



CENTRE DE RECHERCHE DMSP

**Petit abrégé pour mieux comprendre  
la notion de méthode expérimentale  
et ses enjeux méthodologiques**

Maud Herbert  
**Cahier n°340**  
Décembre 2004

**Maud Herbert**

Doctorante  
Centre de recherche DMSP  
Université Paris-Dauphine  
Place du Maréchal de Lattre de Tassigny  
75775 Paris Cedex 16  
Tel : 01 44 05 44 59  
Fax : 01 44 05 44 49  
E-mail : [maudherbert@hotmail.com](mailto:maudherbert@hotmail.com)

# **PETIT ABRÉGÉ POUR MIEUX COMPRENDRE LA NOTION DE MÉTHODE EXPÉRIMENTALE ET SES ENJEUX MÉTHODOLOGIQUES**

## **Résumé**

Ce document a pour objectif d'aider le chercheur qui souhaite utiliser un design expérimental dans ses travaux en lui apportant une synthèse des principes fondamentaux qui régissent la méthode expérimentale (et par extension quasi-expérimentale). L'idée est de bien maîtriser les différentes notions de contrôle et les menaces aux validités en distinguant celles qui peuvent être plus ou moins circonscrites par le chercheur de celles sur lesquelles il a peu de prises mais auxquelles il doit néanmoins accorder de l'attention.

**Mots clés** : méthode expérimentale et quasi-expérimentale, principe du contrôle, menaces à la validité.

## **Abstract**

This paper aims at providing a brief overview of the experimental method's principles to researchers who face for the first time the use an experimental design. The purpose is also to guide them to a better understanding of the basic notions of control and threats to validities. Thus, they could really distinguish the threats that can be more or less circumscribed by the researcher from those that remain "out of control" but nevertheless have to be taken into account in the research process.

**Key words** : experimental and quasi-experimental designs, what control means, threats to validities.

## INTRODUCTION

Au cours d'un processus de recherche, après le temps de la conceptualisation théorique vient souvent le moment où la réflexion abstraite doit être confrontée à la « réalité » empirique. Le chercheur aborde alors l'étape de la mesure où il va recueillir des observations sur le terrain et pouvoir tester les relations théoriques qu'il a préalablement établies.

Formulée sous forme d'hypothèses de recherche les relations entre variables peuvent être causales du type « Si X alors Y ». Une relation causale simple peut être la suivante : « si le téléspectateur est de bonne humeur alors son attitude envers la publicité sera favorable ». Mais qu'elles soient simples ou complexes, il n'existe qu'une seule méthode réellement appropriée pour tester les relations causales : la méthode expérimentale.

Ce document récapitule les fondamentaux de l'expérimentation en sciences sociales, il synthétise les règles d'usage de la méthode expérimentale (partie A). Il est important de préciser dès maintenant qu'à l'origine, expérimentation rime avec laboratoire, le milieu fermé idéal garant de la « vraie » expérimentation. Or, pour des questions de réalisme, de moyens et de faisabilité, les chercheurs en sciences sociales et en sciences de gestion sont souvent obligés de tester les relations causales sur le terrain. Si sortir du milieu protégé du laboratoire rend le principe du contrôle plus difficile (Bhaskar, 1975), il n'affranchit pas le chercheur des règles fondamentales de la méthode expérimentale. Parmi ces règles, le principe du contrôle et les différentes validités doivent être pleinement intégrés par le chercheur-expérimentateur, ce document s'attache à définir ces notions (partie B) dont la maîtrise dans le cadre de la méthodologie expérimentale peut aider le chercheur à comprendre qu'il doit faire face à des sources d'erreur sur lesquelles il n'a que peu de prise et qui menacent pourtant sa recherche (partie C).

### A- Les règles d'or de la méthode expérimentale

Ce document a pour but de présenter rapidement les grandes règles qui régissent la méthode expérimentale. Le propos n'est pas de s'appesantir sur la méthode mais d'en rappeler les principes essentiels afin d'aider le chercheur néophyte à comprendre globalement les bases fondamentales, utiles au développement de son futur travail. Pour une présentation plus détaillée mais tout de même succincte, le lecteur pourra, ensuite, se reporter au chapitre 3 (p.99 à 147 écrit par Aronson, Wilson & Brewer) de l'ouvrage « The Handbook of Social Psychology » volume 1 sous la direction de Gilbert, Fiske & Lindzey (1998).

#### 1- Identifier toutes les variables en jeu et formaliser les relations

L'expérimentation permet de tester les relations causales. Le principe fondamental est de faire varier des éléments d'une situation prédéterminée et de mesurer les conséquences de ces variations. Il s'agit de tester des relations du type « si X varie alors Y varie ». Dans cette relation causale simple, si toutes choses restent égales par ailleurs (notion importante sur laquelle nous reviendrons), X est la variable explicative des effets observés sur Y qui est la variable à expliquer.

Dans la terminologie courante, une variable explicative est appelée variable indépendante ou encore variable « *treatment* » (Cook & Campbell, 1979) et la variable à expliquer est une variable dépendante ou encore variable « *outcomes* » (Cook & Campbell, 1979).

Par définition, la variable dépendante est celle qui dépend de la variation d'autres variables. Elle est l'objet de recherche du chercheur qui sélectionne les variables indépendantes considérées comme la cause possible des effets observés sur la variable dépendante.

Une expérimentation implique la variation d'une ou de plusieurs variables indépendantes. La plupart du temps la variation de la variable indépendante résulte d'une manipulation de cette variable, ce qui est l'essence même de l'expérimentation. Cependant certaines variables indépendantes ne sont pas manipulables, il en va ainsi du genre par exemple, dans ce cas, la variable indépendante est dite invoquée. La variable indépendante qui est manipulée ou invoquée comporte donc au minimum deux modalités. Par exemple, la variable humeur peut en posséder deux : mauvaise humeur vs. bonne humeur. Il faut noter que le nombre de modalités se détermine au moment où le chercheur opérationnalise la variable. Par exemple, l'humeur peut prendre plusieurs autres modalités comme humeur maussade, constante ou versatile, tout dépend de la littérature mobilisée et des hypothèses de la recherche. Une variable indépendante pourrait adopter de très nombreuses modalités, par exemple s'il s'agit d'une variable de catégorie socioprofessionnelle (opérationnalisation d'une variable niveau de vie par exemple), la nomenclature de l'INSEE prévoit 42 catégories non agrégées (INSEE, 2003). En fait, la décision revient au chercheur. Dans le cadre de l'expérimentation, il lui incombe de déterminer le nombre optimal de modalités car, comme nous le verrons par la suite, leur quantité détermine le nombre de cellules du plan expérimental.

Ces deux types de variables sont les variables fondamentales directement identifiables de l'expérimentation mais il existe un autre type de variable à prendre en compte car elles peuvent influencer les effets mesurés. Il s'agit des variables dites « externes » ou variables contextuelles.

Le chercheur doit en effet discerner les variables qui pourraient exercer une influence sur la relation causale et dont la présence pourrait conduire à une autre interprétation des effets observés sur la variable dépendante. Par exemple, le jugement de l'individu envers la publicité en général pourrait être une variable externe venant influencer ou occulter une relation causale entre variable indépendante (l'humeur) et dépendante (attitude envers la publicité).

Ces variables, dites aussi variables parasites, doivent être contrôlées, c'est à dire soit mesurées puis comprises dans l'analyse statistique soit maintenues constantes ou homogènes de façon à annuler leur influence sur la variable dépendante (pendant l'expérimentation ou post hoc). La relation de cause à effet est moins perturbée dans le cas d'un maintien constant (Lynch, 1982). L'évocation des variables de contexte et de leur contrôle met en évidence une des caractéristiques voire la caractéristique la plus intéressante de la méthode expérimentale : la possibilité d'isoler la relation de cause à effet dans un milieu « fermé », c'est à dire d'annuler les interférences des variables externes pour ne mesurer que le lien entre variable indépendante et dépendante. Il est évident qu'en recherche comportementale il existe un grand nombre de variables exogènes et qu'il est difficile voire impossible de toutes les prendre en compte (Cronbach, 1975 ; Calder, Phillips & Tybout, 1981, 1982). Cependant, l'identification préalable des variables externes est importante et doit faire l'objet d'une attention particulière de la part du chercheur. Il doit hiérarchiser l'importance de ces variables parasites pour essayer d'éliminer au mieux leur influence (Aronson et al., 1998). Cronbach (1975) conseille de mesurer le plus de variables parasites possibles pour ne pas se retrouver démuné au moment de l'analyse, c'est ce qu'il appelle la prise de mesure « hors hypothèses ».

Une fois que l'ensemble des variables est identifié, il convient de formaliser les relations entre ces variables, souvent sous la forme d'un modèle qui permet d'avoir une vue globale des liens entre variables et de spécifier leur nature. En effet, certaines variables peuvent être médiatrices, modératrices ou quasi-modératrices. Rappelons qu'une variable médiatrice est

une variable par laquelle une variable indépendante va influencer la variable dépendante : « X1 influence X2 qui influence Y » où X2 est la variable médiatrice. Par exemple, la variable satisfaction peut être une variable médiatrice entre la qualité perçue et l'intention de ré-achat. Une variable modératrice est une variable qui module l'effet de la variable indépendante sur la variable dépendante dans un sens positif ou négatif. Pour conclure qu'il y a un effet modérateur, un effet d'interaction entre la variable indépendante et la variable modératrice doit exister. En d'autres termes, l'action de chaque variable indépendante sur la variable dépendante doit être différente selon les modalités de la variable modératrice\*. Il peut aussi exister un effet simple de la variable modératrice sur la variable dépendante, dans ce cas, il convient de parler de variable quasi-modératrice. Pour une explication détaillée et illustrée des processus modérateurs et médiateurs, le lecteur peut consulter Chumpitaz Caceres et Vanhamme (2003).

Une fois que toutes les variables sont spécifiées, que la nature de chacune est définie et que les liens qui les unissent sont identifiés sous la forme d'hypothèses rédigées, le chercheur passe à une étape cruciale de l'expérimentation : le plan expérimental. Le plan expérimental est soumis à des obligations rigoureuses qu'il convient de rappeler car le chercheur ne peut commencer à se soucier de la validité des effets mesurés que si ces règles sont respectées.

## **2- Monter un plan expérimental en examinant l'orthogonalité des variables indépendantes, vérifier si la randomisation est possible.**

Le plan expérimental, aussi appelé plan d'expérience, formalise le dispositif de l'expérimentation. Il définit quelles variables indépendantes sont sélectionnées et comment elles sont manipulées, combien de variables il faut contrôler et comment le faire.

Au moment de construire son plan d'expérience, le chercheur doit avoir déterminé quelles unités expérimentales il veut interroger. L'unité expérimentale est le spécimen qui représente le mieux la population à laquelle il souhaite pouvoir extrapoler les résultats de sa recherche. Par exemple, en marketing, une unité expérimentale peut être un point de vente ou un consommateur lambda. Le plan expérimental permettra de déterminer le nombre d'unités expérimentales à soumettre à l'expérimentation.

Enfin, le plan expérimental permet surtout de visualiser l'affectation des traitements aux unités expérimentales, à savoir qui va être exposé à telle ou telle condition expérimentale (exemple de traitement dans une condition expérimentale : un scénario susceptible de manipuler l'humeur dans un sens positif). Le terme « visualiser » est choisi à dessein car le plan d'expérience se matérialise souvent sous la forme d'un tableau récapitulant les différentes modalités de la ou des variables indépendantes, chaque case du tableau représentant une condition expérimentale déterminée.

Le plan d'expérience existe dès qu'une seule variable indépendante est manipulée, mais cette situation est rare car souvent deux ou plusieurs variables indépendantes peuvent varier simultanément. Pour tester leurs effets sur la variable dépendante et les effets d'interaction que nous avons évoqués précédemment, il faut utiliser un plan factoriel. Le plan factoriel complet résulte du croisement entre l'ensemble des modalités de toutes les variables

---

\* Si l'action de la variable indépendante est inversée selon les modalités de la variable modératrice, l'interaction est dite croisée. Si l'action de la variable indépendante a plus ou moins d'effet selon les modalités de la variable modératrice, l'interaction est dite convergente. Enfin si l'effet de la variable indépendante peut s'annuler pour une des modalités de la variable modératrice, il s'agit d'un effet de suppression. (*suppressor*)

indépendantes manipulées, ce qui définit tous les traitements testés dans l'expérimentation. Par exemple, si le modèle comprend 3 variables indépendantes V1, V2, V3 et que chacune a deux modalités, le plan factoriel complet définira  $2 \times 2 \times 2$  conditions expérimentales soit 8 conditions différentes.

Il est évident que la quantité de conditions et donc d'unités expérimentales est directement proportionnelle au nombre de variables indépendantes manipulées et aux modalités associées à ces variables, par conséquent un plan factoriel complet peut vite s'avérer « lourd » à mettre en place. D'où la possibilité d'utiliser des plans factoriels fractionnaires qui diffèrent du plan complet en affectant plusieurs traitements à une même unité expérimentale.

En plus des variables indépendantes, le plan expérimental peut prendre en compte les variables externes qu'il faut contrôler. Selon le nombre de facteurs contrôlés (un, deux, trois ou plus de trois) le plan expérimental est respectivement qualifié de plan aléatoire en bloc (complet ou incomplet), de plan en carré latin, de plan en carré gréco-latin, de plan en carré hyper-gréco-latin. Pour une présentation rapide de ces différents plans dans le cadre de la recherche en marketing, le lecteur intéressé se reportera à l'ouvrage *Market* (Evrard, Pras, Roux et al., 1997, p. 216 à 223).

La condition *sine qua non* du bon déroulement d'une expérimentation et de la validité d'un plan factoriel quel qu'il soit, est que l'ensemble des variables indépendantes soient réellement indépendantes, c'est à dire qu'elles n'aient pas d'influence les unes sur les autres afin de pouvoir isoler les effets que chacune d'entre elles a sur la variable dépendante. Si les variables indépendantes sont corrélées, il est impossible d'envisager de les manipuler dans un même plan expérimental (Keppel, 1991). L'orthogonalité des variables est une condition importante que tout chercheur doit vérifier avant de construire un plan factoriel.

Même si cela n'est pas toujours possible, l'affectation aléatoire des unités expérimentales est une autre des règles d'or de l'expérimentation car elle garantit un meilleur contrôle des effets indésirables.

Cook & Campbell (1979) utilisent un exemple significatif pour expliquer les bienfaits de l'affectation aléatoire des unités expérimentales : soit une expérimentation agricole où le but est de tester l'efficacité de nouvelles techniques. Dans ce cas, les variables indépendantes peuvent être par exemple le type de graines ou les engrais utilisés. La variable dépendante est la productivité à l'hectare et les unités expérimentales des parcelles de terrain de la même taille dans une ferme. Comment peut-on être sûr que la hausse de la productivité est bien due à l'utilisation du nouvel engrais si toutes les parcelles traitées avec cet engrais sont situées au sud de l'exploitation et donc plus ensoleillées que les unités traitées avec un ancien engrais ?

Un des moyens d'éviter ce type de question est d'affecter les parcelles de façon aléatoire au traitement par le nouvel engrais (ce qui est appelé randomisation). Les parcelles traitées avec le nouvel engrais et celles non-traitées seront donc dispersées dans l'exploitation et, s'il y a suffisamment de parcelles, on peut dire qu'une parcelle moyenne ayant subi le traitement est comparable à une parcelle moyenne n'ayant pas subi le traitement. Par conséquent, les résultats ne seront pas liés à la localisation géographique, à la nature du sol, au degré d'ensoleillement. Il est évident que chaque parcelle de terrain est différente d'une autre, comme chaque être humain interrogé est différent d'un autre, le principe de la randomisation est simplement de lisser ces différences dans une moyenne pour permettre la comparaison de deux états et l'imputation à une cause. L'affectation aléatoire est la condition des « vraies » méthodes expérimentales, elle apporte une réponse précise à la question « est-ce que le traitement a des effets ou non ? » (Cronbach, 1988). Pour une présentation complète de la théorie de l'affectation aléatoire et des ses applications, le lecteur peut consulter les pages 247 à 277 de l'ouvrage *Experimental and Quasi-experimental Designs* (Shadish, Cook & Campbell, 2002).

Cependant, la randomisation parfaite, en suivant une table des nombres au hasard n'est pas toujours possible surtout dans les recherches de terrain en sciences sociales, on parle alors de quasi-expérimentation. Les situations où la randomisation n'est pas possible sont fréquentes, par exemple lorsque les variables indépendantes sont invoquées. Le temps est aussi un enjeu important, la randomisation consomme beaucoup de temps et d'effort, si la nature des variables change avec le temps, l'affectation aléatoire n'est même pas conseillée (Cook & Campbell, 1979 ; Shadish, Cook & Campbell, 2002). En fait le choix de la randomisation résulte d'un arbitrage subtil entre l'accessibilité du terrain et ses caractéristiques, le temps imparti pour mener la recherche, les coûts humains et financiers, les questions éthiques liées à la manipulation et surtout les enjeux liés aux résultats de l'expérimentation. Le lecteur souhaitant approfondir ces points se reportera aux pages 341 à 386 de l'ouvrage *Quasi-experimentation ; design & analysis issues for field settings* (Cook & Campbell, 1979).

Il faut noter que ces auteurs rappellent que l'affectation aléatoire n'est pas une règle du tout ou rien. Le chercheur en sciences sociales connaît plus ou moins son terrain et s'il ne peut randomiser dans les règles de l'art (en suivant une table des nombres au hasard) il peut s'arranger pour choisir les unités qui reçoivent le traitement. Par exemple, randomiser 1000 étudiants est compliqué, affecter aléatoirement 50 groupes de 20 étudiants devient plus accessible. Par conséquent, même dans des conditions difficiles, il est possible de s'acquitter d'un certain degré de contrôle qui améliorera toujours la qualité des inférences (Cook & Campbell, 1979).

Nous avons déjà utilisé plusieurs fois le terme « contrôle », il est indissociable de la méthode expérimentale. Il est impossible pour le chercheur de comprendre le principe des menaces à la validité et donc le problème posé par les artéfacts de la demande s'il n'a pas compris pleinement le sens du terme contrôle et les différentes facettes de la validité.

## **B- Comprendre le principe du contrôle et de la validité pour prendre la mesure des menaces qui peuvent peser sur les résultats d'une expérimentation**

### **1. Les modes de contrôle**

Dans le langage expérimental, contrôler a un sens général qui veut exprimer que le chercheur doit tout faire pour s'assurer que les inférences sont valides, c'est à dire que les liens de cause à effet entre variables sont testés sans interférences. Concrètement, il y a quatre modes de contrôle dans le cadre de l'expérimentation. Nous avons déjà plus ou moins évoqué ces sens mais en dresser la liste et les préciser permettra d'éviter toute confusion.

#### *1. Le contrôle par affectation aléatoire*

Il s'agit du contrôle des variables indépendantes par l'affectation aléatoire des unités expérimentales, qui fut évoqué dans le paragraphe précédent.

#### *2. Le contrôle de la situation expérimentale*

Pour illustrer ce mode de contrôle, prenons l'exemple simple d'une expérimentation en chimie. Dans ce cas, le contrôle de la situation expérimentale se fait en stérilisant les éprouvettes afin qu'aucun résidu ne vienne perturber la réaction entre deux corps chimiques. Il est évident qu'en sciences sociales les choses ne sont pas aussi simples, il n'y a pas de contrôle total de la situation expérimentale puisqu'un certain nombre de données de l'expérimentation (et en premier lieu le sujet humain) sont « incontrôlables ». Par contre le design et les procédures expérimentales peuvent être contrôlés et le devoir du chercheur est d'éliminer un maximum de sources externes de

variance pour isoler au mieux la relation de cause à effet entre variable indépendante et dépendante. Il apportera une attention particulière à l'opérationnalisation de ses variables, à ses procédures expérimentales et aux tests des manipulations en prenant soin de ne pas se contenter uniquement de vérifier le succès de la manipulation mais aussi qu'elle correspond bien au construit et n'induit pas, de façon malencontreuse, la manipulation de variables parasites (Perdue & Summers, 1986). Ces recommandations peuvent paraître triviales mais en réalité, le chercheur ne peut pas se permettre d'injecter, par un manque d'attention, des effets indésirables qui fausseraient l'imputation des effets à une cause. Nous verrons dans cette section qu'un expérimentateur scrupuleux doit déjà faire face à suffisamment de biais et d'erreurs indépendantes de sa volonté.

### 3. *Le contrôle par le groupe de contrôle*

Le troisième mode de contrôle fait référence au raisonnement scientifique qui veut que « si X alors Y » implique que « si non X alors non Y » (Boring, 1969). Dans ce contexte, le terme contrôle est associé au principe du groupe de contrôle ou groupe témoin. Rappelons qu'un des designs les plus fréquents en recherche sociale (notamment quand l'affectation aléatoire n'est pas parfaite) est le design suivant :

$$\begin{array}{ccc} O_1 & X & O_2 \\ Oc_1 & & Oc_2 \end{array}$$

Dans ce cas, il y a deux groupes dont un groupe témoin (Oc) qui ne subit pas le traitement (symbolisé par X), les deux groupes sont mesurés avant et après ce traitement. Ce type de contrôle se fonde sur le fait que les variables externes vont influencer les deux groupes de la même façon, par conséquent le différentiel entre les deux groupes isolera l'effet imputable au traitement.

### 4. *Le contrôle par « prise de conscience »*

Nous avons baptisé ainsi le mode de contrôle qui consiste à repérer et essayer de prendre en compte (de façon active ou passive) les sources d'interférences pouvant troubler la relation de cause à effet. Nous avons déjà parlé de ce type de contrôle en évoquant les variables parasites dont l'influence doit être prise en compte dans l'analyse statistique. Cette prise en compte se traduit soit par le maintien constant de la variable externe de façon à neutraliser les effets de sa variation, soit par sa simple mesure et son incorporation dans le traitement et l'analyse des données. Dans leur ouvrage, Cook & Campbell (1979) soulignent que le chercheur doit contrôler en prenant conscience de la menace que certaines caractéristiques de la situation expérimentale peuvent faire peser sur la validité des mesures. Afin d'illustrer leur propos, ils expliquent que le sens du terme contrôle s'exprime dans les phrases du type « *Nous avons essayé de contrôler la possibilité que les répondants comprennent l'hypothèse de recherche en leur donnant une fausse explication de l'objectif de recherche au début de l'expérimentation* » (sic p.8). Dans le cadre de notre recherche, cet exemple a une résonance particulière et pour mieux comprendre pourquoi, il faut le resituer dans le champ de la validité de la recherche.

## 2. **Les facettes de la validité, quel enjeu pour la recherche ?**

Le concept de validité introduit par Campbell & Stanley (1963, 1966) puis développé par Cook & Campbell (1979) est sans doute la source la plus intarissable de débats méthodologiques entre chercheurs en sciences sociales. Parce qu'il reflète les postures épistémologiques, ce concept donne régulièrement lieu à des productions d'articles



controversés. En comportement du consommateur, le débat enflammé Calder/Lynch en témoigne (Calder et al., 1981, 1982, 1983 ; Lynch, 1982, 1983 ; Berkowitz & Donnerstein, 1982, McGrath & Brinberg, 1983). Cette polémique portait plutôt sur la validité externe et il faut rappeler qu'il existe quatre types de validités définies par Cook & Campbell (1979).

Les deux premiers types portent sur l'ensemble des conclusions concernant la relation causale entre les variables manipulées et mesurées, il s'agit de :

- ❑ la **validité statistique** qui juge de la pertinence des inférences statistiques à partir des effets observés.
- ❑ La **validité interne** qui s'inquiète de savoir si les effets observés peuvent bien être attribués à une relation causale précise. (*i.e.* les effets sont-ils bien dus aux variables étudiées et non à d'autres variables).

Les deux autres types de validité ne se préoccupent plus des relations entre les variables opérationnalisées mais plutôt de ce qu'elles représentent théoriquement. Nous y retrouvons :

- ❑ La **validité de construit** qui s'explique comme la généralisation possible des résultats empiriques à des construits et des relations théoriques d'un degré d'abstraction supérieur. Ce qui englobe la convergence entre la définition conceptuelle, l'opérationnalisation (ou validité de construit de l'effet) et la pertinence nomologique des relations entre concepts (ou validité de construit de la cause).
- ❑ La **validité externe** qui concerne la « généralisabilité » des mesures, c'est à dire la possibilité de transposer les relations causales à des populations, des situations et des moments différents. Il est à noter que la validité externe est assez complexe, il ne s'agit pas simplement de juger du degré de réalisme de l'expérimentation ou de représentativité de la population étudiée, mais aussi de considérer la robustesse des relations causales de la recherche (*i.e.* leur degré d'indépendance aux variables contextuelles) (McGrath & Brinberg, 1983).

Pour chaque type de validité, Cook & Campbell (1979) ont établi une liste des « menaces » qui peuvent leur porter préjudice. Notre objet n'est pas de présenter ces quatre listes qui sont explicitées des pages 39 à 80 dans l'ouvrage *Quasi-experimentation ; design & analysis issues for field settings* (Cook & Campbell, 1979) mais bien de comprendre comment ces validités s'articulent entre elles, si certaines sont prioritaires ou plus importantes que d'autres et quels sont les enjeux pour la recherche.

Il est évident qu'aucun type de validité ne peut être négligé. Prise indépendamment l'une de l'autre, chaque validité, si elle ne peut être garantie ou confortée, posera des problèmes au chercheur, à des degrés divers et à des moments différents dans son processus de recherche. De plus, pour les auteurs de la typologie, les validités fonctionnent comme des vases communicants : le renforcement d'un type de validité entraîne l'affaiblissement d'un autre. Par exemple, utiliser un échantillon homogène sur une caractéristique constante peut consolider la validité interne mais amoindrir la validité externe.

Dans ces conditions, il est illusoire de s'attendre à ce qu'une recherche, même si elle ne concerne qu'une seule et simple relation causale, soit irréprochable à tous points de vue. Par conséquent, il revient au chercheur d'arbitrer et de gérer ses priorités. Cook & Campbell (1979) parlent de « *trade-off* » et hiérarchisent l'importance des validités en fonction des objectifs et du type de recherche. Ils distinguent les recherches théoriques des recherches appliquées. Dans le premier cas, la validité interne prévaut sur toutes autres validités, il faut que les relations causales soient sûres pour pouvoir être extrapolées aux concepts théoriques. Par ordre décroissant d'importance les recherches théoriques devraient donc porter attention surtout à la validité interne et à la validité de construit, viendrait ensuite la validité statistique

puis seulement la validité externe. Au contraire dans les recherches dites appliquées, la priorité doit être accordée aux résultats, (*i.e.* le traitement a-t-il bien les effets escomptés ?) et à leur capacité à être extrapolés (*i.e.* les effets sont-ils reproductibles ?). Dans ce cas, le chercheur accordera en priorité de l'importance à la validité interne et à la validité externe, puis ensuite à la validité de construit de l'effet puis à la validité statistique et enfin à la validité de construit de la cause.

La distinction entre recherche théorique et appliquée n'est pas très séduisante car elle sous-entend une scission entre deux « philosophies » de la recherche. Il est sans doute préférable de ne pas opposer théorie et pratique et de parler de raisons et d'objectifs différents servant de fondements à la réflexion méthodologique du chercheur engagé dans un processus de recherche. En comportement du consommateur et dans la recherche marketing en général, la conceptualisation comme les implications pratiques jouent des rôles souvent d'égale importance et il est donc nécessaire que le chercheur apprenne à définir ses priorités en termes de validités. La hiérarchisation des validités paraît dès lors comme inévitable et sans doute nécessaire, comme l'ont compris Calder et al. (1982, 1983) ; Lynch, (1982, 1983); McGrath & Brinberg, (1983).

Toutefois, nous pouvons noter que la validité interne revient toujours en première position dans la hiérarchie, quels que soient les enjeux de la recherche. De plus, elle rejoint la validité externe sur une préoccupation particulière : le contrôle des variables contextuelles. En effet, en des termes simples, la validité interne permet de s'assurer que l'on a bien mesuré la relation causale que l'on voulait mesurer et elle seule. Quand à la validité externe, elle est forte si les résultats sont robustes, c'est à dire répliquables, ce qui n'est possible que si les effets observés ne sont pas liés à une situation contextuelle (c'est à dire à un dispositif expérimental particulier, à un moment dans le temps, à un échantillon de sujets spécifique). Ces variables contextuelles, Lynch (1982) les appelle « *background factors* » en expliquant qu'il s'agit des facteurs liés au sujet expérimental ou au dispositif expérimental.

Cette analyse permet de constater qu'il existe des effets indésirables qui pèsent sur la relation causale et qui sont dus à des facteurs exogènes. Il est possible de reconsidérer les listes de menaces de Cook & Campbell (1979) en tenant compte des variables contextuelles liées au design expérimental et au sujet/groupe expérimental.

### **C- Les effets indésirables sur la relation causale : une approche par le degré de contrôle sur les facteurs exogènes.**

L'analyse des menaces sur les validités interne et externe (Cook & Campbell, 1979) permet de catégoriser les effets indésirables liés au dispositif expérimental et au sujet/groupe expérimental selon le degré de contrôle que le chercheur peut exercer pour s'en prémunir. Ainsi, certains effets peuvent être plus ou moins maîtrisés par le chercheur alors qu'une autre catégorie lui échappe partiellement voire totalement. Pour cette dernière, l'expérimentateur ne peut que constater l'existence de la menace et la possibilité d'interaction avec l'effet principal testé.

#### **1. Effets plus ou moins « contrôlables » par le chercheur**

**L'effet d'instrumentation** : il est lié à la qualité des instruments de mesure et à la manière dont ils sont utilisés. Ainsi, un effet artificiel peut être obtenu si l'intervalle de réponses change entre le début et la fin d'une échelle ou si l'instrument est modifié entre le pré-test et le post-test. L'effet d'instrumentation peut aussi apparaître à cause de la familiarité à

l'instrument acquis par l'expérimentateur entre pré-test et post-test. Dans ces cas, l'effet d'instrumentation est maîtrisable à condition de ne pas changer d'instrument ni de mode de réponse en cours d'expérimentation ou encore de maintenir la vigilance de l'expérimentateur (ou d'éviter un expérimentateur unique pour qu'il n'y ait pas de phénomène de lassitude).

**L'effet de sélection :** il existe lorsque les groupes expérimentaux qui sont comparés sont trop différents, dans ce cas c'est la différence qui peut être à l'origine d'une variation et pas la relation causale étudiée. Il y a aussi effet de sélection lorsque le chercheur prélève certains sujets pour les éliminer de l'analyse. La seule façon de se prémunir radicalement de cet effet est l'affectation aléatoire des sujets sans toucher aux unités au moment de l'analyse (sauf en procédant de façon aléatoire). Nous avons vu que cette condition était difficile à remplir dans le cadre « naturel » du terrain néanmoins le chercheur peut s'efforcer de randomiser partiellement et éviter la sélection intentionnelle au moment de l'analyse statistique.

**L'effet de régression statistique :** cet effet peut rendre difficile les comparaisons entre groupe faible et fort lorsque leur différence est très marquée au départ. Dans ces conditions, les écarts à la moyenne du groupe fort vont être statistiquement plus faibles et les écarts du groupe faible statistiquement plus forts. Il est en effet plus facile pour un sujet faible de réaliser un progrès significatif alors qu'un sujet qui part d'un niveau élevé aura plus de mal à augmenter son score. Par exemple, dans le cadre d'une recherche pédagogique, si, après un enseignement adapté, les gains réalisés par les élèves d'un groupe faible sont plus importants que les gains réalisés par le groupe fort, il reste à démontrer que cette différence de gain est bien due à des différences d'efficacité du traitement pédagogique et non à l'effet de régression (de Landsheere, 1992). Selon Cook & Campbell (1979), pour éviter ce type d'effet, il faut s'assurer de la fiabilité de la mesure par un test-retest en s'assurant que la moyenne du groupe interrogé n'est pas radicalement différente de la population où ce groupe a été sélectionné.

**La diffusion d'informations entre groupes :** Si les traitements impliquent l'utilisation d'informations spécifiques qui diffèrent d'un groupe à l'autre (comme des scénarii par exemple) et que les groupes échangent ces informations entre eux, la mesure de la relation causale peut être affectée et invalidée. Pour éviter cet inconvénient, les groupes peuvent être confinés en laboratoire. Sur le terrain, le chercheur peut aussi isoler ses groupes (par éloignement géographique, recueils simultanés d'information).

**La comparaison entre groupes traités de façon inégale : recherche de compensation et découragement :** La recherche de compensation a deux facettes, elle intervient lorsqu'un traitement est défavorable et que le groupe qui le subit s'en trouve désavantagé. D'une part, l'expérimentateur a tendance à vouloir revaloriser le groupe desservi en rétablissant une certaine équité. D'autre part, si le groupe se rend compte qu'il est comparé à un groupe mieux traité, ces sujets vont essayer de compenser cet état de fait en se montrant plus « combattifs » qu'ils ne le sont vraiment<sup>α</sup>. Dans les deux cas, les mesures sont faussées : dans le premier cas, l'effet va être artificiellement amoindri et dans le second faussement renforcé.

Par ailleurs, si les groupes se comparent il peut aussi y avoir un effet de découragement chez le groupe qui se rend compte qu'il est défavorisé ou qu'il ne lui arrive rien de particulier (groupe de contrôle). Dans ce cas, l'isolement est le meilleur moyen d'éviter les conséquences de la comparaison sociale et l'expérimentateur pourra prévoir des compensations qui viendront soulager le groupe défavorisé une fois l'expérimentation terminée.

---

<sup>α</sup> Saretsky (1972) appelle cet effort « l'effet John Henry » du nom d'un mineur qui ayant appris qu'il était comparé à une machine d'extraction s'est littéralement tué à la tâche pour que son rendement soit meilleur que celui de l'outil.

**Les interactions entre traitements :** Si le sujet expérimental est exposé à des traitements successifs différents, il n'est pas possible d'isoler ses réponses à un traitement du contexte des autres traitements. La manière la plus efficace de mesurer les effets de chaque traitement est de faire en sorte qu'un groupe ne soit soumis qu'à un seul traitement, c'est à dire qu'il y ait autant de groupes que de conditions expérimentales (communément appelé design inter-sujets). Sinon le chercheur prendra soin d'analyser les effets d'un traitement avant de mesurer les effets d'un nouveau traitement sur les mêmes sujets.

**Les attentes de l'expérimentateur ou effet Pygmalion :** Les travaux de Rosenthal (1966, 1969, 1973) ont montré que l'expérimentateur peut volontairement ou inconsciemment influencer les réponses des sujets dans un sens favorable à l'hypothèse de sa recherche. Dans les expérimentations en marketing, les mêmes effets peuvent se faire ressentir (Venkatesan, 1967). Pour Cook & Campbell (1979) cette tendance peut être contrecarrée à différents niveaux. D'abord en minimisant les contacts entre expérimentateur et sujets expérimentaux. Ensuite, il est possible de déléguer le recueil des mesures à plusieurs expérimentateurs qui ne sont pas au courant des hypothèses de recherches et par conséquent ne peuvent induire les réponses. Il est aussi possible d'observer les expérimentateurs afin de détecter et réduire les comportements d'attente. Enfin, il est possible d'intervenir à un niveau supérieur en faisant réaliser les analyses de données par des individus autres que les expérimentateurs qui les ont collectées (Rosenthal & Rosnow, 1991).

## **2. Effets qui échappent au contrôle du chercheur**

**Effet de test ou familiarité au test :** Si un même phénomène est mesuré plusieurs fois auprès d'un même sujet, sa première réponse influencera les suivantes. Dans ce cas, il est possible d'enregistrer une variation de réponse qui n'est pas due à la manipulation d'une variable mais qui apparaît parce que le sujet répond en s'appuyant sur le souvenir de ses réponses passées. Par exemple, on observe qu'un individu qui répond deux fois à un test de QI va améliorer sa performance au second test. Cette familiarité au test ne peut être évitée si le chercheur doit prendre des mesures répétées sur un même échantillon (dans une situation de test-retest par exemple).

**Effet d'histoire :** Ce sont toutes les modifications de l'environnement du participant qui ont lieu après le pré-test et avant la prise de mesure finale. Elles ne sont pas liées au phénomène étudié mais peuvent influencer son attitude, son comportement et donc ses réponses. Ces événements sont incontrôlables par le chercheur, le laboratoire permet un certain isolement mais ce n'est pas le cas pour les recherches de terrain. Rien ne peut garantir qu'un événement contextuel ne va pas impacter les réponses entre pré et post-test. Le chercheur ne peut que se montrer attentif aux impacts éventuels pour les prendre en compte dans l'analyse de ses résultats.

**Effet(s) de maturation :** Il s'agit des modifications que le temps peut entraîner sur le sujet expérimental et qui sont indépendantes de la relation causale mais peuvent l'influencer. Le sujet peut avoir faim, froid, son humeur peut varier, il peut éprouver un sentiment de lassitude. Si l'expérimentation est étalée dans le temps, il peut simplement vieillir, s'assagir, accumuler des expériences qui changent sa vision des choses. Ses réponses sont alors influencées par ces facteurs et non par la manipulation de la variable indépendante.

***Mortalité expérimentale*** : Pour expliquer la mortalité expérimentale, il est possible de parler de biais de sélection indépendant de la volonté du chercheur. En effet, il arrive que des sujets arrêtent de répondre au cours d'une expérimentation ou alors refusent de participer à nouveau lors d'un deuxième questionnement. De facto, ils sabordent la distribution aléatoire ou modifient la composition des groupes, biaisant ainsi le principe de la comparaison.

***Effet dû à l'angoisse de l'évaluation*** : Les travaux de Rosenberg (1965,1969) ont montré que les répondants pouvaient manifester une crainte d'être mentalement évalués par l'expérimentateur. En conséquence, ils essaient de démontrer qu'ils sont des individus compétents et équilibrés afin d'obtenir une évaluation positive. Selon Cook & Campbell (1979), ce type de comportement risque de créer un effet de confusion dans les analyses mais ils déclarent ne pas pouvoir en prendre la réelle mesure : *«l'étendue et l'impact de ce type d'orientation ne sont pas clairs [...] cependant, il est possible que des effets du traitement soient dus à des répondants qui ont voulu se présenter à l'expérimentateur sous un jour qui leur permettait d'espérer une évaluation favorable de sa part»* (p. 67).

Notons qu'un parallèle peut être dressé entre ce type de comportement et la notion de biais de désirabilité sociale (Crowne & Marlowe, 1960) qui se définit comme une tendance de l'individu à vouloir, par ses réponses, présenter une image de soi irréprochable aux yeux de la société.

**Effet Hawthorne et le fait de vouloir deviner l'hypothèse de recherche :** En sciences comportementales, le phénomène connu sous le nom d'« effet Hawthorne » fait référence aux résultats, positifs ou négatifs, qui ne sont pas dus aux facteurs expérimentaux, mais à l'effet psychologique que la conscience de participer à une recherche et d'être l'objet d'une attention spéciale exerce sur le sujet ou sur le groupe expérimental. Il tire son nom de Hawthorne, faubourg de Chicago, où un groupe de chercheurs industriels (Mayo et al., 1924- 1932) avait pu observer une hausse de productivité chez les ouvriers de la *Western Electric Company* due au fait qu'ils se sentaient flattés d'être les sujets d'observations constantes. Réagissant à cette attention, ils adaptaient leur comportement de façon à apporter des données « positives » pour plaire aux expérimentateurs. Cette « bonne volonté » altéra tous les résultats de l'étude. Dans la liste des menaces, l'effet Hawthorne est associé à l'envie manifestée par les sujets d'une expérimentation de comprendre les hypothèses de recherche (*hypothesis-guessing*). Pour Cook & Campbell (1979) deviner l'hypothèse de recherche intervient même si le sujet ne peut se comparer à un autre groupe ou à un voisin, il s'agit d'un processus où « *le répondant n'est au courant que de son traitement et persiste à essayer de découvrir ce que l'expérimentateur veut mettre en évidence par son travail* ». Dans ces conditions, Cook & Campbell admettent que le chercheur peut toujours masquer l'hypothèse de recherche (en la rendant difficile à découvrir ou en fournissant au sujet des explications différentes du réel objectif de recherche) mais trouvent que ces méthodes sont finalement d'un intérêt très relatif car « *de toute façon, les répondants ne sont pas passifs et ils peuvent toujours élaborer leurs propres hypothèses, qu'elles coïncident ou non avec celle de l'expérimentateur* » (p.66). Et ces auteurs de conclure : « *malgré les discussions intensives sur les effets de confusion qui résulteraient de cette volonté de fournir à l'expérimentateur des mesures qui l'arrangeraient, il n'y a pas vraiment de preuve de l'existence d'un effet Hawthorne dans les recherches de terrain ni même en laboratoire [...] Toutefois, nous manquons de travaux théoriques et empiriques sur les conditions dans lesquelles vouloir deviner l'hypothèse (a) existe, (b) à cause du traitement, (c) se traduit par un comportement spécifique qui (d) peut conduire à des conclusions erronées sur les effets du traitement (e) dans le cadre des recherches de terrain.* » (p.67).

En 2002, Shadish, Cook & Campbell regroupent l'angoisse de l'évaluation, le fait de deviner l'hypothèse et d'adopter un comportement de complaisance et l'effet placebo sous le terme « réactivité à la situation expérimentale ». Ils proposent un certain nombre de méthodes pour essayer de réduire ces problèmes<sup>α</sup> : (1) mesurer la variable dépendante en dehors du contexte expérimental, (2) mesurer les effets de la manipulation de manière décalée dans le temps (3) éviter les pré-tests qui donnent des indications sur les résultats attendus (4) utiliser le design de quatre groupes de Solomon (5) standardiser et réduire les interactions entre l'expérimentateur et participant (6) répondre à la curiosité des participants en lui expliquant un faux objectif (techniques de dissimulation) (7) utiliser des sujets qui expliquent comment ils pensent devoir répondre aux manipulations (8) réduire l'angoisse de l'évaluation en assurant l'anonymat et la confidentialité des réponses.

---

<sup>α</sup> Pour l'instant nous « listons » leurs recommandations, nous aurons l'occasion de revenir sur ces méthodes dans ce chapitre.

## Bibliographie

- Aronson, E., Wilson, T.D. & Brewer, M.B. (1998). Experimentation in Social Psychology, dans Gilbert, D.T., Fiske S.T. & Lindzey G., *The Handbook of Social Psychology*, vol I, 4ème éd. Oxford University Press, 99-147.
- Berkowitz, L. & Donnerstein, E. (1982). External Validity is More Than Skin Deep, *American Psychologist*, 37, n°3, 245-257.
- Bhaskar, R. (1975). *A realist theory of science*. Leeds, England: Leeds.
- Boring, E.G. (1969) Perspective: Artifact and Control, 1-10 dans Rosenthal, R. & Rosnow, R.L. (1969), *Artifact In Behavioral Research*. New York: Academic Press.
- Calder, B.J., Phillips L.W., Tybout, A.M. (1982), The Concept of External Validity, *Journal of Consumer Research*, 9, 240-244.
- Calder, B.J., Phillips L.W., Tybout, A.M. (1981) Designing Research for Application, *Journal of Consumer Research*, 8, 197-207.
- Calder, B.J., Phillips L.W., Tybout, A.M. (1983), Beyond External Validity, *Journal of Consumer Research*, 10, 112-114.
- Campbell, D.T. & Stanley, J.C. (1963) *Experimental and quasi-experimental designs for research*. Chicago : RandMcNally.
- Chumpitaz Cacerès R. et Vanhamme J. (2003). Les processus modérateurs et médiateurs : distinction conceptuelle, aspects analytiques et illustrations, *Recherche et applications en Marketing*, 18, 2, 67-100.
- Cook, T.D & Campbell, D.T. (1979), *Quasi-Experimentation, Design & Analysis Issues for Field Settings*. Houghton Mifflin: Boston.
- Cronbach, L.J.(1975) Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116-127.
- Cronbach, L.J.(1988). Five perspectives on validity argument, 3-17 dans Wainer, H; & Braun, H. (édts) *Test Validity*, Hillsdale Erlbaum.
- Crowne, D., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 4, 349-354.
- De Landsheere, G. (1992). *Dictionnaire de l'évaluation et de la recherche en éducation*. PUF : Paris.
- Evrard, Y., Pras, B., Roux, E., Choffray, J.M., Dussaix, A-M., Claessens. M. (1997). *Market, études et recherches en marketing*, 2<sup>nde</sup> édition. Editions Nathan.
- Keppel, G. (1991). *Design and analysis : a researcher's handbook*. (3ème éd). Englewood Cliffs, Prentice-Hall.
- Lynch, J.G. Jr. (1982), On the External Validity of Experiments in Consumer Research, *Journal of Consumer Research*, 9, 225-239. (Erratum in March, 1983, p. 455).
- Lynch, J.G. Jr. (1983), The Role of External Validity in Theoretical Research, *Journal of Consumer Research*, 10, 109-111.
- Mc Grath, J.E. and Brinberg D. (1983), "External Validity and the Research Process: A Comment on the Calder/Lynch Dialogue," *Journal of Consumer Research*, 10, 115-124.

- Perdue, B. C. & Summers J.O; (1986), Checking the Success of Manipulations in Marketing Experiments, *Journal of Marketing Research*, 23, 317-326.
- Rosenberg, M.J. (1965). When Dissonance Fails: On Eliminating Evaluation Apprehension from Attitude Measurement. *Journal of Personality and Social Psychology*, 1, 1, 28-42.
- Rosenberg, M.J. (1969). The Conditions and Consequences of Evaluation Apprehension, 280-348 dans Rosenthal, R. & Rosnow, R.L., *Artifact In Behavioral Research*. NY: Academic Press.
- Rosenthal, R. & Rosnow, R.L. (1975) *Primer of Methods for the Behavioral Sciences*. Wiley.
- Rosenthal, R. & Rosnow, R.L. (Eds) (1969). *Artifact In Behavioral Research*. Academic Press: New York.
- Rosenthal, R. (1966). *Experimenter effects in behavioural research*. New York: Appleton-Century-Crofts.
- Rosenthal, R. (1973). *On the social psychology of the self-fulfilling prophecy : further evidence for Pygmalion effects*. M53, New York: MSS Publication.
- Saretsky, G. (1972). The OEOPC experiment and the John Henry effect. *Phi Delta Kappan*, 53, 579-581.
- Shadish, W.R., Cook T.D. & Campbell D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton-Mifflin.