des aspects personnels et environnementaux ; ce qui permet au chercheur/à la chercheuse de mener une description approfondie et détaillée du cas sur lequel se concentre la recherche. Si le sujet d'étude se consacre aux institutions sociales, alors chaque institution sociale est considérée comme un cas, tandis que les individus deviennent de simples parties ou facteurs impliqués dans la formation du cas.

L'étude de cas est un examen approfondi d'une situation ou d'un cas spécifique. L'idée principale d'une étude de cas est d'étudier un cas de manière détaillée et approfondie, à travers toutes les méthodes appropriées et disponibles. Il peut y avoir une variété de questions pour les études de cas, mais l'objectif général reste de parvenir à une compréhension la plus complète possible du cas social analysé.

Pour parvenir à étudier les modèles sociaux et leur évolution historique, la méthode de l'étude de cas peut se consacrer à une étape spécifique de l'histoire de la société ou à toutes les étapes que la société a traversées. Les outils de cette méthode comprennent l'entretien personnel, l'étude des documents et registres officiels, des notes personnelles, l'analyse des statistiques et des données quantitatives disponibles, etc.

### **Échantillonnage**

L'examen de l'échantillon est une partie essentielle et très importante de la recherche scientifique, eu égard à la difficulté d'atteindre tous les membres de la société (en raison du coût élevé et du temps requis). Le chercheur la chercheuse doit choisir un échantillon ou un sous-groupe de la population étudiée qui l'aide à comprendre les schémas et la dynamique de la population concernée. Par conséquent, un échantillon peut être défini comme un segment ou une partie de la communauté étudiée qui porte les caractéristiques de cette communauté et qui la représente en ce qui concerne l'objet de la recherche. Ainsi, les définitions suivantes peuvent être fournies:

- **Société étudiée:** Tous les individus, événements ou observations qui constituent le sujet de la recherche. Par exemple: les élèves des écoles en RDC.
- **Échantillon:** Un groupe partiel de la population étudiée. Par exemple: 200 élèves de 10 écoles en Algérie répartis dans différentes régions.
- **Unité d'échantillonnage:** L'un des individus ou observations choisis dans l'échantillon. Par exemple: un élève d'une école de la wilaya d'Oran.

La sélection précise et appropriée de l'échantillon donne des résultats largement similaires à ceux que l'on peut obtenir lors de l'étude de la population étudiée. Une erreur dans la sélection d'un échantillon peut conduire à des résultats de recherche inexacts ou complètement faux.

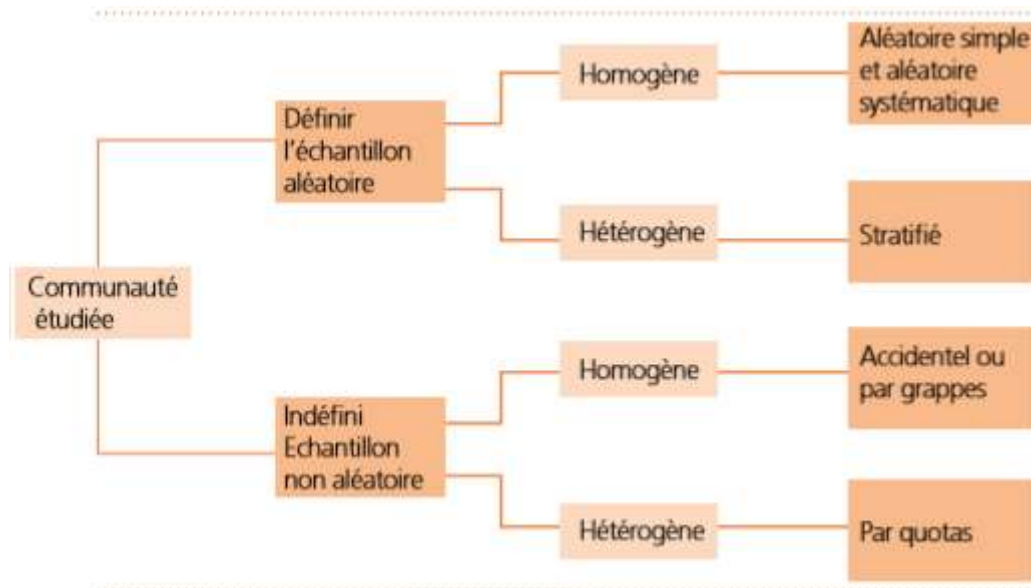
### **Types d'échantillons**

Il existe deux principaux types d'échantillons<sup>3</sup> : les échantillons aléatoires et les échantillons non aléatoires.

Le type d'échantillon est choisi en fonction de la capacité à déterminer la population de l'étude. Si la population étudiée est définie et connue, le chercheur/la chercheuse doit prélever un échantillon aléatoire pour qu'il soit représentatif. Par exemple, si le chercheur/ la chercheuse mène une étude sur les élèves des écoles publiques, la population étudiée sera définie car il peut obtenir une liste de tous les élèves inscrits à l'école publique. Par conséquent, le chercheur/la chercheuse doit, dans ce cas, prélever un échantillon aléatoire

Si la population étudiée n'est pas spécifiée, le chercheur/la chercheuse doit choisir un échantillon non aléatoire. Cela se produit dans les cas où il est difficile de connaître le nombre total des membres de la communauté qu'il/elle veut étudier. Par exemple, si le chercheur/ la chercheuse s'intéresse à l'étude du phénomène de l'abus de drogues, il est difficile de déterminer le nombre de consommateurs de drogues parce que ces listes n'existent pas et que peu de gens déclarent consommer de la drogue (dans les hôpitaux, par exemple), en particulier dans les pays qui criminalisent la consommation de drogues. Par conséquent, le chercheur/la chercheuse ne peut pas choisir un échantillon aléatoire parce que la population étudiée

n'est pas définie et l'accès à celui-ci n'est pas facile. Dans un tel cas, le chercheur/la chercheuse a recours au choix d'un échantillon non aléatoire.



- **Échantillonnage aléatoire** : Les échantillons aléatoires sont ceux dans lesquels il y a une chance égale pour chaque élément de la population étudiée de faire partie des unités de l'échantillon. Un échantillon aléatoire est choisi lorsque la population étudiée est définie et connue et la méthode de sélection est non sélective car les items sont choisis au hasard en fonction du type d'échantillon, en tenant compte de l'homogénéité et de l'hétérogénéité dans la communauté.

Si la population étudiée est définie et homogène, un échantillon aléatoire simple peut être prélevé (en établissant des tableaux de chiffres pour tous les membres de la population étudiée et en tirant le nombre d'échantillons requis par lot aléatoire). Un échantillon aléatoire systématique peut aussi être prélevé (en prenant, par exemple, un élément sur cinq dans le tableau des numéros de série de la population étudiée).

Dans le cas où la population étudiée est hétérogène, un échantillon aléatoire stratifié doit être prélevé. Dans ce type d'échantillon, la population étudiée peut être divisée en groupes ou strates en fonction de cette variance. Par exemple, lorsque vous étudiez le travail bénévole parmi les étudiant.es universitaires, on constate qu'il vaut mieux diviser les étudiant.es en strates par année scolaire, ou selon le genre (hommes ou femmes), etc. Lorsque cela est possible, l'échantillon aléatoire simple est le meilleur type d'échantillons

car il donne la meilleure représentation de la population étudiée. Les résultats de la recherche sur la base d'échantillons aléatoires peuvent être généralisés pour parler de la population étudiée.

- **Échantillonnage non aléatoire** : L'échantillonnage non aléatoire est utilisé lorsqu'il est difficile d'accéder à un échantillon aléatoire parce que la population de recherche n'est ni déterminée ni définie. Par exemple, nous ne connaissons pas le nombre exact et les caractéristiques de la population des combattant.es de l'Organisation État Islamique en Irak et au Levant (EIIL). Nous ne pouvons donc pas prendre un échantillon aléatoire si nous voulons les étudier. Nous pouvons donc prélever un échantillon accidentel (de commodité) ou un échantillon par grappes si nous considérons que la population EIIL est homogène, ou bien un échantillon de quota si nous décidons de diviser les combattant.es de l'EIIL en Arabes et en étrangers. Les échantillons non aléatoires sont caractérisés par le fait qu'ils ne donnent pas la même possibilité à tous les membres de la population étudiée d'apparaître dans l'échantillon. Par conséquent, nous ne pouvons pas généraliser sur la base d'échantillons non aléatoires.

**Choisissez un type d'échantillonnage pour chacune de ces recherches :**

1. Étude sur les combattant.es en Syrie	A. Aléatoire simple
2. Étude sur les travailleurs/travailleuses du textile enregistré.es dans un syndicat en Egypte	B. Aléatoire systématique
3. Étude sur les étudiant.es universitaires au Liban	C. Stratifié
4. Étude sur les étudiant.es de l'Université libanaise	D. Accidentel ou par grappes
5. Étude sur l'homosexualité en Jordanie (comparaison hommes/femmes)	E. Par quotas

1. (D ou E) ; 2. (A ou B) ; 3. (C) ; 4. (A ou B) ; 5. (E)

### **Étapes d'examen de l'échantillon**

Le processus d'examen des échantillons passe par plusieurs étapes :

1. Définir clairement et précisément la population de l'étude en fonction des caractéristiques qui distinguent ses membres car



cela affecte le nombre de membres de l'échantillon et la qualité de l'échantillon que le chercheur/la chercheuse doit choisir ;

2. Déterminer les membres de la population d'origine à étudier. Le chercheur/la chercheuse doit si possible les organiser en tableaux (selon les numéros de série) car cela facilite la sélection d'un échantillon représentatif de la communauté.
3. Déterminer les variables de l'étude afin de contrôler le plus grand nombre possible de variables non étudiées.

### **Déterminer le nombre approprié d'échantillons**

Le nombre approprié d'individus dans l'échantillon est déterminé en fonction de plusieurs critères:

- Homogénéité ou hétérogénéité de la population: Plus l'homogénéité entre les membres de la population est grande, plus le nombre nécessaire à la représentation est petit, et vice versa.
- La méthode de recherche utilisée : Les études d'enquête nécessitent le plus grand nombre possible de membres de la communauté, tandis que les études expérimentales ou qualitatives dépendent d'un plus petit nombre d'individus.
- Le degré de précision requis : Si l'étude nécessite un haut degré de précision, par exemple s'il y a des décisions à prendre en fonction des résultats de la recherche.

Nous aurons besoin d'un plus grand nombre d'individus de l'échantillon pour garantir la confiance nécessaire pour généraliser les résultats.

Il est très important d'attirer l'attention sur le fait que le nombre d'individus dans l'échantillon n'est pas toujours le facteur le plus important dans la sélection des échantillons. Malgré l'importance du nombre, la représentation de l'échantillon par rapport à la population étudiée demeure le facteur le plus important pour déterminer l'échantillon approprié. Par exemple, vous constaterez que la plupart des sondages électoraux dépendent de la taille d'un échantillon de mille personnes, quel que soit le nombre d'électeurs.

### **Qu'est-ce que la recherche qualitative?**

Caractéristiques de la recherche qualitative

Outre le fait de collecter et d'analyser la documentation, la recherche qualitative s'appuie sur des entretiens ou des observations sur le terrain naturel de la vie sociale quotidienne.

Parmi les caractéristiques de la recherche qualitative figurent les éléments suivants :

- La recherche qualitative accroît notre compréhension de tout phénomène social dont nous ne savons guère de choses ou corrige notre connaissance de phénomènes sociaux répandus mais non expliqués avec précision.
- La recherche qualitative permet d'accéder à des informations approfondies difficiles à exprimer de manière quantitative ou statistique.
- La recherche qualitative se caractérise par une flexibilité et une ouverture aux variables car les entretiens et les observations ne sont pas codifiés ou standardisés pour tous les cas.
- Dans la recherche qualitative, le chercheur n'a pas de contrôle préalable sur le domaine de la recherche et les méthodes de collecte d'informations.

### **Stratégie de recherche qualitative: une étude de cas**

L'étude de cas est l'une des stratégies les plus importantes de la recherche qualitative. Cette stratégie est basée sur l'exploration et l'examen d'une situation spécifique ou d'un cas particulier.

L'idée de base dans une étude de cas est d'étudier un cas (ou peut-être plusieurs cas) d'une manière détaillée et précise, en utilisant tous les moyens appropriés et disponibles.

Il peut y avoir une variété d'objectifs ou de questions pour l'étude de cas mais le but général est d'atteindre la compréhension la plus complète et la plus globale possible de ce cas.

L'étude de cas se distingue des autres types de recherche en ce qu'elle approfondit la compréhension d'un cas particulier dans son état et contexte naturels, sans se préoccuper de généraliser les résultats à d'autres cas. Les avantages de l'étude de cas sont qu'elle met la lumière sur un sujet spécifique en approfondissant un cas (ou plusieurs cas spécifiques). et se

spécialise subtilement sur la question. Les études de cas comportent une richesse importante de connaissances scientifiques car elles analysent les détails de cas et éclairent les interactions sociales et les dynamiques qui ne peuvent être saisies par la plupart des autres stratégies de recherche.

### **Méthodes de collecte d'informations dans la recherche qualitative**

Il existe quatre méthodes de base pour collecter des informations dans la recherche scientifique qualitative: entretien individuel, entretien de groupe, observation et analyse de documents.

**Entretien individuel** : Les entretiens individuels sont généralement utilisés pour des sujets sensibles, des expériences personnelles ou pour approfondir la compréhension des opinions et des positions des individus dans la société car à travers l'entretien, le chercheur/ la chercheuse, peut identifier les idées, les sentiments et les points de vue des autres. De plus, il/elle peut reconstruire des événements sociaux à travers les réponses obtenues lors d'entretiens individuels. Cette méthode repose sur l'établissement de la confiance entre le chercheur/la chercheuse et le/la participant.e à la recherche afin d'assurer la fiabilité et l'exactitude des réponses.

#### **Il existe deux types d'entretiens :**

- ❖ Un entretien structuré dans lequel on pose une série de questions préparées à l'avance au participant/à la participante. Tous les participant.es reçoivent les mêmes questions, dans le même ordre et de la même manière. Le rôle du chercheur/de la chercheuse est neutre et de par sa nature, ce type d'entretien se concentre sur des questions rationnelles et non émotionnelles. Les réponses peuvent être des modèles prédéfinis (avec peu de diversité de et des réponses variées). Des questions ouvertes peuvent également être utilisées pour permettre au participant/à la participante de répondre comme il/elle le souhaite, sans restriction ni protocole.
- ❖ Un entretien semi-structuré est un entretien non codifié avec des questions ouvertes et approfondies. Le rôle du chercheur/la chercheuse est alors plus proche de celui d'un.e régulateur/régulatrice de dialogue que d'un équivalent. Ce type d'entretien permet au chercheur/à la chercheuse de comprendre la pensée et le

comportement du/de la participant.e sans abandonner les idées préconçues ou les classifications du chercheur/de la chercheuse, ce qui peut limiter les déclarations et l'interaction du/de la participante.

**Groupe de discussion (focus group):** Lorsque le chercheur/la chercheuse travaille lors de l'entrevue avec un groupe de participant.es en même temps, il s'agit d'un groupe de discussion. Dans ce type d'entretien, le rôle du chercheur/de la chercheuse est de gérer et de faciliter le dialogue, et sa tâche est d'enregistrer l'interaction qui a lieu entre les participant.es. Des compétences pour gérer le dialogue et orienter la discussion dans la direction souhaitée sont alors requises.

Le groupe de discussion peut faire ressortir des aspects du cas étudié qui peuvent ne pas apparaître dans d'autres types d'entretiens, et ce, grâce à ce qui résulte de l'interaction entre les opinions des participantes.

Le groupe de discussion peut faire ressortir des aspects du cas étudié qui peuvent ne pas apparaître dans d'autres types d'entretiens, et ce, grâce à ce qui résulte de l'interaction entre les opinions des participant.es.

**Observation :** Dans la méthode d'observation, l'étendue de la participation du chercheur/de la chercheuse avec les membres de l'étude varie selon le type et la nature de la recherche. L'étude peut adopter l'observation participative, de sorte que le chercheur/ la chercheuse rejoint le groupe cible et partage sa vie quotidienne. Le chercheur/la chercheuse peut aussi s'éloigner des participant.es et agir en spectateur/spectatrice pendant la période de collecte d'informations.

L'observation peut être quantitative et structurée, ou qualitative et non structurée.

- ✓ Lors de l'observation quantitative, le chercheur/la chercheuse observe et recueille des informations numériques grâce à des outils préparés à l'avance (par exemple, le nombre d'élèves qui interagissent en classe, le nombre de personnes déplacées franchissant les frontières, le calcul du temps nécessaire pour terminer le passage de la frontière, etc.).
- ✓ Quant à l'observation qualitative, elle est moins structurée car le chercheur/la chercheuse n'utilise pas de classifications et de modèles prédéterminés mais enregistre plutôt ses observations

automatiquement et ouvertement, de sorte qu'il/elle enregistre la réalité telle qu'elle se produit. L'idée de base ici est que la classification et la description à laquelle les informations résultant de l'observation sont soumises apparaîtront après la collecte et l'analyse des informations, plutôt que de figurer dans l'information pendant le processus d'observation.

**Analyse du contenu** : La collecte et l'analyse de documents est l'un des principaux moyens d'accéder à l'information dans la recherche qualitative car les documents historiques ou contemporains peuvent être une source importante de recherche. L'importance de cette méthode de recherche s'accroît car la documentation est devenue l'une des caractéristiques fondamentales des sociétés modernes. Grâce aux documents existants, le chercheur/la chercheuse peut étudier et analyser une réalité et obtenir des résultats importants et utiles sans avoir besoin de recourir à des recherches sur le terrain.

### **Rédaction des questionnaires d'entretien ou d'enquête**

Le chercheur/la chercheuse doit suivre des règles et des protocoles précis lors de la rédaction des questionnaires d'entretien ou d'enquête, notamment:

- La clarté dans la formulation de la question ;
- La concision : toute question ne doit pas dépasser une ligne et il ne doit pas y avoir d'introduction à la question car cela pourrait affecter la réponse du/de la participant.e ;
- La présence d'une variable dans chaque question : chaque question doit aborder un aspect (variable) du sujet pour maintenir l'exactitude. Chaque question doit faire l'objet d'une seule interrogation.
- Habituellement, le questionnaire commence par des questions générales avant d'entrer dans les détails et les questions directement liés à la question de recherche ainsi qu'aux hypothèses.
- Lorsque le chercheur/la chercheuse commence à rédiger le questionnaire ou les questions de l'entretien, il/elle doit s'efforcer de réduire les risques d'influence sur la réponse des répondants. Par conséquent, l'étape de rédaction des questions est considérée comme l'une des étapes les plus sensibles et les plus difficiles.

Il existe trois sources principales qui peuvent influencer les réponses:

**L'effet de la formulation des questions et le choix du vocabulaire:** Il ne fait aucun doute que la façon dont la question est formulée a un impact sur la réponse. Le vocabulaire que nous choisissons pour la question peut pousser le/la participant.e à chercher la réponse dans une direction spécifique. C'est un problème grave mais il n'a pas été soigneusement étudié. Par exemple, la question suivante a été posée au même échantillon, et voici les résultats des réponses:

A) Pensez-vous que les États-Unis devraient permettre une rhétorique anti-démocratique?

– 21% des réponses sont pour la liberté d'expression

B) Pensez-vous que les États-Unis devraient empêcher la rhétorique anti-démocratique?

– 39% des réponses sont pour la liberté d'expression

Par conséquent, vous pouvez voir comment le vocabulaire choisi dans nos questions peut affecter les réponses.

De plus, dans les questionnaires, les chances que nous donnons au répondant/à la répondante dans les réponses peut influencer son choix. Dans l'exemple ci-dessous, on voit que dans le premier cas, les probabilités de réponse à la question : Combien d'heures regardez-vous la télévision par jour ? « Entre une demi-heure et une heure et demie » ; adressent un message indirect au/à la participant.e, à savoir que ce laps de temps est une durée en moyenne "normale" passée à regarder la télévision quotidiennement. Dans ce cas, la plupart des participant.es répondent généralement qu'ils/elles regardent la télévision tous les jours entre une demi-heure et une heure et demie (réponses du milieu).

Dans le second cas, vous voyez que les possibilités suggèrent que le temps « normal » moyen passé à regarder la télévision quotidiennement se situe entre une heure et demie et deux heures et demie. Ici, nous constatons que si le/la participant.e est exposé à ces probabilités au lieu des probabilités du premier cas, le nombre moyen d'heures quotidiennes passé à regarder la télévision augmentera automatiquement dans les réponses. Cette augmentation est précisément due à l'effet de la formulation des probabilités

et pas nécessairement à une augmentation réelle des heures de télévision.

**Exemple : Combien d'heures regardez-vous la télévision par jour ?**

Possibilités de réponse dans le cas 1	Possibilités de réponse dans le cas 2
<input type="radio"/> Moins d'une demi-heure	<input type="radio"/> Moins d'une heure et demie
<input type="radio"/> Entre une heure et une heure et demie	<input type="radio"/> Entre deux heures et deux heures et demie
<input type="radio"/> Entre une heure et une heure et demie	<input type="radio"/> Entre deux heures et deux heures et demie
<input type="radio"/> Plus d'une heure et demie	<input type="radio"/> Plus de deux heures et demie

L'effet de l'ordre dans la succession des questions ou des réponses :  
Tout comme la formulation des questions et réponses affecte le résultat de la recherche, il en est de même pour l'ordre dans la succession des questions et réponses. Dans l'exemple ci-après, nous constatons que l'ordre des questions (du général au spécifique, ou vice versa) peut affecter la réponse. Si le/la participant.e à la recherche est une personne relativement heureuse mais éprouve des problèmes conjugaux, l'ordre des questions peut influencer ses réponses. Si nous l'interrogeons d'abord sur son niveau de bonheur général, il/elle peut répondre qu'il/elle est heureux/se ou même très heureux/se. Cependant si nous commençons par la question sur le bonheur conjugal, l'élément négatif dans la réponse peut affecter sa mesure de son bonheur général parce qu'il l'aura incité à réfléchir à la relation conjugale dans sa mesure du bonheur général.

**Exemple** : l'ordre dans la succession des questions (effet d'ordre)

Question 1 Comment décririez-vous votre niveau de bonheur général ?	Question 2 Comment décririez-vous votre niveau de bonheur conjugal ?
Très heureux	Très heureux
Heureux	Heureux
Pas heureux	Pas heureux
Pas heureux du tout	Pas heureux du tout

Question 3 Comment décririez-vous votre niveau de bonheur général ?	Question 4 Comment décririez-vous votre niveau de bonheur conjugal ?
Très heureux	Très heureux
Heureux	Heureux
Pas heureux	Pas heureux
Pas heureux du tout	Pas heureux du tout

Il en va de même pour l'ordre des réponses (pas seulement les questions). Si les probabilités sont positives (très heureux) ou négatives (pas du tout heureux), cela peut affecter les choix des participantes.

**Exemple** : l'ordre dans la succession des réponses

Question 1 Comment décririez-vous votre niveau de bonheur général ?	Question 2 Comment décririez-vous votre niveau de bonheur conjugal ?
Très heureux	Très heureux
Heureux	Heureux
Pas heureux	Pas heureux
Pas heureux du tout	Pas heureux du tout

**L'effet du chercheur sur les réponses** : Enfin, le chercheur/la chercheuse doit également considérer sa propre influence sur les réponses et essayer de la réduire. On sait que la présence du chercheur/de la chercheuse a un impact sur la manière dont les réponses sont choisies et sur la fiabilité des informations collectées car l'apparence du chercheur/de la chercheuse et la manière dont il/elle se présente peuvent affecter les réponses et la manière dont les participant.es interagissent avec les questions de recherche. Les participant.es répondent généralement en fonction de ce qu'ils/elles pensent



que le chercheur/la chercheuse veut entendre, ou encore en fonction de ce qui est socialement acceptable. Les réponses ne reflètent donc pas nécessairement la vérité. De ce fait, il est important que le chercheur/la chercheuse essaie de rassurer l'interlocuteur sur le fait qu'il n'y a pas de jugement de valeur sur les participants. De même il doit le convaincre de la nécessité de donner la réponse la plus proche de la vérité et non de ce qui est forcément compatible avec les valeurs de la société.

Par exemple, si le sujet de la recherche porte sur les relations sexuelles pré-nuptiales dans les sociétés arabes, et que l'équipe de travail compte deux chercheuses, l'une voilée et l'autre vêtue d'une robe courte, alors il y a une forte probabilité que cela affecte les réponses que chacune des chercheurs/chercheuses recueillera. La tenue vestimentaire de la chercheuse peut en effet inciter les participants à répondre dans une direction spécifique pour « satisfaire » la chercheuse ou pour être en harmonie avec les valeurs dominantes dans la société.

### **L'analyse des données**

La phase de collecte des données recueille généralement une quantité importante et variée de données, notamment des transcriptions d'entretiens, des notes de terrain et des commentaires préliminaires, en plus de nombreux documents divers liés au sujet de recherche. À première vue, il semble que l'analyse, l'interprétation et l'extraction des significations du matériel collecté soit une tâche ardue. Le chercheur/la chercheuse se trouve souvent dans un état de frustration face à l'énorme quantité de données. Cependant, avec l'expérience, la richesse informationnelle devient une source de créativité et d'analyse approfondie.

La phase d'analyse des données comprend leur organisation et leur répartition en unités qui peuvent être traitées et synthétisées afin de rechercher des modèles et des tendances permettant de découvrir ce qui est important et ce qui peut être appris de ces données.

Il y a six étapes à suivre pour analyser les données qualitatives : transcrire les entretiens, organiser les données, désagréger les données, enregistrer les données, identifier les modèles et les tendances et vérifier les résultats.

- **Transcription des entretiens:** Il est très important que les entretiens soient rédigés afin de faciliter le processus d'analyse. Si l'entretien est enregistré, il est possible de télécharger et d'écrire uniquement les parties importantes directement liées à la recherche.
- **Organisation des données:** À ce stade, le chercheur/la chercheuse dispose d'une grande quantité de données, entre entretiens, observation, les documents, etc. Il/elle a également fait quelques observations initiales lors de la collecte d'informations. Ces informations doivent être organisées et arrangées de manière à y revenir rapidement et à les traiter de manière à faciliter leur analyse. Il n'y a pas de modèle d'organisation unique mais le chercheur/la chercheuse peut organiser les données en fonction de la méthode de collecte des informations (observation, entretien ou documents). Le classement peut aussi s'opérer, en fonction des individus avec qui il/elle a mené la recherche, ou autrement, selon ce que le chercheur/la chercheuse jugera approprié pour lui et pour la méthode qu'il adoptera dans l'analyse.

Cette organisation peut être faite manuellement, en plaçant les données dans des fichiers. Sinon, des programmes automatisés peuvent être utilisés pour les classer et les indexer sur l'ordinateur. Il existe également des programmes de recherche qualitative dédiés qui facilitent l'organisation des données et les processus d'analyse, tels que NVIVO.

- **Codage des données :** Lors de la lecture initiale des données, le chercheur/ la chercheuse commence à pratiquer un système de classification qui l'aide lors de l'analyse. Ce type de classification consiste à donner des titres aux informations contenues dans les données collectées Cette classification consiste à donner un titre ou un nom aux parties qui, selon le chercheur/la chercheuse, sont significatives dans sa recherche. Les questions de recherche sont considérées comme un facteur fondamental dans la définition et l'orientation du système de classification car le chercheur/la

chercheuse élabore des questions de recherche à la lumière des sujets et des facteurs importants de sa recherche.

- **Enregistrement des données** : Après l'étape de classification, le chercheur/ la chercheuse doit relire les données et enregistrer ses observations après avoir établi dans son esprit une structure pour ce système de classification. Il procède ainsi après avoir donné un titre, en fonction du sujet traité aux nombreuses unités de données dont il/elle dispose et que des points qui représentent des repères aient commencé à voir le jour, bien que pas totalement clairs au début de leur formation et pas visibles lors de la collecte initiale des informations.

Ces notes se présentent sous la forme de questions qui mènent à des recherches plus poussées, que ce soit à partir des informations disponibles ou pour la recherche d'informations complémentaires. Lesdites notes peuvent revêtir la forme d'enregistrement des relations entre les catégories qui ont été définies mais qui doivent être vérifiées.

Plus la lecture est répétée, plus la probabilité de découvrir de nouveaux éléments dans les données est grande. Le chercheur/la chercheuse doit donc lire ses données plusieurs fois et ne pas se contenter d'une ou deux lectures. Chaque fois qu'il y a de nombreuses questions sans réponse ou qu'il n'est pas possible de construire un bon système de classification, cela est un indicateur de la faiblesse de l'échantillon et du besoin de disposer de plus de données.

- **Déterminer les tendances et les modèles** : La détermination des tendances et des modèles est un type de classification mais il s'agit d'un niveau d'abstraction plus élevé. Par conséquent, certains chercheurs/chercheuses l'appelleront codage axial parce qu'il fait tourner les catégories sur un seul axe. D'autres le nommeront familles de codage car il regroupe un certain nombre de catégories dans une même famille. On peut également la qualifier de classification déductive (par opposition à la classification descriptive).

Une fois la classification ouverte effectuée et les notes placées dessus, les données classifiées sont relues pour classer à nouveau les catégories sous la forme de modèles et de tendances à un niveau d'abstraction supérieur à la classification ouverte qui consiste à attribuer des titres pour les unités d'information.

Ce type de classification nécessite une réflexion approfondie et une lecture attentive en vue de trouver des relations et de faire des comparaisons entre les ensembles de données. Ainsi, le chercheur déterminera les modèles et tendances qui ont été formés à partir de la classification des données et commencera à intégrer certains d'entre eux tout en les comparant.

- **Recouplement** : À la fin de la phase d'analyse, le chercheur/la chercheuse peut revenir aux études précédentes pour vérifier les résultats de son étude. Contrairement à la méthode de recherche quantitative, l'accent doit être mis sur les études antérieures à ce stade et non au début de la recherche (car la recherche quantitative sélectionne des hypothèses). Le recouplement durant cette étape diffère de celui qui s'opère dans la recherche quantitative où le chercheur/la chercheuse arrive avec des hypothèses qu'il/ elle veut examiner. Les hypothèses de la recherche qualitative naissent généralement du processus d'analyse et non avant lui. Le processus d'analyse est un processus échelonné qui se poursuit jusqu'au dernier moment de la rédaction du rapport de recherche.

### **Erreurs courantes dans la formulation des résultats de la recherche qualitative**

Il faut veiller à ne pas tomber dans les trois erreurs courantes lors de la formulation des résultats d'une recherche qualitative:

- 1. Utilisation de pourcentages** : Il n'est pas approprié d'utiliser des pourcentages lorsque l'on contredit les résultats d'une recherche qualitative car l'échantillon utilisé par ce type de recherche est généralement limité et non représentatif.

- 2. Généralisation**: La généralisation n'est pas requise lors de la

rédaction des résultats d'une recherche qualitative car un échantillon limité ne représente pas l'ensemble de la population.

**3. Ne pas discuter des défis:** Une erreur courante lors de la rédaction d'études de recherche est de ne pas discuter des défis qui pourraient affecter les résultats ou l'analyse. Le chercheur/la chercheuse devrait consacrer un paragraphe dans son étude aux défis et obstacles.

### **Qu'est-ce que la recherche quantitative?**

#### **Caractéristiques de la recherche quantitative**

La recherche quantitative repose sur la collecte et la mise en rapport d'informations et de faits qui peuvent être quantifiés et mesurés ou de faits sociaux qui peuvent être convertis en chiffres, statistiques et données graphiques. Ce type de recherche est basé sur la mesure des opinions à travers une enquête, un questionnaire ou bien la mesure du comportement par l'observation et la collecte d'informations enregistrées. La recherche quantitative est particulièrement importante pour les études démographiques ou les études liées au vote et au comportement politique ou social en général.

Quelles sont les hypothèses de recherche ?

La recherche scientifique quantitative est basée sur les tests d'hypothèses. Une hypothèse est une supposition, une affirmation ou une prédiction qui décrit la relation possible ou attendue entre les variables de la recherche. On peut donc dire qu'une hypothèse scientifique est une explication possible d'une question de recherche qui peut être testée statistiquement. Habituellement, il existe plusieurs hypothèses dans la recherche scientifique quantitative, et le chercheur/la chercheuse sélectionne des hypothèses en fonction des informations et des données collectées. Si l'on suppose, par exemple, que le taux d'échec aux examens officiels en Algérie a considérablement augmenté cette année et que des chercheurs/chercheuses ont voulu analyser ce phénomène, les hypothèses peuvent examiner la relation entre plusieurs variables sur la base des théories et analyses présentées précédemment. Par exemple, on peut choisir l'effet du nombre d'heures d'étude, la difficulté des examens, la compétence des professeurs, la situation politique, etc. sur les résultats aux examens.

L'hypothèse comprend généralement deux variables de base (pas plus) :

- La variable indépendante qui est la cause. Par exemple, lors de l'étude des résultats d'examens, la variable indépendante peut être le nombre d'heures d'étude.
- La variable dépendante qui est affectée par la variable indépendante et en est le résultat, comme les examens officiels.

La variable indépendante dans une recherche peut être la variable dépendante dans une autre.

### **Quels sont les avantages des hypothèses ?**

Parmi les avantages des hypothèses, il y a le fait qu'elles contribuent à :

- Déterminer de manière exacte l'étendue du problème de recherche ;
- Organiser la recherche car les hypothèses représentent l'élément de base du sujet de recherche ;
- Relier la question de recherche aux théories présentées ;
- Guider le chercheur/la chercheuse dans le choix du type d'observations, de procédures et de questionnaires ;
- Guider le chercheur la chercheuse dans l'analyse des résultats et expliquer scientifiquement les phénomènes ;
- Tester les méthodes statistiques appropriées afin de tester les relations entre les variables de recherche.

### **L'hypothèse nulle et l'hypothèse alternative**

La recherche quantitative comprend deux types d'hypothèses:

- **Hypothèse nulle:** C'est l'hypothèse qui considère qu'il n'y a pas de relation entre la variable indépendante et la variable dépendante. Par exemple, l'hypothèse nulle considère qu'il n'y a pas de relation entre le nombre d'heures d'étude et le résultat à l'examen.
- **Hypothèse alternative:** C'est l'hypothèse de la recherche qui considère qu'il existe une relation entre la variable indépendante et la variable dépendante. Par exemple, l'hypothèse alternative considère que les résultats à l'examen sont susceptibles de s'améliorer avec l'augmentation du nombre d'heures de préparation pour l'examen. La méthode quantitative consiste à tester l'hypothèse nulle et à essayer de l'infirmer. La logique scientifique est basée sur l'infirmer des

hypothèses, et non sur leur affirmation. La connaissance scientifique est exacte mais non absolue (il n'y a pas de vérité à 100% dans la science) car il y a toujours une marge d'erreur dans la recherche scientifique.

### **Caractéristiques des hypothèses**

L'hypothèse devrait être

- **Spécifique** : Chaque hypothèse doit être spécifique et tester la relation entre une variable indépendante et la variable dépendante. Ainsi, les hypothèses scientifiques ne peuvent pas examiner plus qu'une relation d'une variable à la fois. Par exemple, l'hypothèse suivante ne peut pas être considérée comme une hypothèse scientifique car elle n'est pas spécifique et comprend plus d'une variable :

« En raison de leur statut de réfugié.es et de l'importance des appareils électroniques pour communiquer avec la famille et les amis, les Syrien.nes dépensent aujourd'hui plus qu'auparavant pour ces appareils. »

Il y a dans cette hypothèse plus d'une variable indépendante (l'effet du statut de réfugié.e, l'importance des appareils électroniques pour la communication) et le libellé n'est ni spécifique ni précis car l'expression « plus qu'auparavant » est vague et non quantifiable.

- **Précise et mesurable**: L'hypothèse doit être précise pour décrire la relation entre la variable dépendante et la variable indépendante. L'utilisation d'expressions inexactes et pas claires doit être évitée. Par exemple si l'hypothèse est la suivante: « Les Syrien.nes dépensent beaucoup pour les appareils électroniques. » Comment « beaucoup » est-il mesuré ? Ici également, il est difficile de mesurer la variable, car l'hypothèse est imprécise et non scientifiquement mesurable. Il en va de même pour l'hypothèse suivante :

- Les Syrien.nes dépensent plus en appareils électroniques qu'auparavant.

À quoi correspond « auparavant » ? Est-ce le mois dernier, l'année dernière ou la dernière décennie ? La réponse peut changer complètement en fonction de la période de temps que nous adoptons pour désigner le « auparavant »,

car cette hypothèse n'est ni claire, ni précise.

- **Falsifiable** : Le philosophe Karl Popper (2005) dit que la connaissance scientifique est sujette à la falsifiabilité car la science ne traite pas de faits fiables à 100% mais introduit plutôt de nouvelles connaissances avec une marge d'erreur ; ce qui permet de progresser dans les connaissances et la science. Par exemple, la recherche de l'existence de la vie sur la planète Terre n'est pas une recherche scientifique parce qu'il s'agit d'une connaissance évidente. La recherche de l'existence de la vie sur d'autres planètes est cependant une recherche qui nécessite une méthode scientifique pour la découvrir ou la nier.
- **Liée à un phénomène social et non pas à une relation personnelle** : L'hypothèse doit refléter un modèle ou une relation sociale générale et non une relation personnelle ou entre certains individus. Donc, si l'hypothèse est la suivante :
  - J'ai dépensé plus en appareils électroniques que mon voisin.

Cette hypothèse est mesurable, testable et falsifiable mais ce n'est pas une hypothèse scientifique pour une étude en sciences sociales parce qu'elle n'ajoute rien à notre connaissance de la société.

#### Exercice

Choisissez la bonne hypothèse :

- Les Syrien.nes dépensent beaucoup en appareils électroniques.
- J'ai dépensé plus en appareils électroniques que mon voisin turc.
- En raison de leur statut de réfugiés et de l'importance des appareils électroniques pour communiquer avec la famille et les amis, les Syrien.nes dépensent aujourd'hui plus qu'auparavant sur ces appareils.
- Les Syrien.nes dépensent plus en appareils électroniques qu'auparavant.
- Les Syrien.nes ne dépensent pas plus en appareils électroniques que les Turc/Turques.

La dernière hypothèse est la bonne réponse, car elle est spécifique, liant une variable indépendante (nationalité) et une variable dépendante (dépenses en appareils électroniques). De plus, cette hypothèse est précise, mesurable et falsifiable.

#### Variable quantitative et variable qualitative

Il existe quatre types de variables qui peuvent être utilisées dans la



recherche scientifique:

- **La variable quantitative/numérique:** Par exemple, âge, prix, nombre de manifestations, etc.
- **La variable qualitative/catégorielle :** par exemple, sexe (homme/femme), nationalité (allemande/française/britannique), etc.

Il existe deux types de variables qualitatives :

- **La variable qualitative nominale :** Il s'agit d'une variable qualitative qui n'est pas hiérarchisée, telles que : le sexe ou la nationalité, car il n'y a pas de hiérarchie ou d'ordre pyramidal dans ces catégories. Les réponses à la variable nationalité peuvent donc être présentées comme suit : (Allemand/ Français/Britannique), ou bien (Français/Britannique / Allemand) sans que cela n'affecte le résultat de la recherche.
- **La variable qualitative ordinale:** C'est une variable qualitative hiérarchique qui répond à un ordre pyramidal spécifique. Par exemple, les variables relatives à l'opinion. Les réponses se présentent comme suit : tout à fait d'accord, d'accord, pas d'accord, pas du tout d'accord ; et la hiérarchie de ces réponses ne peut pas être modifiée.

Le type de variables affecte les types de statistiques qui peuvent être utilisées dans l'analyse des données.

### **Statistiques descriptives ou statistiques inférentielles**

Il existe deux types de statistiques de base

- **Statistiques descriptives:** La description peut se faire de différentes manières. Il y a d'une part la description par l'image (les diagrammes, le dessin avec des points, les colonnes, les courbes, les secteurs circulaires) et d'autre part la description par le nombre, c'est-à-dire la description d'un groupe de données en nombres. Cette description peut aussi s'articuler autour d'une certaine valeur (centralisation) ou être en désaccord sur une certaine valeur (dispersion).
- **Statistiques différentielles:** Elles sont liées au test de l'hypothèse et de la relation des variables entre elles (causale ou corrélation) et sont basées sur une analyse de régression.

# ANALYSE AVANCEE DES DONNEES STATISTIQUES

## Analyse en Composantes Principales (ACP)

### 1-Définition

Analyse en Composantes Principales (ACP) est :

- Une méthode **descriptive** a pour objectif l'analyse des tableaux de données qui ne comportent pas des structure préalable (aucune distinction ni entre variable ni entre individu)
- Le but principal est de **résumer** l'information contenue dans un tableau composé d'un nombre élevé de ligne et de colonnes

→ Un outil statistique de synthèse de l'information

→ Un outil très important pour traiter les données quantitatives

### 2-Principe

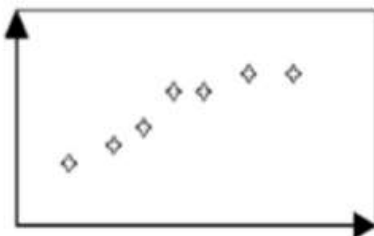
Analyse en Composantes Principales (ACP) permet de :

- Résumer les informations contenant dans un tableau en  $n$  individus et  $p$  variables
- Remplacer les  $p$  variables avec  $q$  nouvelles variables avec  $q < p$

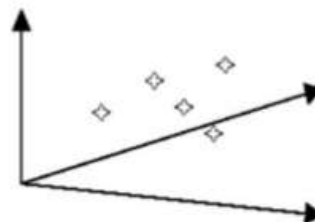
D'un point de vue géométrique Le nuage de points représentant les données s'inscrit dans un espace de  $P$  dimensions, puisque chaque point représente un individu par rapport à  $X_1, X_2, \dots, X_n$

→ Il est difficile de visualiser les relations existant entre les variables dès que  $p > 3$

Si la dimension  $P= 2$   
Il est facile de présenter le nuage de points



Si la dimension  $P > 3$   
Il est difficile de présenter le nuage de points

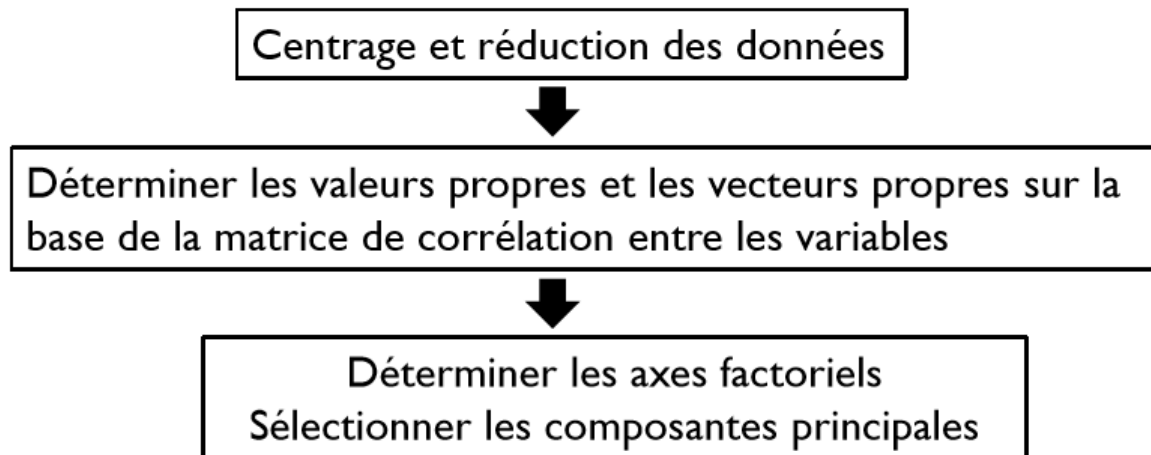


### 3- Démarche

Soit :  $n$  individus caractérisés par  $p$  variables métriques Ces données

sont présentées dans un tableau appelé la Matrice des données de dimension  $n \times p$

→ Les étapes pour déterminer la composante principale:



→ **Centrage et réduction des données:**

Soit :  $n$  individus caractérisés par  $p$  variables métriques

- Les  $p$  variables sont de nature différente, pour homogénéiser les unités, les  $p$  variables seront centrées et réduites,
- Les données sont centrées et réduites signifie que pour chaque variable la moyenne est nulle ( $\bar{X} = 0$ ) et la variance égale à 1 ( $\sigma = 1$ )

Matrice Centrée Réduite est obtenue par la formule suivante :

$$x_{ij} = \frac{x_{ij} - \bar{X}}{\sigma_i}$$

→ Centrage et réduction des données:

La Matrice des variances covariances permet de mesurer la liaison linéaire qui peut exister entre un couple de variables statistiques

$$\begin{vmatrix} \text{Var X1} & \text{Cov (X1,X2)} & \text{Cov (X1,X3)} \\ \text{Cov (X2,X1)} & \text{Var X2} & \text{Cov (X2,X3)} \\ \text{Cov (X3,X1)} & \text{Cov (X3,X2)} & \text{Var X3} \end{vmatrix}$$

Si  $\text{Cov}(X_2, X_1) = 0 \rightarrow$  les variables  $X_1$  et  $X_2$  sont indépendantes  
 Si  $\text{Cov}(X_2, X_1) \neq 0 \rightarrow$  les variables  $X_1$  et  $X_2$  sont dépendantes  
 (existe une relation linéaire entre les variables)

Obtenue par la formule suivante :

$$V = \frac{1}{n} MC * MC^t$$

$MC$  : Matrice Centrée  $\Leftrightarrow MC : x_{ij} = (x_{ij} - \bar{X})$

$MC^t$  : Matrice Transposée

Matrice des corrélations entre variables permet d'analyser les relations bilatérales entre les variables :

Obtenue par la formule suivante :

$$U = \frac{1}{n} CR^t * CR$$

Avec

$CR$  : Matrice Centrée Réduite

$CR^t$  : Matrice Centrée Réduite Transposée

### Cas d'application

Une étude consiste à déterminer les facteurs de la localisation internationale d'une marque. Soit le tableau des données suivant:

	IDE	Taux croissance économique (%)	Taux d'inflation (%)
Pays A	300	2	6
Pays B	450	2	4
Pays C	950	8	2
Pays D	700	7	5

### Travail à faire:

1. Calculer la moyenne et l'écart type des variables

2. Déterminer la Matrice Centrée Réduite
3. Déterminer la Matrice des variances covariance
4. Déterminer la Matrice des corrélations entre variables
5. Déterminer le polynôme caractéristique
6. Calculer les valeurs propres
7. Calculer et interpréter l'inertie des axes factoriels
8. Déterminer les vecteurs propres orthogonaux associés aux valeurs propres
9. Calculer et interpréter la corrélation des variables avec les composantes principales
10. Calculer et interpréter la contribution CONTR

**Solution**

- 1- Calculer la moyenne et l'écart type des variables

**Définition :**

**La moyenne** est un outil de calcul **permet de résumer** une liste de valeurs numériques en un seul nombre réel sans tenir compte de l'ordre de la liste.

$$\bar{X} = \frac{1}{n} \sum x_i$$

**Définition :**

L'écart type est un outil de calcul permet de mesurer **la dispersion** des valeurs d'un échantillon. C'est la racine carrée de la variance :

$$\sigma = \sqrt{\text{variance}}$$

Avec la variance est la moyenne des carrés des écarts à la moyenne :

$$V = \frac{1}{n} \sum (x_i - \bar{X})^2$$

Avec

$x_i$ : les valeurs de la variable

$\bar{X}$  : la moyenne de la variable

	IDE	Taux croissance économique (%)	Taux d'inflation (%)
Pays A	300	2	6
Pays B	450	2	4
Pays C	950	8	2
Pays D	700	7	5
<b>Moyenne</b>	<b>600</b>	<b>4,75</b>	<b>4,25</b>
<b>Ecart type</b>	<b>247,50</b>	<b>2,77</b>	<b>1,48</b>

Calcul: Pour la variable  $x_1$

$$\bar{X}_1 = \frac{300 + 450 + 950 + 700}{4} = 600$$

$$\sigma_1 = \sqrt{\frac{(300 - 600)^2 + (450 - 600)^2 + (950 - 600)^2 + (700 - 600)^2}{4}} = 247,5$$

## 2- Déterminer la Matrice Centrée Réduite (MCR)

$$\text{MCR : } x_{ij} = \frac{x_{ij} - \bar{X}}{\sigma_i} \quad \left| \begin{array}{ccc} -1,21 & -0,99 & 1,18 \\ -0,80 & -0,99 & -0,16 \\ 1,41 & 1,17 & -1,50 \\ 0,40 & 0,81 & 0,50 \end{array} \right|$$

Calcul: Pour la variable  $x_1$

$$\left| \begin{array}{l} x_{11} = \frac{x_{ij} - \bar{X}}{\sigma_i} = \frac{300 - 600}{247,5} = -1,21 \\ x_{21} = \frac{x_{ij} - \bar{X}}{\sigma_i} = \frac{450 - 600}{247,5} = -0,8 \\ x_{31} = \frac{x_{ij} - \bar{X}}{\sigma_i} = \frac{950 - 600}{247,5} = 1,41 \\ x_{41} = \frac{x_{ij} - \bar{X}}{\sigma_i} = \frac{700 - 600}{247,5} = 0,4 \end{array} \right|$$

3- Déterminer la Matrice des variances covariances :  $V = \frac{1}{n} MC * MC^t$

Matrice centrée (MC):	-300	-2,75	1,75
$x_{ij} = (x_{ij} - \bar{X})$	-150	-2,75	-0,25
	350	3,25	-2,25
	100	2,25	0,75

Calcul: Pour la variable x1

$$x_{11} = (x_{ij} - \bar{X}) = 300 - 600 = -300$$

$$x_{21} = (x_{ij} - \bar{X}) = 450 - 600 = -150$$

$$x_{31} = (x_{ij} - \bar{X}) = 950 - 600 = -350$$

$$x_{41} = (x_{ij} - \bar{X}) = 700 - 600 = -100$$

$$V = \frac{1}{n} MC * MC^t$$

$$V = \frac{1}{4} \begin{vmatrix} -300 & -2,75 & 1,75 \\ -150 & -2,75 & -0,25 \\ 350 & 3,25 & -2,25 \\ 100 & 2,25 & 0,75 \end{vmatrix} * \begin{vmatrix} -300 & -150 & 350 & 100 \\ -2,75 & -2,75 & 3,25 & 2,25 \\ 1,75 & -0,25 & -2,25 & 0,75 \end{vmatrix}$$

$$V = \begin{vmatrix} 22503 & 45007 & -26253 & -7501 \\ 11252 & 5627 & -13127 & -3752 \\ 26253 & -13127 & 30629 & 8751 \\ -7501 & -3751 & 8751 & 2501 \end{vmatrix}$$

#### 4- Déterminer la Matrice des corrélations entre variables

$$U = \frac{1}{n} CR^t * CR$$

$$U = \frac{1}{4} \begin{vmatrix} -1,21 & -0,80 & 1,41 & 0,40 \\ -0,99 & -0,99 & 1,17 & 0,81 \\ 1,18 & -0,16 & -1,50 & 0,50 \end{vmatrix} * \begin{vmatrix} -1,21 & -0,99 & 1,18 \\ -0,80 & -0,99 & -0,16 \\ 1,41 & 1,17 & -1,50 \\ 0,40 & 0,81 & 0,50 \end{vmatrix}$$

$$U = \begin{vmatrix} 1 & 0,99 & -0,8 \\ 0,99 & 1 & -0,6 \\ -0,8 & -0,6 & 1 \end{vmatrix}$$

(X1; X2) = 0,99 Forte corrélation positive entre IDE et Taux de Croissance  
Taux de Croissance augmente → IDE augmente

(X1; X3) = -0,8 Forte corrélation négative entre IDE et Taux d'inflation  
Taux d'inflation augmente → IDE diminue

#### 5- Déterminer le polynôme caractéristique

$$\text{Det} | U - \lambda I | =$$

Avec

$$U = \begin{vmatrix} 1 & 0,99 & -0,8 \\ 0,99 & 1 & -0,6 \\ -0,8 & -0,6 & 1 \end{vmatrix} \quad I = \begin{vmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{vmatrix}$$

$$\text{Det} | U - \lambda I | = \begin{vmatrix} 1-\lambda & 0,99 & -0,8 \\ 0,99 & 1-\lambda & -0,6 \\ -0,8 & -0,6 & 1-\lambda \end{vmatrix}$$



6- Calculer les valeurs propres de  $\lambda$

$$\text{Det} | U - \lambda I | = 0 \quad \text{Det} \begin{vmatrix} 1-\lambda & 0,99 & -0,8 \\ 0,99 & 1-\lambda & -0,6 \\ -0,8 & -0,6 & 1-\lambda \end{vmatrix} = 0$$

$$\text{Det} \begin{vmatrix} 1-\lambda & 0,99 & -0,8 \\ 0,99 & 1-\lambda & -0,6 \\ -0,8 & -0,6 & 1-\lambda \end{vmatrix} = 0$$

$$\begin{cases} \lambda^2 - 3\lambda + 2,98 = 0 \\ \lambda = 0 \\ \lambda_1 = 2,35 \\ \lambda_2 = 0,65 \end{cases}$$

7- Calculer et interpréter l'inertie des axes factoriels

→ L'inertie de l'axe factoriel:

Le pourcentage (%) d'inertie exprimé par un axe factoriel permet d'évaluer la quantité d'information contenue dans cet axe :

$$\text{Inertie d'un axe} = \frac{\text{Valeur propre correspondante}}{\text{somme des valeurs propres (Inertie totale)}}$$

$$\text{Axe 1 : } \lambda_1 = 2,35$$

$$\text{Inertie Axe 1} = \frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{2,35}{2,35 + 0,65} = 0,78$$

→ Cet axe contient 78% des informations

Axe 2:  $\lambda_2 = 0,65$

$$\text{Inertie Axe 2} = \frac{\lambda_2}{\lambda_1 + \lambda_2} = \frac{0,65}{2,35 + 0,65} = 0,22$$

→ Cet axe contient 22% des informations

**8- Déterminer les vecteurs propres orthogonaux associés aux valeurs propres**

Axe 1 :  $\lambda_1 = 2,35$  → On cherche  $\vec{a} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$  tel que :

$$U\vec{a} = \lambda_1 \vec{a}$$

$$\begin{vmatrix} 1 & 0,99 & -0,8 \\ 0,99 & 1 & -0,6 \\ -0,8 & -0,6 & 1 \end{vmatrix} * \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2,35 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\begin{cases} 1x + 0,99y - 0,8z = 2,35x \\ 0,99x + 1y - 0,6z = 2,35y \\ -0,8x - 0,6y + 1z = 2,35z \end{cases}$$

$$\begin{cases} 1x + 0,99y - 0,8z - 2,35x = 0 \\ 0,99x + 1y - 0,6z - 2,35y = 0 \\ -0,8x - 0,6y + 1z - 2,35z = 0 \end{cases}$$

$$\begin{cases} x = 0,57 \\ y = 0,54 \\ z = 0,64 \end{cases} \Rightarrow U = \begin{pmatrix} 0,57 \\ 0,54 \\ 0,64 \end{pmatrix}$$

Axe 2 :  $\lambda_2 = 0,65$   $\rightarrow$  On cherche  $\vec{a} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$  tel que :

$$U\vec{a} = \lambda_2\vec{a}$$

$$\begin{vmatrix} 1 & 0,99 & -0,8 \\ 0,99 & 1 & -0,6 \\ -0,8 & -0,6 & 1 \end{vmatrix} * \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,65 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\begin{cases} 1x + 0,99y - 0,8z = 0,65x \\ 0,99x + 1y - 0,6z = 0,65y \\ -0,8x - 0,6y + 1z = 0,65z \end{cases}$$

$$\begin{cases} 1x + 0,99y - 0,8z - 0,65x = 0 \\ 0,99x + 1y - 0,6z - 0,65y = 0 \\ -0,8x - 0,6y + 1z - 0,65z = 0 \end{cases}$$

$$\begin{cases} x = 0,53 \\ y = 0,99 \\ z = 0,34 \end{cases} \Rightarrow V = \begin{pmatrix} 0,53 \\ 0,99 \\ 0,34 \end{pmatrix}$$

**9- Calculer et interpréter la corrélation des variables avec les composantes principales**

→ Consiste à la projection des individus sur les axes principaux Avec :

MCR \* U ⇒ Projection sur axe 1

MCR \* V ⇒ Projection sur axe 2

$$\text{MCR} \begin{vmatrix} -1,21 & -0,99 & 1,18 \\ -0,80 & -0,99 & -0,16 \\ 1,41 & 1,17 & -1,50 \\ 0,40 & 0,81 & 0,50 \end{vmatrix} \quad U = \begin{pmatrix} 0,57 \\ 0,54 \\ 0,64 \end{pmatrix} \quad V = \begin{pmatrix} 0,53 \\ 0,99 \\ 0,34 \end{pmatrix}$$

MCR \* U ⇒ Projection sur **axe 1**

$$\begin{vmatrix} -1,21 & -0,99 & 1,18 \\ -0,80 & -0,99 & -0,16 \\ 1,41 & 1,17 & -1,50 \\ 0,40 & 0,81 & 0,50 \end{vmatrix} * \begin{pmatrix} 0,57 \\ 0,54 \\ 0,64 \end{pmatrix}$$

MCR \* V ⇒ Projection sur **axe 2**

$$\begin{vmatrix} -1,21 & -0,99 & 1,18 \\ -0,80 & -0,99 & -0,16 \\ 1,41 & 1,17 & -1,50 \\ 0,40 & 0,81 & 0,50 \end{vmatrix} * \begin{pmatrix} 0,53 \\ 0,99 \\ 0,34 \end{pmatrix}$$

Résultat de projection :

	AXE 1	AXE 2
I1	-0,46	-1,22
I2	-1,09	-1,46
I3	0,48	1,4
I4	0,99	1,2

Ce tableau permet de calculer la contribution des axes

### → La contribution des individus

Pour calculer la contribution de chaque individu à l'inertie, on utilise la formule suivante :

$$CONTR = P_i \frac{\|x_i\|^2}{I_i}$$

$x_i$  : Valeur CR

$P_i$  : Poids  $\frac{1}{n}$

$I_i$  : La somme des valeurs propres ( $\sum \lambda_i$ )

	X1	X2	X3	CONTR
Pays A	0,12	0,08	0,17	<b>0,37</b>
Pays B	0,05	0,08	0,01	<b>0,14</b>
Pays C	0,17	0,11	0,19	<b>0,47</b>
Pays D	0,1	0,05	0,02	<b>0,17</b>
	<b>0,44</b>	<b>0,32</b>	<b>0,39</b>	

Calcul

$$(x1 ; Pays A) \quad CONTR = \frac{1}{4} * \frac{\|1,21\|^2}{3} = 0,12$$

$$(x1 ; Pays B) \quad CONTR = \frac{1}{4} * \frac{\|0,8\|^2}{3} = 0,05$$

$$(x1 ; Pays C) \quad CONTR = \frac{1}{4} * \frac{\|1,41\|^2}{3} = 0,17$$

$$(x1 ; Pays D) \quad CONTR = \frac{1}{4} * \frac{\|0,4\|^2}{3} = 0,1$$

$$\sum \lambda_i = 3$$

	X1	X2	X3	CONTR
Pays A	0,12	0,08	0,17	<b>0,37</b>
Pays B	0,05	0,08	0,01	<b>0,14</b>
Pays C	0,17	0,11	0,19	<b>0,47</b>
Pays D	0,1	0,05	0,02	<b>0,17</b>
	<b>0,44</b>	<b>0,32</b>	<b>0,39</b>	

→ Le tableau permet de déterminer la contribution des individus dans l'analyse:

- Le pays **C** contribue de **0,47** (47%) pour expliquer le phénomène
- Le pays **A** contribue de **0,37** (37%) pour expliquer le phénomène

→ Le tableau permet de déterminer la contribution des variables dans l'analyse:

- La variable X1 contribue de 44% dans l'analyse
- La variable X2 contribue de 32% dans l'analyse
- La variable X3 contribue de 39% dans l'analyse

10- Calculer et interpréter la contribution CONTR

→ La contribution **CONTR** des axes est calculée par la formule suivante :

$$\text{Axe 1 : } \lambda_1 = 2,35 \quad \Rightarrow \quad \text{CONTR} = P_i \frac{\|x_i\|^2}{\lambda_1}$$

$$\text{Axe 2 : } \lambda_2 = 0,65 \quad \Rightarrow \quad \text{CONTR} = P_i \frac{\|x_i\|^2}{\lambda_2}$$

→ La contribution **CONTR** des axes

	AXE 1	AXE 2	CONTR AXE 1	CONTR AXE 2
Pays A	-0,46	-1,22	<b>0,02</b>	<b>0,57</b>
Pays B	-1,09	-1,46	<b>0,12</b>	<b>0,82</b>
Pays C	0,48	1,4	<b>0,02</b>	<b>0,75</b>
Pays D	0,99	1,2	<b>0,1</b>	<b>0,55</b>

**Calcul**

$$(Axe 1 ; Pays A) \text{ CONTR} = \frac{1}{4} * \frac{\|0,46\|^2}{2,35} = 0,02$$

$$(Axe 1 ; Pays B) \text{ CONTR} = \frac{1}{4} * \frac{\|1,09\|^2}{2,35} = 0,12$$

$$(Axe 2 ; Pays A) \text{ CONTR} = \frac{1}{4} * \frac{\|1,22\|^2}{0,65} = 0,57$$

$$(Axe 2 ; Pays B) \text{ CONTR} = \frac{1}{4} * \frac{\|1,46\|^2}{0,65} = 0,82$$

→ La contribution **CONTR** des axes

	AXE 1	AXE 2	CONTR AXE 1	CONTR AXE 2
Pays A	-0,46	-1,22	<b>0,02</b>	<b>0,57</b>
Pays B	-1,09	-1,46	<b>0,12</b>	<b>0,82</b>
Pays C	0,48	1,4	<b>0,02</b>	<b>0,75</b>
Pays D	0,99	1,2	<b>0,1</b>	<b>0,55</b>

**Interprétation**

- Pour Axe 1, le point déterminant est 0,12 (12%)
- Pour Axe 2, le point déterminant est 0,82 (82%)

## **METHODES ET TECHNIQUES D'ETUDES DES ATTITUDES ET DES OPINIONS)**

### **I. Le sondage, les possibilités et les limites**

Comme tout outil de mesure, le sondage offre des possibilités et des limites. Un usage approximatif ou improvisé du sondage peut alimenter l'idée un peu facile voulant que « l'opinion publique n'existe pas » ou qu'on puisse « faire dire n'importe quoi aux chiffres », ce qui est parfois vrai des études mal conçues. On doit éviter de laisser entendre que tout ce qui se chiffre renvoie forcément à une vérité. On peut assez facilement être mystifié par l'idée que la vie sociale peut être mesurée avec la précision qu'offrent les sciences de la nature, ce qui est rarement le cas. En effet, contrairement aux réalités physiques qu'on suppose stables, la société génère continûment son propre sens, ce qui rend improbable l'idée d'une connaissance définitive de la réalité sociale. Les acteurs sociaux construisent leur propre interprétation du monde. Cela étant, à l'intérieur d'un contexte donné, le recours à la statistique permet d'atteindre un niveau de validation et de généralisation que ne peuvent atteindre d'autres méthodes d'enquête.

De façon caricaturale, on a pu prétendre que l'analyse quantitative était fondée sur l'idée d'une société prédéfinie qui impose à ses membres leurs croyances et leurs pratiques : références, institutions, rôles et fonctions. Le sondage s'appuierait dans cette perspective sur une conception socio déterministe du monde. Celle-ci s'opposerait à une approche plus interactionniste de l'action sociale, qui postule au contraire que les acteurs génèrent quotidiennement le sens de leur vie personnelle et de leur activité : système de sens et logiques d'action auxquels seule la recherche qualitative donnerait accès. S'opposeraient ainsi de façon mécanique des approches plus explicatives et compréhensives du monde social. Cette dichotomie renvoie cependant à une opposition qui ne rend compte ni des réalités collectives, ni des destinées individuelles, ni de la dialectique



des rapports sociaux. Cela étant, un certain nombre de postulats sous-tendent l'analyse quantitative et, plus particulièrement, la pratique du sondage en tant que méthode de recherche.

## **II. Quelques postulats**

Le recours au sondage est fondé sur l'idée que dans un contexte ou un champ social donné des récurrences sont observables. On entend par là que certaines pratiques, certaines attitudes et certaines opinions sont partagées. Ce postulat est aussi vrai pour l'étude d'une collectivité nationale que pour un sous-groupe ou une catégorie particulière à l'intérieur de cette société. Ainsi, les données de l'enquête sociale générale, périodiquement menée par Statistique Canada, démontrent que sur certains aspects les citoyens québécois entretiennent un point de vue différent de ceux d'autres provinces<sup>8</sup>. Cela étant, à l'intérieur de la société québécoise elle-même, certains groupes se distinguent les uns des autres. On a par exemple constaté depuis longtemps que les citoyens les moins scolarisés, et dont la condition sociale est généralement plus précaire, entretiennent des attentes plus élevées à l'égard de la justice et sont, par voie de conséquence, plus critiques à l'égard du système de justice.

L'analyse quantitative postule également que les faits sociaux sont en lien les uns avec les autres et renvoient à des logiques plus générales. Ainsi distingue-t-on souvent les variables dites indépendantes (cf. le genre, l'âge, le revenu ou la scolarité) des variables dépendantes (les attentes des répondants en matière de justice ou leur confiance à l'égard des juges ou des praticiens). Entre ces variables dépendantes et indépendantes se glissent souvent des variables dites intermédiaires (cf. l'expérience juridique des citoyens, la nature de cette expérience). Les rapports entre ces variables sont d'abord analysés dans le cadre de croisements les mettant en relation les unes avec les autres (on parle alors d'analyse bivariée). On sait par exemple que l'expérience juridique des répondants (on parle parfois également d'informateur ou de participant au sondage) est fonction de leur âge et que moins un individu est âgé, plus mince

sera vraisemblablement son expérience juridique. Plusieurs variables peuvent cependant se trouver combinées (par exemple la scolarité, le revenu, le genre et le rapport au droit et à la justice). On parle alors d'analyse multivariée. C'est notamment le cas des analyses de variance et de l'analyse multifactorielle de correspondance.

### **Qu'est-ce qui se mesure ?**

Comme on l'a dit, contrairement à la distinction stricte qu'on propose souvent entre méthodes quantitative (de nature plus factuelle) et qualitative (plus centrée sur l'étude des représentations sociales), le sondage peut proposer des mesures de faits (l'expérience, l'âge, la pratique religieuse, etc.), des mesures d'opinions (ce que l'informateur croit vrai ou faux, bon ou mauvais) ou des mesures d'attitude, c'est-à-dire la propension de l'informateur à réagir d'une façon ou d'une autre dans un contexte particulier. Dans tous les cas, il s'agit toujours de perceptions. Ainsi, un répondant indiquant avoir été victime de discrimination au cours des deux dernières années peut avoir vécu cette expérience il y a quatre ans, mais en avoir gardé un souvenir vif. Ces imprécisions s'atténuent cependant en fonction du nombre de répondants interrogés et les règles de la statistique permettent un certain contrôle de ces variations (marge d'erreur, taux de signification, mesure de dispersion intergroupes, etc.). Cela étant, on privilégiera généralement la mesure des faits sur celle des attitudes et la mesure des attitudes sur celle des opinions, étant entendu que ces variables sont généralement reliées les unes aux autres. Une longue fréquentation des données de sondage révèle la cohérence des individus.

Dans tous les cas, il faut éviter d'exiger des informateurs qu'ils décrivent leur réaction à l'égard d'une situation atypique, inusitée ou invraisemblable, qu'ils prennent position sur des problèmes nécessitant des compétences spécialisées qu'ils n'ont pas ou sur des enjeux sans importance réelle pour eux. Vous éviterez ainsi de créer artificiellement des faits sociaux. Cette question est abordée ultérieurement dans ce texte.

#### **IV. Population et échantillon**

Il arrive, mais c'est assez rare, qu'une étude puisse être conduite auprès de tous les membres d'une population donnée : les étudiants d'une même classe, les membres d'une petite association ou d'une même organisation. De façon plus générale cependant, lorsque la population étudiée est importante, un tel sondage ne peut être réalisé qu'à partir d'un échantillon tiré de cette population.

Idéalement l'échantillon doit être représentatif de la population étudiée. De manière à s'assurer de cette représentativité, il doit être constitué de façon aléatoire, c'est-à-dire de manière à ce que chaque membre de cette population ait une chance égale et non nulle de participer à l'étude. À partir d'une liste ordonnée, formée des membres de cette population (numérotés 1 à N), on peut établir un tel échantillon en recourant à des sites Internet établissant de façon aléatoire une suite de nombres situés à l'intérieur de cette fourchette.

Plus votre échantillon est important, plus la marge d'erreur diminue. Ainsi, un échantillon de 1 000 répondants présente une marge d'erreur de 3,16 %, 19 fois sur 2012. Il s'agit de l'échantillon privilégié par les firmes de sondage. Il constitue pour cette raison un standard dans le domaine des sondages menés auprès de la population. Il est parfois judicieux de fonder votre étude sur un échantillon plus important. C'est notamment le cas lorsqu'on prévoit analyser la réalité de segments très minoritaires au sein de la population étudiée.

En contrepartie, si votre échantillon est très faible, il faut évaluer sérieusement l'intérêt de mener une enquête par sondage. Ainsi, on peut s'interroger sur la précision d'un sondage mené auprès de 250 répondants au sein d'une population importante et dont la marge d'erreur est de plus de 15 %.

Bien que votre échantillon soit établi sur une base aléatoire, il arrive que les caractéristiques des répondants rejoins ne correspondent pas exactement à celui de la population étudiée. L'échantillon doit alors être redressé pour tenir compte de ces

disproportions. On peut pondérer l'échantillon en tenant compte des caractéristiques de la population en s'appuyant sur le recensement canadien le plus récent. Cette opération est généralement réalisée à l'aide d'une variable de pondération tenant compte de l'âge, du sexe et de la région des répondants. En diminuant le poids de certains répondants et en augmentant le poids de certains autres, on rétablit la correspondance entre l'échantillon et la population étudiée.

### **La construction d'un questionnaire**

Toute étude quantitative doit être fondée sur une stratégie d'analyse exigeante. Il s'ensuit que le chercheur doit savoir par avance quel usage il entend faire de chaque question posée, c'est-à-dire de chaque variable qu'il entend étudier.

Ainsi, contrairement à plusieurs projets de nature qualitative qui comportent une part d'exploration nécessaire et qui exigent, une fois l'enquête réalisée, un très important travail d'analyse, le travail qui précède l'établissement d'un questionnaire constitue la plus grande part de votre recherche. Il y a très peu de place ici pour l'improvisation, compte tenu des coûts généralement associés aux enquêtes de ce type. Ce qu'on appelle parfois fishing expedition est absolument à proscrire. Les hypothèses que vous entendez valider doivent l'emporter sur votre simple curiosité. L'idée qu'une question improvisée sur un coin de table vous ouvrira la porte d'une découverte inédite est tout simplement mal fondée, même s'il arrive que les réponses contre-intuitives recueillies lors d'une enquête quantitative ouvrent la porte à un nouveau questionnement. Dans ce domaine comme dans beaucoup d'autres, c'est en cherchant à comprendre une chose qu'on en découvre une autre. Mais il faut d'abord savoir ce qu'on cherche. S'agissant du sondage, aucune étude ne peut être conçue sans un fort travail théorique préalable.

Il faut d'abord définir les variables qu'on entend mesurer et établir le plus précisément possible les dimensions susceptibles de préciser cette mesure. Évaluer la confiance des citoyens dans le système de justice est possible dans le cadre d'une question simple :

Faites-vous très confiance, plutôt confiance, peu ou pas des tout confiances au système de justice ? Mais cette mesure mérite généralement d'être précisée à l'aide d'autres questions portant sur la légitimité de l'institut judiciaire, l'impartialité des juges, l'équité des jugements, l'accessibilité du service juridique, la transparence du processus judiciaire, etc. Ces dimensions ont souvent fait l'objet d'études antérieures et sont parfois l'objet de cadres théoriques complets auxquels vous pouvez puiser. Là comme ailleurs, il faut éviter de « partir à la pêche ». C'est généralement une erreur de supposer qu'on est le premier à s'intéresser à une question.

De même, il faut, chaque fois que la chose est possible, exploiter les questions et les échelles établies dans le cadre d'enquêtes antérieures. Leur intérêt a souvent déjà été démontré et le recours à ces variables permet d'envisager une éventuelle comparaison des données. Tout projet de sondage exige encore ici un bon inventaire de la littérature.

Il peut cependant arriver que vous deviez définir vos propres variables et, par extension, forger vos propres questions. On aborde plus loin les règles qui doivent guider leur composition. Dans tous les cas, il faut éviter d'improviser. Chaque question doit idéalement être fondée sur une hypothèse, c'est-à-dire sur la mise en relation éventuelle de cette variable avec une autre. Bien qu'une exploitation plus inductive des données reste possible, votre sondage doit d'abord être construit sur un plan d'analyse prédéfini et s'appuyer sur un solide cadre théorique.

La plupart des projets de questionnaire sont trop longs. Deux facteurs doivent vous guider : le risque d'abandon et les coûts de l'enquête. Un questionnaire standard, administré au téléphone, ne devrait jamais dépasser les 20 minutes, même lorsque le sujet suscite l'intérêt de vos informateurs. Autrement, les répondants mettent fin à l'entretien avant la fin du questionnaire, et ce, quelles que soient les modalités de la cueillette : papier, Internet, face-à-face, etc. Il est donc préférable d'envisager un questionnaire court et très stratégiquement

conçu. Si ces arguments ne vous ont pas convaincu, rappelez-vous simplement que, dans le cadre d'un sondage téléphonique administré par une firme spécialisée, chaque minute d'interview représente des coûts de 1 000 \$. La retenue vous apparaîtra tout à coup comme une vertu cardinale.

Tout questionnaire doit être précédé d'un très court préambule présentant le sujet de la recherche. Il est contre-indiqué d'exposer vos hypothèses en détail, sauf si vous tenez absolument à biaiser votre échantillon. Il faut rappeler à l'informateur que les données recueillies sont confidentielles, qu'ils peuvent refuser de répondre à une question, qu'il ne s'agit pas d'un test et que toutes les réponses sont valables, puisque c'est leur opinion et leur expérience qui comptent. Dans tous les cas, soyez bref. Indiquez, lorsque c'est le cas, que le projet est mené dans le cadre d'une recherche universitaire. S'informer des habitudes, des valeurs, des pratiques ou des opinions de quelqu'un d'autre constitue toujours un exercice délicat. Introduisez le tout par les questions les plus simples et les plus générales. Vous passerez graduellement à des sujets plus précis concernant leur opinion personnelle, leurs attitudes, puis leurs expériences. C'est la raison pour laquelle les questions de nature sociodémographique sont toujours posées à la toute fin du questionnaire : sexe, revenus personnels, date de naissance (et non âge), niveau de scolarité, etc. Il s'agit d'informations que vos informateurs hésiteront à fournir au départ.

Le motif d'abandon le plus courant est la redondance des formes. Une succession de questions structurées sur le même motif ennueie rapidement l'informateur. Au contraire, il faut que, par la succession des questions, le répondant soit toujours interpellé. Les premières questions, on l'a dit, sont toujours les plus simples. Elles ne visent souvent qu'à introduire le sujet. Il ne faut pas hésiter à recourir alors à des échelles binaires. Par exemple : Personnellement, croyez-vous que nous vivons dans une société juste ? (Oui, Non). On peut par la suite exploiter des échelles plus complexes : Diriez-vous

que vous êtes très d'accord, plutôt d'accord, plutôt en désaccord ou tout à fait en désaccord que les juges cherchent généralement à être justes ? Comme on l'a indiqué, la structure des questions doit varier de manière à éviter un effet de répétition.

Au cours du processus conduisant à la composition d'un questionnaire, on regroupe souvent les questions par thèmes. Chacun de ces blocs thématiques renvoie à des variables et des dimensions spécifiques : conscience du droit, expérience judiciaire, connaissance des lois, confiance à l'égard du système de justice, sentiment de compétence à l'égard du droit, capacité financière à recourir aux services d'un avocat, attentes à l'égard des juges ou des praticiens, etc. Il faut parfois éviter de regrouper dans un même ensemble toutes les questions reliées à une même variable (par exemple à l'un ou l'autre de ces thèmes). Panacher les questions est une façon simple de tromper la lassitude de l'informateur. On évitera également alors les « effets socratiques ». On entend par là le risque de voir un informateur – longuement questionné sur un thème – se forger de solides convictions sur des enjeux qui, en définitive, l'intéressaient assez peu. On risque alors de créer de toute pièce des compétences ou des opinions qui n'existent tout simplement pas. Dans tous les cas, il faut se rappeler que les enquêtes que nous menons n'ont pas pour but de former ou d'informer l'informateur sur une matière particulière, mais de connaître ses opinions, ses attitudes et ses pratiques réelles. Pour la même raison, il faut éviter de placer les questions dans un ordre qui oblige l'informateur à répondre systématiquement dans le même sens. Ainsi, si vous posez à la suite l'une de l'autre la question Considérez-vous très utile, plutôt utile, peu ou pas du tout utile que les citoyens connaissent les lois qui régissent la vie en société suivie de celle demandant si Personnellement, ils considèrent connaître très bien, plutôt bien, plutôt mal ou très mal ces lois ?, la réponse à la seconde de ces questions se situera forcément dans le droit fil de la première. Ces questions doivent être placées à des endroits différents du

questionnaire et, dans tous les cas, il faut vous poser à vous-même cette question cruciale : Avez-vous beaucoup, assez, un peu ou pas du tout besoin de cette information ? Il faut donc savoir se sonder soi-même. Étant établi qu'il est difficile de conclure quoi que ce soit à partir d'une seule mesure, il ne faut pas se priver de poser la même question de diverses façons tout au long du questionnaire. Les individus tiennent des positions beaucoup plus cohérentes qu'on le croit, contrairement à l'opinion paternaliste qu'on rencontre dans le milieu académique. De façon générale, les répondants ont les attitudes de leurs opinions. Ainsi, la confiance plus ou moins grande que les citoyens placent dans les tribunaux est généralement corrélée avec celle qu'ils mettent dans la compétence et l'impartialité des juges. Ils sont de même (et pour les mêmes raisons) plus ou moins disposés à voir leurs problèmes pris en charge par le tribunal, en cas de besoin ou de nécessité. Valider les réponses reçues les unes par les autres vous permet de confirmer la validité de votre terrain. En contrepartie, il peut arriver que certaines incohérences surprennent. Il faut alors prendre le temps d'évaluer si ces contradictions apparentes ne sont pas tout simplement le fait d'un mode de penser qui, bien que tout à fait cohérent, diffère de celui qu'on prêtait à nos informateurs. Vous venez alors peut-être de mettre le doigt sur une réalité importante sous-estimée par la littérature.

Dans tous les cas, le questionnaire doit faire l'objet d'un pré-test. Comme on le verra maintenant, il est essentiel de faire passer le questionnaire auprès de quelques dizaines de sujets pour valider la formulation et la compréhension de vos questions.

### **Terrains et cueillette**

Les premiers sondages menés par Gallup dans les années 1930 étaient réalisés par des intervieweurs spécialisés dans le cadre d'entrevues personnelles (face-à-face). Ils circulaient d'une maison à l'autre suivant un plan d'échantillonnage des adresses civiques, des quartiers et des villes qui garantissait le caractère aléatoire de la cueillette. L'arrivée du téléphone dans tous les ménages a par la suite



fait du sondage téléphonique le grand standard dans le domaine de la recherche. Tirés aléatoirement du bottin téléphonique de chaque région, ou (plus tard) générés informatiquement, les numéros de téléphone retenus fournissaient un échantillon également représentatif. Le sondage postal a longtemps également été utilisé, malgré des taux de réponse souvent assez faibles. Aujourd'hui, Internet offre d'autres possibilités.

Les travaux réalisés par les spécialistes de la recherche quantitative ont depuis longtemps révélé que dans la mesure où ces enquêtes sont menées auprès d'un échantillon équivalent, tiré aléatoirement, ils fournissent des données d'égale valeur. Il s'agit surtout de choisir la modalité la plus adaptée à la nature de votre enquête. S'agissant de sondages menés auprès de l'ensemble de la population adulte du Québec (ou du Canada), la communauté scientifique favorise encore largement le sondage téléphonique parce que c'est le seul à être fondé sur un échantillon aléatoire. Deux difficultés se présentent cependant. La première tient au taux de réponse toujours plus faible de ces enquêtes. Obtenir des taux de plus de 40 % aujourd'hui n'est possible que si on consent à poursuivre l'enquête sur de très longs mois. Sur des questions brûlantes d'actualité, un rapide changement de contexte (la publicité entourant un événement relié au sujet de l'enquête, par exemple) risque toujours de venir contaminer l'échantillon. Dans le cadre d'enquêtes menées sur une période de deux ou trois semaines, les taux de réponse ne sont tout au plus que de 15 à 20 % aujourd'hui. La deuxième difficulté tient aux budgets associés à ces sondages. Une enquête menée auprès de 1 000 répondants représente des coûts de 25 000 à 30 000 \$. En contrepartie, la plupart des firmes de sondage conduisent périodiquement des enquêtes omnibus. Elles sont composées de questions issues de diverses sources : universités, entreprises, gouvernement. On peut y placer, pour un coût raisonnable, de 7 à 8 questions susceptibles de répondre aux besoins d'une recherche exploratoire. Le fichier fourni par le prestataire de

service comprend alors, en plus des réponses à nos propres questions, les variables socio-économiques qu'on retrouve dans tous les sondages et une variable de pondération tenant compte de la nature de l'échantillon.

Le sondage Internet peut être envisagé dans certaines situations particulières, notamment si l'enquête est menée auprès des membres d'une population dont vous possédez toutes les adresses Internet. C'est notamment le cas lorsque l'enquête vise à recueillir l'opinion ou l'expérience d'un groupe relativement homogène : membres d'un ordre professionnel, étudiants d'une certaine cohorte, diplômés d'un département ou d'une faculté, professionnels œuvrant dans les différents établissements d'une commission scolaire ou d'un CIUSSS, membres de grandes associations, bénéficiaires d'un service spécifique ou consommateur de certains biens, etc. Pour des montants très raisonnables, il est alors possible de mener une enquête assez représentative de la population étudiée en vous assurant de la complicité et de la participation des responsables de ces institutions ou de ces réseaux et en faisant les relances nécessaires pour vous assurer d'obtenir un bon taux de réponse. Certains logiciels spécialisés (Survey Monkey et Limesurvey, par exemple) permettent la mise en ligne de votre questionnaire et la production d'un chiffrier susceptible d'être exploité par un logiciel de traitement de données quantitatives, ou de façon plus artisanale par un logiciel Excel.

Certaines firmes de sondage proposent la conduite d'enquêtes menées par Internet auprès d'un sous-échantillon de la population – lui-même formé de plusieurs milliers de répondants potentiels – souvent présenté comme représentatif. Les membres de ces réseaux de répondants sont donc périodiquement appelés à participer à l'une ou l'autre des enquêtes menées par ces firmes et sont récompensés de diverses manières pour leur contribution. Il ne s'agit évidemment pas d'un échantillon aléatoire. La représentativité de ces enquêtes est par conséquent douteuse et n'est appuyée sur aucune marge d'erreur. On

peut par ailleurs supposer que l'habitude de répondre systématiquement à ces enquêtes fait de ces participants des « répondants professionnels » dont le profil correspond de moins en moins à celui du citoyen ordinaire. Finalement, la participation à des activités de ce type exige un bon niveau de littéraires, ce qui pose problème au sein d'une population comptant 53 % d'analphabètes fonctionnels.

Les sondages papier représentent souvent une opération lourde si l'enquête vise un grand nombre de répondants. Ils se justifient notamment dans le cadre d'enquêtes plus modestes, menées de façon continue auprès des bénéficiaires ou des usagers d'un service offert par un organisme communautaire ou un service gouvernemental. Ils sont cependant de plus en plus en rares. Les sondages « papier » menés par courrier postal sont par ailleurs presque totalement disparus.

C'est également le cas des enquêtes menées en personne (face-à-face), qui permettent rarement la constitution d'un échantillon représentatif, sauf s'il s'agit d'interroger un nombre limité de personnes regroupées pour un temps dans un certain lieu. Cela étant, il peut être nécessaire de procéder de cette manière dans le cadre d'enquêtes menées auprès de personnes peu scolarisées ou dont la langue d'origine n'est ni le français ni l'anglais et dont le témoignage ne peut être reçu qu'oralement. Dans plusieurs cas, la tablette peut remplacer le papier. Sur une autre échelle, notamment dans le cas d'enquêtes menées auprès de la population en général, il faut éviter les sondages « de trottoir » ou « de centres d'achat ». Ils sont sans aucune valeur scientifique, même sur une base exploratoire.

### **La formulation des questions**

La rédaction d'une bonne question de sondage est à la fois un art et une science. Il s'agit d'un savoir-faire qui ne peut s'acquérir que dans la foulée d'une succession d'erreurs. Quelques règles doivent cependant vous servir de guide. La première, et la plus importante, consiste à bien distinguer la question que vous vous posez de celle

que vous poserez. Trop de chercheurs espèrent une réponse directe à une question sophistiquée construite sur des considérations absolument étrangères aux préoccupations de la population étudiée. C'est tout simplement maladroit. Il faut distinguer vos propres préoccupations de chercheurs de celles, souvent très quotidiennes, de vos informateurs. Pour cette raison, les données de faits sont souvent les plus précises et les plus fiables, même s'il s'agit de perceptions personnelles. La meilleure façon de distinguer les genres est de définir précisément vos variables (ce que vous avez besoin de savoir pour répondre à votre question de recherche) et de n'envisager la rédaction du questionnaire qu'une fois toutes ces variables (et leurs indicateurs) définies. Ce n'est qu'une fois ce tableau dressé que vous pourrez réfléchir à la formulation d'une question dont la fonction n'est pas tant de répondre directement à la question que vous vous posez que de recueillir les informations qui vous permettront d'y répondre.

*Ainsi, ne posez jamais de question du genre : Croyez-vous que l'accès au droit et à la justice soit fonction de la condition sociale des individus ? Demandez plus simplement : Si vous deviez poursuivre quelqu'un ou si vous deviez vous défendre à la cour, auriez-vous personnellement les moyens financiers de le faire ? Puis informez-vous de leur revenu personnel. Vous saurez tout*

Il s'agit ici d'un cas évident, mais il doit en être ainsi de tous vos questionnements. Plusieurs variables nécessitent d'être ventilées en fonction de plusieurs dimensions. Ainsi, demander si vos informateurs ont une expérience de la cour n'est pas suffisant pour circonscrire cette variable. Il faut plutôt demander : *Êtes-vous déjà allé dans une cour, un tribunal ou devant un tribunal administratif en tant que personne impliquée, témoin, observateur ou personne qui en accompagnait une autre ?*, en laissant à l'informateur la possibilité de répondre à plus d'une option. Vous parviendrez alors à tenir compte des différentes dimensions possibles de l'expérience judiciaire des répondants.

### **La concordance des questions et des échelles**

Une difficulté rencontrée dans plusieurs sondages est l'incohérence des options proposées, compte tenu de la nature de ce qu'on entend mesurer. Il arrive ainsi qu'en réponse à des questions de connaissance ou de compétence, les options proposées portent plutôt sur la pratique des informateurs ou sur leur attitude. C'est le cas dans cet exemple déjà cité, concernant la pratique des consommateurs : Oui, et je ne le fais pas (habituellement). On tente encore ici de mesurer deux choses à la fois (une connaissance et une pratique). Par ailleurs la réponse comprend une parenthèse susceptible d'être prise en compte (on mesure alors une attitude) ou non (on mesure alors une pratique). Les questions étant généralement construites au cours d'une assez longue période de gestation, ces glissements sont fréquents et doivent être corrigés.

### **Les questions ouvertes**

Incapables de situer précisément les paramètres d'une question, on peut avoir tendance à laisser certaines questions ouvertes, de manière à ce que l'informateur puisse s'exprimer librement. Cette option ne présente pas toujours le même intérêt.

Dans le cadre d'un sondage téléphonique, c'est à proscrire. Tous ces témoignages seront d'abord transcrits par l'intervieweur qui, par économie, sera inévitablement appelé à les traduire dans ses propres termes. Il s'ensuit toujours une standardisation implicite des commentaires reçus et ceux-ci deviennent alors inutilisables, même une fois codés et regroupés thématiquement. On est en effet alors amené à surcoder des contenus déjà codés (par l'intervieweur). C'est alors aussi coûteux qu'inutile.

Par contre, dans le cadre d'une enquête menée auprès des usagers ou des bénéficiaires d'un service, c'est très utile. Ces commentaires permettent d'améliorer, sinon d'enrichir significativement, les services offerts et permettent la conception de questions plus susceptibles de mesurer l'expérience et les besoins des bénéficiaires et usagers de ces mêmes services.

Il en va de même des sondages réalisés en ligne. Si le sujet de l'enquête présente un intérêt pour ceux qui y participent, on peut véritablement bénéficier de commentaires riches, souvent utiles pour la suite de la recherche. Ainsi, dans une enquête menée auprès des diplômés de la Faculté de droit de l'Université de Montréal, les commentaires rédigés par les anciens de la Faculté ont permis au comité de nuancer de façon intéressante les réponses fournies dans le cadre du questionnaire. Si la plupart de ces diplômés se déclaraient très satisfaits de la formation reçue de la Faculté, les témoignages rédigés « au long » révélaient que cette formation devrait être de nature plus interdisciplinaire aujourd'hui ; un aspect laissé dans l'ombre par l'enquête elle-même.

### **L'analyse des données**

Les stratégies qui doivent guider l'analyse des données de sondage exigeraient un texte spécifique<sup>17</sup>. Indiquons pour l'heure que faute de connaissances suffisantes dans le domaine de la statistique, de nombreux fichiers de sondage restent sous-exploités par les chercheurs. Les équipes de recherche et les institutions intéressées à la réalisation d'enquêtes de ce type s'arrêtent essentiellement aux tendances centrales, c'est-à-dire à l'étude des fréquences statistiques. Ainsi, le fichier une fois pondéré, on se penche spontanément sur les réponses reçues pour chaque question. On insiste sur « les gros chiffres ». Si 90 % des personnes interrogées affirment que les coûts des services professionnels constituent la principale raison qui les forcerait à se présenter seules à la cour, on confirme ce que tout le monde suppose spontanément sans pouvoir vraiment l'affirmer. Il s'agit donc de constats majeurs et validés.

On doit cependant tenir compte également de données plus nuancées. Pour le chercheur, constater que 50 % des citoyens entretiennent une certaine méfiance à l'égard du système de justice est une véritable aubaine. La question est alors de savoir « qui dit quoi ? » L'âge ou l'expérience d'un procès sont-ils des variables indépendantes ou intermédiaires susceptibles d'expliquer ces

variations ? Les femmes et les hommes se distinguent-ils sous cet aspect ? Vos hypothèses de départ vous guideront dans l'analyse de vos données, d'où la nécessité d'établir dès le départ vos variables et vos dimensions, et ce, bien avant de définir vos questions. Un bon plan d'analyse est essentiel à une exploitation systématique de vos données. Les logiciels SPSS (commercial) et R (en libre accès) sont conçus pour faciliter cette exploitation<sup>18</sup>. Si vous en êtes à vos premières armes dans le domaine de l'analyse statistique, il ne faut pas hésiter à vous adjoindre un sociologue ou un politologue rompu à ces opérations. Vous apprendrez beaucoup. Dans tous les cas, il ne faut pas sur interpréter vos données. Vous risqueriez d'y perdre votre crédibilité. Tenir compte des marges d'erreur et plus encore des taux de signification fait partie des règles de l'art. Il ne faut pas exclure de vous inscrire à l'un ou l'autre des cours de recherche statistique proposés par les départements de sciences sociales. C'est à tort que les juristes hésitent à entreprendre ces excursions en terrain étranger qui élargiraient significativement leur intérêt et leur méthode de recherche. Indiquons finalement l'intérêt des analyses de variance qui permettent souvent de repérer et de synthétiser des variables latentes qui, sans avoir fait l'objet d'une mesure spécifique, ressortent de la combinaison de plusieurs variables de l'enquête. L'analyse multi variée – notamment les analyses multifactorielles – permet ainsi d'explorer de façon inductive les données recueillies et de pousser les connaissances vers de nouvelles directions de recherche, une fois tirées (plus déductivement) les premières conclusions de votre enquête.