

Institut des Sciences et Pratiques d'Education et de Formation

Université Lumière - Lyon 2

Licence de Sciences de l'Education

UNITE D'ENSEIGNEMENTS 5 – COMPLEMENT DE LA MAJEURE

METHODOLOGIE DE LA RECHERCHE EN EDUCATION

Arnaud SIMEONE

PLAN DE RECHERCHE

Polycopié n°4

Support de cours - 2005 / 2006

Document uniquement destiné à l'enseignement

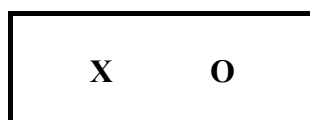
SOMMAIRE

| | |
|--|-----------|
| INTRODUCTION | 3 |
| 1 – LES PLANS DE RECHERCHE CLINIQUES OU PRE-EXPERIMENTAUX. | 3 |
| 1.1 – Le plan post-test à groupe unique. | 4 |
| 1.2 – Le plan pré-test / post-test à groupe unique. | 6 |
| 1.3 – Le plan post-test avec groupe témoin non équivalent (<i>ex post facto</i>). | 8 |
| 1.4 – Le plan post-test avec multiples groupes d'étude ou témoins non équivalents. | 10 |
| 2 – LES PLANS DE RECHERCHE QUASI-EXPERIMENTAUX. | 11 |
| 2.1 – Les plans de recherche avec groupe témoin non équivalent. | 12 |
| 2.1.1 – Le plan pré-test / post-test avec groupe témoin non équivalent. | 12 |
| 2.1.2 - Le plan pré-test / post-test avec groupe témoin non équivalent avec au moins deux pré-tests. | 15 |
| 2.1.3 - Le plan pré-test / post-test avec groupe témoin non équivalent avec mesure analogue au pré-test. | 17 |
| 2.1.4 - Le plan pré-test / post-test avec groupe témoin non équivalent avec échantillons séparés pour pré-test et post-test. | 19 |
| 2.2 – Les plans de recherche avec cohortes. | 21 |
| 2.2.1 – Le plan avec cohortes et comparaison simple. | 22 |
| 2.2.2 - Le plan avec cohortes et démantèlement du traitement. | 24 |
| 2.2.3 – Le plan avec cohortes et pré-test / post-test. | 25 |
| 2.2.4 – Le plan avec cohortes multiples selon un cycle récurrent. | 27 |
| 2.3 – Les plans de recherche à séquence temporelle interrompue. | 28 |
| 2.3.1 – Le plan à séquence temporelle interrompue unique. | 29 |
| 2.3.2 – Le plan à séquence temporelle interrompue avec groupe témoin non équivalent. | 30 |
| 2.3.3 - Le plan à séquences temporelles multiples avec traitement décalé. | 31 |

certain d'entre eux proposent de pratiquer des évaluations avant et après la confrontation au traitement, à la situation ou au phénomène (*pré* et *post-test*), ils n'associent pas toujours un groupe contrôle aux groupes de recherche. En fait, ils ne permettent pas d'effectuer suffisamment de mesures comparatives pour s'assurer totalement de la validité des résultats observés. Quand cela est possible, ces plans sont donc à éviter, car ils sont fortement susceptibles d'être influencés par les multiples biais menaçant la validité interne.

1.1 – Le plan post-test à groupe unique.

Description :



Dans ce plan de recherche, un groupe d'individus est soumis à un traitement ou est confronté à un événement ou à une situation (**X**). Une ou plusieurs mesures (**O**) sont appliquées en une seule session sur les individus de ce groupe, toujours après l'exposition au traitement, à l'événement ou à la situation. Cependant, la durée de la période située entre cette exposition et la ou les mesures n'est pas fixée et peut considérablement fluctuer d'un protocole de recherche à l'autre. Dans certains cas, la mesure peut avoir lieu juste après l'exposition au traitement, à l'événement ou à la situation. Dans d'autres cas, le délai entre l'exposition et la mesure peut être de plusieurs années. Par ailleurs, la durée de ce délai n'est pas forcément équivalente pour tous les individus du groupe étudié.

Analyse :

La principale caractéristique de ce plan est qu'il n'intègre pas explicitement de Variable Indépendante (V.I.). Seule une ou plusieurs Variables Dépendantes (V.D.) sont identifiables. Cependant, le chercheur peut souhaiter **contrôler** une ou plusieurs variables, qui serviront ensuite à décrire le groupe d'individus participant à la recherche. Ces variables n'ont alors qu'un rôle indicatif et ne peuvent pas servir à constituer des sous-groupes de participants ou à produire des comparaisons.

Ce plan peut être utilisé pour **recueillir** des opinions, des comportements déclarés ou des souvenirs, **évaluer** des connaissances, des caractéristiques démographiques, sociologiques

¹ de Landsheere, G. (1992, 2^{ème} édition). Dictionnaire de l'évaluation et de la recherche en éducation. Presses Universitaires de France : Paris.

ou psychologiques, ou encore *décrire* des comportements ou des conduites. Par contre, il *ne permet pas d'expliquer* l'occurrence de ces comportements ou cognitions. L'usage de ce plan se justifie lorsqu'il n'est pas possible de poser *a priori* une hypothèse qui souligne l'existence d'une relation entre une V.I. et une V.D. Il est donc très souvent utilisé dans le cadre de recherches exploratoires, qui examinent des situations, des groupes d'individus, des pratiques ou des phénomènes n'ayant jamais fait l'objet d'études auparavant. La plupart du temps, celles-ci visent à décrire un phénomène ou une situation en émergence, ou encore une pratique innovante, dont il convient d'établir un *descriptif*, avant de pouvoir élaborer des hypothèses. Il est aussi employé par défaut, lorsque l'objet de la recherche, les circonstances ou l'avancée de la technologie (*les contraintes de la recherche*) ne permettent pas de mettre en place un autre type de plan.

Validité :

Ce plan de recherche est le plus simple possible. En première analyse, sa mise en pratique peut paraître peu contraignante et son attrait important. C'est d'ailleurs l'un des plus utilisés en Sciences Humaines et Sociales. Néanmoins, ses limitations sont très importantes.

La première difficulté liée à ce type de protocole réside en l'absence de pré-test, ce qui ne permet pas d'évaluer véritablement si l'exposition au traitement, à l'événement ou à la situation (**X**) a eu un effet sur la mesure (**O**). La deuxième difficulté provient du fait qu'il n'y a pas de groupe de comparaison ou de groupe témoin, ce qui ne permet pas de savoir si l'*éventuel* effet sur la mesure est bien dû à l'exposition au traitement, à l'événement ou à la situation plutôt qu'à des facteurs autres (voir la partie consacré à la validité interne d'un protocole de recherche dans le polycopié n° 3). L'interprétation des résultats d'une recherche utilisant ce plan est toujours très délicate. En fait, ce plan de recherche ne permet pas d'affirmer que les caractéristiques, les reports verbaux ou les comportements des participants tel qu'ils sont évalués au post-test sont bien la conséquence du traitement, de l'événement ou de la situation.

Le principal intérêt des recherches utilisant ce type de protocole est de permettre la genèse d'hypothèses qui pourront ultérieurement être testées à l'aide de plans de recherche plus sophistiqués. Un autre intérêt est permettre de réaliser une description qualitative très fouillée, à l'aide d'un grand nombre de V.D., sur un petit échantillon, voire un seul individu. Dans ce dernier cas, la validité du plan de recherche provient de la multiplicité, et potentiellement de la convergence, des mesures effectuées.

Exemples :

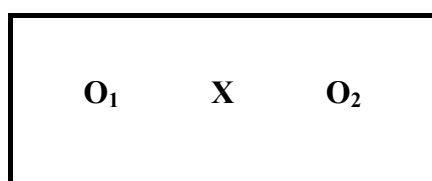
1) Un chercheur se demande comment des élèves de Cours Moyen utilisent l'outil Internet dans le contexte de la classe. Pour se faire, il réalise des observations des pratiques de l'Internet (**O**) sur un groupe d'élèves placés dans une situation où ils sont susceptibles d'utiliser cet outil (**X**).

2) Un chercheur se demande quel est le niveau de connaissance des règles du code de la route des élèves de Terminale. Pour se faire, il passe un questionnaire d'évaluation des connaissances des règles du code de la route (**O**) à un vaste échantillon d'élèves scolarisés en classe de Terminale (**X**).

3) Un chercheur se demande quelle opinion des élèves de primaire ayant participé à une classe verte ont vis-à-vis du tri sélectif des déchets. Pour se faire, il effectue des entretiens sur ce thème (**O**) auprès d'une dizaine d'élèves ayant participé au cours de l'année scolaire à une classe verte (**X**).

1.2 – Le plan pré-test / post-test à groupe unique.

Description :



Dans ce plan de recherche, un seul groupe d'individus est évalué avant (**O₁**) et après (**O₂**) l'exposition à un traitement, un événement ou une situation (**X**). Ce plan permet de mesurer le changement associé à cette exposition. Comme précédemment, la durée des périodes situées entre l'exposition (**X**) et les évaluations (**O₁** et **O₂**) n'est pas fixée, et il n'est pas même obligatoire que ces deux périodes aient une durée équivalente.

Analyse :

Ce plan de recherche est préférable au plan post-test à groupe unique parce qu'il offre un point de comparaison : le pré-test. Bien évidemment, pour que cette comparaison soit possible, il est nécessaire que l'outil de mesure employé soit similaire au pré-test et au post-test.

Validité :

L'intérêt principal de ce plan est qu'il permet de vérifier l'existence d'une fluctuation dans une mesure (V.D.) effectuée avant et après l'exposition à un traitement, un événement ou une situation.

Cependant, l'absence de groupe de comparaison ou de groupe témoin ne permet pas d'affirmer que les changements constatés sont bien dû à l'exposition au traitement, à l'événement ou à la situation, et donc d'éliminer l'existence d'hypothèses rivales. Les capacités heuristiques de ce plan sont donc limitées.

Remarques :

Toutes les situations de recherche n'autorisent pas ou ne permettent pas la mise en place d'un pré-test. Soit parce que l'événement dont on teste l'effet n'était pas prévisible (par exemple, un acte de vandalisme dans un établissement scolaire ou un événement politique international), soit parce que le groupe d'individus étudiés est déjà inséré dans la situation dont on souhaite tester l'impact (par exemple, la participation à une activité sportive ou l'intégration dans une association), soit parce que le traitement dont on désire évaluer l'effet est déjà commencé lorsque la recherche débute (par exemple, un programme d'inclusion scolaire ou une méthode pédagogique d'apprentissage de la lecture). Dans ces circonstances, il est parfois possible d'utiliser certaines données d'archives permettant de documenter l'état des participants précédent le traitement, l'événement ou la situation. Cependant, même dans les cas les mieux documentés, cette solution est rarement satisfaisante. D'une part, les données d'archive ne permettent pas de répondre exactement au questionnement du chercheur, car elles n'ont pas été recueillies dans ce but. D'autre part, la plupart du temps l'outil initial de recueil des données n'est pas disponible ou même décrit, ce qui rend difficile sa duplication et son utilisation lors du post-test. La conception et la mise en place d'un réel pré-test sont donc toujours préférables à cette approximation.

Exemples :

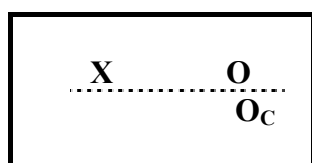
1) Un chercheur souhaite vérifier si un programme éducatif tutorial a un effet sur le développement des capacités langagières d'élèves de première section de maternelle (2 à 3 ans). Pour se faire, il effectue tout d'abord une estimation des niveaux de compréhension et de production langagière d'un large groupe d'élèves (O_1). Cette première évaluation lui permet de repérer et de sélectionner les 30 élèves dont le niveau de compétences langagières est évalué comme le moins élevé. Ensuite, après avoir obtenu l'autorisation des parents, il intègre

ces élèves dans le programme éducatif, qui consiste en la participation à une série d'ateliers sur durée de huit mois à raison d'une heure par jour (**X**). Enfin, à la fin de ce programme, il réalise une nouvelle évaluation des compétences langagières de ces élèves à l'aide du même outil (**O₂**).

2) Un chercheur réalise une enquête sur la perception que des élèves de Sixième ont de leur propre risque d'être victime d'un acte de violence à l'école (**O₁**). Deux mois après la fin de cette étude, un acte de violence extrêmement médiatisé a lieu dans un collège de la région. Le chercheur se demande si cet événement (**X**) a eu un impact sur la perception que les collégiens ont du risque d'être victime d'un acte de violence. Pour se faire, il effectue une nouvelle estimation du risque perçu d'être victime d'un acte de violence à l'école (**O₂**), à l'aide du même outil et sur le même échantillon.

1.3 – Le plan post-test avec groupe témoin non équivalent (*ex post facto*).

Description :



Dans ce plan de recherche, un groupe d'individus est soumis à un traitement ou est confronté à un événement ou à une situation (**X**), puis une ou plusieurs mesures (**O**) sont appliquées en une seule session sur les individus de ce groupe. Aucune évaluation n'a été effectuée ou aucune donnée d'archive n'est disponible en guise de pré-test. C'est ce qu'indique l'expression ***ex post facto*** (*après le fait*). Par contre, une évaluation est réalisée sur un groupe de comparaison (**O_C**), au même moment et de la même façon que le groupe ayant été soumis au traitement. Il faut souligner que le groupe de comparaison est ***qualifié de non équivalent***, parce que les individus qui le composent n'ont pas été sélectionnés à partir de la même population que le groupe ayant été soumis au traitement. Plus précisément, il n'y a pas eu préalablement au traitement assignation au hasard d'un ensemble d'individus dans l'une ou l'autre des deux conditions : le groupe de traitement (**O**) et le groupe contrôle (**O_C**). Dans le schéma, ce phénomène est illustré par la présence du trait pointillé.

Analyse :

Ce type de plan est utilisé lorsqu'un chercheur doit faire une évaluation non préalablement prévue ou planifiée d'un traitement qui a déjà été appliqué. Parfois, il n'a même pas contrôlé l'administration de ce dernier. Ces pratiques sont courantes dans les secteurs de l'éducation et du travail social, où un service ou une pratique peut être rendu disponible bien avant que le commanditaire ou le concepteur ne soit prêt à en faire l'évaluation, ou même qu'il n'en ressente la nécessité. Comme ce plan ne comprend pas d'évaluation pré-test, c'est la comparaison entre les mesures obtenues *ex post facto* sur les groupes expérimental (**O**) et contrôle (**O_C**) qui va permettre vérifier si le traitement (**X**) a bien eu un effet.

Validité :

Bien qu'il soit parfois le seul moyen possible d'estimer dans quelle mesure un service, un traitement ou une politique a atteint ses objectifs, ce plan comporte des limites importantes. Tout d'abord, l'absence de pré-test dans ce plan ne permet pas d'évaluer véritablement si l'exposition au traitement, à l'événement ou à la situation (**X**) a eu un effet sur la mesure (**O**). Ensuite, le fait qu'il n'y ait pas de groupe de comparaison équivalent ne permet pas de savoir vraiment si l'éventuel effet sur la mesure est dû à l'exposition au traitement ou à un biais dans la sélection des participants. Cependant, il est possible de limiter le degré d'incertitude vis – à – vis de la validité des résultats en recueillant des informations supplémentaires sur les sujets des deux groupes. De fait, si sur des dimensions pertinentes, susceptibles d'avoir un effet sur la Variable Dépendante, les deux groupes de participants ne se distinguent pas, alors il n'y a *a priori* pas de raison de douter de la fiabilité des différences ou des convergences constatées.

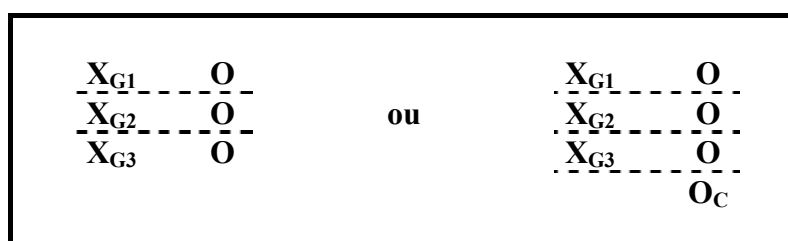
Exemple :

Un responsable d'association d'éducation pour la santé doit rendre compte auprès de ses financeurs de l'intérêt d'un programme d'éducation à la nutrition que son association conduit depuis cinq ans auprès d'une trentaine de classes de cours préparatoire d'écoles primaires de LYON. Malheureusement, aucune phase évaluation n'avait été prévue initialement dans la conception de ce programme éducatif. Ce responsable décide donc de réaliser une évaluation par questionnaire des connaissances et des comportements nutritionnels (**O**) auprès de la totalité des élèves ayant participé à ce programme (**X**). Pour pouvoir estimer dans quelle mesure les comportements et les connaissances déclarées sont le

résultats de la participation au programme d'éducation à la nutrition, il va aussi interroger à l'aide du même questionnaire un groupe d'élèves du même âge issus d'écoles comparables, n'ayant jamais participé à un quelconque programme d'éducation à la nutrition en milieu scolaire (O_C).

1.4 – Le plan post-test avec multiples groupes d'étude ou témoins non équivalents.

Description :



Ce plan intègre la possibilité de comparer l'effet de plusieurs *types* ou *intensités* de traitements, situations ou événements (X_{G1} , X_{G2} , X_{G3}) sur une ou plusieurs mesures groupées (O). Si cette dernière intervient toujours après l'exposition au traitement, à l'événement ou à la situation, la durée de la période située entre cette exposition et la ou les mesures n'est pas fixée. Ce plan offre aussi la possibilité d'introduire un groupe contrôle (O_C). Cependant, comme dans le plan précédent, les différents groupes de sujets ne sont pas considérés comme équivalents, et il n'y a pas de possibilité d'effectuer une mesure en pré-test.

Analyse :

Ce plan est juste une variation plus complexe du plan *ex post facto*. Il est utilisé quand il s'avère nécessaire de réaliser l'évaluation non préalablement planifiée d'une série de traitements, de situations ou d'événements qui sont déjà survenus. Si ces traitements, situations ou événements peuvent se distinguer par leur *nature* (par exemple, différentes méthodes d'apprentissage de la lecture) ou leur *intensité* (par exemple, différentes durées de sessions de soutien), ils restent liés à une même dimension.

De fait, le but étant de comparer leurs effets sur une même Variable Dépendante (O), il faut qu'ils soient comparables. Ce plan ne comprenant pas de pré-test, c'est la comparaison entre les mesures obtenues *ex post facto* sur les différents groupes expérimentaux (O) et contrôle (O_C) qui va permettre vérifier 1) si les traitements (X) ont bien eu un effet, et 2) lequel de ces traitements a eu l'effet le plus souhaitable.

Validité :

Les limites de ce plan sont les mêmes que pour le précédent : le fait que les participants ne soient pas distribués aléatoirement dans les différents groupes et l'absence de pré-test ne permettent pas d'être sûr que les différences constatées entre les groupes de participants sont bien dues à la différence de nature ou d'intensité des traitements. Néanmoins, contenu des contraintes des terrains, ce plan peut quelques fois s'avérer le seul qui puisse être mis réellement en place.

Exemple :

Un programme de réinsertion sociale mené en milieu carcéral est conduit sur différents groupes de prisonniers, qui se distinguent par leur durée de participation à ce programme ($X_1 = 1$ mois, $X_2 = 4$ mois, $X_3 = 8$ mois). Cette durée variant en fonction de leur date de libération conditionnelle, les participants au programme ne sont pas assignés de manière aléatoire dans les groupes. Si le taux de récidive de ces individus (O) fluctue en fonction de la durée du traitement reçu (X_1 , X_2 ou X_3), il sera alors tentant de considérer que le traitement a eu un effet. La possibilité de comparer le taux de récidive des groupes participants au programme à celui d'un groupe de non-participants (O_c) pourrait permettre d'accroître le degré de confiance qu'il est possible d'accorder à ces résultats

2 – LES PLANS DE RECHERCHE QUASI-EXPERIMENTAUX.

Un plan de recherche est qualifié de quasi-expérimental s'il permet de tester l'hypothèse d'une relation entre une variable indépendante et une variable dépendante, sans toutefois respecter toutes les conditions favorables à une inférence causale fiable. De façon plus précise, ce type de plan ***implique la présence d'une variable indépendante***, mais ***ne permet pas de manipuler réellement l'ensemble de ces modalités ou de ces degrés***.

Dans ce cas de figure, les participants ne peuvent pas être assignés de manière aléatoire dans les différents groupes reflétant les modalités de ou des variables indépendantes, car les caractéristiques traduites par ces modalités sont intrinsèquement attachées aux participants. Le chercheur ne peut que constater ou mesurer la présence de ces caractéristiques chez chacun des sujets, et en suite répartir ces derniers dans les différents groupes. En d'autres termes, le chercheur ne construit pas et ne manipule pas les différentes conditions issues de la variable indépendante, il se contente de mesurer chez les participants cette variable préexistante et d'observer l'effet de celle-ci sur une variable dépendante.

Cette catégorie de plans est particulièrement utilisée dans le domaine des sciences de l'éducation, où par exemple, il est peu fréquent que le chercheur puisse totalement déterminer au hasard quels individus participeront à une action éducative et ceux qui en seront exclus, ou que le chercheur puisse décider quand tel groupe d'individus suivront un programme d'enseignement. En éducation ou en formation, il est donc rare qu'une recherche puisse satisfaire tous les critères qui permettent de préserver sa validité interne, et les plans de recherche quasi-expérimentaux apparaissent comme un bon compromis entre d'une part les contraintes d'un terrain et d'autre part, les critères de rigueur scientifique.

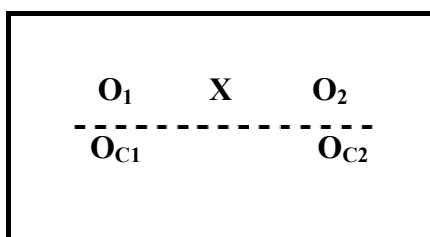
Il existe trois grandes catégories de plan de recherche quasi-expérimentaux : les plans de recherche à groupes non équivalents, les plans de recherche avec cohortes et les plans de recherche à série temporelle interrompue.

2.1 – Les plans de recherche avec groupe témoin non équivalent

Ce type de plans est particulièrement lorsqu'il n'est pas possible d'assigner aléatoirement les participants aux différentes conditions de la recherche et qu'il est possible de se replier sur un groupe de comparaison qui n'est cependant pas nécessairement équivalent au groupe soumis au traitement, à l'événement ou à la situation.

2.1.1 – Le plan pré-test / post-test avec groupe témoin non équivalent.

Description :



Ce plan comprend deux groupes distincts, l'un ayant été exposé au traitement, au phénomène ou à la situation (O) et l'autre pas (O_C), ces deux groupes étant évalués avant et après l'intervention (X). Ces groupes sont considérés comme non équivalents, parce que non constitués sur la base d'une répartition aléatoire.

Il est possible de complexifier ce protocole de base en ajoutant un ou plusieurs groupes d'intervention, lorsque l'on souhaite tester simultanément l'influence de plusieurs traitements, phénomènes ou situation, ou celle de plusieurs degrés d'exposition à un même

traitement, phénomène ou situation. Il est également possible de prendre en compte plusieurs groupes témoins, si l'on souhaite par exemple proposer différents groupes de placebo.

Analyse :

Ce plan permet de mesurer l'importance du changement entre un pré-test et un post-test, selon que le groupe de participants ait été ou non exposé à l'intervention. L'analyse des résultats nécessite donc une triple comparaison : 1) entre le groupe soumis à l'intervention et le groupe contrôle et 2) entre le pré-test et le post-test, et 3) entre les écarts constatés entre les groupes au pré-test et ceux observés entre les groupes au post-test.

Validité :

Dans ce type de plan, la présence d'un groupe contrôle et d'une évaluation au pré-test permet de mieux répondre aux problèmes de validité interne soulevés par les protocoles pré-expérimentaux. Toutefois, le fait que les groupes n'aient pas été constitués sur la base d'une assignation aléatoire à partir de la même population, c'est-à-dire qu'ils soient non équivalents, peut poser des problèmes d'interprétation. Ces derniers seront plus ou moins aigus selon que le chercheur sera dans la capacité de démontrer la proximité initiale entre le ou les groupes contrôle et le ou les groupes exposés à l'intervention, et ceci à partir de variables jugées pertinentes. A cet égard, la comparaison entre les groupes traitements et contrôles au pré-test est particulièrement importante. Plus le chercheur pourra montrer que les groupes traitements et contrôles sont semblables au pré-test, plus le protocole de recherche sera considéré comme valide sur le plan interne. Les techniques d'échantillonnage par quotas ou les procédures d'appariement visant à constituer des groupes similaires sur certaines dimensions peuvent permettre a priori d'accentuer la proximité des groupes traitements et contrôles.

Néanmoins, même ainsi, tous les problèmes d'interprétation ne sont pas résolus, parce que cette procédure ne tient compte que des potentielles différences observées au pré-test et non de la possibilité que ces différences s'accroissent avec le temps, en dehors de tout effet de l'intervention. Trois types de biais peuvent venir parasiter la relation entre la variable indépendante et la variable dépendante.

Tout d'abord, il est possible que les participants d'un groupe deviennent plus compétents dans la dimension évaluée par la V.D. que les participants des autres groupes pendant la période de l'intervention, et ce indépendamment de cette intervention. Ce phénomène peut intervenir soit à cause d'une ***interaction sélection – maturation***, soit à cause d'une ***interaction sélection – expérience***. Dans le premier cas, les deux groupes de

participants peuvent montrer des courbes d'évolution ou de développement divergentes dues à des facteurs *endogènes* les distinguant, ce qui pourrait rendre compte partiellement ou entièrement d'un changement plus important dans un groupe que dans un autre. Dans le second cas, les participants d'un groupe peuvent avoir été exposés au cours de l'intervention à des facteurs *exogènes*, tels que des expériences relationnelles stimulantes ou valorisantes, autres que celles prévues par le traitement et qui au final pourraient expliquer la progression plus importante de ce groupe par rapport à l'autre. Ces deux phénomènes interactifs ont une plus grande probabilité de se manifester dans les contextes éducatifs ou d'apprentissage que dans tout autre contexte, parce les participants les plus avancés cognitivement et socialement sont ceux qui seront les plus susceptibles d'être stimulés, de bénéficier de l'intervention, et au final d'accentuer leur avance eu égard au contenu donné. Ils ont aussi une chance plus importante de survenir dans les situations où les participants peuvent s'autosélectionner à recevoir le traitement (par exemple, sur la base d'un volontariat), parce que dans ce cas, les participants soumis à l'intervention se démarqueront par leur plus forte motivation au changement. Enfin, il est clair que ces phénomènes interactifs sont plus susceptibles d'apparaître lorsque l'intervalle entre le pré-test et le post-test est grand.

Ensuite, il est possible que le caractère non équivalent des groupes étudiés produise un *biais de régression statistique différentielle*, c'est-à-dire que la régression statistique à la moyenne peut se manifester de façon plus marquée dans un groupe que dans un autre.

Par exemple, si un chercheur recrute pour un programme de lutte contre l'illettrisme des élèves de primaire sur la base de leurs résultats à un test de lecture et d'écriture, en ne retenant que ceux qui ont obtenus les scores les plus bas, et qu'il constitue un groupe contrôle à l'aide d'une procédure d'appariement sur la base des principales variables socio-démographiques, il est clair qu'il peut s'attendre à ce que les participants du groupe exposé à l'intervention améliorent leurs scores au post-test de manière significativement plus importante que ceux du groupe contrôle, et ceci indépendamment de tout effet de l'intervention éducative.

Enfin, il est aussi possible que l'outil choisi pour estimer l'amplitude du changement (la Variable Dépendante) puisse être plus ou moins apte à différencier les individus selon leur niveau ou leur position dans la dimension évaluée. Par exemple, il est notoirement connu que les tests d'intelligence sont généralement plus sensibles et plus fiables pour détecter des différences dans la zone médiane de cette dimension que dans ces extrémités. Ainsi, un même niveau de progression aura moins de chance d'être détecté si au départ l'individu est positionné sur les extrémités de ce type d'échelle de mesure que s'il se situe au centre. Des

phénomènes de tassement ou de plafonnement de la mesure peuvent survenir. Par exemple, si la moyenne d'un groupe à un test de mathématiques est dès le départ de 18 / 20, sa marge *mesurable* de progression sera moins grande que celle d'un groupe qui a obtenu une moyenne de 12 / 20 (Effet plafond).

Exemple :

Une infirmière travaillant au sein d'un Point Santé universitaire est chargée d'évaluer l'impact d'un dispositif d'aide au sevrage tabagique sur la réussite et le maintien de l'arrêt de la consommation du tabac. Pour ce faire, elle va procéder en plusieurs étapes. Tout d'abord, lors des bilans médicaux obligatoires de 1^{ère} année, elle va interroger à l'aide d'un questionnaire les étudiants fumeurs sur leurs connaissances et leurs attitudes vis-à-vis de ce produit, ainsi que sur leur niveau de consommation et leur niveau de motivation et d'intention déclarée d'arrêter cette consommation (Pré-test, O_1 et O_{C1}). A la suite de la passation de ce questionnaire, elle va proposer à certains étudiants fumeurs de participer au programme d'aide au sevrage mise en place par l'université, cette participation se faisant sur la base du volontariat. Ensuite, les étudiants volontaires seront convoqués pour participer au programme, et leur assiduité sera contrôlée (X). Enfin, quelques semaines plus tard, l'infirmière demandera aux étudiants assidus au programme de répondre une seconde fois au questionnaire utilisé lors du pré-test (O_2). Parallèlement, sur la base de leurs réponses au premier questionnaire, à l'aide d'une procédure d'appariement, elle va constituer un deuxième groupe d'étudiants fumeurs non participants au programme, globalement similaires à ceux du groupe engagé dans l'action de santé, tant au niveau de leur niveau de consommation initiale que leur intention déclarée d'arrêter cette consommation. Elle proposera aussi à ce deuxième groupe d'étudiants de répondre une seconde fois à son questionnaire (O_{C2}).

2.1.2 - Le plan pré-test / post-test avec groupe témoin non équivalent avec au moins deux pré-tests.

Description :

| | | | |
|----------|----------|-----|----------|
| O_1 | O_2 | X | O_3 |
| O_{C1} | O_{C2} | | O_{C3} |

Ce plan de recherche est une variation plus rare du précédent. Il se distingue par l'adjonction d'au moins une mesure supplémentaire de pré-test, réalisée à une occasion différente de la première. Généralement, les durées séparant d'une part les deux pré-tests et d'autre part le second pré-test et le post-test sont similaires.

Analyse :

Quand le contexte opérationnel le permet (c'est-à-dire principalement lorsque l'administration du traitement peut être retardée sans entraîner de problèmes éthiques ou de difficultés financières insurmontables), ce type de protocole de recherche permet un meilleur contrôle de la validité des résultats. Néanmoins, le chercheur doit s'assurer qu'a priori, la répétition des mesures n'a pas d'effet particulier sur la variable dépendante.

Validité :

Ce plan offre plusieurs avantages au niveau de la validité interne. Tout d'abord, la comparaison de la différence entre les deux mesures de pré-test pour le groupe de traitement ($O_2 - O_1$) et pour le groupe contrôle ($O_{C2} - O_{C1}$) permet de mieux apprécier si un phénomène de sélection-maturation a cours avant l'exposition au traitement (X). Ensuite, la comparaison des différences entre les deux mesures de pré-test et entre la seconde mesure de pré-test et la mesure de post – test, pour le groupe de traitement $[(O_2 - O_1) - (O_3 - O_2)]$ et pour le groupe contrôle $[(O_{C2} - O_{C1}) - (O_{C3} - O_{C2})]$ permet d'apprécier les parts du changement dues au traitement et aux phénomènes de sélection-expérience ou de sélection maturation. Enfin, la comparaison des résultats aux deux pré-tests permet également d'évaluer l'importance d'un éventuel effet de régression statistique.

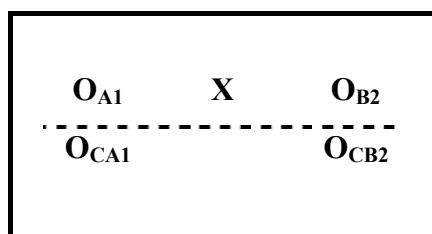
Exemple :

Un chercheur souhaite mettre en place un dispositif d'aide au développement des compétences logico-mathématiques chez des enfants inscrits en Cours Préparatoire. Il sélectionne trois écoles primaires parmi l'ensemble des écoles de la ville LYON pour les intégrer dans son programme, et il recrute trois autres écoles pour servir de groupes témoins. Cependant, ce chercheur s'attend aussi à ce que ce type de compétences puissent se développer ontologiquement chez les enfants de cette classe d'âge, sans compter que les écoles sélectionnées ne sont toutes situées dans des quartiers socio-économiquement équivalents. Afin d'évaluer l'importance des éventuels biais dus aux phénomènes de régression statistique et de sélection des participants, notre chercheur décide de mettre en

place un plan de recherche où les compétences logico-mathématiques des groupes traitement et contrôle seront tout d'abord évalués lors d'un premier pré-test (O_1 et O_{C1}) quelques temps après la rentrée, puis lors d'un deuxième pré-test trois mois après (O_2 et O_{C2}). C'est seulement après ce second pré-test que les élèves du groupe traitement seront véritablement intégrés dans le programme éducatif, et ceci pour une durée de trois mois. Durant cette période, ceux du groupe témoin ne sont soumis à aucune intervention particulière. A la fin du programme éducatif, l'ensemble des élèves (groupe traitement et groupe témoin) est une troisième fois évalué lors d'un post-test (O_3 et O_{C3}).

2.1.3 - Le plan pré-test / post-test avec groupe témoin non équivalent avec mesure analogue au pré-test.

Description :



Ce plan de recherche est utilisé lorsqu'il n'est pas possible d'utiliser un outil de mesure totalement identique au pré-test et au post-test. Le chercheur est alors contraint d'employer deux outils de mesures différents, mais dont les scores doivent néanmoins être substantiellement corrélés. En d'autres termes, ils doivent évaluer la même chose. Dans le schéma présenté ci-dessus, les lettres **A** et **B** font références à ces mesures différentes mais cependant *analogues*. En dehors de cette distinction, ce plan de recherche est identique au plan pré-test / post-test avec groupe témoin non équivalent.

Analyse :

Ce plan peut être employé dans deux types de situation de recherche :

- quand il n'est pas possible de répéter la mesure sans être confronté à un biais de familiarisation / apprentissage avec la mesure,
- quand le traitement consiste en l'acquisition de nouvelles connaissances ou de nouvelles pratiques non encore inscrites au répertoire de l'individu. Dans ce cas, les connaissances ou les réponses sont avant le traitement tellement étrangères à la population

étudiée qu'un pré-test portant *stricto sensu* sur les connaissances ou les pratiques à acquérir n'aurait pas de sens.

Il est aussi possible d'avoir recours à ce plan de façon rétrospective, afin de consolider la mise en évidence d'un résultat obtenu à l'aide d'un plan de recherche pré-expérimental de type *ex post facto*. De fait, l'adjonction à ce dernier type de plan d'une ou plusieurs mesures analogues obtenues avant le traitement (par exemple, en ayant recours à des données d'archives) peut permettre de s'assurer de la plus ou moins grande équivalence des groupes témoin et traitement, favorisant ainsi l'inférence causale. En d'autres termes, l'utilisation d'une mesure analogue au pré-test permet de vérifier si le groupe expérimental et le groupe témoin sont équivalents en ce qui concerne une variable conceptuellement associée à la variable dépendante.

Validité :

En plus des limites associées au plan pré-test / post-test avec groupe témoin non équivalent, la validité interne de ce plan de recherche dépend considérablement de la force de la relation entre les deux mesures **A** et **B**. Plus la corrélation entre la mesure analogue **A** et la mesure dépendante **B** est forte, plus ce plan s'apparente au plan précédent. Plus cette corrélation est faible, plus l'interprétation des résultats sera équivoque.

Exemples :

1) Un professeur de français voudrait tester l'efficacité d'une méthode innovante d'apprentissage de l'orthographe chez des élèves de 6^{ème}. Pour ce faire, il teste d'abord la compétence en orthographe de 4 de ces classes de 6^{ème} à l'aide d'une dictée de 600 mots (**A**). Ensuite, pendant un semestre, il va utiliser la méthode innovante avec deux de ces classes (**X**), et une méthode plus classique avec les deux autres. A la fin du semestre, il va à nouveau tester la compétence en orthographe de l'ensemble des élèves de ces quatre classes, avec une épreuve de dictée différentes, mais de difficultés et de longueur équivalentes (**B**).

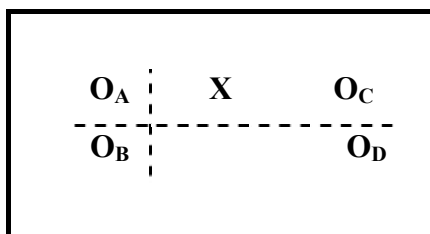
2) Un professeur de Sciences de la nature souhaite mettre en place une intervention éducative visant l'acquisition de notions touchant la protection de l'environnement. Cependant, les élèves concernés par ce projet n'ayant jamais été exposés à ces notions, il ne semble pas pertinent de les interroger sur ce domaine lors du pré-test. Le professeur de Sciences de la nature décide donc de pré-tester deux de ces classes de Seconde à l'aide d'un Q. C. M. concernant les connaissances générales en Sciences de la nature (**A**), de faire

participer l'une de ces classes à l'intervention éducative (**X**), et de tester à nouveau les deux classes avec un Q. C. M. traitant de la notion de protection de l'environnement (**B**).

3) Une psychologue de l'éducation a été intégrée dans une équipe socio-éducative, qui applique depuis deux ans un programme de stimulation des compétences cognitives (**X**) auprès d'un groupe d'enfants déficients intellectuels. L'équipe lui demande d'évaluer l'impact de ce programme chez les enfants intégré dans ce programme. La psychologue n'a pas de difficultés à trouver dans l'institut d'accueil un groupe d'enfants comparables, ne participants à aucun programme de ce type. Par contre, aucune évaluation préalable des compétences n'ayant été anticipée par les praticiens de l'équipe socio-éducative, elle ne dispose pas de mesures pouvant faire office de pré-test. Néanmoins, en explorant les dossiers médico-psychologiques des enfants, elle constate que tous les enfants de l'institut sont soumis à une batterie de tests de leurs compétences psycho-cognitives lors de leur admission. Elle considère alors que ces tests sont suffisamment analogues à ceux qu'elle envisageait d'appliquer en post-test (**B**) qu'ils peuvent être envisagés comme un pré-test (**A**) crédible.

2.1.4 - Le plan pré-test / post-test avec groupe témoin non équivalent avec échantillons séparés pour pré-test et post-test.

Description :



Ce plan de recherche est employé lorsque le chercheur croit que le simple fait d'administrer un pré-test risque d'influer sur le post-test de façon telle que l'interprétation des résultats dégagés ne pourra être qu'équivoque. De fait, par exemple, le fait d'évaluer une dimension subjective (par exemple, le bien – être ou la satisfaction au travail) peut parfois attirer l'attention du participant sur ce point, et modifier radicalement son point de vue, en dehors de tout impact d'un traitement consécutif. Pour pallier à ce phénomène, le chercheur peut sélectionner des échantillons différents 1) pour les groupes témoins et expérimentaux **et** 2) pour le pré-test et le post-test, soit en tout **quatre groupes** différents de participants. Ces distinctions sont illustrées dans la figure ci-dessus par les traits pointillés verticaux ou

horizontaux et par l'utilisation de lettres différentes pour chaque groupes de pré-test et de post-test.

Analyse :

Dans ce plan de recherche pré-test / post-test avec groupe témoin non équivalent avec échantillons séparés pour pré-test et post-test, la grande difficultés consiste à trouver ou constituer quatre groupes de participants indépendants et cependant globalement comparables. Si cette difficulté est dépassée, il convient remarquer que seul le groupe traitement évalué en post-test doit effectivement être soumis au traitement, à la situation ou au phénomène, ce qui en pratique peut grandement faciliter la mise en place de ce devis de recherche. Dans le même ordre d'idée, on peut signaler que seul le groupe contrôle évalué en post-test ne doit pas avoir été confronté d'une manière ou d'une autre au traitement, à la situation ou au phénomène.

Validité :

Toute la validité de ce plan de recherche repose sur les postulats 1) que les différents groupes de participants sont comparables et 2) qu'ils sont représentatifs de la population étudiée. Par conséquent, le contexte d'application qui serait le plus valide est celui où les individus d'une même population sont assignés aléatoirement aux groupes de pré-test et aux groupes de post-test. Toutefois, même lorsque le groupe pré-test et le groupe post-test sont comparables, il faut garder à l'esprit que le groupe expérimental et le groupe témoin ne le sont pas (c'est-à-dire O_A versus O_B et O_C versus O_D). Toutes les limites signalées dans le cas du plan pré-test / post-test avec groupe témoin non équivalent s'appliquent donc ici. De plus, ce plan de recherche ne permet pas de juger directement de l'équivalence des groupes post-tests à partir de la mesure au pré-test. En d'autres mots, la validité interne de ce plan est faible, et il est clair qu'il ne doit être utilisé qu'en dernier recours. La seule solution pour que ce plan ait une certaine validité est de constituer avec précaution les différents groupes afin d'éviter les biais d'échantillonnage, et de s'assurer que les groupes pré-tests et post-tests soient toujours comparables après le processus de sélection. A cette fin, le chercheur devra intégrer dans son outil de mesure des variables pertinentes pour s'assurer que les groupes du pré-test sont comparables aux groupes du post-test.

Exemple :

Un chercheur souhaite tester l'efficacité d'un programme d'amélioration de la qualité de vie à l'école. Cependant, il sait que le simple fait d'attirer l'attention des individus sur leur

qualité de vie suffit à modifier la perception qu'ils en ont. Il ne peut donc pas utiliser les mêmes participants au pré et au post-test. Pour limiter l'effet de ce biais, il va constituer à l'aide d'un tirage aléatoire deux groupes de participants équivalents. Ces deux groupes vont participer au programme (sans pour autant que leur participation soit explicite). Un questionnaire d'évaluation de la qualité de vie sera proposé aux participants de l'un des groupes quatre semaines avant le début du programme (**A**), et aux participants de l'autre groupe deux semaines après la fin de ce programme (**B**). Parallèlement, dans une autre école, deux autres groupes de participants seront aléatoirement constitués : l'un se verra proposer le questionnaire de qualité de vie pendant la même période que le premier groupe de traitement (**O_B**), l'autre au même moment que le second groupe de traitement (**O_D**). Aucun des participants de ces deux groupes analogues n'est intégré dans une quelconque intervention éducative.

2.2 – Les plans de recherche avec cohortes.

Lorsque la nature de la problématique ou les conditions d'exercice de la recherche ne permettent pas de constituer un groupe de comparaison qui ait été non exposé à la variable indépendante au moment même où cette variable était présentée au groupe traitement (par exemple, dans le cas de programmes de prévention ou de politiques gouvernementales appliquées à grande échelle et qui visent l'ensemble de la population), il est possible d'utiliser des plans de recherche avec cohortes. Le terme de *cohorte* fait référence à des groupes d'individus (par exemple, des promotions scolaires) qui se succèdent de façon cyclique dans des organisations humaines (par exemple, l'école) ou, plus généralement, à un groupe d'individus ayant vécu un événement de vie commun pendant une période variant de un à dix ans.² Ainsi, les élèves fréquentant la première année d'une école donnée constituent une cohorte qui sera remplacée par une nouvelle cohorte d'élèves l'année suivante. De même, un groupe d'élèves gendarmes intégrés après concours en école de gendarmerie peut être considéré comme une cohorte, distincte du groupe d'élèves entré l'année précédente.

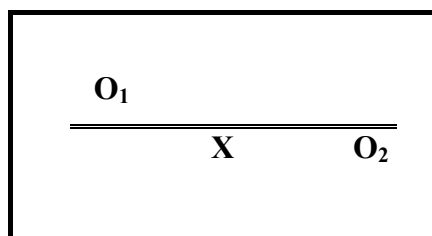
Dans un plan de recherche, l'existence de cohortes peut être mise à profit selon deux approches. D'une part, une approche dite *transversale* permet d'étudier l'effet de différences liées à l'âge sur le développement. D'autre part, une approche quasi-expérimentale permet de comparer plusieurs cohortes, certaines ayant reçues un traitement ou ayant été exposées à un phénomène ou à une situation, alors que d'autres non.

² Glenn, N. (1977). *Cohort analysis*. Beverly Hills (C.A.) : Sage Publications.

Cette dernière approche considère que les cohortes qui se suivent d'une année à l'autre ne diffèrent que minimalement les unes des autres (à l'exception bien sûr de l'exposition au traitement, au phénomène ou à la situation qui fait l'objet de l'évaluation), de telle sorte que l'une peut servir de comparaison à l'autre. De fait, une très grande partie de la validité de ce plan de recherche repose sur le postulat que deux cohortes consécutives provenant d'un même milieu ne sont pas plus différentes l'une de l'autre que deux groupes contemporains provenant de milieux différents. Ce postulat peut être discuté.

2.2.1 – Le plan avec cohortes et comparaison simple.

Description :



Dans le schéma ci-dessus, le double trait indique que les deux groupes de participants (les deux cohortes) ne peuvent pas être considérés comme tout à fait équivalents. De fait, ils n'ont pas été constitués à l'aide d'une procédure de répartition par tirage aléatoire à partir d'une même population. Le O_2 correspond aux résultats obtenus sur la cohorte ayant été exposée au traitement, à la situation ou au phénomène étudié. Ces résultats sont comparés à ceux obtenus sur une cohorte précédente (O_1), évaluée par exemple un an plutôt, et qui n'a pas été soumise à ce traitement, à cette situation ou à ce phénomène.

Analyse :

Le principe de ce plan de recherche est de comparer deux groupes distincts de participants sur la base d'évaluations décalées dans le temps. La première est réalisée avant la survenue d'un événement, d'une situation ou d'un traitement et le groupe concerné n'y a donc pas été exposé (il peut même ne pas avoir conscience ou connaissance de la possibilité d'occurrence de ce traitement, de cet événement ou de cette situation). La seconde est effectuée *a posteriori*, sur un groupe de participants qui a été confrontés au traitement. Les éventuelles différences entre ces deux mesures permettent d'estimer si le traitement (X) a eu, ou non, un effet et quelle peut être son amplitude. Dans ce type de plan, il arrive très souvent

que les chercheurs souhaitent tester l'impact d'un phénomène qu'ils n'ont pas toujours anticipé (par exemple, un événement politique ou un « *accident* » de l'histoire). Si l'évaluation post-phénomène (**O₂**) ne comporte pas de difficultés majeures, ces chercheurs n'ont par contre que très rarement eu l'opportunité de réaliser eux – mêmes une mesure sur une cohorte antérieure (**O₁**). L'obstacle principal à la mise en place de ce type de plan réside donc dans la possibilité de trouver une évaluation préalable concernant la thématique ciblée et réalisée sur la population étudiée (par exemple, en explorant des archives).

Validité :

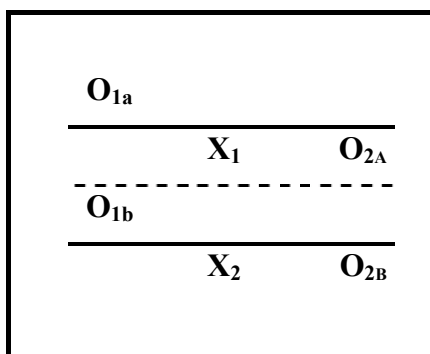
La validité de ce plan de recherche peut être remise en cause pour des raisons liées à la sélection et à l'expérience vécue des participants. Par exemple, certaines cohortes scolaires peuvent comporter une majorité d'âinés alors que d'autres peuvent comprendre une proportion plus importante de benjamins. Or, il a été montré que la position dans l'ordre des naissances est un facteur qui peut, entre autre, influencer l'acquisition du vocabulaire. Ainsi, selon l'objet d'étude, le chercheur devra contrôler *a posteriori* que ses cohortes sont au moins équivalentes par rapport aux principales variables susceptibles de venir parasiter la relation entre les variables indépendante (**X**) et dépendante (**O₁** et **O₂**).

Exemple :

Un sociologue souhaite estimer l'impact de la loi sur la laïcité dans les écoles, collèges et lycées publics (loi n° 2004 – 228 du 15 mars 2004) sur la perception que les élèves de Terminale ont du climat social dans leur établissement. Etant donné que cette loi est appliquée au même moment à l'ensemble des établissements scolaires publics du territoire, il n'est pas possible que ce chercheur trouve un groupe d'élèves n'ayant pas été confronté à ce changement législatif, et qui pourrait servir de point de comparaison. Néanmoins, ce chercheur trouve qu'une enquête sur le climat scolaire auprès des élèves de Terminale est réalisée tous les ans depuis une décennie dans les établissements du département par un laboratoire de l'I.N.R.P. (Institut National de Recherche Pédagogique). Il va se servir de ces données d'enquêtes obtenues auprès de cohortes précédentes (**O₁**) comme point de comparaison, auquel il va confronter les résultats de sa propre enquête (**O₂**), effectuée après la mise en application de la loi.

2.2.2 - Le plan avec cohortes et démantèlement du traitement.

Description :



Le plan de recherche avec cohortes et démantèlement du traitement est globalement similaire au précédent : les résultats obtenus sur deux cohortes (O_1 et O_2) sont comparés, la première n'ayant pas été exposée au phénomène (X), alors que la seconde l'a préalablement été. Cependant, le traitement, la situation où l'événement a été ici démantelé, c'est-à-dire que chaque cohorte O_1 et O_2 a été subdivisée en sous-groupes, chacun d'eux étant exposés à un degré ou à une modalité différente de la variable indépendante (X). Dans le schéma ci-dessus, X_1 et X_2 font référence à des différences de degré ou de modalité d'exposition à la variable indépendante. Les indices a et b font eux références aux sous-groupes issus de la subdivision des cohortes. Il faut souligner que le nombre de sous-groupes issus de cette subdivision n'est en aucun cas limité : tout dépend de l'hypothèse testée.

Analyse :

Relativement analogue au plan précédent, l'intérêt principal du plan de recherche avec cohortes et démantèlement du traitement est de permettre le test de l'effet présumé de différents niveaux ou modalités de traitement. Si la variable indépendante a un effet, il est possible de supposer 1) qu'il existera un écart entre les résultats obtenus sur les deux cohortes O_1 et O_2 et 2) que l'écart entre les deux conditions (X_1 et X_2) sera plus élevé dans les sous-groupes issus de la cohorte O_2 que dans les sous-groupes issus de la cohorte O_1 .

Validité :

Si l'écart entre les deux conditions (X_1 et X_2) est plus élevé dans les sous-groupes issus de la cohorte O_2 que dans les sous-groupes issus de la cohorte O_1 , ce résultat peut être considéré comme concluant, parce qu'il laisse peu de place aux hypothèses rivales, telles que par exemple un effet dû à un changement de procédure ou d'instrumentation, ou encore à un

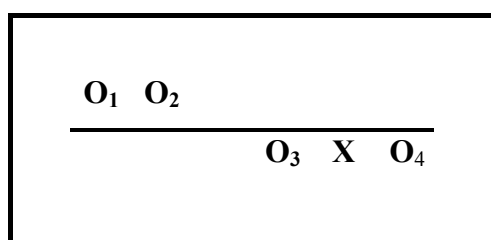
effet dû à l'expérience vécue. La validité interne de ce plan est donc supérieure à celle du plan avec cohortes et comparaison simple.

Exemple :

Le ministère de l'éducation de la République de Moldavie Orientale³ rend obligatoire l'enseignement des arts plastiques et de l'histoire de l'art dans les établissements du secondaire. Cependant, il ne précise pas les modalités pédagogiques et didactiques de cet enseignement, qui sont laissées à l'appréciation des professeurs. Un sociologue voudrait estimer l'impact de ce nouvel enseignement (**X**) sur les pratiques culturelles (**O**) des élèves de lycée (par exemple, leur fréquence de visite d'expositions ou de musées). Comme point de référence, il dispose de données d'enquête sur les pratiques culturelles, recueillies l'année précédente sur sa population d'étude (ce type d'enseignement n'était pas alors proposé). Néanmoins, comme il n'y a pas d'homogénéité du traitement **X** (plusieurs types de pratiques pédagogiques plus ou moins participatives sont repérés), notre chercheur va distinguer plusieurs types d'enseignement des arts plastiques et de l'histoire de l'art (**X**₁, **X**₂,, **X**_n), selon le degré de participation active de l'élève. Il fait l'hypothèse que plus l'élève sera amené à une participation active dans la classe, plus sa pratique culturelle hors les murs sera importante.

2.2.3 – Le plan avec cohortes et pré-test / post-test.

Description :



Il est parfois possible de combiner les caractéristiques (intérêts et limites) du plan avec cohortes et du plan pré-test / post-test. Cette situation se présente lorsque 1) il n'est pas possible de trouver un sous-groupe de la population d'étude qui puisse échapper au traitement dont on souhaite estimer l'effet (**X**), bien que 2) ce traitement et son occurrence soient tout à fait prévisibles et anticipables. Ce plan nécessite d'opérer une mesure de la Variable Dépendante (**O**₃) avant la survenue du traitement (**X**), et de réitérer cette mesure après son

³ Pays fictif.

administration (**O₄**). Dans l'idéal, deux mesures de la Variable Dépendante (**O₁** et **O₂**) devront avoir été effectuées préalablement sur les participants de la cohorte précédente, l'intervalle de temps séparant **O₁** et **O₂** étant équivalent à celui qui séparera **O₃** et **O₄**. Ce plan est aussi dit « *à cycles institutionnels* », parce qu'il est généralement utilisé afin d'estimer les effets d'un changement dans le cycle d'activités d'une institution.

Analyse :

La logique d'analyse des données récoltées à l'aide de ce plan de recherche est similaire à celle du plan pré-test / post-test avec groupe témoin. La différence est que la cohorte antérieure sert ici de groupe de comparaison (groupe témoin) à la cohorte qui a été soumise au traitement, à l'événement ou à la situation (**X**).

Validité :

La validité interne de ce plan de recherche repose sur la capacité du chercheur à démontrer que la cohorte du groupe témoin est équivalente à la cohorte ayant été exposé au traitement (**X**) en ce qui a trait aux variables jugées pertinentes par rapport à l'objet de l'étude. Elle est notamment fondée sur la comparaison des mesures de pré-test (**O₁** et **O₃**), celles-ci devant dans l'idéal ne pas se différencier.

Néanmoins, même si les mesures effectuées en **O₁** et **O₃** sont d'un niveau équivalent, la question de l'effet de l'expérience vécue n'est pas réglée. De fait, il se peut que les participants de la première cohorte aient vécu entre le pré-test et le post-test des événements particuliers qui expliquent les résultats obtenus. Ce type de variables parasites peut d'autant plus se manifester si le délai entre le pré-test et le post-test est important. Pour cette raison, il peut être utile d'étendre le plan de recherche en intégrant plusieurs cohortes avant celle exposée au traitement (**X**), qui seront considérées comme de multiples groupes de comparaison. Cette procédure permet de vérifier si la cohorte immédiatement antérieure est atypique, eu égard aux variables considérées.

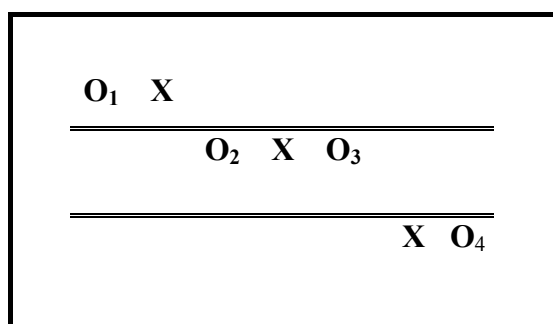
Exemple :

Un Institut de Formation en Soins Infirmiers a pour projet de mettre en place l'année prochaine un tutorat à distance à l'aide des T.I.C.E. des étudiants de première année en stage professionnel. L'objectif de ce dispositif est de limiter le stress des élèves infirmiers lors de leurs premiers contacts avec le terrain professionnel. Pour financer ce projet, l'institut candidate sur un appel à projet du ministère de l'éducation, qui exige entre autre qu'une

évaluation de l'impact de cette modalité pédagogique soit prévue. L'année précédant la mise en place du projet, l'équipe des enseignants de cet institut décide donc d'effectuer deux mesures du stress des étudiants, avant (O_1) et après (O_2) leur stage. Ces résultats seront comparés à ceux obtenus l'année suivante lors de mesures réalisées aux mêmes périodes (O_3 et O_4), la seule différence étant que les étudiants auront bénéficié du tutorat à distance (X) lors de leur stage.

2.2.4 – Le plan avec cohortes multiples selon un cycle récurrent.

Description :



Ce plan de recherche est utilisé pour évaluer l'effet d'une pratique ou d'un phénomène régulier, implantée depuis de nombreuses années. Il prend en considération trois cohortes. La première est constituée d'individus évalués avant le traitement (O_1). La deuxième est formée d'individus évalués avant (O_2) et après (O_3) le traitement l'année suivante (c'est-à-dire, par exemple, lors d'une deuxième année). La troisième est constituée de participants évalués en post-test (O_4), lors d'une troisième année.

Analyse :

L'effet du traitement X peut être observé à trois occasions : 1) en comparant O_2 et O_3 , 2) en comparant O_1 et O_3 , et 3) en comparant O_2 et O_4 .

Validité :

Tout d'abord, il est possible de vérifier si O_1 est équivalent à O_2 . Dans le cas contraire, cela remet en cause le postulat d'une équivalence entre les cohortes. De même, si O_3 est différent de O_4 , cet effet peut suggérer la non équivalence des cohortes ou la présence d'un effet du contexte historique qui génère différences dans l'expérience vécue du traitement (X). Par contre, comme O_2 et O_3 concernent les mêmes individus, les différences entre ces deux

observations ne peuvent s'expliquer par un biais de sélection. La validité interne de ce plan est donc plus préservée que celles de la plupart des autres plans avec cohortes. Cependant, l'utilisation de plus d'une mesure (pré-test et post-test) peut parfois poser des problèmes.

Exemple :

Une journée d'information sur les risques de la consommation d'alcool au travail est menée chaque année auprès des Terminales d'un lycée professionnel. Le directeur de l'établissement voudrait avoir une estimation de l'impact de cet événement sur la perception des risques des élèves. Pour se faire, il pratique d'abord à l'aide d'un questionnaire une évaluation du risque auprès d'une première cohorte d'élèves, avant que ceux-ci aient participé à cette journée (**O₁**). Puis, l'année suivante, sur une autre cohorte d'élèves, il effectue à l'aide du même outil des mesures avant (**O₂**) et après (**O₃**) la journée d'information. Enfin, la troisième année, il réalise une autre estimation de la perception du risque sur une nouvelle cohorte, après que celle-ci n'ait participé à l'événement (**O₄**).

2.3 – Les plans de recherche à séquence temporelle interrompue.

Un plan de recherche présente une séquence temporelle lorsque plusieurs observations sont répétées pendant une période de temps donnée. Ces mesures peuvent se rattacher à un même groupe d'individus, ou se rapporter à un objet plus global, comme par exemple la classe, l'école ou le quartier, c'est-à-dire un espace dans lequel des individus différents se succèdent avec le temps (des cohortes). La procédure de séquence temporelle peut contribuer à deux types d'études. D'une part, elle peut permettre de décrire l'évolution naturelle ou le développement d'un phénomène. On parle alors d'étude longitudinale. D'autre part, elle peut aussi permettre d'évaluer l'effet dans le temps d'un événement ou d'un traitement particulier. Cette stratégie quasi-expérimentale s'appuie essentiellement sur l'analyse des tendances observées avant et après la survenue de l'événement ou du traitement en question, et plus particulièrement sur l'observation d'une discontinuité entre la tendance observée avant le traitement ou l'événement, et celle observée après. Ce type de protocole suppose donc qu'il soit possible de déterminer précisément le moment de l'apparition de l'événement ou de l'application du traitement.

2.3.1 – Le plan à séquence temporelle interrompue unique.

Description :

| |
|--|
| O₁ O₂ O₃ O₄ X O₅ O₆ O₇ O₈ |
|--|

Ce plan de recherche est le protocole à séquence temporelle le plus élémentaire. Il est comprend un seul groupe de participants, qui est évalué à plusieurs reprises, avant (**O₁, O₂, O₃, O₄**) et après (**O₅, O₆, O₇, O₈**) l'exposition au traitement, à l'événement ou à la situation (**X**).

Analyse :

La fonction de ce plan de recherche est principalement de permettre l'évaluation de l'effet de l'exposition à un traitement, un événement ou une situation par la comparaison des tendances observées avant et après (ou pendant) le traitement dans le même groupe de participant. Fondamentalement, ce plan de recherche n'est en fait rien d'autre qu'un plan pré-test / post-test à groupe unique, auquel on aurait ajouté plusieurs mesures de la Variable Dépendante, afin de pouvoir évalué la présence et l'importance de biais potentiels liés à la maturation, à l'utilisation de plus d'une mesure ou au phénomène de régression statistique. Dans la mesure où ces biais sont présents, ils devraient apparaître lors de l'examen de la tendance avant l'exposition au traitement **X** (mesures de **O₁ à O₄**), et donc ne pas être confondus avec l'effet dû stricto sensu à ce traitement.

Validité :

Les principales menaces sur la validité interne de ce plan de recherche ont pour origine 1) l'expérience vécue et 2) l'éventuelle fluctuation de l'instrument de mesure. D'une part, il est possible que des phénomènes particuliers, extérieurs à l'intervention **X** elle-même, se produisent entre **O₄ à O₅**, et qu'ils soient responsables des changements observés dans la tendance. D'autre part, lorsque l'étude s'étend sur une longue période de temps, il se peut que la définition opérationnelle ou la sensibilité de la variable dépendante risque de se modifier, entraînant un phénomène de fluctuation de l'instrument de mesure. Dans un même ordre d'idée, il est possible que la procédure d'utilisation de l'outil de mesure s'altère sensiblement de mesures en mesures, entraînant une illusion de changement avant et après l'intervention.

Pour terminer, le chercheur doit s'assurer 1) que l'exposition au traitement **X** n'ait pas entraîné une déperdition des participants à l'étude, de sorte que la composition du groupe avant et après le traitement ne soit plus équivalente (*biais de sélection des participants*) et 2) que la répétition des mesures n'attire pas de manière excessive l'attention des participants sur l'objet de l'étude, générant chez ces derniers des préoccupations ou des intérêts qui ne se seraient jamais manifestés autrement (*biais dû à la répétition de la mesure*).

Exemple :

Un chercheur en Sciences de l'éducation voudrait tester l'effet de la participation à un programme de prévention du V.I.H. (**X**) sur les comportements sexuels déclarés des étudiants (**O**). Pour se faire, au cours des quatre mois précédant l'intervention préventive, il interroge à trois reprises un groupe d'étudiants à propos de leurs comportements sexuels (**O₁, O₂, O₃**), ceci à l'aide d'un questionnaire. Puis, à la suite de l'intervention, il réitère à trois reprises la passation de son questionnaire (**O₄, O₅, O₆**). Il espère ainsi pouvoir 1) observer si la tendance des réponses au questionnaire se différencie avant et après l'intervention, 2) contrôler l'effet d'un éventuel événement extérieur à l'intervention sur la Variable Dépendante (par exemple, la révélation publique de la séropositivité d'un personnage célèbre : le fameux **effet « Majic Johnson »**), et 3) évaluer la persistance dans le temps de l'effet de l'intervention **X**.

2.3.2 – Le plan à séquence temporelle interrompue avec groupe témoin non équivalent.

Description :

| | | | | | | | | |
|-----------------------|-----------------------|-----------------------|-----------------------|----------|-----------------------|-----------------------|-----------------------|-----------------------|
| O₁ | O₂ | O₃ | O₄ | X | O₅ | O₆ | O₇ | O₈ |
| O_{c1} | O_{c2} | O_{c3} | O_{c4} | | O_{c5} | O_{c6} | O_{c7} | O_{c8} |

Ce plan de recherche comprend un groupe expérimental et un groupe contrôle, qui sont évalués à plusieurs reprises, aux mêmes moments, avant (respectivement, **O₁, O₂, O₃, O₄** et **O_{c1}, O_{c2}, O_{c3}, O_{c4}**) et après (respectivement, **O₅, O₆, O₇, O₈** et **O_{c5}, O_{c6}, O_{c7}, O_{c8}**) l'exposition au traitement, à l'événement ou à la situation (**X**).

Analyse :

Ce plan est relativement analogue au précédent. Il permet juste une comparaison supplémentaire avec les tendances observées dans un groupe non équivalent. C'est à dire de comparer 1) l'évolution des tendances de $[O_1 \text{ à } O_4]$ et de $[O_{c1} \text{ à } O_{c4}]$ et l'évolution des tendances de $[O_5 \text{ à } O_8]$ et de $[O_{c5} \text{ à } O_{c8}]$.

Validité :

Si ce plan de recherche présente globalement les mêmes limites que le précédent, il permet cependant une meilleure appréhension des éventuels facteurs parasites associés à l'expérience vécue et aux fluctuations de l'instrument de mesure. Plus le chercheur pourra démontrer que le groupe exposé au traitement **X** et le groupe contrôle sont exposés aux mêmes sources de variation (en dehors bien sûr des différences associées au traitement), plus les différences de tendances observés entre les deux groupes pourront donner lieu à interprétation.

Exemple :

Chaque année, une classe de troisième d'un lycée technologique participe à concours d'invention, qui nécessite que ces membres s'investissent intensément dans le projet pendant une durée d'un mois. Cependant, le proviseur de l'établissement se demande si la progression des élèves dans les autres matières enseignées ne souffre pas de ce surinvestissement. Il décide donc de comparer la tendance de progression d'acquisition des savoirs dans deux classes équivalente, la seule donnée les distinguant étant que l'une participe au concours alors que l'autre non. Si les tendances de progression d'acquisition des connaissances dans les autres matières sont comparables entre les classes, avant et après la participation au concours, le proviseur pourra raisonnablement en conclure que l'engagement de la classe dans le projet ne perturbe pas l'acquisition des connaissances dans d'autres domaines.

2.3.3 - Le plan à séquences temporelles multiples avec traitement décalé.

Description :

| | | | | | | | | |
|-------|-------|-------|-------|----------|-------|----------|-------|-------|
| O_1 | O_2 | O_3 | O_4 | O_5 | O_6 | X | O_7 | O_8 |
| O_1 | O_2 | O_3 | O_4 | X | O_5 | O_6 | O_7 | O_8 |

Dans ce plan, tous les groupes de participants ont été exposés au traitement, à l'événement ou à la situation. Néanmoins, ils l'ont été à des moments différents. C'est un type de protocole particulièrement utile lorsqu'il n'est pas déontologiquement acceptable qu'un groupe soit privé du traitement **X**. Celui-ci est juste *différé* pour l'un des groupes.

Analyse :

Ce plan de recherche est un cas particulier du protocole précédent. Ici la tendance observée dans le groupe qui reçoit en premier le traitement **X** est comparée avec celle observée dans le groupe qui ne l'a pas encore reçu. En d'autres termes, la comparaison portera essentiellement sur l'évolution des mesures entre d'une part [**O₄** et **O₅**] pour le premier groupe et d'autre part [**O₄** et **O₆**] pour le deuxième groupe.

Validité :

Les obstacles à la validité de ce plan sont globalement ceux reprochés au plan précédents. L'avantage de celui-ci est qu'il permet de tester la généralisation des résultats à plus d'une population, dans des contextes différents, et à des moments différents (c'est-à-dire d'accroître la confiance en la validité externe de ce plan), tout en offrant un certain contrôle des facteurs menaçant la validité interne.

Exemple :

Un chercheur souhaite tester l'efficacité d'un programme scolaire de prévention de la violence au sein des établissements mis en place au niveau d'une académie entière. Ce programme vise notamment à augmenter les attitudes égalitaires et la croyance en sa capacité à ne pas recourir à la violence, ainsi qu'à diminuer l'occurrence d'attitudes de contrôle abusif d'autrui et d'attitudes d'agression. Il concerne toutes les classes de troisième de l'académie, et il consiste en deux séances de 75 minutes de discussion en petits groupes autour des thèmes pré-cités. Toutes les classes de troisième participent au programme, mais faute de personnel qualifié, toutes ne reçoivent pas l'intervention au même moment de l'année. Cette organisation permet au chercheur de pratiquer tout au long de l'année une observation répétée des comportements agressifs dans les établissements scolaires concernés (une grille de recensement précise et détaillé est renseignée chaque mois de l'année par le C.P.E. de chaque établissement), et d'observer l'effet de l'intervention sur l'occurrence de ces comportements, cette dernière se produisant à des moments différents de l'année selon les établissements.

3 – LES PLANS DE RECHERCHE EXPERIMENTAUX.

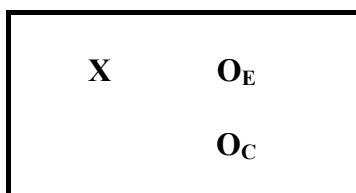
Un plan de recherche expérimental n'est pas forcément un *design* complexe, nécessitant la mise en place d'un dispositif lourd. De fait, pour être qualifié d'expérimental, un plan de recherche doit simplement réunir trois caractéristiques : 1) l'existence d'au moins une Variable dépendante, 2) la présence d'au moins une Variable Indépendante, et 3) la possibilité de *manipuler* cette Variable Indépendante, c'est-à-dire d'assigner de manière aléatoire les participants de l'échantillon à l'une des conditions expérimentales créées par la manipulation de la variable indépendante. Deux grandes formes de plans de recherche expérimentaux sont identifiées : les **plans à groupes indépendants** et les **plans à groupes dépendants**.

3.1 – Les plans à groupes indépendants.

La principale caractéristique des plans à groupes indépendants est qu'ils peuvent intégrer autant de *groupes de participants* que de *groupes de traitements*.

3.1.1 – Le post-test seul avec groupe témoin.

Description :



Dans ce plan, un groupe (le groupe expérimental identifié par l'indice _E) est exposé à un traitement ou à une situation (**X**), alors qu'un autre groupe ne l'est pas et sert de groupe témoins (le groupe témoin ou contrôle identifié par l'indice _C). A la suite du traitement, les deux groupes sont mesurés sur la variable dépendante (**O**). Comme pour les plans pré et quasi – expérimentaux, **O** peut être composé de plusieurs mesures (plusieurs indicateurs), c'est à dire de plusieurs variables dépendantes.

Analyse :

Comme les participants sont assignés au hasard dans l'un ou l'autre des groupes, ceux-ci sont donc considérés comme équivalents au départ et différents seulement selon qu'ils aient reçu ou non le traitement (**X**). En comparant les résultats obtenus en **O_E** et en **O_C**, ce plan

permet d'observer si le groupe expérimental se différencie du groupe contrôle par rapport à la dimension (**X**) considérée. De fait, si $O_E \neq O_C$, et que l'expérience n'est pas biaisée, la seule explication possible de cette différence est la présence ou non du traitement (**X**). Il faut souligner que le groupe témoin n'est pas nécessairement un groupe qui ne reçoit aucun traitement : il peut s'agir d'un groupe qui reçoit un traitement *placebo*, qui n'est pas susceptible a priori d'avoir un effet sur la dimension évaluée par **O**.

Validité :

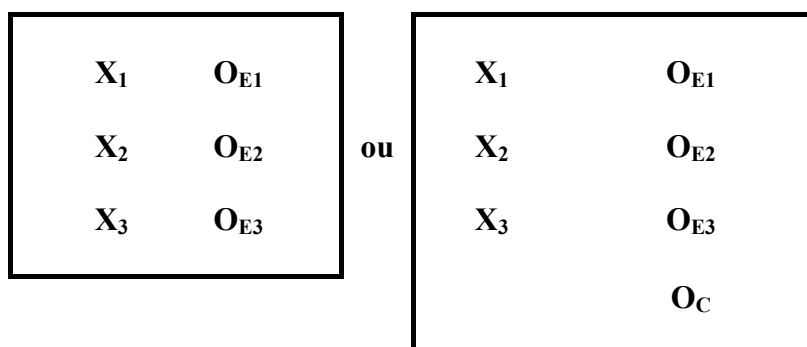
Si les participants sont réellement assignés de manière aléatoire dans les deux conditions (expérimentale et contrôle), il n'y a pas de raison de penser que la *validité interne* de la recherche puisse être menacée. Par ailleurs, comme ce plan ne comprend pas de pré-test, il n'y pas de raison de penser que la *validité externe* puisse être menacée par la répétition de la mesure ou par une interaction prétest – traitement (**X**). Ce plan, qui est l'un des plus simples dans sa structure, est probablement aussi l'un des plus fiables.

Exemple :

Un chercheur souhaite vérifier si un contexte de co-présence (**X**) a une influence sur la performance dans une tâche de résolution de problème (**O**). A partir de son échantillon de participants, il constitue deux groupes sur la base d'un tirage aléatoire : chacun des participants du premier groupe (**E**) va devoir travailler en situation de co-présence (une personne se tient à ses côtés dans la pièce où il travaille), alors que chacun de ceux du deuxième groupe (**C**) va devoir travailler en situation d'isolement (il est seul dans la pièce où il travaille). La tâche à réaliser est une série de petits problèmes de logique, et la Variable Dépendante (**O**) le temps mis pour les résoudre (en secondes). L'outil de mesure est un chronomètre déclenché et arrêté par le participant. Si le temps moyen chronométré pour O_C est supérieur à celui chronométré pour O_E , alors le chercheur pourra conclure que la situation de co-présence accroît la rapidité d'exécution de ce type de tâche.

3.1.2 – Le plan à variable indépendante ou catégorique.

Description :



Ce plan de recherche n'est pas fondamentalement différent du précédent. Simplement, la Variable Indépendante (X) peut être ici catégorique (elle peut comporter plusieurs modalités de traitement) ou continue (elle peut comporter différents degrés d'un traitement). Même si dans l'illustration ci-dessus il n'y a que trois traitements (X_1 , X_2 , X_3), il n'y a théoriquement pas de limitation au nombre de traitements différents que l'on peut souhaiter tester. En fait, il peut y avoir autant de traitements que le chercheur le juge utile. Comme précédemment, les participants sont assignés au hasard dans l'un ou l'autre des groupes, ce qui rend les groupes équivalents entre eux, sauf sur la variable indépendante (X). Enfin, ce type de plan n'exclut pas la présence d'un groupe témoin (O_C), même si l'accent est ici mis sur la comparaison entre les différents niveaux de traitement.

Analyse :

Ce plan de recherche est particulièrement approprié lorsque le test de l'hypothèse nécessite de comparer les effets de différentes catégories de traitements sur une mesure ou d'examiner l'effet d'une variable indépendante continue sur une variable dépendante. Il permet d'observer si les groupe expérimentaux se différencient entre eux et se différencient du groupe contrôle par rapport à la dimension considérée, en comparant les résultats obtenus en O_{E1} , O_{E2} , O_{E3} et en O_C . De fait, si $O_{E1} \neq O_{E2} \neq O_{E3} \neq O_C$, et que l'expérience n'est pas biaisée, la seule explication possible de cette différence est la présence ou non du traitement (X).

De plus, avec ce type de plan, il est possible de n'établir que des différences entre certains groupes et pas entre d'autres. Toutes les combinaisons sont possibles. Par exemples, $O_{E1} = O_{E2} = O_{E3} \neq O_C$ ou $O_{E1} = O_{E2} \neq O_{E3} = O_C$

Validité :

Les critères de validité sont similaires à ceux du plan précédent.

Exemple :

Un chercheur souhaite vérifier si le niveau de bruit dans un environnement de travail (**X**) influe sur les comportements observés (**O**) dans une tâche nécessitant une coopération entre les membres d'un groupe. Il répartit aléatoirement les individus composant son échantillon en quatre groupes, qui correspondent à quatre niveaux différents de bruit dans l'environnement de travail : 30, 60, 90 et 120 décibels. Les participants travaillent par groupe de cinq, autour d'une table circulaire. La tâche réalisée et les conditions de réalisation de cette tâche sont globalement identiques pour tous les participants. Seul le niveau de bruit auquel ils sont exposés varie (**X₁**, **X₂**, **X₃**, **X₄**). Pendant leur épreuve, le chercheur recense le type et le nombre de leurs comportements de coopération à l'aide d'une grille d'observation. Il suppose qu'en moyenne 1) plus l'environnement est bruyant, moins les comportements de coopération seront nombreux, et 2) que ce phénomène sera d'autant plus vrai pour les comportements verbaux.

3.1.3 – Le plan avec variables concomitantes.

Description :

| | | |
|----------|----------------------|-----------------------|
| C | X₁ | O_{E1} |
| C | X₂ | O_{E2} |
| C | X₃ | O_{E3} |

Ce plan est très similaire aux deux précédents. La seule différence est l'introduction d'une variable concomitante (**C**). Cette variable (**C**) est mesurée avant la mise en place du ou des traitements. Elle fait référence à des facteurs de personnalité ou à des caractéristiques individuelles. Il peut s'avérer quelques fois important de tenir compte de différences individuelles (variables modératrices ou médiatrices) de départ lorsqu'on évalue une dimension (**O**) après traitement (**X**).

Analyse :

Il s'agit ici de constater si les différents groupes expérimentaux se différencient entre eux par rapport à la dimension considérée (comparaison des résultats obtenus en O_{E1} , O_{E2} et O_{E3}), compte tenu de l'influence de la variable **C**. Ce type d'analyse complexe nécessite souvent le recours à *l'analyse de régression multiple*. Ce type d'analyse permet surtout d'observer quels sont les poids respectifs de **X** et de **C** dans les différences entre O_{E1} , O_{E2} et O_{E3} .

Validité :

Pour que la validité de ce plan soit bonne, il faut éviter que la mesure de la variable concomitante n'interagisse avec le traitement. Pour éviter toute contamination, il convient de procéder à la mesure de **C** bien avant la mise en place du traitement. Pour le reste, les menaces sur les validités interne et externe sont relativement réduites avec ce type de plan.

Exemple :

Un professeur de Français Langue Etrangère souhaite tester l'efficacité de deux approches pédagogiques innovantes (**X**), auprès d'un échantillon d'individus adultes nouveaux migrants souhaitant acquérir les bases du français utilitaire (**O**). Il répartit les individus inscrits à ses cours de façon aléatoire dans trois conditions expérimentales, selon la pédagogie auquel ils sont exposés : Pédagogie innovante 1 (X_1), Pédagogie innovante 2 (X_2), Pédagogie traditionnelle (X_3). Le niveau de départ en français avant le début du traitement est testé et utilisé comme variable concomitante (**C**). Au bout de dix semaines de cours, le professeur effectue une analyse de co-variance sur les résultats d'un nouveau test de français, afin d'observer lequel des trois groupes a en moyenne le plus progressé dans cette discipline.

3.1.4 – Le plan factoriel pour groupe indépendant.

Description :

| | | |
|-------|-------|----------|
| A_1 | B_1 | O_{E1} |
| A_1 | B_2 | O_{E2} |
| A_2 | B_1 | O_{E3} |
| A_2 | B_2 | O_{E4} |

Un plan est dit factoriel lorsqu'il permet de tester plus d'un facteur en même temps. Dans l'exemple ci-dessus, un plan factoriel de type 2×2 est présenté. Cela veut dire que les effets de deux facteurs (**A** et **B**) sont testés ensemble. Ici, chacun des facteurs (Variables Indépendantes) comporte deux degrés. Il y a donc deux niveaux pour le facteur **A** (**A₁** et **A₂**) et deux niveaux pour le facteur **B** (**B₁** et **B₂**). Chaque participant n'est exposé qu'à un niveau du facteur **A** et qu'à un niveau du facteur **B**. Comme précédemment, les participants sont évalués (**O_E**) après avoir reçu leur combinaison de traitements (**A_i x B_j**). Comme il s'agit d'un plan à groupe indépendant, il y a autant de groupe de participants que de combinaisons possibles de traitement (Ici $2 \times 2 = 4$). Le plan de type 2×2 est le plus simple. Cependant, il peut y avoir plus de deux facteurs (V.I.) manipulés en même temps (par exemple, $A_i \times B_j \times C_k$) ou plus de deux degrés pour un facteur (par exemple $2 \times 2 \times 3$). L'avantage du plan factoriel est qu'il permet d'examiner l'effet de plusieurs facteurs, et l'effet combiné de ces facteurs.

Analyse :

Il s'agit de constater si les différents groupes expérimentaux se différencient entre eux par rapport aux facteurs considérés et à la combinaison de ces facteurs (une réduction sur les facteurs est souvent nécessaire). Ce type d'analyse complexe nécessite souvent le recours à *l'analyse de variance*.

Validité :

Comme pour les plans de cette catégorie, l'assignation aléatoire des participants dans les différentes conditions expérimentales permet de préserver la validité interne. Et comme il n'y a pas de pré-test ou de variable concomitante, la validité externe n'est pas menacée.

Exemple :

Un nutritionniste souhaite tester l'efficacité de différentes approches d'éducation à la nutrition (**X**) auprès d'un échantillon d'individus adultes obèses désirant réduire leur poids (**O**). Deux types d'approches l'intéresse particulièrement : l'auto-contrôle (**A**) et le contrôle externe (**B**). Il répartit de façon aléatoire les participants de son étude dans quatre conditions expérimentales, selon le traitement auquel ils sont exposés : auto-contrôle seul (**X₁ : A₁B₀**), contrôle externe seul (**X₂ : A₀B₁**), auto-contrôle et contrôle externe (**X₃ : A₁B₁**), groupe témoin (**X_C : A₀B₀**). Au bout de dix semaines de traitement, le nutritionniste effectue une analyse de variance afin d'observer lequel des quatre groupes a perdu en moyenne le plus de poids (Variable Dépendante).

3.2 - Les plans à groupes dépendants.

Dans ce type de plan, le ou les mêmes groupes de participants peuvent recevoir l'ensemble des traitements possibles de manière séquencée.

3.2.1 - Le pré-test / post-test avec groupe témoin.

Description :

| | | |
|----------|---|-----------|
| $O_{E'}$ | X | $O_{E''}$ |
| $O_{C'}$ | | $O_{C''}$ |

Ce plan diffère du post-test seul avec groupe témoin par l'ajout d'un pré-test au groupe témoin ($O_{C'}$) et au groupe expérimental ($O_{E'}$). Par la suite, le traitement est administré au groupe expérimental et une mesure est réalisée après traitement (X) pour les deux groupes ($O_{C''}$ et $O_{E''}$). Ce plan est souvent appelé aussi avant-après, et il particulièrement utilisé lorsque l'on souhaite évaluer un changement après un traitement quelconque.

Analyse :

Dans ce type de plan, il s'agit le plus souvent de calculer une différence de scores entre le pré et le post-test (Avant-Après), et de comparer les résultats obtenus entre les groupes expérimentaux [$O_{E''} - O_{E'}$] et témoins [$O_{C''} - O_{C'}$] (par exemple, en procédant à un test statistique comme le t de Student). Il faut souligner qu'il est possible de complexifier ce plan en introduisant une Variable Indépendante (X) catégorielle ou continue.

Validité :

La validité interne est la même que pour un plan post-test seul avec groupe témoin. Par contre, comme pour le plan avec variable concomitante, la validité externe est menacée par la présence d'une pré-mesure.

Exemple :

Un chercheur souhaite tester l'effet du phénomène de la dissonance cognitive sur les évaluations posées par les enseignants de français. Sous prétexte de contribuer aux corrections d'un concours d'entrée dans une grande école, il demande à un échantillon de professeurs de

français de corriger un paquet de copies (fictives), en les informant qu'une procédure de double correction a été mise en place pour accroître la fiabilité des notes données et l'équité du concours (O_E). Une semaine après rendu son paquet de copies corrigées, chaque professeur participant à l'étude reçoit à nouveau son paquet de copies. Dans une première condition, le paquet est accompagné des notes données par l'autre correcteur, et les évaluations des deux correcteurs sont quasi-identiques (*condition de dissonance faible*). Dans une deuxième condition le paquet est aussi accompagné des notes données par l'autre correcteur, mais les évaluations du second correcteur sont systématiquement plus élevées que celles du professeur (*condition de dissonance forte*). Enfin, dans une troisième condition, le paquet n'est pas accompagné des notes de l'autre correcteur (*condition contrôle*). Quelque soit la condition, le chercheur propose au professeur, s'il le souhaite, de réévaluer les notations qu'il a posées, en tirant éventuellement bénéfice de cette semaine de recul. Notre chercheur fait l'hypothèse que les professeurs intégrés dans la deuxième condition (*dissonance forte*) qui seront les susceptibles de modifier leurs notations.

3.2.2 – Le plan factoriel pour groupe dépendant (avec variables injectées).

Description :

| | | | |
|-------|-------|-------|-------|
| A_i | B_j | C_k | O_E |
|-------|-------|-------|-------|

Dans ce type de plan, un groupe de participant est soumis à toutes les combinaisons possibles des différents traitements. Ce plan est principalement utilisé dans les méthodologies impliquant l'utilisation de scénarios (par exemple, une matrice d'allocation de ressources ou un protocole de jugement ou un questionnaire).

Analyse :

Ce type d'analyse complexe nécessite souvent le recours à *l'analyse de variance*. A la base, une réduction sur les facteurs est nécessaire pour évaluer l'effet de chaque facteur et l'effet de chaque combinaison de facteur

Validité :

Comme pour les plans expérimentaux précédents, l'assignation aléatoire des participants dans les différentes conditions expérimentales permet de préserver la validité

interne. Et comme il n'y a pas de pré-test ou de variable concomitante, la validité externe n'est pas menacée.

Exemple :

Un chercheur se demande comment de jeunes enfants de 4 à 5 ans jugent moralement les actes délictueux. Pour se faire il propose à un échantillon de 25 enfants de cette classe d'âge d'attribuer une punition plus ou moins sévère (**O**) à l'auteur d'un acte transgressif, cet acte pouvant se distinguer par 2 facteurs : sa gravité (3 niveaux de gravité → **A₃**) et son intentionnalité (3 niveaux d'intentionnalité → **B₃**). Chaque enfant doit donc juger de la punition à attribuer en fonction de $3 \times 3 = 9$ actes délictueux. Une réduction sur les facteurs permettra d'évaluer 1) quel facteur a eu le plus de poids dans la détermination de la sévérité de la punition à attribuer (la gravité ou l'intentionnalité), et 2) comment ces facteurs s'associent entre eux.

EN RESUME ET POUR CONCLURE,

un ensemble non exhaustif de plans de recherche est décrit dans ce document. Chacun d'entre eux nécessite un mode d'analyse et d'interprétation qui lui est propre, et satisfait à des degrés divers aux exigences des critères de validité interne ou externe. Si le choix du plan de recherche est bien entendu influencé en partie par les conditions d'exercice et de faisabilité de la recherche, le concepteur du protocole de recherche doit aussi se poser la question de l'adéquation entre les objectifs de son étude (les hypothèses à tester) et les caractéristiques des différentes possibilités de plan à sa disposition.

Pour des raisons didactiques, ces plans de recherche ont été classifiés en trois grandes catégories : les plans pré-expérimentaux (ou cliniques), les plans quasi-expérimentaux, et les plans expérimentaux. Néanmoins, les possibilités de combinaisons entre les types de plan, à l'intérieur de chaque catégorie ou entre catégories, sont multiples, et le chercheur doit faire preuve d'inventivité plutôt que de chercher à appliquer des recettes pas toujours pertinentes. Il ne devrait donc pas hésiter à concevoir son propre protocole en fonction 1) des principes à respecter pour établir une inférence causale valide, 2) des objectifs de son étude, et 3) des possibilités qui lui sont offertes par le contexte opératoire.