

4

Phases d'une recherche exploratoire

LES EXPLICITATIONS PRÉLIMINAIRES

- 1 La description du problème contextualisé
- 2 L'explicitation des préconceptions
 - 2.1 *Le rôle professionnel*
 - 2.2 *L'épistémologie*
 - 2.3 *La conception de l'éducation*
 - 2.4 *Les orientations méthodologiques*
 - 2.5 *Les orientations théoriques privilégiées*
 - 2.6 *Les critères retenus pour la validité*
 - 2.6.1 La crédibilité
 - 2.6.2 La rationalité
 - 2.7 *Les visées de la recherche*
- 3 Le protocole de diffusion
- 4 La concordance entre les préconceptions, les visées, le problème et les options méthodologiques

Comme première phase d'une recherche exploratoire, l'explicitation des préliminaires fait appel à la plupart des questions épistémologiques et méthodologiques abordées dans les deux premières parties. Sa particularité réside dans l'effort que fournit le chercheur pour en réaliser une synthèse qui campe le décor à l'intérieur duquel le problème sera par la suite conceptualisé et opérationnalisé (chapitre 17). Les préliminaires comportent trois éléments : la description du problème, l'explicitation des préconceptions et le choix du protocole de diffusion.

1 La description du problème contextualisé

La description du problème auquel la recherche tentera d'apporter une réponse constitue la première étape de toute recherche exploratoire. Le tableau 6.1 résume les principaux éléments qui en font partie.

- | |
|--|
| 1.1 Énoncé local dans les catégories et les perspectives des acteurs |
| 1.2 Contexte de la situation dans laquelle baigne le problème |
| 1.3 Implications du problème sur les acteurs et le milieu |
| 1.4 Conséquences du problème |
| 1.5 Priorités à envisager lors de la résolution du problème |
| 1.6 Contraintes du contexte délimitant les solutions du problème |

Tableau 6.1 : Éléments de contextualisation d'une description du problème.

Si la recherche pédagogique part d'un problème pratique – or, la pédagogie est une pratique sociale –, le chercheur ne peut ignorer que les acteurs utilisent une terminologie qui leur est propre, une terminologie de terrain, pour poser leur problème. Quelle est cette terminologie ? En quels termes le problème est-il perçu et décrit par ceux qui le vivent ? La recherche de cette terminologie par le chercheur et la reformulation du problème dans ces termes sont importantes pour plusieurs raisons. D'abord, si le chercheur est capable de reformuler le problème dans des mots que les acteurs comprennent, il est probable qu'il a été capable de comprendre leur problème. Ensuite, la maîtrise de cette terminologie est essentielle pour que le chercheur puisse traduire ses hypothèses de solution dans la terminologie des acteurs, afin que ceux-ci puissent les comprendre et évaluer la pertinence de ces hypothèses par rapport aux problèmes qu'ils vivent. Enfin, si le chercheur veut pouvoir diffuser son travail de sorte qu'il soit utilisé par les acteurs, il faut qu'il puisse aussi en présenter une version qui utilise leur terminologie.

La compréhension du problème tel qu'il se pose sur le terrain de l'action pédagogique implique aussi que le chercheur connaisse bien le contexte de la situation où le problème est posé. Comme une phrase est souvent mal comprise quand elle est isolée de son «cotexte», un problème de recherche est aussi souvent mal posé quand on ne tient pas compte de son contexte. Dans les éléments du contexte, le chercheur doit s'enquérir des implications et des conséquences du problème sur les acteurs et sur leur milieu. Par exemple, l'aménagement d'une bibliothèque que les élèves n'ont pas le temps de fréquenter a probablement moins d'impact que l'implantation d'une technologie. En effet, si celle-ci exige que les enseignants consacrent beaucoup de temps à préparer leurs interventions, elle peut leur donner l'impression de ne pas avancer dans leur programme et de consacrer l'image d'un professeur sans rendement ! L'importance des conséquences sur les acteurs et leur milieu peut aussi déterminer quelles sont les priorités à envisager lors de la solution du problème. Par exemple, si pour les enseignants le problème de l'évaluation est crucial, et si ce sont les mécanismes de compréhension qui intéressent le chercheur, alors qu'il trouve que l'évaluation est une question sans intérêt constituant plus un obstacle qu'une aide, peut-être le chercheur doit-il d'abord envisager de s'attaquer à l'évaluation avant d'aborder la compréhension. Il y a des priorités académiques qui ne sont pas des priorités pour la pratique, et il faut parfois accepter de régler rapidement et efficacement des questions prioritaires pour les praticiens avant d'aborder les problèmes qui intéressent le théoricien.

Mais une fois le problème bien cerné et contextualisé avec ses priorités, une solution rationnelle n'est pas immédiatement abordable.

2 L'explicitation des préconceptions

Avant d'entreprendre une recherche en éducation, tout chercheur devrait se poser quelques questions de manière à expliciter ses préconceptions. Cette explicitation est d'abord à son propre bénéfice avant d'être à l'avantage de la compréhension de ses écrits par les autres. En effet, une majorité des choix méthodologiques est reliée aux conceptions implicites que chacun a de la recherche, de l'éducation, de son rôle, etc. Le chercheur gagnera en temps et en rigueur s'il prend le temps de réfléchir à ses préconceptions, à mieux établir ses choix, afin que les décisions qu'il va forcément devoir prendre, soient plus rationnelles et qu'il ne soit pas surpris lorsqu'on l'interrogera sur les justifications de ses options. Rien de plus pénible pour un chercheur que de ne pas pouvoir répondre lorsqu'on

lui demande d'expliquer pourquoi il a utilisé telle procédure : hésiter, «patiner» sur une demande de justification, constitue une perte de crédit du chercheur pour l'interlocuteur.

Même en recherche exploratoire, le chercheur ne part jamais de rien quand il aborde un problème. S'il veut procéder avec quelques raisons, il doit expliciter et organiser pour lui-même le savoir, les techniques, les conceptions, les attentes et les valeurs qu'il tient sur ce sujet. Cette explicitation lui permettra d'enrichir ce que seules l'intuition et une démarche par essais et erreurs lui auraient permis d'envisager. Mais en outre, le dévoilement des préconceptions permet aussi au lecteur potentiel de savoir à quoi s'en tenir. Il peut partager ou non les préconceptions du chercheur, mais il peut aussi, en les sachant, mieux évaluer la valeur des écrits, car il dispose alors des cadres de référence à partir desquels les énoncés ont été formulés. En disposant de ces cadres de référence, le lecteur peut plus facilement comprendre les procédures de la recherche et le style du rapport.

Dans le cas de la recherche appliquée, l'interrogation doit porter non seulement sur les préconceptions du chercheur, mais aussi sur celles des praticiens avec lesquels il collabore et pour lesquels il entreprend la recherche. En effet, si les conceptions et les attentes des praticiens sont radicalement différentes de celles du chercheur, la probabilité de rejet par les praticiens sera élevée. Car dans beaucoup de cas, si le praticien n'utilise pas un produit de la recherche, ce n'est pas à cause de son inefficacité. C'est plutôt parce que les bases, la philosophie, les valeurs, impliquées par ce produit, ne correspondent pas à la philosophie, aux attentes et aux valeurs du praticien. Il faudra donc en tenir compte dans la délimitation du problème et dans le choix des solutions qui seront proposées. Cela rappelle une des caractéristiques fondamentales de la recherche appliquée : elle ne supporte pas les absolus, elle est souvent le résultat de compromis, d'équilibres entre des forces qui s'opposent.

Enfin, comme la recherche est essentiellement une affaire d'écriture, en plus des préconceptions, le chercheur doit clarifier ses cibles, car bien des choix, entre autres en termes de style, de niveau de vocabulaire, d'argumentation, de moyens de communication dépendront des objectifs du chercheur.

L'explicitation des préconceptions devrait comporter, pour chacun des partenaires, chercheurs comme acteurs, les éléments indiqués au tableau 16.2.

1. La représentation du rôle professionnel
2. L'épistémologie – conception de la recherche scientifique, de son rapport au réel et à la vérité
3. La conception de l'éducation
4. Les orientations méthodologiques
5. Les orientations théoriques privilégiées
6. Les critères de validité de la recherche
 - sur le plan de la crédibilité
 - sur le plan de la rationalité
7. Les visées de la recherche
 - ses enjeux (politique, pragmatique, académique)
 - le type de théories (fonction et niveau)
 - les intentions de recherche et leur démarche
 - le type de pertinence
 - les lieux de résonance et les canaux de communication

Tableau 16.2 : Éléments des préconceptions à expliciter.

Sans que les listes de questions qui suivent soient exhaustives, leur examen devrait aider le chercheur à clarifier ses préconceptions et ses visées. La base de certaines de ces questions a déjà été discutée dans les deux premières parties; ce qui suit n'est alors qu'un rappel. Des listes de catégories illustrent la variété des choix pour chacun des éléments.

2.1 Le rôle professionnel

La représentation que le chercheur se fait de son rôle professionnel peut correspondre à l'une des images suggérées au tableau 16.3. Toutes les images du rôle ont leur place, aucune n'est forcément moins bonne que les autres. Ces rôles ne sont pas forcément figés ni fixés : un chercheur peut, dans certains cas, vouloir remplir un rôle et décider d'en remplir un autre dans d'autres circonstances ou dans une autre étape de sa carrière. Cependant, en fonction des contextes, chacun de ces rôles peut avoir des exigences particulières, tant sur le choix des théories et des outils que sur l'organisation concrète de la démarche. L'exemple le plus clair des contraintes du rôle sur l'organisation d'une recherche est le statut d'étudiant au 2^e ou au 3^e cycle. L'essentiel des choix qu'il effectuera seront contraints par des limites institutionnelles : les thèmes de recherche et les compétences de son tuteur ou directeur, les limites temporelles accordées par ces programmes d'étude, les conceptions du mémoire et de la thèse dans leur rapport à l'apprentissage et au développement du savoir, etc.

- L'orateur ou l'organisateur de congrès scientifiques qui fait de la recherche pour y justifier sa présence, pour obtenir une subvention de voyage, pour être élu président d'une association, devra adopter le style du show.
- L'intellectuel, le producteur de discours, de réflexions cherchera non seulement un effet de vérité, mais aussi un effet de sagesse.
- Le chercheur de solutions aux problèmes de la pratique sera non seulement à la chasse aux problèmes, mais aussi prêt à faire des solutions de compromis.
- L'enseignant qui entreprend une recherche pour améliorer l' (son) enseignement devra être lucide et humble, mais aussi courageux, pour accepter jusqu'au bout de voir ce qui est à améliorer, et pour accepter de le faire.
- Le développeur de savoir, de connaissances, le curieux, celui qui voudrait aller au-delà de ce qui a été dit, devra au préalable se doter d'une très large culture, pour qu'on ne lui dise pas qu'il ne fait que redécouvrir le bouton à quatre trous.
- L'expert au service des praticiens devra acquérir le langage, les catégories des praticiens et accepter de partager le savoir et de trahir la pureté des théories.
- Le conférencier auprès d'associations de praticiens devra vulgariser, à moins qu'il ne veuille devenir leur gourou.
- L'initiateur, qui veut former les plus jeunes en les faisant pratiquer la recherche, devra n'aborder que des problèmes accessibles aux jeunes pour tenir compte de leurs connaissances et de leur expérience.
- Le formateur modèle qui veut former par le modèle qu'il se propose d'être devra être disponible, très explicite, sinon transparent dans ses choix et capable d'expliquer, de faire voir et de faire comprendre ses procédures.
- Le rédacteur de publications scientifiques, l'écrivain prolixe devra se constituer une banque de citations, de références et se constituer des connexions dans les comités de rédaction.
- Etc.

Tableau 16.3 : Quelques conceptions de rôle.

2.2 L'épistémologie

Que l'arrière-fond épistémologique soit plus ou moins explicite, la recherche s'effectue en référence à une conception ou à une préconception de ce que doit être la recherche de la connaissance et du type de connaissance à élaborer. En sachant que les conceptions d'un chacun ne se réduisent pas à de simples catégories, il est éclairant de se demander quelles seraient les catégories qui se rapprocheraient le plus de la conception qui fonde la pratique du chercheur. Si ces conceptions n'impliquent pas nécessairement

le choix de certaines stratégies de recherche, elles imposent parfois des contraintes qui excluent l'utilisation de certaines procédures. Il importe donc d'être conscient des options épistémologiques. Le tableau 16.4 présente quelques illustrations de conceptions rencontrées chez des chercheurs en éducation.

- L'optimisation des pratiques pédagogiques traditionnelles pour en développer les acquis.
- La causalité locale contextuelle ou la recherche d'une compréhension par les raisons, les motifs, les calculs qui ont conduit les sujets; des raisons semblables étant censées jouer dans des contextes semblables.
- Le constructivisme pour lequel la connaissance est conçue comme progressivement élaborée, construite par l'esprit humain en interaction avec son environnement.
- La critique dialectique qui veut mettre en évidence les déterminismes économique-sociaux; la validité de la recherche dépend de son pouvoir de fournir aux individus un meilleur contrôle social, une émancipation.
- L'explication scientifique ou positivisme croit en la possibilité de la recherche de finir un jour par expliquer l'univers ou des portions de celui-ci (comment ça marche) et par édifier un formalisme dont les prédictions ne seraient pas prises en défaut.
- L'historico-herméneutique tente de comprendre la dynamique des événements par l'intentionnalité engagée dans l'histoire des interactions du sujet avec son environnement.
- Le moralisme selon lequel les théories de la connaissance et de l'intervention pédagogique sont limitées par une éthique ou une morale, etc.
- L'objectivisme qui prétend qu'il y a des réalités à connaître, culturellement ou socialement données, qui existent indépendamment de la pensée du sujet; il faut en connaître les contraintes pesant sur l'ajustement écologique et efficace des sujets.
- Le relativisme subjectif selon lequel il n'y a pas d'absolu, tout est subjectif : l'expérience, la connaissance et les démarches pour y parvenir.
- Le pragmatisme éclectique, dit aussi instrumentalisme, stipule que les connaissances et les modalités de recherche ne valent que par leur utilité : la validité de la connaissance dépend de son pouvoir de résoudre les problèmes de la pratique.
- Etc.

Tableau 16.4 : Quelques formes d'épistémologie.

2.3 La conception de l'éducation

La recherche en éducation s'appuie non seulement sur une épistémologie, mais elle est aussi orientée par les conceptions de l'éducation. En effet, la constitution de l'objet de recherche n'est pas indifférente aux valeurs qui sont impliquées par la conception, sinon par l'idéologie pédagogique du chercheur. Les problèmes seront construits différemment et la recherche de solutions sera délimitée par la manière dont le chercheur et ses commanditaires, lorsqu'il y en a, conçoivent le rapport éducatif. Sans être exhaustive, la liste du tableau 16.5 devrait permettre une réflexion à propos des implications de ces conceptions sur les pratiques de recherche

- Pédagogie techno-behavioriste : l'enseignement se fait par l'application technique des connaissances de la psychologie scientifique sur l'apprentissage.
- Pédagogie classique : la formation vise à transmettre un idéal sur le plan des connaissances, des valeurs et des conduites.
- Pédagogie constructiviste : la pédagogie consiste à confronter l'élève à des situations problématiques par la solution desquelles il construit ses connaissances.
- Pédagogie coopérative : l'éducation vise la socialisation et se construit par la participation des étudiants.
- Pédagogie d'avant-garde : la formation cherche à donner les connaissances les plus avancées avec les dernières techniques (à la fine pointe).
- Pédagogie de la découverte : l'enseignement se fait par présentation de contenus permettant à l'élève de faire des investigations et des découvertes personnelles.
- Pédagogie démocratique : l'éducation est centrée sur les interactions dans le groupe, la coopération visant l'émancipation et la démocratisation - entre autres de l'accès aux études.
- Pédagogie humaniste : l'éducation est non-directive, centrée sur les potentialités individuelles, sur la créativité et la spontanéité, elle vise le développement harmonieux et intégré de la totalité de l'individu.
- Pédagogie objectiviste : la formation est rationnelle, centrée sur les moyens d'ajuster la transmission du contenu à apprendre avec les capacités actuelles des élèves et les attentes socioculturelles.
- Pédagogie sensible : l'éducation est basée sur l'expérience sensorielle, la perception : tout commence avec les sens.
- Etc.

Tableau 16.5 : Quelques conceptions de l'éducation.

particulières à chaque chercheur. Notons encore que la prévalence d'une de ces conceptions est surtout affaire de mode et de culture : aucune d'entre elles n'est intrinsèquement meilleure que les autres.

2.4 Les orientations méthodologiques

Étant donné sa formation antérieure et les habiletés qu'il a pu développer, tout chercheur est limité dans ses orientations méthodologiques. Il vaut mieux se limiter à bien faire ce que l'on sait faire et chercher des collaborations spécialisées pour compléter ses lacunes, que de vouloir tout faire et de le mal faire. Reconnaître ses orientations méthodologiques, c'est donc souvent reconnaître ses limites, accepter de ne pas s'engager dans des recherches pour lesquelles on ne dispose pas des outils nécessaires et, en conséquence, se donner la possibilité de choisir judicieusement ses collaborateurs. C'est aussi pouvoir annoncer que les contraintes de nos aptitudes et des collaborations possibles orientent ce qui sera produit et expliquent que certaines approches ne sont pas suivies. Le chercheur n'a pas à avoir honte de ses orientations, de ses préférences et de ses spécialisations méthodologiques : quand il les respecte, elles indiquent sa force et balisent sa crédibilité. Une fois de plus, la liste du tableau 16.6 n'est qu'une incitation à la réflexion.

2.5 Les orientations théoriques privilégiées

Chaque chercheur ayant des connaissances limitées aux acquisitions de son passé, il aborde les problèmes avec un bagage théorique délimité et auquel il aura évidemment recours de manière privilégiée. Les premières idées qui viennent en tête pour désigner et conceptualiser le problème, avant même d'étudier et d'acquérir de nouveaux concepts, proviennent des cadres théoriques que le chercheur privilégie, qu'il maîtrise. L'énoncé des orientations théoriques privilégiées constitue donc un élément important pour comprendre la dynamique d'une recherche. Les orientations théoriques étant aussi nombreuses que les profils de formation des chercheurs, tout inventaire serait trop restrictif pour envisager d'en faire une présentation.

- Expérimentation ou quasi-expérimentation : modèle de recherche où l'on tente de contrôler toutes les variables en jeu, y compris les formes d'intervention du chercheur et de réactions des sujets.
- Édu- ou psychométrie : recherche consacrée à la mise au point et à la validation d'outils de mesure; elle implique une théorie de la mesure du ou des traits visés.
- Étude corrélacionnelle : recherche établissant des relations de concomitance entre catégories, entre codages catégoriels ou numériques, entre mesures de variables.
- Inférence statistique, sondage : étude ayant pour fin d'inférer, avec une marge d'erreur, les caractéristiques d'une population à partir des réponses obtenues auprès d'échantillons.
- Analyse de documents invoqués : analyse (numérique ou non) de matériels produits en dehors du contexte de la recherche. Ex. : archives, documents officiels, lois, statistiques nationales, correspondances, notes de cours, bulletins scolaires.
- Entrevue clinique et examen systématique : entrevues individuelles dans lesquelles les interventions du chercheur sont déterminées soit par les réactions des sujets (entrevue clinique), soit selon un plan pré-établi alors que la forme des réactions des sujets est libre (manipulation systématique).
- Enquête ouverte et par questionnaires : recherche auprès d'échantillons (raisonnés ou non) utilisant des questions ouvertes (enquête) ou fermées (questionnaires).
- Récit de pratique et histoire de vie : recherche utilisant les récits propres aux acteurs comme principale source d'information.
- Étude de cas simple ou croisée : exploration multimodale et multidisciplinaire d'un cas ou de cas croisés (contrastés) mettant en évidence la complexité d'une dynamique.
- Analyse de systèmes souples : analyse systémique ou écosystémique inspirée des travaux de Checkland ou de Bronfenbrenner (système humain intentionnel, complexe...).
- Analyse formelle : examen, à partir de postulats, d'axiomes et de règles logiques, de la cohérence d'une théorie, afin d'énoncer ses implications et ses extensions.
- Analyse critique : analyse de la validité, de la cohérence et des implications d'un discours à partir d'une métathéorie utilisée comme référence.
- Analyse conceptuelle : examen des relations d'un concept avec les autres concepts d'un discours et examen de leur utilisation respective afin de dégager la structure conceptuelle (la théorie latente ou manifeste) du discours.

- Analyse inférentielle : à partir de la mise en évidence de l'implicite d'énoncés interprétatifs ou prescriptifs, inférence de nouveaux énoncés explicitant les exigences ou les conséquences.
- Analyse comparative : comparaison de données quantitatives ou qualitatives obtenues à partir de sites (individus, écoles, régions, pays, systèmes) différents.
- Ethnographie qualitative : immersion et observation, plus ou moins participante, des conduites ou des systèmes (symboliques) d'échanges dans un milieu donné.
- Méthodes descriptives quantitatives : observations systématiques à partir de grilles, mesures à partir de questionnaires, d'échelles ou de tests.
- Etc.

Tableau 16.6 : Exemples d'orientation méthodologique.

2.6 Les critères retenus pour la validité

La valeur que le chercheur attribue à son travail ou celle qu'il cherche à lui faire reconnaître dépend de critères dont le choix est déterminé par plusieurs facteurs. Certains de ceux-ci correspondent à la conception qu'il a de son rôle, d'autres dépendent de préférences, d'habiletés et d'autres facteurs personnels et institutionnels. Comme déjà mentionné au chapitre 5, l'adhésion à ces critères et le souhait de les voir appliqués dans le jugement de la recherche, nécessitent le plus souvent qu'on en tienne compte dès le début de la planification de la recherche: il faut prévoir et mettre en place les moyens de les respecter. Il importe donc que le chercheur identifie les critères qu'il souhaite respecter.

2.6.1 La crédibilité

Pour mémoire, les principaux critères de crédibilité sont :

- La conformation aux théories à la mode;
- La conformation aux attentes du milieu;
- L'apport de solution nouvelle, originale;
- La résolution d'un problème ;
- L'utilité pratique, professionnelle ou sociale;
- La nouveauté de la réflexion, de l'analyse ;
- La nouveauté de la théorie ;
- La consistance du programme de recherche;

- Le consensus académique;
- La publication dans une revue avec arbitrage;
- Le rayonnement (conférence, interview à la radio-TV, séminaire à l'étranger).

2.6.2 La rationalité

Les critères de rationalité, tels qu'envisagés au chapitre 5, peuvent se regrouper en différentes classes.

A. La *validité théorique* dépend de trois examens auxquels on devrait soumettre les énoncés :

- a) Veut-on respecter la logique formelle (analyse du raisonnement) dans la suite des énoncés qui forment les théories, ou veut-on plutôt la sacrifier au profit d'une socio-psychologie de l'argumentation rhétorique ?
- b) Veut-on que les hypothèses résistent au test de la consistance interne ou de la non-contradiction entre les concepts et les déductions tout en respectant la règle de la clôture sémantique; ou bien estime-t-on légitime d'introduire après coup des postulats *ad hoc* pour sauver les apparences ou pour proposer de nouvelles interprétations (exploration) ?
- c) Les théories sur lesquelles on se base ou que l'on propose comme explication apportent-elles plus que les explications antérieures et sont-elles plus explicites; ces théories sont-elles communicables ou ésotériques ?

B. Le *programme* de recherche est-il *cohérent* : y a-t-il un lien entre le but, les intentions de la recherche, la manière de poser le problème (c.-à-d. les orientations théoriques) et les techniques utilisées pour constituer les données, les traiter, interpréter les résultats et valider ces interprétations ?

C. La *problématique* est-elle *pertinente* ? L'épistémologie du chercheur, sa conception de l'éducation et le modèle implicite à la démarche de recherche choisie respectent-ils les contraintes propres à la situation éducative ? Les composantes de la situation éducative, que ce soit au niveau de la situation matérielle, des acteurs, des représentations cognitives et sociales, des valeurs et des jeux de pouvoir sont-elles prises en compte ?

D. Les *modèles* déduits de *théories* pour pouvoir appliquer celles-ci à des *objets*, correspondent-ils aux théories dont ils découlent ? Sont-ils une bonne représentation de l'objet ou en construisent-ils une fiction par

simplification excessive ou par ajout de caractéristiques que l'objet ne possède pas ? Les théories et les modèles décrivent-ils bien l'objet, l'événement en question, ou parlent-ils en fait d'un autre objet ? A-t-on bien défini l'ensemble des éléments constitutifs de la théorie et du modèle, y compris les présupposés ?

E. Le chercheur a-t-il envisagé sa *responsabilité* sur le plan:

- 1° des implications des manipulations de la recherche en regard de l'éthique et de la déontologie professionnelle ?
- 2° des manipulations envisagées dans la recherche en regard des normes comportementales, des coutumes et des usages de la société ?
- 3° de la prudence à l'égard des effets à long terme et des effets secondaires ?
- 4° de l'explicitation du conflit d'intérêts possible entre les visées du chercheur et la pertinence écologique des conclusions et des applications de la recherche ?

2.7 Les visées de la recherche

Le but, les objectifs et le type de pertinence visés par la recherche sont aussi à prendre en considération avant la planification. En effet, comme l'ont montré des travaux récents¹, l'atteinte des visées souhaitées par la recherche implique qu'on y ait pensé avant et qu'on ait planifié certains éléments de la recherche en fonction de ces visées.

Parmi les visées à identifier, on envisagera d'abord, en s'aidant des catégories établies au chapitre 3, *les enjeux* que le chercheur poursuit. Les enjeux (nomothétique, politique, pragmatique et ontogénique) ont leurs exigences, leurs risques et leurs milieux. Ensuite, il s'agira de préciser quelles *fonctions théoriques* (description, interprétation, prescription, etc.) le chercheur souhaite que ses discours remplissent et le *niveau nomologique* (description, compréhension, explication, formalisation) qu'il espère atteindre. Une autre manière de traiter ce problème consiste à poser la question de la portée visée par le discours : prétendra-t-il à quelque généralité, limitera-t-il sa portée à des explications locales ou se tiendra-t-il dans l'intermédiaire, c'est-à-dire une explication contextualisée avec possibilité de transfert à des contextes semblables ?

1 Par exemple : Huberman M., Gather-Thurler M. et E. Nufer, *La mise en pratique des recherches scientifiques* : Étude de la dissémination des résultats du programme national de recherche «Éducation et vie active». Rapport de synthèse, Faculté de psychologie et des sciences de l'éducation, Université de Genève, 1988.

Deux autres éléments, peut-être un peu terre à terre, mais vitaux en termes de crédits aux yeux des organismes qui permettent aux chercheurs de vivre financièrement, doivent aussi être examinés avant d'entreprendre la recherche : soit la forme de pertinence et la résonance visées. La première doit souvent être déclarée : la plupart des formulaires lui réservent une place. La seconde, proche de l'image que le chercheur se fait de son rôle professionnel, vaut peut-être mieux être tue, mais le chercheur doit en être conscient s'il ne veut pas rater sa carrière.

La *forme de pertinence* de la recherche est souvent une question posée sans naïveté par les organismes subventionneurs. Cette question se rapproche de celle des enjeux. Parmi les cas possibles, on peut envisager

- une recherche gratuite, la recherche pour le plaisir de trouver;
- une pertinence pratique, professionnelle (solutions, moyens pour les enseignants);
- une pertinence sociale, pour l'amélioration, le changement social;
- une pertinence théorique (développement, raffinement des connaissances);
- ou tout tenter : à la fois une pertinence théorique et sociale ou socioprofessionnelle.

Les organismes subventionneurs posent la question de la pertinence, non seulement pour identifier l'enjeu de la recherche, mais plus souvent pour amener le chercheur à annoncer comment il s'y prendra pour atteindre cette pertinence. Dans bien des cas, la réponse réside dans le choix des moyens de communiquer les résultats de la recherche, soit le protocole de diffusion.

Le dernier élément est aussi lié à la diffusion : en général, la recherche est entreprise en visant une certaine *résonance*. Par là, on entend non seulement les lieux où le chercheur essaiera de faire connaître ses travaux, mais aussi les bénéfices secondaires ou les retombées espérées à la suite de leur diffusion. Par exemple, quand il s'agit d'une recherche réalisée dans le cadre universitaire, le chercheur peut compter sur certaines retombées : à un extrême, il peut prétendre à une plus ou moins haute résonance académique (promotions, titres, participation à des comités académiques, postes dans des associations ou sociétés savantes, invitations à prononcer des conférences, invitations à publier, etc.) ; à l'autre extrême, il peut tendre à une résonance dans la communauté (consultations par les syndicats, par des organismes politiques ou autres, demande de services par la collectivité, interview dans des magazines, etc.).

Le choix du milieu de résonance a son importance dans la mesure où les différents milieux sont parfois sans passerelle, se méprisent plus ou moins et tiennent à des règles et à des standards incompatibles, entre autres quant aux manières de communiquer les résultats et les conclusions. Il est des cas où le souhait de viser plus d'une résonance exige de prévoir des protocoles de diffusion adaptés à chacune d'entre elles ou de répartir dans le temps et sur plusieurs recherches les résonances à atteindre. À ce titre, il est clair que les exigences académiques que doivent rencontrer les candidats au doctorat pour obtenir leur titre (un premier gain en capital de recherche qui ouvre des portes) ne les laissent pas choisir impunément leur milieu de résonance. Quelles que soient leurs sympathies et leur idéologie, et selon les institutions, une résonance visée dans la communauté locale risque fort de leur attirer des ennuis ou de les conduire à construire une recherche boiteuse tant aux yeux des uns que des autres.

3 Le protocole de diffusion

Le choix d'un protocole de diffusion découle, on vient de le noter, de plusieurs facteurs, dont : le rôle professionnel, les critères de validité et les visées de la recherche. Une fois que ces éléments ont été identifiés, ils commandent une stratégie de diffusion : par quels canaux les résultats seront diffusés. Le choix d'un moyen précis de communication doit se faire dès le départ, dans la mesure où le chercheur doit connaître les règles qui y sont en vigueur, s'appropriier le style et le vocabulaire, et repérer les références qui y sont à l'honneur pour les citer dans la rédaction de sa problématique et dans l'argumentation de ses solutions. Sans cela, le travail entrepris lors de la construction du cadre conceptuel risque d'être inefficace, et le chercheur devra consacrer des énergies supplémentaires à de nouvelles lectures et à une refonte de son cadre conceptuel, au moment où un comité de thèse ou de rédaction exigera des corrections qui auraient pu être anticipées si le protocole de diffusion avait été précisé dès cette phase. Compte tenu des aspects stratégiques de l'écriture de la recherche (chapitre 12), la mise au point du protocole de diffusion fait donc partie des premières phases de la planification d'une recherche.

4 La concordance entre les préconceptions, les visées, le problème et les options méthodologiques

Enfin, le chercheur doit accepter de confronter ses enjeux, le type de théories et le niveau nomologique recherchés, la pertinence et la résonance visées afin d'examiner les incompatibilités possibles entre, d'une part,

chacune de ces visées, et entre ces visées et les préconceptions du chercheur. Car il y a des incompatibilités. Par exemple, il sera difficile de résoudre des problèmes aux enjeux pragmatiques et de construire en même temps des théories formelles; de même, il sera difficile de maintenir un positivisme crédible en faisant de l'ethnographie ou des études de cas, tout en tenant un discours interprétatif et en prétendant construire une description de dépendances fonctionnelles.

Mais l'examen des incompatibilités doit aller plus loin et envisager les relations entre la formulation du problème, une fois que les préconceptions et les visées ont été identifiées, avec les orientations méthodologiques et théoriques acquises ou envisagées. Ainsi, Bouchard et Gélinas² insistent sur la nécessité de commencer toute préparation d'une recherche par l'identification de la position du chercheur sur trois axes bipolaires : ceux de l'épistémologie (allant du constructivisme à l'empirisme), des intentions (connaître-expliquer, agir et changer) et de la méthodologie (de qualitative, présence sur le terrain, à quantitative, expérimentation en laboratoire). La convergence en un point des projections de la position du chercheur sur ces axes, donc leur confrontation, devra être résolue avant de délimiter le problème de la recherche et d'opter pour une méthodologie. Afin d'aider le futur chercheur à visualiser cette exigence, ils ont imaginé un cube dont les trois plans permettraient de situer l'épistémologie, l'intention et le lieu, alors que le problème et la démarche pour le résoudre devraient se placer à l'intérieur du cube, au centre de convergence des projections faites à partir des trois plans.

L'exigence d'une solution aux contradictions dans les explicitations préliminaires, telle qu'illustrée par le modèle du cube de Bouchard et Gélinas, est essentielle et, dans certains cas, suffisante. Dans d'autres cas, l'attribution d'une certaine validité à la recherche exige une confrontation plus analytique, portant sur l'ensemble des catégories envisagées pour les présupposés et les visées par rapport à la formulation du problème et aux choix méthodologiques. La reconnaissance, par un milieu choisi, de la valeur d'une recherche permet au chercheur d'amasser un «capital de recherche» qui portera des intérêts et qui lui permettra de se mettre en position de faire ultérieurement des gains plus importants. La clarification des préconceptions et des contradictions entre les choix permet d'augmenter les possibilités de cette reconnaissance. Autant prévenir que guérir !

2 Bouchard Y. et A. Gélinas, Un modèle alternatif de formation des futur chercheurs, dans «La formation des jeunes chercheurs aux méthodes qualitatives», *Revue de l'Association pour la recherche qualitative*, Volume 3, printemps 1990, p. 121-141.

LE CADRE CONCEPTUEL ET MÉTHODOLOGIQUE

- 1 Le cadre conceptuel ou le cadre théorique**
- 2 Le cadre théorique**
- 3 Le cadre conceptuel et méthodologique**
 - 3.1 *La charpente conceptuelle***
 - 3.2 *L'élaboration des hypothèses de recherche***
 - 3.3 *L'opérationnalisation : choix et construction d'outils***
 - 3.3.1 La sélection des échantillons**
 - 3.3.2 La construction des outils**
 - 3.3.3 Le contrôle de la qualité des données**
 - 3.3.3.1 *Le rapport des indices au réel :
la vraisemblance ou fidélité***
 - 3.3.3.2 *Le rapport des indices aux concepts :
la pertinence ou la validité***
 - 3.3.3.3 *L'intention et l'extension des données :
la contextualité et la transférabilité***
 - 3.4 *La planification***
 - 3.5 *La constitution du matériel de base
et la révision du modèle initial***

Lorsque la situation, dans laquelle un problème se pose, a été décrite à partir du terrain et que les préconceptions du chercheur autant que des acteurs ont été cernées (chapitre 16), le problème peut être clairement énoncé : c'est la formulation de la *problématique* qui constitue souvent, sous ce titre ou sous celui d'introduction, le premier chapitre d'un mémoire ou d'une thèse. Mais, avant de pouvoir reformuler le problème en des termes techniques qui permettront d'envisager des solutions méthodologiques, il faut passer par l'étude de ce qui a été écrit sur le sujet ou sur des sujets semblables, par l'analyse des théories et des méthodologies qui ont été utilisées. Une fois les préliminaires explicités, on aborde le deuxième chapitre d'une rédaction, habituellement intitulé *cadre conceptuel* ou *cadre théorique*.

1 Le cadre conceptuel ou le cadre théorique

La tradition universitaire suggère à l'étudiant qui rédige un mémoire ou une thèse, d'intituler le deuxième chapitre : «cadre théorique» ou «cadre conceptuel». Dans une perspective *nomothétique et vérificative*, on parlera de *cadre théorique* puisqu'il s'agit d'abord de procéder à une analyse spéculative, à la fois conceptuelle, critique et inférentielle, de la «littérature¹» couvrant le domaine de recherche. Cette analyse a pour but de formuler une hypothèse qui sera opérationnalisée (les concepts, inobservables, sont traduits en variables observables ou mesurables), puis mise à l'épreuve. Dans une perspective *nomothétique exploratoire*, on ne parlera plus de cadre théorique ni de déduction d'hypothèse à partir de lois; on construira un *cadre conceptuel* afin de poser des questions. Les réponses à celles-ci devront permettre, par induction, de formuler des hypothèses. Certains

1 L'usage du mot «littérature» est discuté par plusieurs qui lui préfèrent le mot «documentation» lorsqu'il s'agit de décrire l'ensemble des écrits savants sur une question. Le terme documentation s'applique très bien dans le domaine des sciences pures et naturelles; il s'applique moins bien dans le domaine des sciences humaines, en particulier en sciences de l'éducation. En effet, une partie importante des écrits savants de ces dernières disciplines est constituée de travaux spéculatifs (réflexions, analyses, interprétations, recommandations, opinions). De plus, les travaux empiristes sont, la plupart du temps, encadrés de passages spéculatifs : réflexions et interprétations d'une certaine réalité pour justifier la pertinence sociale et théorique de la recherche que l'on décrit ensuite ; analyses, inférences interprétatives, opinions et recommandations dans les conclusions. Comme ces écrits sont plus proches, dans leur contenu et dans leur style, de l'érudition que de la documentation scientifique, nous suivons l'usage courant dans nos disciplines. Il s'agit bien d'examiner la littérature, c'est-à-dire «la bibliographie d'une question; l'ensemble des ouvrages publiés sur cette question», dictionnaire *Robert*.

appellent «cadre conceptuel et méthodologique», ce que d'autres nomment «programme de recherche²». Dans l'un et l'autre cas, il s'agit de formuler le problème de manière suffisamment technique pour que des réponses soient méthodologiquement accessibles.

La *recherche appliquée* comporte aussi une phase d'élaboration conceptuelle, parce qu'elle implique la plupart du temps une enquête (analyse de besoins, analyse de la demande, analyse de situations) qui, à sa manière, débouche aussi sur une induction d'hypothèses : celles portant sur la situation telle qu'elle est, celles se rapportant à la situation souhaitable, et celles qui proposent des hypothèses de solution pour réduire l'écart entre ce qui est et ce qui est souhaitable. Comme cette enquête doit être menée selon les normes scientifiques en usage dans le domaine, il s'agira d'examiner comment les savoirs de la discipline, ou des analogies avec d'autres disciplines, permettent de poser le problème d'une manière technique afin d'en formuler une compréhension plus riche, plus adaptée, ou nouvelle, permettant des hypothèses de solutions satisfaisantes. Aussi, sans que cela soit toujours identifié comme tel, les recherches appliquées comportent une phase de recherche spéculative et l'élaboration d'un cadre conceptuel et méthodologique.

- | |
|--|
| <ol style="list-style-type: none"> 1. Construction de la charpente conceptuelle (ou théorique) 2. Inférence des postulats (hypothèses) méthodologiques 3. Opérationnalisation : choix et construction d'outils 4. Planification des opérations 5. Constitution du matériel de base et ajustements du cadre conceptuel et méthodologique initial |
|--|

Tableau 17.1 : Les phases d'un cadre conceptuel et méthodologique.

En bref, les chercheurs élaborent un cadre théorique lorsqu'ils veulent vérifier des hypothèses déduites, ils construisent un cadre conceptuel et méthodologique lorsqu'ils ont besoin de se donner des outils pour induire des hypothèses. D'une manière générale, l'élaboration du cadre théorique ou conceptuel comporte quatre étapes, comme indiqué au tableau 17.1. Chacune sera examinée plus loin en prenant comme perspective celle de l'élaboration d'un cadre conceptuel plutôt que théorique. Au cours de la

2 Antoine Baby préfère l'expression «programme de travail». Voir Baby A., À travers le chaos épistémologique ou comment la théorie des deux sacs permet de faire un bilan de la recherche qualitative, *Revue de l'Association pour la recherche qualitative*, volume 6, hiver 1992, p. 11. Miles et Huberman utilisent «cadre conceptuel», nous préférons «cadre conceptuel et méthodologique» dans la mesure où les choix méthodologiques sont impliqués par le cadre conceptuel.

description du processus, les variantes particulières au cadre théorique et à la recherche appliquée seront indiquées, principalement aux phases 2 et 3. Mais il faut d'abord insister sur quelques spécificités du cadre théorique qui ne comporte pas la phase 5, parce que le cadre théorique ne peut plus être modifié une fois que l'opération de constitution des données est commencée.

2 Le cadre théorique

La composition du cadre théorique suit différentes modes selon les sous-domaines de recherche, sinon selon les équipes de recherche et les institutions. En bref, il s'agit d'effectuer une revue critique et synthétique des écrits scientifiques du domaine, habituellement appelée revue ou recension de la littérature ou de la documentation. Il faut d'abord repérer les principaux éléments de cette littérature auxquels on applique les stratégies de la recherche spéculative avec plus ou moins de rigueur et de raffinement technique. Comme indiqué au chapitre 8, le but de cette revue est de mettre en évidence un point faible de la théorie pour lequel on espère proposer un ou plusieurs nouveaux énoncés théoriques (thèses). On testera ensuite ce ou ces nouveaux énoncés en le (ou les) confrontant aux thèses anciennes, c'est-à-dire en montrant que les nouveaux énoncés sont plus forts parce qu'ils permettent des applications ou des déductions (des hypothèses) plus pertinentes que les énoncés issus du point faible dépisté lors de la revue de la littérature.

La revue de la littérature doit donc non seulement être conceptuelle, elle doit aussi être critique, tant sur le plan de la validité logique des énoncés théoriques que de la méthodologie des recherches sur lesquelles ces énoncés fondent, provisoirement, leur validité. C'est, entre autres, à partir de l'analyse critique des rapports entre les concepts théoriques et leurs manifestations empiriques (leurs opérationnalisations), que le chercheur pourra, par inférence, formuler les énoncés théoriques complémentaires, qui constituent sa thèse et qu'il va mettre à l'épreuve. Comme cette mise à l'épreuve de la nouvelle thèse doit être solide, on déduira non seulement, sous forme d'hypothèses, à quelles prédictions observables cette thèse devrait aboutir dans une observation sur le terrain ou en laboratoire, mais aussi quelles autres prédictions (hypothèses alternatives) les anciennes thèses produiraient dans une situation identique. En effet, la valeur d'un nouvel énoncé est relative aux prédictions qu'il permet, comparativement aux prédictions permises par les énoncés contestés, et non pas à la probabilité de voir sa prédiction moins rare sous l'effet du hasard. On ne teste pas une théorie

contre une différence nulle produite par le hasard, hasard qui, dans ce cas, n'est que le fantôme, ou l'ombre, de la théorie testée. On teste une théorie nouvelle contre une autre théorie, concurrente ou plus ancienne. Si le cadre théorique ne permet pas d'aboutir à la mise à l'épreuve simultanée de plusieurs (au moins deux) hypothèses précises déduites d'énoncés théoriques concurrents, le chercheur doit conclure que les théories disponibles dans ce domaine sont trop faibles pour mener avec profit une recherche vérificative et qu'il est plus profitable, pour le développement des connaissances, de poursuivre la recherche par des travaux exploratoires. La stratégie trop souvent répandue qui consiste à mettre à l'épreuve une soi-disant hypothèse (qui n'est souvent qu'une intuition) en l'opposant à l'occurrence aléatoire de son contraire n'est qu'une caricature de la démarche scientifique.

Habituellement le chapitre intitulé *cadre théorique* se termine par la formulation des énoncés théoriques nouveaux ou complémentaires (la thèse), des prédictions qui en sont déduites (les hypothèses) et des prédictions déduites des anciens énoncés théoriques (les hypothèses alternatives) pour la situation d'épreuve envisagée. Dans la perspective vérificative, la finale du cadre théorique doit être très précise : les hypothèses sont présentées non seulement sous la forme de relations entre concepts, mais on doit aussi les présenter sous la forme de relations entre variables observables en indiquant la situation exacte dans laquelle ces variables seront observées. Si ces précisions ne sont pas données dès le cadre théorique, le lecteur pourra toujours douter de l'objectivité de la vérification.

Le cadre théorique doit donc présenter, dans sa conclusion, comment les hypothèses déduites seront testées. Mais la justification de l'opérationnalisation et la préparation des outils seront détaillées dans un troisième chapitre, *méthodologique*, chapitre habituellement distinct du cadre théorique. Ce chapitre méthodologique décrira

- 1° comment les concepts théoriques sont opérationnalisés, c'est-à-dire traduits en variables manipulables et observables (dont on peut au moins obtenir une trace); et
- 2° quelle est la variation à observer (l'événement) ou la manipulation à construire (simulant la relation hypothétique entre les variables) pour mettre ainsi les hypothèses à l'épreuve.

La partie méthodologique discutera ensuite le choix des techniques de constitution et de traitement des données, techniques qui doivent être cohérentes et pertinentes par rapport à la manipulation et aux variables retenues comme opérationnalisation des hypothèses.

3 Le cadre conceptuel et méthodologique

Si on modélise cette partie de la recherche selon un schéma semblable pour les recherches inductives et déductives (tableau 17.1), les quatre premières phases du cadre conceptuel et méthodologique d'une recherche exploratoire sont réparties en deux blocs dans la recherche vérificative : les phases 1 et 2 pour le cadre théorique, les phases 3 et 4 pour le chapitre méthodologique. Dans la recherche appliquée, surtout dans la recherche de développement et dans la recherche-action rationnellement planifiée, on peut aussi suivre une structure similaire. Les variantes «vérificatives» et «appliquées» seront indiquées en comparaison avec le modèle «exploratoire».

3.1 La charpente conceptuelle

La charpente conceptuelle (ou cadre conceptuel au sens restreint) est d'abord un instrument de travail rationnel pour le chercheur. La construction de cette charpente devrait permettre une formulation technique du problème telle que des hypothèses, et pas seulement des intuitions, puissent être opérationnalisées. Cette opérationnalisation implique qu'à chaque concept on puisse faire correspondre une variable, dont une trace directe ou indirecte soit accessible, et que la relation présumée entre les variables puisse être, ou bien appréhendée (exploration : comment la percevoir), ou bien décrite (vérification).

Dans la tradition universitaire, la charpente conceptuelle est composée essentiellement par

- 1° une reprise succincte de la description factuelle du problème tel qu'il se présente avec son contexte;
- 2° les connaissances théoriques et méthodologiques dont le chercheur dispose sur le problème au moment où il l'aborde et celles qu'il extraira de son analyse des écrits;
- 3° le réseau conceptuel exprimant le problème construit à partir des éléments qui précèdent en ayant soin d'établir la correspondance des concepts théoriques avec la terminologie des praticiens.

Dans les publications scientifiques où l'espace est restreint, on se limite à l'exposé du troisième élément qui constitue le modèle initial du problème, et dont les phases d'élaboration sont décrites au tableau 17.2.

<ol style="list-style-type: none"> 1. Résumé factuel et contextualisé du problème 2. Bilan des acquis théoriques et méthodologiques pertinents 3. Recherche bibliographique à partir des acquis <ul style="list-style-type: none"> auteurs clefs mots clefs revues spécialisées cascade rétrograde 4. Analyse de la littérature <ul style="list-style-type: none"> critique conceptuelle des écrits critique méthodologique des écrits 5. Construction du réseau conceptuel exprimant le problème 6. Construction du modèle conceptuel initial établissant la correspondance entre la terminologie du terrain (la praxis) et le réseau conceptuel

Tableau 17.2 : Construction de la charpente conceptuelle.

La constitution du premier élément de cette charpente, la description factuelle et contextualisée du problème, a été discutée au chapitre précédent. Le second élément, *les connaissances* dont le chercheur dispose au moment où il aborde le problème, est une donnée résultant de la plus ou moins grande formation du chercheur dans le domaine du problème. Ces connaissances font partie des préconceptions du chercheur par rapport au problème. Elles sont un élément de son idéologie et constituent, à ce titre, une limite. Mais elles peuvent devenir un acquis lorsque le chercheur, à cette étape, est capable de passer d'une utilisation idéologique des acquis à leur *utilisation instrumentale*. Dans une perspective idéologique, le chercheur utilise ses connaissances pour juger, pour évaluer les discours des autres et pour justifier ses présupposés. Dans une perspective instrumentale, le chercheur va s'en servir comme source d'enquête, comme base de questionnement pour élargir, compléter et vérifier les connaissances nécessaires à la conceptualisation du problème. Dans la perspective instrumentale, les connaissances antérieures permettent au chercheur de se donner des pistes pour repérer et analyser la littérature. Aussi, plus la formation du chercheur dans un domaine de recherche est jeune et peu approfondie, plus il devra consacrer d'énergie au troisième et au quatrième éléments de la charpente : la recherche bibliographique et l'analyse des écrits.

Pour effectuer une *recherche bibliographique*, une tactique assez efficace consiste à fouiller dans les revues spécialisées et dans les traités récents afin, d'abord, de trouver quels auteurs ont écrit sur ce problème. Les auteurs majeurs sont ceux qui ont écrit spécifiquement sur le sujet et

ceux qui sont régulièrement cités par d'autres. Disposant du nom de quelques auteurs (et donc du titre de leurs travaux), on peut ensuite effectuer une recherche informatisée. Cette recherche peut se faire à partir de répertoires informatisés comme ERIC, Psychological Abstracts, etc. Il s'agit de se renseigner pour savoir à quels réseaux informatisés les bibliothèques accessibles sont abonnées, car les banques bibliographiques informatisées ne se recoupent pas toujours et il faudra en consulter plusieurs. Cette première recherche informatisée permet de noter sous quels mots clefs ont été classés les premiers auteurs repérés et leurs travaux concernant le problème. Une fois ces mots clefs trouvés, on les utilise pour recommencer une recherche bibliographique en débutant par les travaux les plus récents et en remontant les années. La cascade rétrograde se complétera en examinant quels auteurs, cités dans les travaux trouvés, pourraient être intéressants et en les consultant par la suite.

Quand faut-il arrêter la recherche bibliographique ? La règle de la *saturation* de l'information, que l'on discutera à propos des échantillons d'informateurs, peut s'appliquer. Mais à cette phase, on prend comme critère que l'information bibliographique recueillie doit être suffisante afin de construire un réseau conceptuel assez complet pour instrumenter la recherche. On a déjà noté qu'un examen absolument exhaustif des écrits est pratiquement impossible à notre époque. Le critère de saturation s'entend donc comme une décision stratégique : c'est au chercheur et à ses coéquipiers d'évaluer la stabilité du réseau conceptuel et la suffisance de sa richesse comme outil de recherche.

Le *réseau conceptuel* exprimant le problème est construit à l'issue de l'analyse de la littérature. Dans la pratique, cette construction s'effectue souvent en plusieurs épisodes qui se chevauchent et se reprennent partiellement. Un premier réseau s'ébauche lors de l'enquête auprès des praticiens, lorsque le chercheur fait une synthèse de leurs représentations du problème. La deuxième forme du réseau se construit lorsque le chercheur ramasse ses connaissances pour interpréter la problématique des praticiens à un niveau plus conceptuel permettant d'enclencher la recherche bibliographique. Enfin, la troisième forme du réseau découle du remaniement progressif de la deuxième forme au cours de l'analyse des écrits. Lorsque, comme on vient de le décrire, la construction du réseau conceptuel s'effectue en cycles qui corrigent les formes précédentes, les cycles aboutissent à un modèle conceptuel initial du problème, qui assure probablement la correspondance entre la formulation du problème issue de la praxis et celle provenant des écrits scientifiques. Mais quelle que soit la stratégie suivie, en cycles ou en étapes tranchées selon les phases du tableau 17.2, cette

correspondance doit toujours être vérifiée, car elle conditionne les possibilités de validation et d'utilisation des résultats de la recherche par les acteurs.

Dans la recherche exploratoire, le *modèle conceptuel initial* comportera, parfois organisé à la manière d'un organigramme, une représentation des différents facteurs et des relations entre ces facteurs qui interviennent dans la formulation du problème. Il tient compte des facettes que ce problème peut prendre, des variables qui peuvent l'influencer et des actions et des rôles que les acteurs impliqués y tiennent et pourraient y tenir. Cette représentation modélisée du problème doit être suffisamment claire pour suggérer les principales questions ainsi que les procédés et les instruments pour obtenir des réponses pertinentes. Cette représentation doit aussi synthétiser, au minimum, tout le savoir disponible à propos du problème en question; sinon ce modèle ne peut pas être un outil utile d'appréhension du problème et le chercheur risque de ne redécouvrir que des connaissances banales aux yeux des autres.

En outre, il est souhaitable que ce modèle conceptuel initial soit *multidisciplinaire*, ou plus précisément «*multiréférencé*» (c'est-à-dire comporter des références théoriques multiples), dans une perspective inductive; il doit l'être aussi dans une perspective d'application. Il ne peut se limiter à une seule théorie (que la recherche ne vérifiera pas). Le chercheur doit envisager ce que diverses orientations théoriques apportent à la conceptualisation du problème, s'il veut pouvoir observer et interpréter l'ensemble des faits qui peuvent se présenter dans la situation. Une seule perspective théorique le laisserait borgne et ne lui permettrait pas de chercher tout ce qu'il y a à voir. Cette nécessité de références multiples est une des difficultés que les chercheurs doivent surmonter, car elle exige une très large culture et un esprit ouvert capable d'éclectisme théorique que certains ne manifestent pas du fait de leur ignorance lorsqu'ils sont encore en formation, ou du fait de leur hyperspécialisation quand ils sont devenus experts. Dans une perspective de découverte, il faut entreprendre la démarche avec un champ de vision élargi. L'objet est complexe et sensible à beaucoup de facteurs : il faut pouvoir envisager ces différents facteurs. Cela implique d'être ouvert à des ensembles théoriques plutôt qu'à une théorie. Le modèle conceptuel initial constitue donc une modélisation composite intégrant de manière dynamique des modèles partiels suggérés par diverses orientations théoriques. Plus encore que dans la recherche exploratoire, les modèles utilisés dans la recherche d'application sont des modèles recourant aux compromis, aux équilibres instables entre des facteurs qui tendent à se neutraliser.

Dans la recherche appliquée, le réseau conceptuel devra permettre de créer deux modèles conceptuels : le modèle de la situation de départ et le modèle de la situation d'arrivée. Dans la plupart des cas, on ne peut se contenter d'un modèle général et de deux simulations, faites à partir de ce modèle, pour les situations de départ et d'arrivée. En effet, dans la majorité des situations éducatives, l'écart entre la situation de départ (état observé : passé ou actuel) et la situation d'arrivée (état visé) implique une modification plus qualitative que quantitative. Il ne s'agit pas seulement de variations continues et homogènes des caractéristiques qui affectent les éléments d'une même liste, mais plutôt d'une variation dans la liste des éléments ou des attributs. En conséquence, on devra souvent déduire du cadre conceptuel deux modèles différents, permettant des simulations concrètes, pour la situation de départ et pour la situation d'arrivée.

3.2 L'élaboration des hypothèses de recherche

La phase suivante consiste à inférer des hypothèses de recherche à partir du modèle du problème construit à la phase précédente. Cette inférence se présente sous trois variantes, selon qu'il s'agit de recherche exploratoire, vérificative ou appliquée. L'inférence est un processus à propos duquel peu de choses sont connues. Il consiste à établir des liens entre deux ou plusieurs domaines, mais les procédés à suivre et les règles à respecter pour que l'inférence soit efficace semblent dépendre du problème à résoudre. Mais, plus la culture du chercheur est large, mieux il maîtrise les domaines sur le plan des concepts et sur le plan des systèmes d'inscription qui y correspondent, plus les inférences méthodologiques possibles seront nombreuses et efficaces.

Dans la recherche exploratoire, le modèle initial du problème de recherche suggère une liste de questions à propos des éléments qui composent le modèle : quelle est la liste des éléments, quels sont leurs attributs, quelles sont leurs relations et leurs interactions, comment évoluent-ils, quelles sont leurs sensibilités à l'environnement et, en retour, comment ils modifient ce dernier, etc. De cette liste de questions on pourra inférer quelques hypothèses méthodologiques : si le modèle est tel, alors en recourant à telles stratégies d'exploration on devrait obtenir des réponses utilisables.

Dans la recherche appliquée, c'est à partir du modèle conceptuel de la situation de départ et de celui de la situation d'arrivée qu'on modélisera l'intervention ou le développement, c'est-à-dire que le chercheur induira des hypothèses stratégiques et qu'il ébauchera des simulations testant la

pondération à donner aux différents éléments qui joueront dans l'intervention. Comme le joueur d'un jeu de société (les échecs, par exemple) modélisera les coups possibles à partir de sa connaissance du jeu et de ses règles, le chercheur simulera les interventions possibles et leurs effets à partir d'une représentation (modélisation) des facteurs qui interviennent dans une intervention, cette modélisation étant inférée du réseau conceptuel bâti dans la première phase (théorique) de la recherche appliquée.

<p>A. Recherche exploratoire</p> <p>A.1 Inférence d'un modèle conceptuel du problème</p> <p>A.2 Inférence des questions de recherche correspondant aux éléments du modèle</p> <p>A.3 Inférence d'hypothèses ou <i>postulats méthodologiques</i> : comment obtenir l'information en réponse aux questions de recherche</p> <p>B. Recherche vérificative</p> <p>B.1 Analyse inférentielle des écrits produisant un cadre théorique comportant une thèse rivale nouvelle face aux thèses faibles ou contestées</p> <p>B.2 Dédution d'une <i>hypothèse de travail</i> face aux <i>hypothèses alternatives</i></p> <p>B.3 Opérationnalisation des hypothèses</p> <p>B.3.1 Concepts → variables → indices ou mesures</p> <p>B.3.2 Relations → observations d'événements → manipulations des contingences</p> <p>C. Recherche appliquée</p> <p>C.1 Modélisation de la situation de départ (état observé) et de la situation d'arrivée (état visé)</p> <p>C.2 Inférence d'<i>hypothèses stratégiques</i> : simulation des stratégies permettant de passer de l'état actuel à l'état visé</p>

Tableau 17.3 : L'élaboration des hypothèses de recherche.

Dans la recherche vérificative, comme nous l'avons signalé au début de ce chapitre, l'élaboration des hypothèses rivales aux thèses faibles ou contestées se fait par inférence à partir du cadre théorique. L'analyse critique ayant permis de dépister les thèses faibles, c'est une phase d'analyse inférentielle qui permettra de proposer de nouvelles thèses à tester. Pour tester la thèse inférée, il faut l'appliquer à un domaine du monde observable auquel les thèses contestées, ses rivales, s'appliquent aussi. Il s'agit, à cette étape, de déduire des hypothèses à partir des thèses. Les hypothèses sont les prédictions observables qu'il faudrait pouvoir enregistrer si les thèses

sont tenables. L'hypothèse de travail correspondant à la thèse proposée, et les hypothèses alternatives correspondant aux thèses contestées, sont des applications déduites à partir du cadre théorique. Une fois les hypothèses posées, encore faut-il les opérationnaliser, ce qui comporte deux aspects. D'abord on convertit les concepts en variables dont on peut observer des indices ou des mesures. Ensuite, on construit les situations dans lesquelles ces indices pourront être observés et on planifie les manipulations qui sont censées provoquer la relation hypothétique entre la variable dépendante (observée) et la variable indépendante (manipulée). Contrairement à ce qui se passe dans la recherche exploratoire et dans la recherche appliquée, le cadre conceptuel et méthodologique est stabilisé dès que l'opérationnalisation de l'hypothèse est terminée. Puisque l'on veut vérifier une hypothèse par une mise à l'épreuve déduite du cadre théorique, l'obtention d'un résultat non prévu dès la première collecte de données doit être interprété comme concluant. Le chercheur ne peut pas modifier ses mises à l'épreuve jusqu'à ce qu'il obtienne ce qu'il veut, sans avoir à reformuler les hypothèses et sans reprendre son analyse critique et son analyse inférentielle du cadre théorique. S'il ne peut montrer du premier coup la faiblesse des hypothèses qu'il conteste, il doit reprendre l'ensemble du travail d'analyse.

3.3 L'opérationnalisation : choix et construction d'outils

La phase précédente, soit l'élaboration des hypothèses de recherche, se termine par l'énoncé de leur forme opérationnalisée : quels indices ou quelles mesures faudra-t-il enregistrer pour entreprendre les analyses et le traitement qui conduiront à de nouvelles hypothèses (induites ou de solution) ou qui permettront de décider en faveur d'une alternative (hypothèse déduite). Une fois le choix des indices ou des mesures effectué (à partir de la charpente et du modèle conceptuels), il faut justifier cette opérationnalisation et construire les outils. Trois éléments doivent être considérés dans cette phase (tableau 17.4).

3.3.1. La sélection des échantillons

Selon les éléments du modèle initial et selon les hypothèses de recherche, la question de la sélection des sources d'information ne se réduira pas au choix des sujets (ou des informateurs). Il faudra envisager la possibilité de choisir le terrain (ou les sites) d'enquête parmi l'ensemble des terrains possibles; il en va de même pour les événements à observer et pour les indices qu'on y repérera et les mesures que l'on prendra.

<ol style="list-style-type: none">1. Sélection des échantillons<ol style="list-style-type: none">1.1. de terrains1.2. d'événements1.3. de sujets1.4. d'indices à observer ou de mesures2. Construction des outils permettant<ul style="list-style-type: none">- la constitution du matériel de recherche (exploration)- la réalisation de l'épreuve (vérification) (+ choix des critères de décision et des marges d'erreur)- l'identification des caractéristiques des situations de départ et d'arrivée (application)3. Choix des critères du contrôle de la qualité des données et des techniques de validation des données
--

Tableau 17.4 : Éléments constitutifs de l'opérationnalisation.

Le problème d'une sélection d'échantillons se pose donc autant pour le terrain où l'on se rendra, pour les événements que l'on y observera, pour les informateurs que l'on interrogera et pour les indices et les mesures que l'on notera. Choisir parmi des possibles ne peut se faire de manière rationnelle que si le modèle conceptuel fournit des indications assez précises pour inférer les caractéristiques que doivent posséder les échantillons de terrains, d'événements, d'informateurs d'indices et de mesures, pour permettre la constitution de données fidèles et valides.

La solution souhaitable au problème de l'échantillonnage tient compte de plusieurs exigences. La première exigence stipule que l'information des données réponde aux questions de la recherche. La deuxième implique que la constitution des données tienne compte du caractère complexe du problème. La troisième réclame que l'on puisse estimer le degré d'approximation du réel que les données atteignent.

Le réponse à la première exigence consiste à effectuer des *échantillonnages raisonnés ou théoriques*, c'est-à-dire inférés à partir du modèle conceptuel, et non pas des échantillonnages au hasard ou des échantillonnages cafétéria (c'est-à-dire, selon les occasions qui se présentent). Les échantillons seront choisis parce que leurs caractéristiques qu'ils possèdent permettent d'espérer obtenir efficacement une information pertinente. Dans le cas des échantillons d'informateurs, les techniques privilégiées, pour cette raison, seront l'échantillonnage théorique en boule de neige et l'échantillonnage en cascade contrastée (voir chapitre 14). Par

ailleurs, ces informateurs doivent avoir vécu certains événements qui ne se produisent que dans certains lieux, si l'on veut obtenir un témoignage factuel plutôt qu'une opinion construite par ouï-dire.

La réponse à la deuxième exigence, dont il fut déjà tenu compte dans la construction *multiréférencée* du modèle conceptuel initial, implique d'envisager une *stratégie multimodale de constitution des données*. On tentera d'utiliser différentes stratégies et différents instruments pour deux raisons principales. D'abord, la constitution multimodale permet d'obtenir des données qui portent sur des aspects suffisamment variés du problème pour le saisir dans sa complexité. Ensuite, le recours à des stratégies variées prend en considération les variations de la sensibilité aux événements et de la capacité d'expression que l'on rencontre chez les informateurs. Il n'y en a pas deux identiques; chacun aura vécu de manière différente un même événement, chacun en témoignera selon ses modes privilégiés d'expression. Pour pouvoir atteindre ce qu'il y a de commun dans leur expérience, il faut aller chercher l'information selon des modalités qui s'adaptent à leurs particularités. Une cueillette d'information unimodale ne ferait que renforcer l'expression des différences individuelles. Par exemple, l'un est figé par une question ouverte et n'y répondra pas, alors que l'autre est incapable de décider de la réponse dans un questionnaire à choix multiples; pourtant tous les deux ont la même information, mais ils ne fourniront pas les mêmes données si l'enquête n'utilise qu'une seule technique.

La réponse à la troisième exigence impose de prévoir effectuer des *triangulations* et de respecter le principe de saturation. Au sens strict, la triangulation désigne le recoupement de données obtenues à propos d'une même question (thème ou événement) par des sources identiques sur plusieurs dimensions sauf sur une qui varie. Par exemple, on compare les données fournies par le même informateur à propos du même thème mais obtenues à partir de techniques différentes : un extrait de journal de bord (donnée invoquée), un extrait d'entrevue (donnée suscitée), les réponses à un questionnaire fermé ou à un test (donnée provoquée). Au sens technique du mot, «triangulation» viendrait de la trigonométrie. D'autres, plus poétiques, disent qu'il vient de la navigation (maritime ou aérienne) où le pilote est obligé de prendre plusieurs points (c.-à-d. mesure d'orientation définie par l'angle entre le Nord et la visée d'un objet à l'horizon) à partir de différents repères pour croiser les estimations de sa position et ainsi se situer avec précision. De même, le policier enquêteur ne se contente pas d'un seul témoignage pour affirmer qu'un fait s'est produit, car chaque témoignage n'est qu'une perspective sur la réalité. Chaque mesure prise

isolément n'est jamais qu'une approximation et, dans les sciences de la nature, on utilisera une moyenne ou une statistique, c'est-à-dire le résultat combiné de plusieurs mesures, comme étant la meilleure approximation de la mesure d'un objet. Comme le policier dans une enquête qui confronte les dires des témoins, il faut, dans les sciences humaines, trianguler les données, c'est-à-dire recouper des données de diverses sources et obtenues par diverses techniques, afin d'évaluer dans quelle mesure l'information recueillie permet de se faire une idée valable. Le même problème existe avec les données quantitatives. Le chercheur qui construit son outil en respectant la théorie de la mesure aura fait du multimodal sans le savoir ni le dire. Il aura dû trianguler des sources de données pour pouvoir dire qu'il a abouti à une mesure fidèle, car une mesure isolée n'est jamais qu'une approximation. Cela conduit donc à choisir plusieurs techniques de constitution des données et, comme cela a déjà été suggéré au chapitre 3, à chercher une combinaison de données invoquées, suscitées et provoquées.

Par ailleurs, comme il s'agit d'échantillonnage, l'estimation du degré d'approximation du réel renvoie à la question de la taille des échantillons. Combien de terrains, d'événements, de sujets et de mesures faudra-t-il prendre ? On sait que l'idéal serait de prendre l'ensemble des possibles, puisque, dans ce cas, l'échantillon correspond à la population et qu'en conséquence, il n'y a pas d'erreur d'estimation ou d'inférence possible. Mais on sait aussi que cet idéal est utopique. La réponse pratique consiste à limiter la taille d'un échantillon lorsqu'on peut respecter le principe de la *saturation de l'information*. Selon ce principe, il y a saturation lorsqu'aucune information nouvelle n'est apportée par l'ajout à l'échantillon d'un élément nouveau et différent des autres éléments, selon au moins une des caractéristiques importantes retenues pour sélectionner l'échantillon. Autrement dit, le chercheur peut dire qu'il y a saturation lorsqu'il n'apprend plus rien de neuf, même lorsqu'il ajoute à son échantillon un matériel provenant d'une source (terrain, événement, informateur, mesure) qui varie des autres sur au moins une des dimensions importantes dans la constitution d'un échantillonnage raisonné en fonction du problème de recherche et du cadre conceptuel.

3.3.2 La construction des outils

Après avoir sélectionné les techniques de constitution des données, il faut préparer *leur support et le format d'enregistrement* du matériel en fonction de la charpente, des données et des échantillons. Deux situations se rencontrent à cette étape. Ou bien les outils, autrement dit les systèmes

d'inscription, dont on a besoin ont déjà été développés par d'autres chercheurs et sont accessibles. Dans ce cas, on les emprunte ou on les achète selon leur disponibilité sur le marché. Lorsqu'ils n'ont pas encore été développés, ou lorsqu'ils reposent essentiellement sur les habiletés techniques du chercheur lui-même, il doit se les forger. Dans ce cas, les chapitres 13, 14 et 15 auront non seulement présenté un éventail des choix possibles, mais ils auront aussi suggéré les moyens de se les approprier.

Une fois les outils choisis et construits, et avant de planifier les opérations, il faut encore choisir quelles stratégies de validation des données et de contrôle de leur valeur on utilisera. Il existe plusieurs techniques de validation à choisir en fonction du contexte, de leur accessibilité et en fonction du problème.

3.3.3 Le contrôle de la qualité des données

Avant de planifier la réalisation des opérations, il faut anticiper la phase de validation des données. En effet, avant d'entreprendre l'analyse et le traitement des données, on examinera quel statut on peut leur accorder, car la validité des interprétations dépend de la qualité des données. Trois questions, déjà envisagées au chapitre 5, seront abordées : le degré de vraisemblance ou d'artificialité des données, leur pertinence ou la force de leur référence aux concepts et aux théories, et leur représentativité ou leur contextualité, c'est-à-dire la possibilité d'étendre à certains événements, à certains lieux et à certains individus les inférences que l'on fera sur les données recueillies. Mais les réponses qu'on pourra donner à ces questions dépendent de la manière dont les données auront été constituées : en effet, certaines stratégies de collecte du matériel permettent de lui accorder une plus ou moins grande valeur. Il faut donc prévoir, dès cette phase, quelles techniques seront utilisées pour assurer le contrôle de la qualité des données. Idéalement, le chercheur en choisira plusieurs et en tiendra compte dans la planification.

3.3.3.1 *Le rapport des indices au réel : la vraisemblance³ ou fidélité*

Sans analyser les différents indices statistiques permettant une estimation du degré de fidélité, celle-ci se définit, d'une manière générale, par le fait que la trace obtenue reflète la caractéristique étudiée de l'objet avec une marge d'erreur acceptable, tout en sachant qu'une représentation

3 Les termes «fiabilité» et «plausibilité» se rencontrent aussi comme équivalents des autres expressions indiquant le rapport des indices au réel. L'usage du terme fiabilité est particulièrement ambigu dans la mesure où, chez certains auteurs,.../...

parfaite est impossible à obtenir. En effet, le dispositif qui permet l'inscription de la trace réduit toujours l'objet à quelques dimensions, souvent isolées et traduites dans un autre système de signes que celui par lequel elle pourrait être appréhendée dans son contexte naturel. En tant que représentation, une donnée ne peut jamais être qu'une approximation. Il faut d'ailleurs se méfier d'une donnée dont les indices indiqueraient une fidélité parfaite, car celle-ci ne s'obtient que si le système inscripteur est biaisé en faveur d'un résultat artificiel affichant cette perfection, au détriment d'une représentation approchée de la caractéristique de l'objet. Comme on a pu le voir dans certaines pratiques, l'observateur obsédé par la fidélité centre plus son attention sur l'écart potentiel avec la bonne mesure que sur la manifestation de la caractéristique qu'il est censé mesurer. Cela conduit l'observateur à voir ce qu'il pense que l'autre observateur voit, et non pas ce qu'il y aurait à voir.

L'ampleur de la marge d'erreur acceptable n'est pas un absolu. L'infailibilité est impossible, et l'obsession des résultats purs conduit à dépenser son temps et ses énergies au raffinement infini des instruments et à ne jamais rien pouvoir dire de ce que ces instruments devraient un jour mesurer. Le réalisme stratégique consistera donc à utiliser des instruments raisonnablement fidèles et à accepter de limiter la portée des énoncés, c'est-à-dire à rester réservé quant à leur fiabilité, car ils ne pourront jamais être que relatifs aux conditions particulières des recherches à partir desquelles ils ont été formulés. En effet, une mesure fidèle ne l'est que dans son contexte; elle peut être tout à fait biaisée dans un autre contexte de recherche ou lors de son utilisation comme outil de pratique. Par exemple, un test d'intelligence peut être très fidèle et pertinent dans une recherche donnée, mais n'avoir aucune valeur en milieu scolaire comme mesure prédictive ou de classement des performances.

.../... il correspond aussi au problème de l'objectivité et se pose par rapport à l'ensemble de la recherche, allant de la constitution jusqu'à l'interprétation et la vérification; dans ces cas, c'est pratiquement à propos du chercheur lui-même que se pose la question de la fiabilité. Certains spécialistes des méthodes quantitatives rejettent l'utilisation du mot «vraisemblance» comme équivalent de la fidélité, parce que ce terme fait partie d'expressions désignant des techniques ou des notions statistiques, comme le «maximum de vraisemblance». Ce terme a cependant une utilisation courante en recherche qualitative et renvoie très bien à la question posée de savoir si les observations sont «vraies» ou «fausses», si elles sont exactes ou entachées d'erreurs. S'il fallait supprimer du vocabulaire des diverses disciplines tous les termes dont se sont emparés les mathématiciens, on ne se parlerait plus. Nous préférons garder le terme de vraisemblance car il renvoie très bien, étymologiquement, à la question que les méthodologues se posent sous ce terme.

L'aspect métrique (ou statistique) de l'estimation de la fidélité des données relève des théories de la mesure. Miles et Huberman⁴ ainsi que Goetz et Lecompte⁵ ont proposé des stratégies permettant d'assurer une vraisemblance relative aux données, qu'une estimation métrique de leur fidélité puisse ou non être calculée. En bref, les stratégies reprises dans la synthèse qui suit, recourent à deux procédés : la répétition du contenu de l'observation à partir de plusieurs outils et de plusieurs points de vue, et la description détaillée des systèmes inscripteurs, ce qui permettrait d'en dépister les faiblesses. Une première série de techniques, présentée au tableau 17.5, concerne la phase de collecte du matériel.

Une des particularités des méthodes qualitatives tient à ce que le matériel recueilli est souvent d'une abondance et d'une fluidité telles qu'il ne peut pas servir tant qu'on ne l'a pas filtré et condensé. Cela implique que des données utilisables n'existent qu'à partir du moment où une première analyse du matériel a été effectuée : sélection de passages significatifs, transcription, codage, etc. Ces opérations font aussi partie du système inscripteur : l'instrument n'y est autre que le chercheur lui-même. Aussi, certaines techniques, présentées au tableau 17.6, ont été proposées de façon à estimer la fidélité du travail effectué par le chercheur lors de cette phase. Elles sont, adaptées à cette phase particulière, des variantes de la triangulation.

Les méthodes qualitatives suivent une démarche qui parfois se présente comme un cycle dans lequel une phase ultérieure conduit à reprendre une phase initiale. De plus, elles impliquent souvent un aller-retour entre le bureau des chercheurs et les informateurs sur le terrain. Aussi certaines techniques sont récurrentes dans les phases où le système inscripteur joue le rôle essentiel. Les techniques présentées au tableau 17.7 permettent, en outre, un contrôle plus général de la qualité des données, car elles touchent à la fois leur vraisemblance (fidélité) et leur pertinence (validité).

4 Miles, M.B. et A.M. Huberman, *Qualitative Data Analysis, A Sourcebook for New Methods*. Beverly Hills, CA : Sage 1984.

5 Goetz J.P. et M.D. Lecompte. *Ethnography and Qualitative Desing in Educational Research*. Orlando, FL : Academic Press, 1984.

- 1 **Présence prolongée sur le terrain**
Une présence prolongée permet de diminuer l'information artificielle par une meilleure compréhension de son contexte de production.
- 2 **Enregistrement mécanique des données**
Mais les instruments électroniques sélectionnent aussi le matériel et le déforment. De plus, il y a un risque de panne : il faut toujours le doubler en manuscrit.
- 3 **Triangulation des sources et des méthodes**
Exemples : récit convergent de deux répondants, convergence des récits d'un observateur et d'un acteur de la même situation, d'un observateur impliqué et d'un observateur périphérique. Les expressions «chevauchement des données» et «concomitance des méthodes» sont des équivalents à la triangulation. Les techniques qui suivent sont des variantes plus ou moins sophistiquées de la triangulation.
- 4 **Questions multiréférencées**
Technique à envisager lorsque le chercheur n'utilise qu'une seule méthode de constitution, comme le questionnaire ou l'entrevue structurée. Il s'agit d'utiliser des questions issues de perspectives théoriques différentes et complémentaires.
- 5 **Confrontation des réponses données aux situations possibles**
Technique importante lorsque les informateurs peuvent répondre en exagérant ou en inventant (l'histoire du pêcheur qui a pris une sardine de 45 centimètres...). Cette technique implique que le chercheur connaisse bien les caractéristiques essentielles de l'objet et du terrain de la recherche.
- 6 **Contrôle par observations fortuites de la scène sociale**
Observations non annoncées ou indirectes de la scène sociale qui permettent de contrôler les énoncés des informateurs.
- 7 **Confrontation de chercheurs ayant la même formation**
Pour le faire, prévoir, dès la première phase, de donner une formation identique (en groupe) sur l'objet de recherche et sur les méthodes aux enquêteurs, aux analystes, etc.
- 8 **Analyse comparative continue**
Lors d'une présence prolongée sur le terrain, comparer les nouvelles observations et leur codage avec les anciennes observations et leur codage; cela permet de raffiner et d'ajuster les concepts aux réalités des acteurs.
- 9 **Croiser les rôles d'observateur et de participant**
Croiser les réactions et observations de l'enseignant-chercheur et celles de l'observateur-chercheur pour estimer l'effet de leur statut et de leur centration sur des actions différentes.
- 10 **Croiser les sites usés et neufs**
Croiser des sites neufs, moins accessibles, mais essentiels, avec certains sites toujours choisis qui deviennent sursaturés (ex. : l'école-labo où tout le monde va) et donnent des réactions altérées.

Tableau 17.5 : Techniques de contrôle de la vraisemblance : collecte du matériel.

<p>1 Définition opérationnelle des codes (index, répertoire) : descripteurs à basse inférence. Prévoir non seulement un index et un répertoire avec les cas limites, mais aussi la syntaxe de combinaison des codes multiples (important pour le tri informatisé).</p> <p>2 Double codage intra-codeur Après avoir codé l'ensemble du matériel, le chercheur refait le codage de quelques protocoles déjà codés. Le second codage s'effectue sur des protocoles tirés au hasard et ne comportant aucune trace du premier codage.</p> <p>3 Reproduction indépendante (double codage ou contre codage) Un second chercheur refait les codages et les analyses du chercheur principal. Comparaison des codages et des analyses de celui qui a été sur le terrain avec celles de celui qui n'y a pas été. Lorsqu'il y a étude sur plusieurs sites, chaque chercheur peut doubler son collègue et ainsi, tout en ayant l'expérience d'un terrain semblable, contrôler les analyses sans avoir été en contact direct avec les informateurs.</p> <p>4 Contrôle en double insu Variante de la reproduction indépendante : deux chercheurs, n'ayant pas participé à la collecte (aveugles par rapport au terrain) et indépendants l'un de l'autre, analysent le matériel (et les données). N'ayant travaillé que sur les documents, ils confrontent ensuite leurs analyses entre elles.</p> <p>5 Pondérer l'évidence Donner plus de poids aux sources fiables, aux données de meilleure qualité, ce qui implique une critique des sources dans laquelle on confronte :</p> <ul style="list-style-type: none"> - données établies >< données préliminaires - vu ou reporté de première main >< obtenu de seconde main - comportements, activités, observations >< rapports, jugements, sentiments impressions - informateur expérimenté >< informateur novice - collecte officielle en situation formelle >< collecte informelle - données fournies volontairement >< réponses obtenues par des questions ou sous pression - répondant seul avec l'informateur >< répondant en présence d'autres ou en situation de groupe.

Tableau 17.6 : Techniques de contrôle de la vraisemblance : constitution des données (sélection du matériel et codage).

- 1 **Explicitation sémantique intersubjective**
Chacun des participants (les informateurs et les chercheurs) explicite sa dénotation et sa connotation des mots, afin d'établir une communication minimale. L'absence de cette technique explique bien des malentendus entre chercheurs et praticiens.
- 2 **Retranscription des interviews dans les catégories de l'informateur**
Respect du matériel initial pour des données fidèles, valides et contextualisées, qui permet de construire des catégories moins abstraites et de transférer les conclusions vers les praticiens.
- 3 **Occurrence intersite de catégories dérivées des participants**
La concordance des récits à propos d'un même phénomène observé à partir de plusieurs sites est un indice de la fidélité des observations, alors que l'utilisation de descripteurs semblables donne une indication de leur validité.
- 4 **Description détaillée des choix et des présupposés**
 - le statut du chercheur sur le terrain – le choix des informateurs
 - les situations et conditions sociales – les prémisses et concepts analytiques
 - les méthodes personnelles de collecte et d'analyse
- 5 **Évaluation des effets dus au chercheur**
Les influences sur le site (réactions des informateurs à son arrivée), les biais dans la collecte (les instruments choisis limitent le matériel recueilli) et le traitement des données (le type de résultats induit par les techniques de traitement utilisées).
- 6 **Réunion d'analyse intersite**
Les techniques d'investigation et le matériel recueilli sont analysés et comparés pour dépister les différences, les variations d'attitudes et de techniques dues à la pression du terrain sur les chercheurs. (Exemple : tel chercheur ne fait plus les observations de contrôle que d'autres chercheurs continuent à effectuer).
- 7 **Description des procédures et vérification interne des comptes**
Description quotidienne des procédures (chronique, journal de recherche), suivie de leur évaluation par le chercheur ou son équipe. Elle implique la rédaction de lexiques, de dictionnaires, de syntaxes et le recueil des calculs effectués.
- 8 **Vérification externe des comptes**
À la fin de la recherche ou d'une phase importante, et après une vérification interne, un expert extérieur à la recherche examine la chronique de façon critique pour y dépister les failles et confirmer les limites.

Tableau 17.7 : Techniques de contrôle de la qualité des données en cours de recherche.

3.3.3.2 Le rapport des indices aux concepts : la pertinence ou la validité

La validité se préoccupe du fait qu'en assumant une erreur technique minimale, la trace obtenue réfère bien à la caractéristique de l'objet que l'on veut étudier, et non pas à une autre caractéristique. La non-validité, autrement dit la non-pertinence de la trace par rapport à la caractéristique que l'on veut décrire, résulte le plus souvent d'un mauvais choix d'instrument, d'une erreur dans la construction ou dans le fonctionnement de l'instrument, ou d'une mauvaise utilisation de l'instrument. Le mauvais choix d'un indice serait semblable à l'observation de la couleur des yeux pour déterminer le sexe des sujets. L'erreur de construction ou de fonctionnement la plus typique se produit lorsque le chercheur utilise un vocabulaire technique que l'informateur ne comprend pas ou auquel il attribue une autre signification. C'est le cas du malentendu dans lequel le chercheur croit que l'informateur répond à une question alors qu'il répond à une autre. La mauvaise utilisation s'illustre dans l'analyse de la retranscription d'une entrevue hors contexte lorsque, par exemple, le chercheur analyse un segment comme une louange, alors que, s'il disposait du ton de la voix de son informateur, il la comprendrait comme une moquerie ou une critique.

Validité en général

- Analyse comparative continue
- Retranscription des interviews dans les catégories de l'informateur
- Présence prolongée sur le terrain (observation participante)

Validité de contenu

- Pondérer l'évidence

Validité de concordance

- Enregistrement mécanique des données et descripteurs à basse inférence
- Confrontation de chercheurs ayant la même formation
- Contrôle par les pairs ou contrecodage
- Contrôle en double insu
- Vérification externe des comptes
- Contrôle de l'effet de sélection de site
 - Réunion d'analyse intersite
 - Occurrence intersite de catégories dérivées des participants
 - Répétition

Validité conceptuelle de convergence

- Triangulation des méthodes

Tableau 17.8 : Techniques communes de contrôle de la vraisemblance et de la pertinence.

Comme la validité s'intéresse à l'autre face du même problème que la fidélité et que celle-ci est un élément nécessaire (mais non suffisant) à la validité, les méthodologues considèrent qu'une bonne partie des techniques d'élaboration et de contrôle de la fidélité servent aussi à la validité. La liste du tableau 17.8 reprend les *techniques communes à la fidélité et à la validité*.

Certaines stratégies de recherche, extraites des suggestions de Miles et Huberman, s'adressent plus particulièrement aux différents types de validité habituellement distingués dans la tradition psychométrique. Les *techniques spécifiques de contrôle de la validité*, présentées au tableau 17.9, tentent d'en établir un équivalent.

Validité de concordance
Confirmation par les locaux
Validité prédictive
Vérification locale d'une prédiction ou d'une rétroduction
Validité conceptuelle
Apport conceptuel des indices
Vérification locale d'un réseau causal
Analyse des cas exceptionnels, marginaux, déviants
Variations concomitantes
Pertinence par rapport aux indices rivaux et complémentaires

Tableau 17.9 : Techniques spécifiques de contrôle de la pertinence.

Pour la validité de concordance, on suggère la *confirmation par les locaux*. Il s'agit de vérifier dans quelle mesure un ou des informateurs confirment le codage et les analyses effectués par le chercheur. Cette vérification peut se faire en fournissant aux informateurs locaux une synthèse des observations, en leur demandant de l'évaluer et d'indiquer les corrections souhaitées. La vérification du codage implique que le chercheur communique sa grille de codage et vérifie que l'informateur la comprend.

Pour la validité prédictive, on peut procéder à la *vérification locale d'une prédiction ou d'une rétroduction*. Lorsqu'un contrôle longitudinal est possible, on vérifie la valeur d'un indice par les prédictions qui y sont associées. La validité est estimée auprès des répondants locaux qui ont vécu les événements prédits. On peut procéder de même pour une rétroduction, c'est-à-dire pour la détermination après coup du facteur ou des facteurs à l'origine d'un événement du passé. La procédure de validation comporte quatre étapes :

- 1° Établir la prédiction avec précision;
- 2° Dresser la liste des facteurs favorables et défavorables;
- 3° Le temps venu, envoyer deux enveloppes aux informateurs :
 - 3a) Dans la première enveloppe, on définit la variable prédictive et la prédiction, on demande d'évaluer la réalisation de la prédiction en spécifiant la situation actuelle, et d'expliquer la description de la situation en listant les principaux facteurs contributifs;
 - 3b) La seconde enveloppe comporte la prédiction et la liste estimée des facteurs favorables et défavorables; on demande d'évaluer la pertinence de chaque facteur en tenant compte de la situation réalisée et en justifiant l'évaluation;
- 4° Retour des enveloppes aux analystes.

La validité conceptuelle peut être évaluée par différentes techniques. On peut d'abord évaluer *l'apport conceptuel des indices*, en comparant les nouvelles relations entre indices et concepts à la suite de la mise au point de son instrumentation : quel est le nouveau degré de précision et de clarification atteint, quelles nouvelles distinctions permet le nouvel instrument et quelles difficultés et ambiguïtés élimine-t-il ? On peut aussi effectuer une *vérification locale d'un réseau causal*. À cette fin, on procure aux informateurs une synthèse des observations faites sur le site, le réseau causal (qui est l'hypothèse induite) et son explication, c'est-à-dire, la narration reliant les indices, les concepts et leur signification. On leur demande d'indiquer les lacunes et les contradictions du réseau et de son commentaire (la narration) et de les corriger de la façon la plus précise possible. La cohérence du sens d'un indice peut aussi être vérifiée en *analysant les cas exceptionnels, marginaux et déviants*. Il s'agit de vérifier le sens d'un indice observé chez des individus par l'absence du même indice chez des individus qui, par leur caractère exceptionnel, marginal ou déviant selon la théorie, ne devraient pas présenter cet indice. Dans une même perspective, on peut vérifier si les valeurs observées pour des indices reliés à des concepts qui *varient en concomitance* selon les théories, ne présentent pas des valeurs contradictoires aux variations annoncées par les théories. Enfin, on peut évaluer la validité d'un indice en confrontant la pertinence des descriptions qu'il produit avec celles obtenues par des indices reliés, d'une part à des concepts (théories) *rivaux*, et d'autre part à des concepts semblables ou *complémentaires*.

3.3.3.3 *L'intention et l'extension des données :* *la contextualité et la transférabilité*

La plupart des chercheurs se soucient de ce que leurs travaux et leurs résultats soient connus du groupe avec qui et pour qui ils ont travaillé. Mais ils sont tentés de généraliser leurs conclusions. Ne voulant pas que leurs travaux «dorment sur des tablettes», ils cherchent souvent à étendre à certains événements, à certains lieux et à certains individus, les inférences faites sur les données recueillies. Toute recherche est confrontée, un jour ou l'autre, avec le problème de sa validité externe, c'est-à-dire de la légitimité de l'extension, de l'extrapolation ou du transfert des énoncés issus de la recherche relativement contextualisée.

Pour le praticien qui souhaite utiliser les résultats de la recherche faite par d'autres, la question de la transférabilité est cruciale. La question a deux faces.

1. La similitude entre le contexte qui a permis telle recherche et la situation actuelle est-elle suffisante pour que l'on puisse espérer que les résultats de cette recherche soient utilisables dans le contexte qui nous occupe présentement ? C'est la question de l'extension des données.
2. Si la similitude n'est pas suffisante, quelles inférences tirer sur les conclusions de la recherche à partir de la différence observée entre les contextes ? C'est la question de l'intention des données; autrement dit, dans quelle limite contraindre les conclusions, ou encore, que veulent exactement et seulement dire les données ?

Afin de répondre à ces questions, certaines techniques ont pour objectif de connaître le mieux possible les caractéristiques des données à partir desquelles les conclusions ont été ou seront tirées. En effet, dans la mesure où l'on connaît très bien le contexte dans lequel les données ont été obtenues, le chercheur et l'utilisateur peuvent plus facilement estimer à quels contextes les résultats sont transférables. Car, paradoxalement, des données très contextualisées (représentatives d'un groupe bien délimité) ont une bonne transférabilité, parce qu'on sait assez bien inférer les limites du transfert qu'elles permettent. Par ailleurs, des données très générales (soi-disant représentatives d'une large population) ont une faible transférabilité car, si elles permettent de tenir des discours généraux, elles échouent lorsqu'on veut appliquer leurs conclusions à un sous-groupe particulier. Lorsque les données sont représentatives d'une trop large population dont on ne connaît que des traits très généraux, les groupes et les individus

auxquels on veut appliquer leurs conclusions se révèlent quasi tous se comporter comme des exceptions !

Les techniques qui suivent sont des suggestions permettant d'établir la contextualité des données.

A. L'échantillonnage théorique ou raisonné permet d'identifier la représentativité des personnes et des situations en fonction de leur expérience de l'événement que l'on veut étudier. Il oblige le chercheur à décrire les traits essentiels et accessoires de ses échantillons et à y inclure les cas typiques et atypiques, les répondants qui s'accordent et ceux qui divergent, les acteurs et les observateurs, etc.

B. L'échantillonnage contrasté en cascade se construit en partant d'un petit échantillon théorique disponible; on demande à chacun de ces informateurs d'introduire l'enquêteur auprès d'un autre informateur qui diverge sur un trait précis. Allant d'un informateur contrasté à un autre, on couvre un éventail des positions différentes, mais identifiées, sur la question. Ce type d'échantillonnage devient progressivement représentatif de la variété des positions rencontrées.

C. Construire un inventaire adéquat des échantillons possibles des groupes, des situations, des événements, afin que le choix ne porte pas sur un échantillon qui se révélerait, après coup, n'être qu'un échantillon exceptionnel ou extrême.

D. Afin de ne pas gonfler ou, au contraire, de ne pas banaliser les conclusions, on veille à ce que les observations portent non seulement sur des événements typiques, mais aussi sur des événements banals ou atypiques. Sinon, on risque de construire une théorie sur des exceptions ou de ne pas voir ce qui est exceptionnel dans le matériel recueilli.

E. Établir la représentativité du matériel, des personnes, de leur discours, des événements, des sites, afin de répondre aux questions : sont-ils rares, les rencontre-t-on souvent, où et quand ? La représentativité peut aussi être établie par comparaison et par contraste : en comparant, sur les dimensions principales, un échantillon avec d'autres échantillons, on peut parfois mieux identifier en quoi il est typique et de quoi il est représentatif.

F. La documentation « épaisse » et détaillée du contexte permet une description exhaustive du contexte local pour que d'autres puissent estimer la parenté entre ce contexte-là et leur terrain d'enquête.

Ces techniques portent sur le contrôle de la qualité des données. Mais les conclusions ne dépendent pas que des données, et la pertinence de leur

transfert peut aussi dépendre des procédés d'analyses, de traitement et d'induction des hypothèses. Aussi, mais cela n'entre pas en compte dans la planification des premières opérations de la recherche, on ne devra pas oublier les techniques spécifiques à la vérification des interprétations.

3.4 La planification

La recherche est une entreprise dont, comme on l'a dit, l'objectif est de produire des énoncés crédibles à partir de traces, ou inscriptions, de l'objet étudié. Dans la plupart des cas, cette entreprise exige la collaboration et la concertation de plusieurs personnes; parmi celles-ci, on trouvera des chercheurs, des consultants, des sources d'information. L'intervention de ces personnes ne se produit pas simultanément et, si l'on veut éviter des temps morts, il convient de planifier la recherche et d'essayer de respecter le plan. En effet, lorsqu'une recherche implique le recours aux enseignants ou aux élèves comme source d'information, il peut être très utile de planifier l'organisation des tâches pour être prêt à les rencontrer à un autre moment que lorsqu'ils sont absents des classes ! Lorsqu'il s'agit d'une recherche entreprise dans le cadre d'études académiques, il importe aussi de planifier la recherche de manière à respecter les délais qui sont imposés et à utiliser les ressources lorsqu'elles sont disponibles. En effet, la recherche exige souvent de mobiliser des ressources matérielles et financières sans lesquelles certaines tâches sont impossibles. Cette planification sera aussi impérieuse lors de recherches subventionnées, commanditées ou dans les recherches appliquées : il s'agira alors d'organiser le travail de manière à produire le matériel ou le rapport dans les délais convenus au contrat.

La planification de la recherche se fait souvent en deux temps : une première lors de l'élaboration du projet, une seconde lorsque le projet est accepté (financé) et qu'on va pouvoir passer à l'action. Cette seconde planification, qui constitue l'organisation concrète de la recherche, est l'occasion de se rendre compte que le temps et le matériel disponible constituent les premières contraintes qui obligent souvent à réajuster un projet idéalement construit sur le plan méthodologique. La réalisation qui en découlera sera peut-être un peu frustrante dans un premier temps, car il faudra couper dans ce que l'on souhaitait faire. Mais, lorsque la recherche aboutit, même amputée des astuces méthodologiques dont on avait rêvé, la satisfaction de l'écrit (guide d'utilisation, rapport, thèse, article) produit dans les temps impartis compense pour les petites frustrations vécues lors des corrections de la planification.

Pour qu'une bonne planification puisse se faire, il est essentiel de porter beaucoup de soin à la première étape de la recherche : la délimitation du problème. Cela étant entendu que les explicitations préliminaires ont déjà été formulées, écrites et discutées avant ou à la suite de la délimitation du problème. À partir de la délimitation et de la construction du cadre conceptuel, il est important que le chercheur modélise non seulement l'objet de sa recherche, mais aussi la démarche de sa recherche (revoir le chapitre 11, Modélisation et simulation, en particulier les tableaux 11.1 et 11.2, ainsi que la figure 11.1, qui illustrent une modélisation et une simulation des premières phases d'une recherche). La planification et la réalisation imposent que l'on dispose d'une bonne représentation de ce qui devra être fait. On ne peut pas se contenter de modéliser la recherche; il faut en faire des simulations : des valeurs concrètes (temps, énergie, personnels, moyens, instruments, argents) devront être affectées à chaque élément du modèle de façon à estimer avec réalisme les possibilités d'effectuer la recherche dans les contraintes temporelles et matérielles qui sont imposées.

Dans le cas d'une recherche appliquée du type recherche-développement, non seulement faudra-t-il modéliser les différentes étapes de la recherche, mais aussi l'objet à développer. Plus encore, il faudra simuler les différentes formes et les réalisations possibles de cet objet à développer, en faisant varier les caractéristiques des dimensions de l'objet et en évaluant la performance (efficacité, effectivité et efficacité) de chacun des modèles simulés. Une fois que les simulations auront produit un modèle de l'objet et de la recherche qui semble respecter les contraintes imposées, on peut considérer le projet de recherche comme réalisable et passer à l'action. Il s'agit alors d'engager les moyens prévus dans la simulation retenue, et de tenir une chronique des événements imprévus, des nouvelles contraintes, et des décisions prises pour les affronter, car l'on sera probablement obligé de s'écarter, en cours de réalisation, du modèle prévu pour arriver à atteindre l'objectif fixé, à résoudre le problème.

Bien qu'un schéma général soit suggéré par l'organisation des cinq derniers chapitres de ce livre, la planification des recherches peut considérer une organisation différente des opérations en fonction des enjeux et des formes principales de recherche.

Dans la recherche visant des enjeux nomothétiques, il serait recommandé de débiter par la clarification des présupposés du chercheur, ce que l'on a appelé les «explicitations préliminaires». Ensuite, on poursuivra par la description du problème et l'élaboration du cadre conceptuel ou théorique. Enfin, on entreprendra la mise en oeuvre méthodologique avec

la constitution des données, leur analyse et leur traitement, l'interprétation des résultats et sa validation pour terminer avec la rédaction du rapport.

Dans la recherche aux enjeux pragmatiques, lorsque c'est la solution apportée au problème qui constitue le critère de validité de la recherche, les premières étapes doivent être inversées. Il s'agit de commencer par la description du problème *in situ* en tenant compte du vocabulaire et des catégories propres aux acteurs sur le terrain, mais aussi de leur contexte, des contraintes qu'ils vivent et de leurs priorités. C'est seulement une fois que le problème aura été bien situé, dans le quotidien, que l'on passera aux explicitations préliminaires en y ajoutant une dimension importante. En effet, lorsque le chercheur n'est pas lui-même à la fois acteur et chercheur, c'est-à-dire lorsqu'il travaille pour d'autres (collègues ou clients), il ne lui suffit pas d'identifier quels sont ses présupposés et ses conceptions, il faut aussi qu'il clarifie et identifie quelles sont les attentes et quels sont les présupposés de ceux avec qui ou pour qui il va travailler. Sinon, la recherche risque fort de manquer de pertinence écologique. Or celle-ci, définie comme la capacité de l'intervention à maintenir et à enrichir les relations des individus avec leur environnement, est une condition essentielle à l'utilisation des résultats de la recherche. Une fois les attentes et les préconceptions réciproques clarifiées, la recherche se poursuit par l'élaboration du cadre conceptuel et méthodologique, suivie par la phase de modélisation de l'objet (si on envisage un développement ou une innovation) et des phases de la recherche.

La recherche aux enjeux politiques se déroule habituellement selon une forme qui se situe entre deux extrêmes. À une extrémité, elle est rationnelle, planifiée et vise des changements à moyen ou à long terme ; à l'autre extrémité, elle est intuitive et réactive, car elle doit répondre à une urgence.

Dans le premier cas, la démarche idéale consiste à clarifier d'abord les attentes et les préconceptions des acteurs et des chercheurs, ce qui constitue la base à partir de laquelle on construit le modèle du but à atteindre. Ensuite, on passe à la description du problème *in situ*, ce qui constitue le modèle de l'origine, c.-à-d. de la situation de départ. On complète cette phase par l'analyse de la situation, c'est-à-dire l'identification des écarts entre les modèles de l'origine et du but. Il s'agira ensuite de modéliser, de simuler et d'évaluer quelles interventions, auprès de qui et dans quelles conditions, permettront de réduire les écarts identifiés lors de l'analyse. C'est seulement après ces phases d'enquête et de conceptualisation que la recherche planifiée passera à l'action.

Dans le cas de la recherche réactive, l'action est souvent mise en place intuitivement, car il s'agit de produire un changement à instaurer impérieusement et immédiatement dans le cours même d'une action qui se déroule et qui menace. Dans ce type d'intervention, c'est l'action qui précède la recherche, celle-ci constituant une réflexion évaluative enclenchée *a posteriori*. Elle prend alors la forme d'un récit rétrospectif des événements qui se sont produits, récit accompagné d'une introspection tentant d'élucider les perceptions et les décisions. Ensuite, une analyse comparant les conditions de la situation critique qui a déclenché l'action, avec les conditions obtenues après l'action, permet d'évaluer l'action et de tirer pour l'avenir les leçons de ce qui s'est passé.

3.5 La constitution du matériel de base et la révision du modèle initial

- | | |
|---|---|
| 1 | Collecte du matériel |
| 2 | Analyse sommaire du matériel recueilli |
| 3 | Confrontation des données fournies avec le modèle conceptuel initial du problème (étape 6 de la phase 1, voir tableau 17.2), soit avec <ul style="list-style-type: none"> - le modèle du problème et la liste des questions de recherche (exploration) - la simulation de la situation de départ et de la situation d'arrivée (application) |
| 4 | Constitution de données complémentaires |

Tableau 17.10 : Constitution du matériel empirique de base et révision du modèle initial.

Dans la recherche exploratoire, le modèle initial du problème, issu du réseau conceptuel, doit aussi être envisagé comme initial, imparfait, à compléter, sinon à reconstruire tout au long de la collecte des données. Un lien constant est maintenu entre le cadre conceptuel et la collecte des données. Comme outil de recherche, le modèle initial est une *charpente provisoire*, qui non seulement sera remplie et progressivement recouverte de matériel, mais qui sera aussi renforcée en cours de route lorsque des contraintes nouvelles ou des matériaux non prévus apparaîtront. Le modèle initial, issu du cadre conceptuel, doit donc être ébauché de la manière la plus complète dans la première phase de la recherche, ce que Gauthier et Baribeau⁶ appellent *la délimitation*. Mais il doit rester ouvert tant que durent la collecte des données et l'analyse préliminaire qui l'accompagne, car il

6 Gauthier F. et C. Baribeau, Traitement de la qualité d'un plan de recherche-action. Dans Cl. Paquette (éd.). *Des pratiques évaluatives*. Victoriaville, QC : Éditions NHP, 1984, p. 285-322.

faut que le chercheur puisse envisager des nouvelles questions, parfois essentielles à la compréhension de son objet, à partir des informations qu'il recueille.

En posant aux informateurs les questions issues de cette première représentation conceptuelle du problème, le chercheur va apprendre des choses, si bien que, la plupart du temps, il se rendra compte que des questions imprévues devraient être posées. Les analyses préliminaires ont pour objectif, mais aussi pour limite, de vérifier si l'échantillonnage choisi produit un matériel complet : quelles questions supplémentaires doivent être ajoutées, quelle question qui n'apporte que des réponses banales et redondantes peut être retirée, quel terrain et quels événements devraient être inclus. Tant que la saturation n'est pas atteinte et qu'on ramasse des données, le modèle initial du problème doit pouvoir changer. La liste de questions sera modifiée, de nouvelles techniques d'investigation apparaîtront nécessaires et seront utilisées. La construction de la charpente va bouger tant que durera la collecte de données. Mais une fois que la collecte de données est terminée, quand on va commencer le traitement, on ne peut plus y toucher. À ce moment-là, le cadre conceptuel est fixé. Le modifier quand le traitement est commencé, c'est risquer d'introduire, dans le modèle initial, les hypothèses auxquelles on cherche d'aboutir, c'est-à-dire qu'on risque d'insérer dans la formulation du problème des éléments qui forcent l'induction des interprétations finales. Si on veut faire de la recherche sérieuse, rigoureuse et crédible, il faut s'interdire d'introduire, en cours de route, dans les énoncés mêmes du problème, des éléments qui appellent les explications qu'on proposera, par la suite, comme trouvées.

Une fois *la collecte terminée, le cadre conceptuel est stabilisé*. Dès que débutent l'analyse et la structuration du matériel, la charpente ne bouge plus : c'est le schéma interprétatif qui commence à se construire, le matériel recueilli grâce à la charpente étant alors aménagé, organisé pour lui donner un sens. Autrement dit, le modèle initial doit être suffisamment complet et souple pour conduire au matériel qui permettra d'induire des hypothèses nouvelles, mais il ne comporte pas ces dernières.

Dans la recherche appliquée, le modèle de la situation de départ n'a de valeur que dans la mesure où il constitue un portrait fidèle de la situation réelle. Il ne peut pas être seulement une construction rationnelle, il doit nécessairement être validé par une confrontation avec des données fiables. C'est la raison pour laquelle, après une première phase de collecte des données, on doit souvent corriger le modèle de la situation de départ et refaire la simulation des stratégies. Mais une fois cette correction faite, le

modèle de l'objet initial doit aussi être fixé. En effet, on ne peut pas changer la représentation de l'objet initial de façon à réduire l'écart avec l'objet final et ainsi à faciliter (illusoirement) l'application des stratégies. Cette réduction artificielle constitue un leurre pour le chercheur et pour ses clients, dans le cas d'un chercheur-consultant. On ne peut pas non plus prendre un portrait obtenu en cours de route et le mettre à la place du portrait initial de façon à minimiser ou à maquiller les changements produits. Si une telle atténuation peut avoir un effet bénéfique à court terme sur l'acceptation d'un programme par un client, à long terme cette atténuation risque d'empêcher toute dissémination et de miner la crédibilité du chercheur.

L'élément clef pour assurer la crédibilité d'une recherche exploratoire (surtout de type qualitatif) ou d'une recherche appliquée est d'en rendre la démarche transparente. Pour ce faire, la chronique (ou journal de recherche) doit rapporter toutes les opérations et toutes les décisions, si possible en indiquant quelles en sont, au moment même, les raisons. Tout ce qui a pu être implicite dans une vérification classique et expérimentale devra être explicite dans le modèle exploratoire qualitatif. Le cadre conceptuel et méthodologique est un des lieux de cette explicitation. La chronique des opérations et des décisions est un autre élément essentiel de l'explicitation qui permet d'assurer, par la transparence de la recherche, sa validité interne.

La validité externe de la recherche dépendra aussi du récit que le chercheur fera des aménagements apportés à son cadre conceptuel et méthodologique. Par ailleurs, c'est le récit des modifications, des aménagements au plan de recherche qui apporte souvent une contribution intéressante à l'histoire de la découverte. Or, en deçà des aspects formels de la rédaction du rapport de la recherche, c'est justement l'histoire de la découverte qui apporte le plus d'arguments pour fonder sa crédibilité aux yeux du lecteur.

Comme le modèle initial se modifie pendant la collecte des données, il est donc essentiel de tenir une chronique des modifications apportées au modèle initial du problème. Cette chronique ne doit pas être un roman anecdotique : elle rapporte seulement les changements effectués, le contexte qui les a produits et les raisons des décisions qui ont conduit à ces modifications. C'est la chronique des modifications prévues à la planification qui terminera ce chapitre et qui, surtout dans la recherche qualitative, permettra aux évaluateurs de dire que la recherche a été transparente et qu'elle est crédible.

LES MÉTHODES D'ANALYSE EXPLORATOIRE

- 1 Les étapes de l'analyse et du traitement
- 2 Les moments de l'analyse des données selon leur format
- 3 Les niveaux d'exploitation
- 4 L'exploitation globale d'un matériel
- 5 L'exploitation analytique
 - 5.1 *L'analyse conceptuelle*
 - 5.2 *L'analyse du contenu*
 - 5.2.1 Les contenus latents
 - 5.2.1.1 *L'analyse phénoménologique*
 - 5.2.1.2 *L'analyse interprétative*
 - 5.2.1.3 *L'analyse structurale de la communication*
 - 5.2.2 L'analyse des contenus manifestes
- 6 Les analyses du discours
 - 6.1 *L'analyse propositionnelle des textes*
 - 6.2 *L'analyse de la structure des textes (analyse nodale)*
 - 6.3 *L'analyse propositionnelle du discours*
 - 6.4 *La nouvelle rhétorique et l'analyse des discours sur l'éducation*
- 7 Les grilles d'analyse et leurs présupposés théoriques

1 Les étapes de l'analyse et du traitement

L'examen de pratiques de recherche en éducation force à opérer des distinctions nécessaires si l'on veut poser le problème de l'analyse de contenu et entrevoir ses solutions. Chronologiquement, il faut considérer trois étapes principales souvent confondues lorsqu'il s'agit d'analyser un matériel qualitatif (invoqué ou suscité).

La première étape consiste à *analyser le matériel* recueilli afin d'en *extraire les données*. En effet, si l'on prend l'exemple d'une entrevue, le matériel recueilli est trop abondant pour pouvoir être aisément saisi, trop riche et trop particulier pour pouvoir être comparé, trop redondant et trop anecdotique pour pouvoir en mettre le sens en évidence. Le matériel de base comporte non seulement de l'information mais aussi du bruit; il faut séparer l'information du bruit avant de pouvoir essayer de mettre en évidence la signification de cette information. Cette première étape recouvre des opérations qu'on appelle habituellement l'analyse de contenu. Elle se termine par le codage du matériel et son inscription sur des fiches, soit des fiches par informateurs, soit des fiches par sites ou par thèmes.

La deuxième étape est l'*examen des données* obtenues à partir de l'analyse du matériel afin d'en *décrire* le contenu. Il s'agit le plus souvent d'exprimer les données dans des tableaux ou des matrices qui font voir comment les données peuvent être représentées : quelle forme, quelle apparence, quelle structure. L'examen des données doit nous faire dire ce qu'on peut voir dans nos données sans avoir à effectuer de transformations sur ces données. Examiner les données, c'est les examiner successivement selon les points de vue qui ont servi à organiser leur production. Par exemple, on regarde si tous les sujets ont répondu à toutes les questions, s'ils ont abordé tous les thèmes et dans quel ordre; s'ils ont utilisé toutes les catégories de réponse possibles ou s'ils ont utilisé un nombre limité d'entre elles. D'un point de vue quantitatif, on dira qu'il s'agit de la description des données : la forme des distributions, les mesures de tendance centrale, de dispersion, de relation, etc. C'est à l'occasion de cette étape que l'on pourra obtenir des réponses aux questions de la vraisemblance (fidélité) et de la pertinence (validité) des données.

Dans la troisième étape, on effectuera des *transformations* sur les données afin de *produire des résultats*. C'est ce que l'on appelle le *traitement*. Dans cette étape, des ensembles de données seront remplacés par des supercodes (par exemple, remplacer les profils des sujets sur n variables par le nom ou par le rang de la classe qui les réunit), ils seront

combinés avec d'autres ensembles et symbolisés par un hypercode (par exemple, si la matrice sur deux variables produit quatre nuages de points, on remplace chacun des points – donc chacun des sujets – par le nom du nuage auquel il appartient). On va ensuite examiner comment ces résultats (ou relations entre supercodes) s'enchaînent, comment ils se distribuent, comment ils se regroupent. Le traitement peut comporter deux phases. Il consiste d'abord à transformer les données en résultats, c'est-à-dire remplacer les données par le résultat d'opérations effectuées sur elles afin de les condenser et de les mettre en relation. Ensuite, il consiste à construire de nouvelles relations et de nouvelles synthèses, plus abstraites, à partir du résultat des transformations.

Dans certaines recherches, l'affichage et la présentation des données organisées selon l'un ou l'autre point de vue permettent déjà d'émettre certaines hypothèses. Il s'agit des cas où des relations apparaissent entre des variables primaires, c'est-à-dire entre des variables telles qu'elles ont été immédiatement appréhendées à partir des procédures utilisées pour la constitution des données. Mais telles hypothèses ne s'obtiennent généralement que lorsque le problème porte sur des quasi-évidences (celles dont il suffit que quelqu'un les énonce pour que tous se demandent comment ils n'y avaient pas pensé avant).

Dans les cas où l'évidence d'une relation n'apparaît pas à l'examen des données, il faut procéder à l'étape du traitement. Comme le traitement transforme les données, on doit alors se poser la question de la légitimité des transformations, c'est-à-dire de savoir si de telles modifications sont permises, si elles sont raisonnables étant donné les caractéristiques du matériel. Il faut ensuite s'interroger sur les biais idéologiques ou théoriques que le chercheur aurait pu introduire dans ses résultats par le choix des transformations utilisées (c'est la question de l'objectivité ou de la neutralité). En outre, il faut aussi reprendre pour les résultats les questions qui s'étaient posées pour les données. En effet, après les transformations que les données ont subies, il faut se demander dans quelle mesure les résultats issus des traitements ont gardé 1° un rapport d'indices valables et plausibles avec le réel (vraisemblance ou fidélité) ; et 2° un rapport d'indices référentiels avec les concepts et les théories utilisés pour énoncer les interprétations (pertinence ou validité).

Notons, dès maintenant, qu'une recherche exploratoire ne se termine pas avec le traitement des données : une fois les résultats produits, il faut les interpréter et valider ces interprétations ; il en sera traité au chapitre 20.

2 Les moments de l'analyse des données selon leur format

L'analyse ne s'effectue pas au même moment et de la même manière si l'on considère les différents types de données, soit les données provoquées d'une part, et les données invoquées ou d'interaction d'autre part.

Lorsque les données sont du type *provoqué*, c'est-à-dire lorsque le format de ces données dépend uniquement du chercheur (questionnaires fermés, tests, manipulations systématiques ou expérimentales), le problème de l'analyse du contenu s'est posé avant que les données ne soient produites par le sujet, lors de la construction de la situation stimulus. Les catégories d'analyse sont construites et validées avant la production des données provoquées. Celles-ci sont contraintes par le format des catégories qui constitue un moule dans lequel le sujet, ou l'informateur, doit couler sa réponse. Comme telle, l'analyse du contenu ne pose donc pas de problèmes au chercheur lorsque les données sont provoquées. Par exemple, dans le cas des tests scolaires, chaque question est l'équivalent d'une rubrique et, lors de la construction du test, le chercheur a attribué des catégories et des valeurs aux réponses dont la liste est fixée. L'éventail des cas permis étant fourni aux sujets par les consignes de l'épreuve, ce sont les sujets qui font correspondre leur pensée aux catégories présentées par le chercheur. Dans le cas des tests et des autres situations provoquant des données, ce sont donc les sujets qui réduisent leurs particularités pour les faire coïncider aux catégories de l'épreuve : ce sont les sujets, et non pas le chercheur, qui procèdent à l'analyse de contenu et qui effectuent le codage.

Le problème de l'analyse de contenu se pose surtout au chercheur pour les données *invoquées*, celles qui préexistent à la recherche et dont le format de production est indépendant du chercheur (archives, articles de journaux, etc.). Il se pose aussi pour les données *d'interaction*, soit celles dont le format dépend à la fois du chercheur et des sujets, puisque résultant de leur interaction (dans les diverses formes d'entrevues). Le chercheur doit interpréter ce matériel pour le classer, le réduire, et le structurer, alors que dans le cas des données provoquées, le problème est transféré aux sujets une fois que l'épreuve a été construite. La solution consiste à établir la liste des rubriques (c.-à-d. les cases à remplir qui correspondent aux questions issues du cadre conceptuel), le catalogue des catégories dans lesquelles on fera rentrer la diversité des réponses fournies par les sujets, et les valeurs ou la pondération attribuées aux catégories.

3 Les niveaux d'exploitation

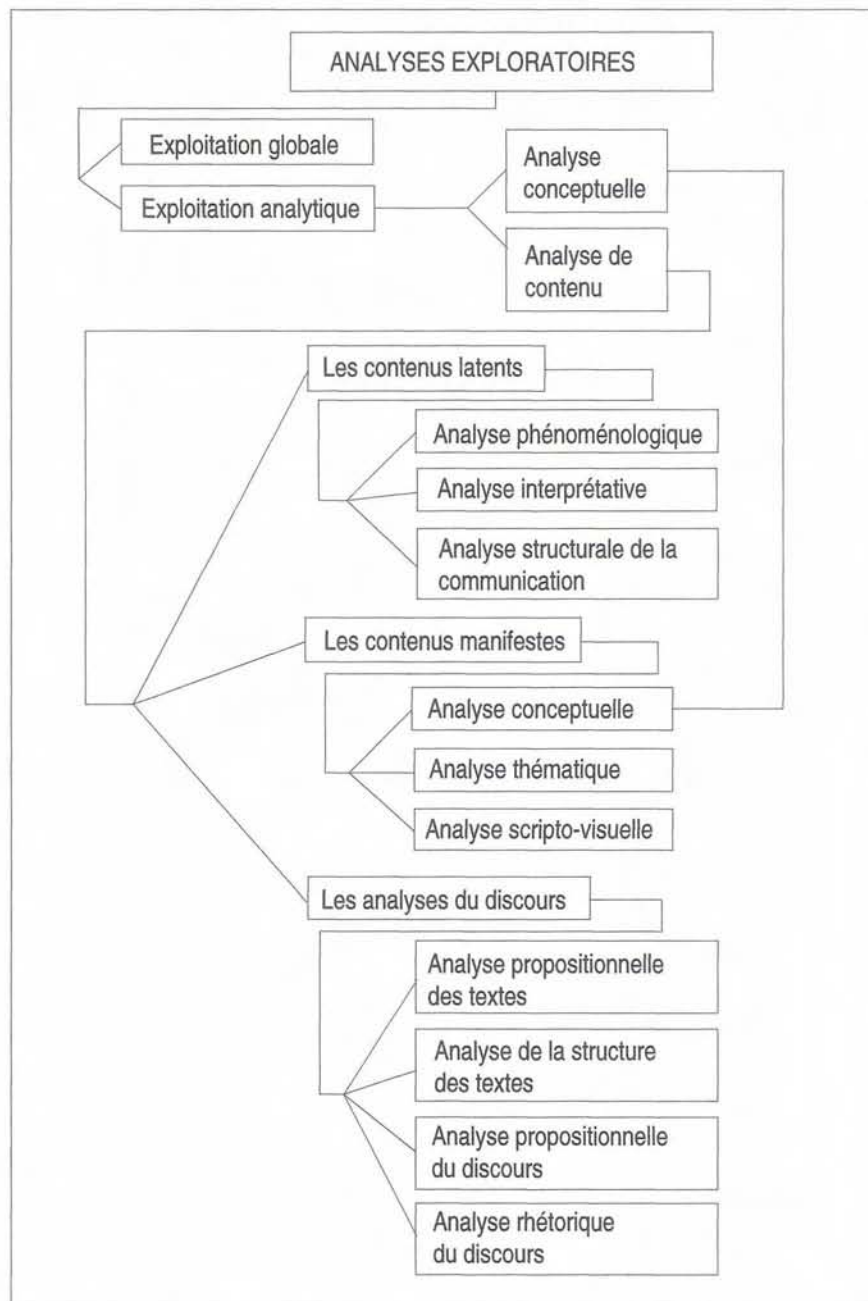


Figure 18.1 : Organigramme des choix méthodologiques dans l'analyse exploratoire.

On rencontre deux niveaux d'analyse, soit une exploitation globale et une exploitation analytique. La rigueur associée à ces deux niveaux d'exploitation est différente, mais l'instrumentation et le temps nécessaire varient aussi. Utiliser l'une ou l'autre constitue donc un choix : ce que l'on gagne avec l'une, on le perd avec l'autre. Comme la figure 18.1 l'indique, d'autres choix devront être faits si l'on privilégie une exploitation analytique. La suite de ce chapitre précise ce que ces choix signifient.

4 L'exploitation globale d'un matériel

Dans l'exploitation globale des données, le chercheur procède par ce qui apparaîtrait comme une intuition esthétique ou politique. Le but est d'élaborer un modèle du sens ou de la signification du texte à analyser afin, le plus souvent, d'en montrer la valeur ou les défauts.

Les phases de la stratégie sont les suivantes :

1° La première phase ne se retrouve pas énoncée dans tous les cas : bien qu'elle soit nécessaire à la suite du travail, elle n'est pas toujours présentée dans le rapport écrit de la recherche. Elle consiste à exposer le *modèle* de l'architecture *du sens attendu* du texte : qu'est-ce qu'il devrait comporter, qu'est-ce qu'il devrait signifier et par quelle argumentation il devrait ou pourrait le faire. Lorsque le modèle reste implicite, le lecteur ou le découvriera ou le reconstruira au fil des phases suivantes.

2° La deuxième phase consiste en un *repérage d'extraits* qui peuvent servir de témoin, sinon de preuve, de la présence ou de l'absence des éléments du modèle présumé. Souvent l'analyste relève les énoncés qui paraissent être des évidences pour l'auteur du texte, c'est-à-dire les pétitions de principes sur lesquelles l'argumentation s'appuie. Par exemple, c'est la technique utilisée par Pierre Gillet¹ dans son analyse du discours expérimental sur l'éducation. Nous citons :

... nous pourrions, maintenant, en tenter l'archéologie et mettre en évidence les formations discursives qui prolifèrent autour de son objet central, car c'est moins dans le traitement objectif de l'Éducation qu'à travers les présumés qui le fondent, que les sujets de ce discours se reconnaissent et s'identifient ; cela revient à dire : à travers «ce qui va de soi», ce qui, pour eux, est évident... ce que nous nommons «formations discursives» (page 193).

1 Gillet P., *Pour une pédagogie*, ou l'enseignant praticien, Paris, PUF, 1987.

D'autres analystes, ayant construit un modèle idéal du discours (phase 1) à partir de leurs propres présupposés, relèvent surtout, en contrepoint aux évidences, le jeu des manques par rapport à leur construction idéale. Alors, ce sont les défauts du discours qui permettront de le caractériser. Dans ces cas, l'intention est «critique» et l'analyse peut friser la caricature sinon la mauvaise foi lorsqu'il est évident, dès le départ, que les présupposés des analystes ne coïncident pas avec ceux de l'auteur. C'est sur une telle base que certaines analyses de programmes et de manuels d'enseignement ont procédé pour aboutir à des condamnations que l'on pouvait anticiper à l'avance étant donné les *a priori* des analystes².

3° La troisième phase consiste en une *reconstruction synthétique du sens* découvert par l'enchaînement des évidences et des manques du discours. L'exposé, souvent polémique, du modèle reconstruit est appuyé par la présentation d'exemples et de contre-exemples argumentatifs bien choisis, qui constituent en eux-mêmes des aveux en même temps que des jugements incontestables. Dans beaucoup de cas, l'analyste n'expose pas la manière dont il a procédé pour cette reconstruction synthétique du modèle du discours, il se contente d'en présenter le résultat.

La comparaison de l'exploitation globale et de l'exploitation analytique du matériel montre que ce type d'exploitation synthétique évite de parcourir la plus grande partie des étapes qui confèrent un minimum de rigueur et de crédibilité aux recherches qualitatives. En effet, si l'on se réfère aux cinq étapes envisagées, on remarque qu'une telle exploitation évite toute analyse des données : pas de codage, pas d'examen des données, pas de traitement ni de validation des interprétations. Dans une telle démarche, le départage du bruit et de l'information est à peine esquissé. Partant d'une sélection d'extraits du matériel, l'analyse aboutit directement aux interprétations et ces dernières ne semblent pas devoir être validées, puisque le parallélisme des plans entre l'interprétation et la sélection des passages parlants paraît évident au chercheur. Mais l'évidence de l'analyste risque bien de ne pas être acceptée par ceux qui ne partagent pas les convictions que véhicule la théorie interprétative à laquelle le chercheur a adhéré et qu'il ne fait que retrouver dans ses interprétations. De plus, il est remarquable que de telles exploitations globales aboutissent rarement à un remaniement de la théorie interprétative : elles ont plutôt l'habitude de mettre en cause la qualité des données ou la sincérité, la conscience ou l'information de l'auteur du texte analysé.

2 Par exemple, certaines analyses «féministes» et «marxistes» des manuels.

Certaines analyses psychanalytiques ou psychosociologiques sur des discours et des pratiques pédagogiques à partir de textes³, d'entrevues⁴ ou d'observations⁵, ont utilisé cette exploitation globale des données. Pour comprendre la démarche des auteurs lorsque le modèle de la première phase est implicite, il faut chercher leurs références dans d'autres travaux ou par les citations de leur bibliographie. S'il ne peut pas découvrir ces références, le lecteur reste à la fois médusé et sceptique devant l'argumentation. La crédibilité accordée au travail dépend alors de la notoriété des auteurs ou du penchant idéologique du lecteur : il l'accepte s'il y croyait déjà à l'avance, sinon il s'en amuse, à moins qu'il ne se choque. Malgré son manque de rigueur, l'exploitation globale des données est assez répandue dans les écrits pédagogiques, parce qu'elle est rapide et qu'elle supporte assez bien les débats politiques ou idéologiques qui animent le champ de l'éducation.

5 L'exploitation analytique

Quand on envisage l'exploitation analytique des données, il faut distinguer entre l'analyse conceptuelle et l'analyse des contenus. L'analyse des contenus s'intéresse à l'information contenue dans un message, alors que l'analyse conceptuelle a pour but de clarifier le sens des mots utilisés pour construire le message. Aussi, dans bien des cas, il faudra effectuer une analyse conceptuelle avant d'entreprendre une analyse de contenu.

5.1 L'analyse conceptuelle

L'analyse conceptuelle n'est pas une analyse de contenu au sens où cette dernière a pour effet de sélectionner l'information du bruit et de la condenser. L'analyse conceptuelle cherche à cerner la dénotation et la connotation d'un concept ou d'une notion en ayant pour but la clarification des énoncés du discours. Comme on en a parlé au chapitre 6, elle est la démarche privilégiée de la recherche spéculative et de la construction du cadre conceptuel ou du cadre théorique. Dans une recherche exploratoire, l'analyse conceptuelle peut même être un préalable à l'analyse de contenu : elle devrait s'effectuer lors de la pré-enquête, au moment où le chercheur prépare son dossier en essayant de mieux connaître son terrain, les occupants de ce terrain et leur langage lorsqu'ils prononcent des énoncés se rapportant à l'objet de la recherche.

3 Par exemple, Rabant C., *La nouvelle illusion pédagogique*, *Orientations*, 1971, p. 29-43 et 87-100.

4 Par exemple, Filloux J., *Du contrat pédagogique*. Paris : Dunod, 1974.

5 Par exemple, Van der Maren J.M., *Fantasme et enseignement*, *Revue Canadienne d'Éducation*, 1987, n°1, p. 31-53.

L'analyse conceptuelle, telle qu'utilisée dans les recherches en éducation, emprunte sa méthodologie aux travaux empiriques de la psycholinguistique et de la sociolinguistique. Sans entrer dans ses détails, relevons qu'elle procède suivant deux stratégies. La première analyse le champ sémantique et relationnel du concept ou de la notion à l'intérieur même du texte. En considérant les concepts à étudier comme pivots, on relève à quels arguments et à quels prédicats ils peuvent être associés, et quelles relations ils entretiennent dans chacune de ces occurrences propositionnelles avec d'autres concepts ou notions. La seconde consiste à établir le champ épisodique du concept ou de la notion, c'est-à-dire le contexte spatio-temporel de son usage. Il s'agit là de repérer quels sont les événements particuliers, quelles sont les anecdotes de la réalité, quels sont les usages concrets auxquels se rattache le concept ou la notion et qui enracinent ses significations dans le quotidien.

L'analyse du champ sémantique est appropriée dans les analyses formelles et théoriques d'un discours, alors que l'analyse du champ épisodique sera plus appropriée aux analyses psychosociales, socio- et psycholinguistiques ou pragmatiques des discours du terrain. Toutes deux impliquent de constituer les listes des occurrences, de construire des tables ou des matrices de co-occurrence, de compter des fréquences, de calculer des indices d'association et de dessiner les arbres ou les réseaux illustrant les champs sémantiques ou épisodiques. Lorsque les chercheurs n'ont pas le temps d'enregistrer les discours en contexte par des entrevues ou des observations participantes, des techniques se rapprochant des questionnaires sont aussi utilisées. Par exemple, des chercheurs ont demandé d'associer par paire des listes de mots, de construire des chaînes d'association de mots ou de mettre en paquets des mots disposés chacun sur un carton puis de relier les paquets par des flèches. L'évolution d'année en année de la carte conceptuelle des étudiants d'un programme a ainsi été établie en présentant sur des fiches les titres des cours⁶. Comme les analyses des champs sémantique et épisodique, ces techniques d'association peuvent aboutir à tracer ce qu'on appelle maintenant une *carte conceptuelle*.

Bien des discussions stériles entre les partisans de théories éducatives pourraient être aplanies, si l'on s'astreignait à effectuer des analyses non

6 Par exemple, Costermans J., Les structures subjectives du lexique, leur genèse et leur évolution : quelques voies d'approche empiriques, *Le langage et l'homme*, 1979, n° 41, p. 3-16. Il y présente diverses techniques et il les illustre en montrant comment les relations entre notions couvrant les matières académiques évoluent entre la première et la quatrième année d'études chez des étudiants en psychologie; il illustre aussi l'évolution du champ sémantique des divers segments du corps, des relations familiales et des adjectifs indiquant les dimensions.

seulement du champ sémantique des combattants, mais surtout de leur champ épisodique ! De nombreux discours semblent diverger alors qu'ils recouvrent des pratiques semblables, et bien des théoriciens utilisent les mêmes concepts avec des significations différentes. Il en est de même pour les observateurs qui étudient un terrain : ils éviteraient sans doute des malentendus s'ils dressaient d'abord la carte conceptuelle des usages de leurs informateurs.

5.2 *L'analyse du contenu*

Au sens restreint du terme, l'analyse du contenu va porter sur deux types de contenus : les contenus *latents* et les contenus *manifestes*. L'attribution du statut latent ou manifeste est indépendante du contenu lui-même : elle n'est que l'expression d'un choix théorique ou idéologique du chercheur. C'est dans ses présupposés que le chercheur a déterminé s'il considère que le discours de ses informateurs est à prendre au pied de la lettre, pour ce qu'il dit simplement, ou bien s'il faut regarder ce discours comme analysable, en droit et en fait, à un double niveau : celui du «dit» et celui du «non-dit». Ce choix découle de la conception que le chercheur se fait de son informateur : son niveau ou son état de conscience ou d'inconscience, son indépendance ou sa dépendance à l'égard des déterminismes sociaux, culturels, économiques, etc.

5.2.1 Les contenus latents

L'analyse des contenus latents a généralement pour but le dévoilement d'une détermination, d'une signification ou d'une structure cachée, non évidente, voire inconsciente. Cette analyse présuppose donc que le discours supporte plusieurs niveaux de message et que plusieurs lectures sont possibles. Elle présuppose aussi que l'émetteur n'a pas fourni explicitement au lecteur ses clés de lecture, non par malice, mais parce que tant l'émetteur que le récepteur sont conditionnés par leur individualité (les particularités de leur histoire) et par leurs contextes socioculturels, et parce qu'ils ne sont pas conscients de ces conditionnements (ils sont aliénés). On rencontre trois modalités principales dans l'analyse des contenus latents : l'analyse phénoménologique⁷, l'analyse interprétative et l'analyse structurale de la communication.

7 À la suite de certains philosophes européens, nous pensons qu'il faut distinguer la phénoménologie et la psychologie phénoménologique (existentielle ou humaniste) : les frontières ne sont pas toujours claires, cela bien que des travaux semblent chevaucher les deux domaines. La première, à l'instar des travaux de Husserl ou de .../...

5.2.1.1 L'analyse phénoménologique

L'analyse phénoménologique, au sens strict et philosophique, est à distinguer des analyses de la psychologie phénoménologique. La première vise ce qui est transcendantal dans la conscience humaine, les fondements de la conscience et de la pensée, la seconde cherche à dégager les structures de base des diverses expériences humaines. Dans les recherches en éducation, les analyses de psychologie phénoménologique sont plus fréquentes que les analyses phénoménologiques au sens strict.

L'analyse phénoménologique présuppose que l'énonciation et la compréhension sont marquées par des caractéristiques personnelles qui, dans la construction par la conscience des phénomènes, nous empêchent de saisir leur dimension existentielle, phénoménologique. Le but de l'analyse phénoménologique est de rejoindre l'existential, c'est-à-dire ce qui est fondamentalement humain dans la construction des phénomènes de conscience. L'analyse phénoménologique est l'outil des philosophes de la phénoménologie, parmi lesquels on note Husserl, Heidegger, Sartre, Merleau-Ponty.

Pour atteindre les processus fondamentaux de construction des phénomènes par la conscience, l'analyse phénoménologique va procéder en trois phases :

- La réduction phénoménologique, c'est-à-dire la mise entre parenthèses ou la suspension des préconceptions, des convictions quant aux événements du quotidien. Il s'agit de se dégager des filtres perceptifs et interprétatifs afin d'accéder à une description «décontextualisée» et «dépersonnalisée» des opérations de la pensée;
- La description soigneuse et complète de ce qui apparaît comme donnée de base, une fois que le matériel a décanté, qu'il a été dégagé du particulier;

.../...Merleau-Ponty, a pour objet d'étude les constructions de la conscience, de la pensée, alors que la seconde, à la manière de travaux américains comme ceux rapportés par Bachelor et Joshi, vise plutôt le dévoilement de la dynamique fondamentale du «vécu» des phénomènes en postulant non pas un quelconque inconscient, mais seulement un particularisme des descriptions de la conscience dû aux interférences socio et proprio-contextuelles. Quoi qu'ils en disent par ailleurs, les exemples apportés par ces auteurs, soit l'étude de la suspicion, de la solitude et de l'empathie, indiquent que la psychologie phénoménologique s'intéresse plus aux états d'âme qu'aux constructions de la conscience. Nous classons les analyses de la psychologie phénoménologique parmi les analyses interprétatives de contenu latent.

- Une réflexion sur les descriptions pour procéder à la réduction eidétique qui consiste à dégager les relations essentielles, fondamentales entre les thèmes, les structures, les récurrences, les significations.

Cette analyse phénoménologique pourra porter sur tous les documents, récits et textes qui rapportent des expériences de conscience. Peu importe, pour la phénoménologie, que le document provienne du chercheur ou d'un autre sujet, parce que leur expérience partage celle des autres en tant qu'êtres humains, et peut permettre de conduire à la compréhension des traits fondamentaux de la conscience humaine⁸.

5.2.1.2 L'analyse interprétative

L'analyse interprétative procède en comparant le discours à analyser avec un modèle ou une fiction idéale issue de l'orientation théorique à laquelle le chercheur adhère. On y considère que le discours à analyser n'est qu'une surface sous laquelle un autre message se cache (le contenu latent) et peut être compris, à condition d'appliquer à la lecture de cette surface une grille ou clef d'interprétation. Cette grille, ou clef d'interprétation, est la trame des relations entre les symboles dont serait constituée la surface du discours. Cette trame est fournie par une théorie, dont un exemple connu est la psychanalyse : au sous-sol du discours conscient existe un discours de l'inconscient, et les mots du conscient seraient des symboles exprimant l'inconscient du sujet.

Dans une première lecture du texte, lecture dite flottante, l'analyste cherche à repérer les passages qui pourraient être des signes, des indices, des symboles de la présence d'une trame théorique. Il cherche donc quels sont les éléments du texte qui sont interprétables à l'évidence, soit les liens signifiants-signifiés les moins contestables dans la théorie interprétative qu'il a choisie. Ensuite, il cherche comment les relations entre ces indices interprétables de la surface peuvent correspondre aux relations entre les éléments de la trame latente de façon à pouvoir faire correspondre le sens

8 Il nous paraît intéressant d'essayer le recours à l'analyse phénoménologique comme démarche dans l'élaboration d'une praxis de l'action éducative, dans la mesure où la praxis se définit comme une réflexion théorisante portant sur les gestes des pratiques quotidiennes. Une telle réflexion devrait permettre l'identification des systèmes indices-repères-actions sur lesquels les praticiens fondent leurs décisions et leurs gestes, sans pour autant pouvoir aisément nommer ces systèmes parce que, comme l'ont montré Delbos et Jorion, ces systèmes se vivent et se travaillent sans se dire. Voir Delbos G. et P. Jorion, *La transmission des savoirs*. Paris : Fondation de la Maison des Sciences de l'Homme, 1984.

de la totalité du texte avec le modèle latent présumé. À cette fin, à partir de la superposition signifiant-signifié des indices déjà dévoilés du texte et de leur interprétation la moins contestable, il procède à des ajustements interprétatifs, à des jeux de langage, de façon à ce que les cases non interprétées du texte de surface puissent correspondre à des clefs d'interprétation de la trame théorique. L'analyse du texte se termine lorsque l'analyste a trouvé les moyens (c.-à-d. les jeux de langage, les équivalences symboliques, etc.) de faire correspondre l'organisation des cases majeures du texte, soit sa structure, avec le modèle latent théorique à dévoiler. Le modèle latent, c'est-à-dire la trame interprétative, peut s'inspirer de diverses théories : psychanalyse⁹, socio-analyse¹⁰, psychosociologie¹¹, psychologie phénoménologique¹², interactionisme symbolique¹³, matérialisme dialectique¹⁴, etc.

Les analyses de la psychologie phénoménologique¹⁵ constituent un cas particulier, entre l'analyse phénoménologique (outil de la philosophie phénoménologique), et l'analyse interprétative. D'une part, elle ne procède pas seulement par la réduction phénoménologique et eidétique effectuée par le chercheur, mais elle procède plutôt par décantation et induction des traits partagés par plusieurs descriptions. Autrement dit, alors que la phénoménologie procède par une réflexion qui tente de dégager ce qui est transcendantal, la psychologie phénoménologique privilégie la comparaison des expériences conscientes en considérant comme fondamentaux les traits de ces expériences qui apparaissent partagés. D'autre part, elle ne postule pas un modèle latent qu'elle tenterait de retrouver dans chacun des discours particuliers, mais elle cherche à construire un modèle général à partir des thèmes récurrents à travers les discours particuliers. On classe cependant les analyses de la psychologie phénoménologique avec les analyses interprétatives, parce qu'elles procèdent suivant des phases proches de celles

9 Par exemple, Janine Filloux, déjà cité.

10 Par exemple, Guigou J., *Les analyseurs de la formation permanente*. Paris : Editions Anthropos, 1979.

11 Par exemple, Van der Maren J.M., Analyse de la relation pédagogique en termes de négociation. *Revue de psychologie et des sciences de l'éducation*, 1967-68, III-2, p. 183-198.

12 Par exemple, Bachelor A. et P. Joshi, *La méthode phénoménologique de recherche en psychologie*, Québec : Presses de l'Université Laval, 1986.

13 Par exemple, les travaux d'Erwin Goffman, dont en ce qui concerne l'éducation *Asile et Stigmates*, publiés à Paris aux Editions de Minuit.

14 Par exemple, Horth R., *L'approche qualitative comme méthode de recherche en sciences de l'éducation*, Pointe-au-Père, Québec : Les Éditions de la Mer, 1986.

15 Giorgi A., *Psychology as a Human Science*. New-York : Harper & Row, 1970.

suivies dans les analyses interprétatives¹⁶ : lecture globale du matériel pour en dégager un sens général, repérage des divers thèmes significatifs dans leur traitement particulier, induction d'une structure fondamentale, vérification et ajustement de la structure induite à partir de chacun des traitements particuliers.

Pour clore provisoirement la discussion sur les différences entre les analyses phénoménologiques et interprétatives, on dira que la phénoménologie construit la compréhension par *réflexion*, que la psychologie phénoménologique l'élabore par *induction*, alors que les interprétations (psychanalytiques ou marxistes, par exemple) l'énoncent par *déduction*.

Deux travers guettent les analyses interprétatives des données. D'abord la projection par laquelle l'analyste choisit la théorie interprétative en fonction de sa propre expérience, tenant celle-ci comme un garant de la validité et de la généralité de la théorie (il s'y retrouve, donc elle est vraie). Ensuite, la toute puissance et le voyeurisme théoriques qui forcent le parallélisme entre le texte et la fiction théorique, soit en attribuant au texte des traits de la théorie qu'il ne partage pas, soit en soustrayant du texte les éléments dont la théorie ne peut pas rendre compte. Autrement dit, les analyses interprétatives se voient souvent reprocher le subjectivisme de l'analyste et les interprétations globalisantes par lesquelles elles parviennent à tout expliquer en recourant à des arguments *ad hoc post hoc*.

5.2.1.3 L'analyse structurale de la communication

L'analyse structurale de la communication¹⁷ s'adresse plus particulièrement aux échanges, à la communication entre partenaires. Elle présuppose que la communication entre partenaires est organisée, structurée, et reçoit un sens par la séquence des interactions, par la position que les partenaires y occupent et par le style qu'ils utilisent. L'ordre dans la séquence des messages donne un sens (ou un non sens) aux paroles des interlocuteurs que ceux-ci ne contrôlent qu'en partie, ce qui peut donner l'impression qu'ils en sont prisonniers. Le but de l'analyse structurale des communications est de dégager la structure et le style des interactions, de repérer la logique des positions et des énonciations afin de découvrir les

16 Voir Alexandra Bachelor et P. Joshi, opus cité, p. 49-92.

17 Certains y verront plutôt une analyse systémique, selon leur référence à une théorie plus systémique que structurale qui peut très bien s'appliquer ici. Mais la technique d'analyse et les présupposés nous paraissent plus structuraux que systémiques.

déterminations propres aux messages et de permettre aux interlocuteurs de mieux contrôler la communication ou de la restaurer. L'analyse structurale vise autant la compréhension des communications que leur thérapie (dont l'approche serait plutôt cognitive).

Sur le plan technique, diverses procédures peuvent être utilisées. Leur choix dépend de la théorie de la communication¹⁸ adoptée par l'analyste. La base de ces techniques consiste à découper l'enregistrement de la communication, non pas en propositions, mais en séquences formant au moins une interaction. Mais l'analyse de la première séquence (l'énoncé de l'émetteur et la réaction du récepteur) tient compte de son chevauchement avec la séquence suivante (réaction du récepteur et nouvel énoncé de l'émetteur) qui est délimitée en faisant glisser l'origine du découpage d'un des partenaires à l'autre; et ainsi de suite. L'unité d'analyse donc est la séquence d'énoncés, l'interaction, mais l'analyste travaille en comparant les séquences d'interaction initiées par chacun des partenaires : ce qui permet de dégager la structure symétrique ou asymétrique, les positions complémentaires, dominantes, etc., de la communication.

Ce type d'analyse, appliqué aux situations éducatives, paraît plus riche que le diagnostic formulé à partir de grilles comme celles de G. de Landsheere et Bayer¹⁹, ou autres grilles, qui réfèrent à un modèle implicite du bon enseignant et n'envisagent que l'un des partenaires à la fois. Il permet de mieux comprendre l'actualité de l'ensemble de la communication afin de l'adapter, soit aux positions souhaitées par les partenaires mais qu'ils ont perdues dans le feu de l'action, soit aux positions prescrites pour chacun d'eux à partir ou d'un modèle pragmatique, ou d'un modèle théorique idéal de la communication éducative.

18 Une des plus publicisées est celle dite de l'école de Paolo Alto dont le rédacteur le plus prolixe est Paul Watzlawick. Voir entre autres : P. Watzlawick et al., *Une logique de la communication*. Paris : Seuil, 1972. Yves Winkin a fait une présentation synthétique des principaux acteurs de cette perspective dans : *La nouvelle communication*. Paris : Seuil, 1981. La grand public connaît plutôt la version simpliste et «vie pratique» de l'analyse des transactions à travers la «méthode Gordon».

19 Voir : de Landsheere G. et E. Bayer, *Comment les maîtres enseignent. Analyse des interactions verbales en classe*. Bruxelles : Organisation des études, Ministère de l'éducation nationale, 1973; et : de Landsheere G. et A. Delchambre, *Les comportements non verbaux des enseignants : comment les maîtres enseignent II*, Paris, Nathan; Bruxelles, Labor, 1979.

5.2.2 L'analyse des contenus manifestes

L'analyse des contenus manifestes présuppose que les énoncés d'un discours sont des unités complètes en elles-mêmes sur lesquelles des opérations peuvent porter. Ces opérations peuvent avoir trois buts ²⁰ :

- Analyser et comparer des énoncés afin de préciser comment un auteur ou un informateur parle d'un thème. C'est l'analyse conceptuelle dont il fut déjà question au chapitre 6 et au paragraphe 5.1 de ce chapitre;
- Analyser les énoncés afin de dégager la structure du texte, la forme et la richesse de son contenu, et les stratégies de son argumentation. C'est l'analyse des textes et des discours dont quelques modèles sont présentés au paragraphe 6;
- Condenser, résumer ou éclairer, systématiser le contenu de la pensée d'un ou plusieurs énonciateurs ou encore examiner l'évolution et l'importance relative de différentes énonciations réparties dans le temps ou dans l'espace²¹. C'est l'analyse de contenu au sens restreint. Elle peut s'effectuer sous deux formes, l'analyse thématique et l'analyse scripto-visuelle. L'analyse thématique cherche à identifier de quoi parle un document par le repérage, le comptage et la comparaison des thèmes, des idées directrices, et des termes pivots. L'analyse scripto-visuelle²² tente de dégager comment est organisée la présentation du message; elle analyse le support physique du message, sa maquette, en relevant l'ampleur de l'espace réservé, la place du texte (première page, éditorial, etc.), le découpage, les titres, le rapport texte-image, la typographie, la redondance (les reprises), etc. Elle présuppose que la valeur attribuée au message est liée aux soins apportés à sa présentation.

20 Pour quelques indications et une illustration sur l'utilisation de ces trois angles d'analyse, voir Kelly M., «L'analyse de contenu », chapitre 13, pages 295 à 315 dans B. Gauthier, *Recherche sociale, de la problématique à la collecte des données*. Sillery, Québec : PUQ, 1984.

21 Pour une discussion générale du problème, en français, voir Bardin L., *L'analyse de contenu*. Paris : PUF, 1977. Elle ne fournit pas une guide technique comme semble le présenter le manuel de Mucchielli; mais ce dernier conduit à s'illusionner par le caractère général de ses prescriptions, car sa technique ne vaut que pour une gamme réduite de problèmes. Voir aussi L'Écuyer R., *Méthodologie de l'analyse de contenu : méthode GPS et concept de soi*. Sillery, Québec : PUQ, 1990.

22 Voir les travaux de Jacobi D., dont : *Textes et Images de la vulgarisation scientifique*, Collection exploration-recherche en sciences de l'éducation. Berne, New-York, Paris : Peter Lang, 1988.

La recherche des thèmes, de ce qu'on en dit et de la manière dont on les présente, suppose que l'analyste aborde les textes à analyser avec des questions plus ou moins explicites à leurs propos. Que la recherche soit exploratoire ou vérificative, ces questions devraient avoir été explicitées dans le cadre conceptuel de la recherche lors de la phase de délimitation du problème. L'analyse du contenu d'un texte s'effectue dans un premier temps par repérage et codage des unités d'analyse, soit le tri des propositions ou des passages du texte jugés intéressants et leur extraction pour composer la liste des passages significatifs. Dans un deuxième temps, il s'agira d'effectuer divers calculs, tant quantitatifs que qualitatifs, sur les unités extraites. L'ensemble des opérations de codage et de traitement fait l'objet du chapitre 19.

6 Les analyses du discours

Quatre méthodes types d'analyse examinent le texte dans son intégralité. Ces méthodes d'analyse ont leur origine dans divers courants psychologiques et linguistiques et sont toutes reliées à un modèle théorique du sujet en tant que producteur de l'information. Les deux premières s'intéressent au texte en tant que document écrit dans une perspective narrative ou documentaire, les deux autres envisagent le discours en tant qu'instrument d'échange et d'influence dans une interaction. Leur examen détaillé est trop long pour que ces méthodes, leurs variantes et leur théorie soient ici exposées. La présentation qui suit se limite à en illustrer la technique, les résultats et leurs présupposés.

6.1 L'analyse propositionnelle des textes

Inspirés des travaux de la psychologie cognitive se rapportant à la compréhension et à la mémoire, plusieurs modèles plus ou moins sophistiqués d'analyse propositionnelle ont été proposés. Ils sont surtout utilisés pour effectuer des analyses comparatives de la production de sujets auxquels on demande d'effectuer un rappel écrit ou verbal d'un texte qu'ils ont lu ou entendu. Dans ces méthodes, l'unité d'analyse est une «micro-proposition» composée essentiellement d'un prédicat et d'un ou plusieurs arguments. Le prédicat est une qualité ou une action attribuée à une chose, à un objet, alors que l'argument est la chose ou l'objet à qui est attribué le prédicat ou auquel le prédicat s'applique. Le modèle de base de l'analyse des micropropositions a été proposé par Kintsch (1974) pour l'analyse des récits écrits en anglais. Le Ny (1979) puis Ska (1983) en ont fait des adaptations pour le récit en français puis pour un texte de consignes

(procédures). Un exemple illustre la démarche. Il s'agit d'un paragraphe extrait d'un article²³ tel qu'analysé par Ska.

Le texte de base se lit comme suit :

Les psychosociologues ont défini l'école comme un lieu de parole à propos du savoir à transmettre. Ainsi ils en ont donné un but : la communication du savoir. L'organisation de l'école en tant qu'institution ordonnée à la bonne marche de l'existence sociale devrait donc être centrée sur la communication. C'est la communication qui en est le but, qui la justifie et qui la réalise. Dès lors l'organisation scolaire devrait consister en un ensemble de facilitations à la communication.

Et l'analyse propositionnelle en est, phrase par phrase :

Les psychosociologues ont défini l'école comme un lieu de parole à propos du savoir à transmettre.

- 1.1 définir (psychosociologues, école)
- 1.2 temps : passé (1.1)
- 1.3 lieu (école)
- 1.4 de (lieu, parole)
- 1.5 à propos (parole, savoir)
- 1.6 à (savoir, transmettre)

Ainsi ils en ont donné un but : la communication du savoir.

- 2.1 conséquence (2.2 , 1.1)
- 2.2 donner (psychosociologues, but, école) = donner (X but , école)
- 2.3 temps : passé (2.2)
- 2.4 communication (but)
- 2.5 de (communication, savoir)

L'organisation de l'école en tant qu'institution ordonnée à la bonne marche de l'existence sociale devrait donc être centrée sur la communication.

- 3.1 conséquence (3.2 , 2.1)
- 3.2 devoir (organisation, 3.4)
- 3.3 mode : conditionnel (3.2)

²³ Van der Maren J.M., Communication contre métacommunication en administration scolaire, *Revue des sciences de l'éducation*, 1981, VIII, 1, p. 169-171.

- 3.4 centré sur (organisation, communication)
- 3.5 de (organisation, école)
- 3.6 en tant que (école, institution)
- 3.7 ordonnée à (institution , marche)
- 3.8 bonne (marche)
- 3.9 de (marche, existence)
- 3.10 sociale (existence)

C'est la communication qui en est le but, qui la justifie et qui la réalise.

- 4.1 but (communication)
- 4.2 de (but, école)
- 4.3 justifier (but, école)
- 4.4 réaliser (but, école)
- 4.5 et (4.3 , 4.4)

Dès lors l'organisation scolaire devrait consister en un ensemble de facilitations à la communication.

- 5.1 conséquence (5.2 , 4.1)
- 5.2 devoir (organisation, 5.5)
- 5.3 mode : conditionnel (5.2)
- 5.4 scolaire (organisation)
- 5.5 consister (organisation, ensemble)
- 5.6 de (ensemble, facilitations)
- 5.7 de (facilitations, communication)

L'analyse propositionnelle des 5 phrases du premier paragraphe de ce texte produit donc 33 micropropositions, qui selon Kintsch, correspondraient au traitement effectué par les sujets pour encoder en mémoire un tel texte. Meyer (1985,1989)²⁴, parmi d'autres auteurs, a développé un modèle plus complexe incorporant à son analyse des micropropositions indiquant la grammaire des cas (distinction entre agent,

24 Meyer, B.J., *Prose Analysis : Purposes, Procedures and Problems, et Constructing the Content Structure*. Chapitres 2 et 10 dans Britton B. K. et J. B. Black (éds.), *Understanding expository text*. Hillsdale, NJ : Lawrence Erlbaum Associates, 1985.
Meyer, B.J., Young, C.J. et B.J. Bartlett, *Memory Improved : Reading and Memory Enhancement Across the Life Span through Strategic Text Structures*. Hillsdale, NJ : Lawrence Erlbaum Associates, 1989.

patient, bénéficiaire, moment, lieu, manière, etc.) et l'articulation entre les éléments du texte (structure causale, procédure, comparaison, description et problème-solution). Son analyse conduit à multiplier par trois le nombre de propositions à coder par rapport au modèle de Kintsch. Frederiksen (1975,1977,1979)²⁵ a proposé une analyse dont l'unité minimale est le concept. Son modèle ressemble à celui de Meyer, mais il introduit des distinctions plus fines et plus précises encore en ce qui concerne les cas et surtout les types de relations entre les concepts.

Si l'on se fie aux analyses comparatives rapportées par Meyer (1985, p. 33), le modèle de Kintsch, qui est le premier, est aussi le plus simple à apprendre et le plus rapide à utiliser; mais il ne traite que du contenu des propositions. Le modèle de Meyer et celui de Frederiksen seraient plus sensibles que celui de Kintsch aux variations dues au développement et donc plus riches pour analyser des rappels produits par des enfants ou par des sujets déficients. De fait, comme leur analyse découpe encore plus finement la proposition, il suffit que le sujet ait retenu un élément (le prédicat, un des arguments, un marqueur de relation) pour que l'on considère que l'unité a été retenue, alors que selon le modèle de Kintsch, le sujet doit avoir rappelé l'ensemble de la microproposition pour qu'elle soit codée comme présente. Le modèle de Frederiksen permettrait, de plus, d'analyser les relations logiques et les inférences produites par les sujets. Pour le psychologue qui travaille sur les processus cognitifs des jeunes enfants, une telle finesse d'analyse peut être utile. Pour le pédagogue qui tente de vérifier si un apprentissage minimal a été réalisé, il se pourrait que ces analyses conduisent à une inflation : en fonction des exigences scolaires, retenir un segment tellement court d'un texte qu'il est sémantiquement non significatif peut être un leurre. On a critiqué dans les années 1970 les épreuves qui portaient sur la mémoire en ne comportant que des syllabes ou des mots sans relations, parce qu'elles paraissaient sans rapport avec les apprentissages scolaires. Il semble que certaines analyses soient tellement fines qu'elles produisent des résultats qui pèchent par excès d'optimisme :

25 Frederiksen, C.H., Acquisition of Semantic Information from Discourse : Effects of Repeated Exposures. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, p. 158-169.

Frederiksen, C.H., Semantic Processing Units in Understanding Text. Dans Freedle R.O. (éd.), *Discourse Processes; Advances in Research and Theory*. Norwood, NJ : Ablex, 1977.

Frederiksen, C.H., Discourse Comprehension and Early Reading. Dans Resnick L.B. et P.A. Weaver (éds.), *Theory and Practice of Early Reading*, (Vol.1.). Hillsdale, NJ : Lawrence Erlbaum Associates, 1979.

les notes de rappel sont élevées, mais les rappels sont dénués de sens. De plus, l'éparpillement des propositions rappelées sur une liste aussi vaste de propositions à coder conduit à des traitements difficiles à réaliser. Perdu parmi les arbres, on ne voit plus la forêt : il faudra y revenir à propos de la taille des unités d'analyse. Il s'agira alors de choisir une méthode juste assez fine pour produire des données utilisables pour le problème que l'on tente de résoudre.

6.2 L'analyse de la structure des textes (analyse nodale)

Pour analyser des rappels de sujets adultes en situation d'apprentissage scolaire, nous avons, à la suite des suggestions de Meyer, développé un modèle d'analyse qui en retient l'idée de mettre en évidence les relations entre les propositions, c'est-à-dire la structure du texte, mais qui se contente de découper le texte en propositions au sens défini par un grammairien comme Grevisse, et non en micropropositions. Chaque unité d'analyse, la proposition, est définie comme l'ensemble «sujet-verbe-compléments» délimité par deux bornes : les signes de ponctuations et les conjonctions. Pour réaliser cette analyse, après le découpage du texte en propositions, il s'agit de repérer la proposition centrale d'un passage ou d'un paragraphe (son noyau), puis les marqueurs de relations qui indiquent la fonction de la proposition dans le texte et comment elle est reliée aux autres propositions. L'analyse aboutit à retenir, pour chaque proposition du texte, un doublet comportant l'indication de la relation inter-propositionnelle et les mots clefs de la proposition. L'exigence pour le codage des rappels peut, si nécessaire, être un peu plus élevée que celle de Kintsch, puisqu'on peut ne compter comme rappelées que les propositions dont le sujet rappelle la relation en plus du contenu. Par ailleurs, en ne considérant que le contenu rappelé, l'examen des relations de chacune des propositions avec les autres peut fournir des hypothèses, par exemple, quant à la manière dont le sujet pense devoir produire un rappel pour l'examineur.

Comparons l'analyse en micropropositions faites tantôt avec l'analyse nodale du même paragraphe. Les 5 phrases analysées en 33 micropropositions correspondent maintenant aux 8 propositions indiquées entre crochets : [...]. Entre ces crochets, la relation interpropositionnelle est indiquée en italique et le résumé du contenu suit les deux points. La phrase originale suit son analyse.

[1. *Définition* : École lieu de parole]

Les psychosociologues ont défini l'école comme un lieu de parole à propos du savoir à transmettre.

[2. *Extension 1 de définition* : But = communication du savoir]

Ainsi ils en ont donné un but : la communication du savoir.

[3...]

L'organisation de l'école

[4. *Extension 2 de définition* : École = institution sociale]

en tant qu'institution ordonnée à la bonne marche de l'existence sociale

[... 3. *Conséquence des extensions 1 + 2* : organisation centrée sur la communication]

devrait donc être centrée sur la communication.

[5. *Cause 1 de conséquence* : Communication = but de l'école]

C'est la communication qui en est le but,

[6. *Cause 2 de conséquence* : communication = justification de l'école]

qui la justifie

[7. *Cause 3 de conséquence* : communication = réalisation de l'école]

et qui la réalise.

[8. *Conclusion des causes de conséquence* : organisation devrait faciliter la communication]

Dès lors l'organisation scolaire devrait consister en un ensemble de facilitations à la communication.

Ce type d'analyse peut aussi être présenté, comme cela est à la mode chez les psychologues, sous la forme d'un organigramme ou d'une *carte* du texte (ils disent une «map») permettant ainsi de visualiser à la fois la structure du texte et l'importance relative des éléments qui en sont rappelés par les sujets. La figure 18.2 illustre une représentation de ce genre pour l'extrait analysé ci-dessus.

Pour certains auteurs, dont les constructeurs du logiciel «Architext²⁶» qui permet de faire de telles analyses, chaque proposition est un noyau relié à d'autres noyaux, soit directement, soit indirectement, implicitement ou explicitement. Ainsi, dans la figure 18.2, les propositions 5, 6, et 7 sont

26 *ArchiText*, logiciel pour l'analyse et la synthèse de l'information, par Steffin S.A. et M. Jennings, chez BrainPower Inc., 1988. Ce logiciel ne fonctionne qu'imparfaitement sur des textes écrits en français : pour l'utiliser il faut supprimer tous les caractères accentués; de plus, la liste des marqueurs (Stop Word List) est évidemment anglaise. On peut y pallier en introduisant dans le texte, devant chaque marqueur ou devant chaque borne de proposition, un signe (p.ex. : «[«) que le logiciel utilisera pour tronçonner le texte et créer les noyaux.

implicitement reliées entre elles, alors qu'elles sont explicitement reliées à la proposition 3 par une relation causale (c'est parce que la communication est un but de l'école, etc., que l'organisation doit être centrée sur la communication...) et à la proposition 8 qui conclut (Dès lors...) à la fois le paragraphe et le bloc des noyaux 5, 6 et 7. La proposition 8 semble reliée directement au bloc des noyaux 5, 6 et 7 et indirectement à la proposition 1 initiale qui définit ce dont il est question, puis à ses extensions (les propositions 2 et 4).

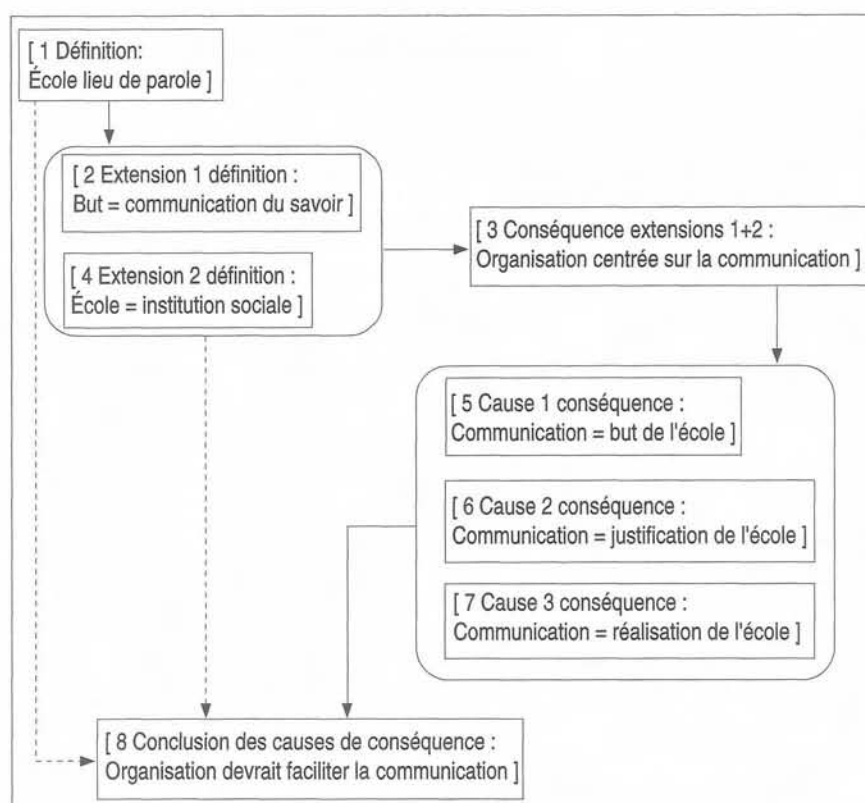


Figure 18.2 : Organigramme ou carte de la structure d'un texte.

Ce type d'analyse des relations interpropositionnelles (ou analyse nodale) peut être utile pour plusieurs raisons. D'abord, pour pouvoir analyser la production des sujets en gardant une vue d'ensemble difficile à maintenir lorsqu'on travaille avec un modèle comme celui de Kintsch ou de Meyer sur des textes réellement utilisés dans l'enseignement postsecondaire. Ensuite, pour observer si et comment les sujets tiennent compte de la

structure des textes, car il peut être important de vérifier dans quelle mesure les sujets ne réduisent pas les différentes formes d'argumentation des textes scientifiques ou théoriques à la simple structure d'une narration. Par ailleurs, lorsque certains effets de mise en page (caractères spéciaux, illustrations, schémas) sont associés à certains noyaux dans une perspective d'optimisation de la présentation, la visualisation de la position des noyaux rappelés permet de faire des analyses intéressantes sur le traitement effectué par les sujets et sur l'effet des éléments visuels ou graphiques adjoints au texte.

Titre : titre ou sous-titre qui marquent le passage à une idée, une thèse

Intertitre implicite : début de paragraphe fonctionnant comme un intertitre

Figure : figure de style pour agrémenter l'exposé (figures ou humour, questions, etc.)

Annonce de X : annonce, introduit, prévient de X

Chapeau-Propos : introduction qui résume

Description : énoncé qui décrit dans le présent, qui énumère les éléments d'un objet, d'un événement

Collection : énonce une collection, une série, une suite (il y a trois x) qui forment un ensemble

Liste : série, suite de, sans l'idée d'un ensemble

Comparaison : évaluation ou présentation en comparaison ou en degré relatif

Expocomp1 : dans une comparaison par contraste en parallèle, exposé du premier élément

Expocomp2 : exposition du 2e élément d'un contraste en parallèle

Analogie : comparaison avec un autre domaine

Évaluation : jugement évaluatif

Thèse : énoncé général central qui sera argumenté, démontré logiquement ou empiriquement

Définition : proposition prédicat qui définit

Argument de X : énoncé à l'appui de, qui justifie, qui légitime

Exposition : exposé d'une idée, d'une thèse

Développement de X : complément, élargissement de X, suite de X, poursuit la même idée que X

Explicitation de X : précise X en explicitant ce qui est implicite, latent, à l'énoncé

Extension de X : proposition associée qui précise, qui complète, qui élargit en situant des lieux d'application, d'utilisation d'une thèse, d'une définition,

Corrélation : étroitement relié, spatialement et logiquement concomitant

Implication : exigence indispensable à respecter, nécessité

<i>Conséquence de X</i> : conséquence inférée ou déduite de X
<i>Contexte de X</i> : contextualisation dans une situation concrète
<i>Spatio-temporel de X</i> : localisation dans l'espace et le temps
<i>Illustration de X</i> : concrétisation de la thèse par un cas
<i>Conclusion de X+X+X</i> : conclusion (synthèse) d'une série d'énoncés
<i>Restriction de X</i> : restriction à une généralisation, conditions mises à une application
<i>Généralisation de X</i> : généralisation de X, absence de restriction
<i>Contraste</i> : relation logique d'opposition
<i>Alternative de X</i> : énoncé d'un autre choix, d'un autre X
Hypothèse : énoncé hypothétique, provisoire et relié à une ou des conditions, annoncé comme tel ou introduit par la conjonction (si alors)
<i>Conditions de</i> : énoncés des conditions dont dépend une hypothèse
<i>Modalisation</i> : nuance, restriction légère, changement minime de sens apporté à un énoncé
<i>Position</i> : annonce du lieu d'un point de vue
<i>Précision de X</i> : précision explicite qui identifie
Procédure : description des étapes d'une procédure
<i>But</i> : objectif, but, finalité, visée
<i>Suite de</i> : conséquence temporelle de X
Cause de X : raison de X
<i>Antécédent de X</i> : préalable (prérequis) explicite ou implicite
<i>Clef de X</i> : cause première, explication finale, raison essentielle d'un raisonnement, etc.
<i>Agent de X</i> : acteur d'une causalité
Problème : énoncé d'une difficulté, d'un problème, d'un obstacle
<i>Solution de problème</i> : solution au problème
<i>Modalité</i> : manière particulière, particularisation (quand, comment ceci, cela) d'une solution ou d'une procédure
<i>Résultat</i> : les résultats d'une recherche, d'une enquête, d'une démarche, les produits

Tableau 18.1 : Liste de fonctions et des relations de liaison pour une analyse nodale.

À l'usage, ce type d'analyse a démontré une autre utilité non négligeable : celle de mettre en évidence la clarté ou la complexité de la syntaxe d'un texte. Lorsqu'on travaille à la formulation et à la reformulation de documents pédagogiques, les difficultés rencontrées à analyser et à dessiner l'organigramme d'un texte fournissent des indices précieux sur la manière de reprendre et de simplifier la construction du texte. Après avoir complété

l'analyse nodale de l'article dont on a présenté un extrait pour fin d'illustration, on savait comment il faudrait le ré-écrire pour que les étudiants le trouvent moins difficile, et cela simplement en modifiant la syntaxe, la structure du texte.

Dans une même perspective, ce type d'analyse peut servir à décortiquer les passages denses de textes que l'on soumet à une analyse conceptuelle. En effet, l'analyse nodale prend trop de temps pour envisager son application à tout un corpus, mais elle est précieuse pour clarifier le sens de certains extraits et pour les comparer. En nous inspirant de Meyer, nous avons mis au point une liste de fonctions applicable à l'analyse des textes habituellement rencontrés dans la littérature en éducation. Cette liste (tableau 18.1) permet l'analyse de texte sur la base des cinq structures de Meyer, soit la description, la causalité, la procédure, la comparaison et le problème, auxquelles nous avons ajouté 3 structures, soit l'évaluation, la thèse et l'hypothèse, ainsi que des éléments généraux de présentation des textes.

6.3 L'analyse propositionnelle du discours

Dans la plupart des cas, les textes pédagogiques, qu'ils soient des outils didactiques ou des documents administratifs, peuvent être considérés comme des discours, c'est-à-dire des textes adressés à quelqu'un. Plus encore que dans le roman, ce sont des intertextes dont le sens résulte de ce qui s'échange entre les interlocuteurs et qui manifestent des enjeux sociaux²⁷. Ghiglione a proposé un mode d'analyse du discours, l'A.P.D.²⁸. Cette analyse propositionnelle du discours tient compte des contrats conversationnels des interlocuteurs et met l'accent sur la forme (syntaxe et argumentation) par laquelle chacun s'exprime à propos de thèmes ou concepts clefs, qu'il appelle les «référents-noyaux». Cette analyse, très méticuleuse, débouche sur une mise en évidence des stratégies utilisées par chacun, non seulement pour respecter ou renforcer les règles du contrat de conversation, mais aussi pour convaincre ou persuader l'autre, pour marquer des points, pour tenir les positions. Comme telle, l'analyse propositionnelle du discours est très longue à mener, mais un logiciel (sur micro-ordinateur IBM, minimum AT

27 Ghiglione R., *L'homme communiquant*. Paris : Armand Colin- collection «U», 1986.

28 Ghiglione R., Matalon B. et N. Bacri, *Les direx analysés, l'analyse propositionnelle du discours*. Paris : Presses Universitaires de Vincennes, 1985.

Ghiglione R., Une méthode d'analyse de contenu : l'analyse propositionnelle du discours. Dans Painchaud G. et M. Anadon (éds.), *Conceptions et pratiques de l'analyse de contenu*. *Revue de l'Association pour la recherche qualitative*, Volume 1 , Hiver 1989, p. 81-117.

avec 20Mg de disque dur) produit des analyses que l'on peut ensuite interpréter à partir de différentes présentations de la nouvelle rhétorique²⁹.

L'analyse propositionnelle du discours et le modèle théorique sous-jacent, présenté dans «*l'homme communiquant*», semblent particulièrement pertinents pour réaliser des analyses des relations professeurs-étudiants. Les résultats présentés par Ghiglione indiquent l'efficacité de sa méthode dans les situations où des interlocuteurs se trouvent liés par un contrat. Certaines des situations analysées semblent assez proches des situations scolaires pour qu'on puisse, non pas transférer simplement les résultats de ces études à l'école, mais utiliser sa technique pour l'analyse des échanges scolaires.

6.4 La nouvelle rhétorique et l'analyse des discours sur l'éducation

La nouvelle rhétorique a aussi d'autres bases qui sont applicables dans la recherche en éducation. Lorsqu'on examine les documents émanant des administrateurs, des ministères ou des penseurs de l'éducation, on ne peut pas ne pas voir une similitude entre leur style et le style juridique. Les consignes données par les enseignants, les procédures qu'ils enseignent sont d'un style qui s'en rapproche. Ces discours pédagogiques sont empreints de légitimation : il leur importe de dire ce qui est bien ou mal, ce qu'il convient de faire ou de ne pas faire, comment il faudrait le faire et avec quelle conséquence (sanction). De plus, ces discours recourent à la rhétorique : ils utilisent un style qui convainc, et dont la persuasion est assez efficace pour qu'on ne puisse pas faire autre chose que d'exécuter l'instruction du message une fois qu'il a été lu ou entendu.

La connexion de la légitimation et de la persuasion dans le discours pédagogique indique une parenté avec le texte de droit qui a conduit Angenot³⁰ à utiliser la rhétorique de Perelman³¹ (développée pour l'analyse

29 Les références à la nouvelle rhétorique de Ghiglione sont principalement :

Grice H.P., *Logique et conversation*. Communications, 1979, 30, p. 57-72.

Grize J.B., *Matériaux pour une logique naturelle*. Neuchâtel. Travaux du Centre de Recherche Sémiologiques de Neuchâtel, 1976, n° 29.

Ducrot O., *Les échelles argumentatives*. Paris : Minuit, 1980.

Ducrot O., *Les mots du discours*. Paris : Minuit, 1980.

30 Par exemple, voir : Germain M. et P. Angenot : Modèles de l'enfant et enjeux éducatifs dans l'école primaire québécoise : quelques techniques d'analyse argumentative. Communication au colloque de l'ARQ à l'UQTR, septembre 1989. Dans son mémoire de maîtrise, M. Germain a utilisé de telles analyses.

31 Perelman Ch. et L. Olbrechts-Tyteca, *Traité de l'argumentation, La nouvelle rhétorique*. Bruxelles : Éditions de l'Université de Bruxelles, 1970, (4^e édition).

des textes de droit) pour analyser les textes pédagogiques produits par le Groupe du Goéland (Écoles alternatives). Les résultats obtenus par ces analyses indiquent que cette méthode pourrait être avantageusement utilisée dans l'analyse des textes idéologiques, des documents théoriques, des règlements, des outils didactiques et des articles de quotidiens se rapportant à l'éducation.

7 Les grilles d'analyse et leurs présupposés théoriques

Les relations qui viennent d'être montrées entre un outil d'analyse des textes et une conception de ce que sont ces textes, suggèrent que le choix d'un instrument n'est jamais indépendant d'une orientation théorique explicite ou implicite. Angenot et Ghiglione indiquent explicitement le lien entre leur instrument d'analyse et la conception (ou théorie) qu'ils ont du texte. Ce n'est pas toujours aussi évident. Aussi le chercheur doit toujours être attentif au fait que le choix d'un instrument implique l'adhésion à une conception ou à une théorie de l'objet auquel l'instrument sera appliqué. C'est une question à se poser au moment du choix des outils, si l'on veut éviter de découvrir que ce choix a introduit, par la théorie qu'il implique, des contradictions avec le cadre conceptuel ou théorique à partir duquel le problème de recherche a été formulé.

LE CODAGE ET LE TRAITEMENT DES DONNÉES

- 1 La détermination des unités d'analyse
 - 1.1 *Les rubriques construites à partir du cadre conceptuel*
 - 1.2 *Les rubriques construites par écho analogique*
 - 1.3 *Les unités de production du sens*
 - 1.4 *Le choix des unités en fonction du problème*
- 2 Les analyses sélectives et le codage
 - 2.1 *Le codage : définition*
 - 2.2 *Les éléments du codage*
 - 2.3 *Les types de codage*
 - 2.4 *Les phases du codage*
- 3 L'examen des données codées
- 4 Le traitement des données codées
 - 4.1 *Le traitement qualitatif des données qualitatives*
 - 4.1.1 La catégorisation
 - 4.1.2 La hiérarchisation et la connexion en réseaux
 - 4.1.3 La conceptualisation et la modélisation
 - 4.2 *Le traitement quantitatif de données qualitatives*
 - 4.2.1 La transformation numérique des données qualitatives
 - 4.2.1.1 *Le codage binaire*
 - 4.2.1.2 *Le codage ordinal*
 - 4.2.2 Le traitement conduisant à des classes :
l'analyse en grappes
 - 4.2.3 Le traitement établissant des relations : les corrélations
 - 4.2.4 La factorisation et la dispersion dans l'espace des axes
principaux
 - 4.3 *Remarques finales sur le traitement*

1 La détermination des unités d'analyse

Quel que soit le niveau d'analyse choisi, un problème de base consiste à déterminer les unités d'analyse : définir sur quels critères certains passages du texte seront retenus et d'autres pas. Notons que cette sélection n'est que provisoire et stratégique. Elle est provisoire parce que, lorsqu'on voudra valider les hypothèses d'interprétations issues des analyses, la première chose à faire consistera à vérifier si ces interprétations, construites sur la portion du matériel sélectionné, fonctionnent aussi bien sur le matériel rebut, c'est-à-dire sur le matériel qui ne nous paraissait pas assez « parlant » pour constituer les unités d'analyse. Cette sélection est stratégique, et c'est en ce sens qu'elle reste un objet de discussion entre analystes, parce qu'elle procède par choix relativement arbitraires pour lesquels aucune règle absolue n'existe. En effet, deux exigences orientent le choix : 1° il y a trop de matériel (si le texte ne comporte que 15 lignes on ne fera pas d'analyse de contenu) et il faut donc sélectionner pour pouvoir condenser; et 2° le réalisme opérationnel conduit à privilégier une méthode de tri cherchant à repérer les passages qui apportent des informations de la manière la plus évidente. Mais l'évidence de l'information reste une question de perception subjective.

La détermination des unités d'analyse s'effectue selon trois modalités différentes. Les deux premières s'appuient sur un modèle de l'objet dont parlerait le texte qu'il faut analyser; c'est l'analyse par rubriques issues du cadre conceptuel ou par rubriques construites par écho analogique. Dans la troisième modalité, les chercheurs s'intéressent plus à la manière dont l'informateur parle de l'objet qu'à ce qu'il dit de l'objet; aussi ces analyses utilisent un modèle de la compréhension ou de la production du sens dans le discours.

1.1 *Les rubriques construites à partir du cadre conceptuel*

Le cadre conceptuel de la recherche comporte souvent une conceptualisation ou une modélisation minimale de l'objet dont parle le texte à analyser. D'une certaine façon, le cadre conceptuel a dressé le contour de l'objet, ses frontières, afin que les résultats de la recherche puissent le remplir. Le cadre conceptuel comporte dans ce cas un postulat méthodologique : si l'on pose certaines questions au texte, on obtiendra de l'information pertinente au problème. Par la liste des questions de recherche, on sait quelles sont les rubriques, c'est-à-dire le nom des cases pour lesquelles le texte devrait fournir des informations. Dès lors, tous les passages du texte qui comportent une information correspondant à une de

ces rubriques seront considérés comme significatifs et retenus comme unités à analyser. Les anecdotes et les passages redondants ne seront pas retenus, bien qu'ils parlent du contenu de la rubrique. Le critère est l'apport d'information au sens de la théorie de l'information : soit une réduction de l'incertitude sur le sujet de la rubrique. Les reprises, les récurrences dans différents paragraphes du texte pourront être comptées comme des indications de la pondération accordée par l'auteur à l'information qu'il livre. Le fait de conduire l'analyse à partir d'une liste de rubriques permettra ultérieurement les comparaisons, les condensations et tous les calculs envisageables lors du traitement. On ne peut réaliser l'analyse comparative de plusieurs documents que si on les confronte à un même modèle, à une liste de rubriques communes.

1.2 Les rubriques construites par écho analogique

Dans certaines recherches, le modèle de l'objet n'est pas disponible avant les analyses préliminaires parce qu'il reste encore à découvrir. Le chercheur a bien une idée de l'objet (sinon, il n'entreprendrait pas la recherche), mais celle-ci découle plus d'une adhésion à une orientation théorique qu'elle ne constitue une hypothèse méthodologique. C'est même plus une intuition qu'il y a là quelque chose à découvrir qu'une hypothèse. Dans ce cas, le modèle de l'objet est inexistant ou n'est pas assez fort pour servir de base à une liste précise de questions. L'analyste ne dispose donc pas de rubriques *a priori*. La lecture du texte s'effectue alors en référence à plusieurs modèles, souvent partiels et analogiques, peu connectés entre eux, perçus comme équiprobables et qui restent latents, c'est-à-dire qu'il n'y en a pas un qui prévaut. Dans cette situation, l'analyste est contraint d'adopter une attitude proche de l'analyse interprétative : sera reconnu comme unité significative tout passage du texte qui éveille un écho, qui suscite une analogie ou une correspondance avec un modèle ou une théorie. Ce sont les passages du texte qui suggèrent les rubriques plutôt qu'ils n'y répondent : ils éveillent des échos dans les connaissances du chercheur. Les éléments qui permettront progressivement d'identifier le modèle de l'objet, ou d'en construire un modèle composite, surgiront donc au fur et à mesure de l'analyse.

La richesse de ce genre d'analyse dépendra de la culture de l'analyste et une bonne stratégie consiste à commencer par une lecture sensible aux échos sur quelques textes, à considérer ces lectures comme préliminaires (il faudra les reprendre par la suite), jusqu'au moment où on peut formuler un modèle provisoire. On considère ce modèle comme stabilisé lorsque la

lecture des textes n'y apporte plus d'éléments nouveaux. Après l'analyse de quelques textes qui servent au débroussaillage du problème et du modèle, on se retrouve comme dans la première situation, soit avec une liste de rubriques; dès lors, le critère de l'information pertinente peut servir à la sélection des unités significatives.

1.3 Les unités de production du sens

Dans la troisième situation, on utilise un modèle hypothétique de la production du texte, c'est-à-dire un modèle de la manière dont l'informateur parle de l'objet. Cela correspond à deux cas : celui où l'on ne dispose pas d'un modèle de l'objet du texte, et celui où l'on préfère examiner le contenant (c.-à-d. la manière dont le sujet parle) plutôt que son contenu. Par exemple, dans certaines recherches didactiques à propos de l'incidence des modes de présentation d'un texte sur sa compréhension par des sujets¹, l'effet des variations dans la mise en page est évalué à partir d'épreuves de rappel. Les rappels qui sont analysés sont induits par des consignes du genre : faire un résumé du texte étudié, expliquer le texte à un ami, écrire un petit article qui commente le texte, etc. Dans de telles études, il semble plus pertinent d'analyser les rappels produits à partir d'un modèle de la compréhension ou de la production du sens, plutôt qu'à partir d'un modèle de l'objet dont le texte traite.

Comme la psycholinguistique fut assez féconde dans les théories de la compréhension et dans celles de la production du sens, le didacticien qui travaille sur des épreuves de rappel a le choix entre différents modèles plus ou moins sophistiqués. Mais le choix d'un modèle n'est pas innocent ni sans conséquence sur les calculs ultérieurement possibles. En effet, comme on l'a vu au chapitre 18 (paragraphe 6, les analyses du discours), une des particularités de ces méthodes est leur exigence d'analyser l'intégralité du texte qui se trouve entièrement examiné et découpé en phrases, en propositions sinon en micropropositions. En conséquence, on constate que plus la taille des unités est petite et le découpage très précis, plus la liste des propositions devient longue et le travail d'analyse laborieux : très vite, il devient impossible de l'effectuer sans recourir à l'informatique.

1 Par exemple, Ska B., Mémoire des consignes et performance à la tâche. *Psychologica Belgica*, XXIII-1, 1983, p. 57-71 (pour l'analyse des rappels).

1.4 Le choix des unités en fonction du problème

Quelle que soit la méthode choisie pour la détermination des unités, la question de la finesse du découpage des unités se pose : est-ce le paragraphe, la phrase (avec quelle ponctuation ?), la proposition, le prédicat ? Certains analystes auraient tendance à miniaturiser les unités et donc à multiplier les entrées des listes, alors que d'autres se contentent de les voir de haut, à en réduire la liste à ce qui saute aux yeux. La juste réponse tient dans l'adéquation de l'outil au problème posé, et le chercheur en éducation doit parfois se rappeler que ses problèmes de recherche ne sont ni linguistiques, ni psychologiques, ni sociologiques.

Une anecdote illustrera ce propos. Le chercheur travaillait avec des enfants qui rencontraient des difficultés majeures dans la compréhension des textes informatifs. Étant donné leurs difficultés, il choisit de faire des analyses propositionnelles en suivant un des modèles les plus sophistiqués de la compréhension, ce qui devait, pensait-il, lui permettre d'identifier la moindre production des élèves. Mais après ces fastidieuses analyses aboutissant à des fichiers gigantesques, des difficultés se présentèrent : à l'oeil, on n'y voyait rien, et les programmes informatiques disponibles ne pouvaient traiter en un bloc une telle masse d'informations. Il fallait donc ou bien tronçonner le texte en différents blocs, ou bien réduire le nombre des propositions en choisissant d'analyser les plus importantes. Sectionner le texte était difficile, car le texte était déjà très court et l'on voyait mal sur quelle base le découper.

Réfléchissant à la réduction du nombre de propositions, on aboutit à ce qui suit. Puisque le texte didactique est multi-style², il semble qu'on puisse *a priori* y distinguer deux parties : le noyau didactique, soit ce que le pédagogue veut enseigner, et la rhétorique didactique, soit les stratégies de communication par lesquelles il va tenter de faire passer ce qu'il veut enseigner. En effet, dans la rédaction d'un texte didactique, le pédagogue utilise plusieurs techniques de communication pour faire passer quelques informations. Il les introduit en les reliant à des connaissances antérieures. Il les explicite par de la redondance, par des paraphrases assurant la compréhension du message, par des illustrations. Il met aussi l'information en contexte en présentant diverses applications, etc. Dès lors, on peut

2 Par exemple, pour J.P. Bronkard le texte pédagogique constitue un texte intermédiaire entre le discours en situation et le discours théorique. Mais, à partir de marqueurs de surface, il ne peut que constater des difficultés de classement, le texte pédagogique se rapprochant tantôt d'un archétype, tantôt d'un autre. Voir Bronkard J.P., *Le fonctionnement des discours*. Neuchatel-Paris : Delachaux et Niestlé, 1985.

procéder à l'analyse des segments du texte après les avoir répartis selon ces deux grandes classes : le noyau et la rhétorique didactique. Il suffit alors de vérifier si l'élève a bien produit les quelques énoncés du noyau didactique pour savoir s'il a compris ce que l'enseignant voulait enseigner. C'est ce que le chercheur finit par faire après un long et très monacal détour.

La situation pédagogique semble donc réclamer l'élaboration d'un modèle qui porte à la fois sur l'objet à transmettre et sur sa communication (c.-à-d. sa construction et sa reconstruction par un sujet considéré autant comme émetteur que récepteur).

2 Les analyses sélectives et le codage

De nombreuses analyses de contenu portent sur des documents où l'analyse intégrale du texte est quasi impossible; par exemple, les entrevues pour lesquelles on a obtenu un matériel long de quelque 75 à 100 pages de texte en simple interligne ! De plus, l'analyse intégrale n'est pas toujours pertinente, comme on vient de le signaler. Tout ce qui y est dit n'est pas intéressant à analyser : l'interlocuteur se répète, il fait des digressions, illustre trois fois son propos pour convaincre l'enquêteur... Il importe donc de sélectionner dans ce matériel abondant ce qui sera à analyser et de coder les passages retenus.

2.1 Le codage : définition

Le codage consiste à accoler une marque à un matériel. Le but du codage est de repérer, de classer, d'ordonner, de condenser pour, ensuite, effectuer des calculs qualitatifs ou quantitatifs, selon que le code utilisé permet l'un et l'autre types de calculs.

Tout matériel peut être codé, c'est-à-dire représenté à l'aide d'un système conventionnel de symboles. On peut coder un matériel très large comme un fichier, un document, ou plus restreint, comme un paragraphe, une phrase ou une proposition dans laquelle on distinguerait le sujet, le verbe, ses modalités, les compléments, etc. L'élément codé est toujours une unité de sens. Cependant, comme on vient d'en discuter, celle-ci peut être prise au sens large, comme un thème, ou au sens restreint, comme l'association des éléments composant une proposition. Dans la majorité des cas, l'unité de sens n'est produite que par une seule source (un auteur). Mais, en référence à une théorie de la communication ou, parfois simplement au bon sens, l'analyste peut aussi décider que l'unité est nécessairement constituée par une séquence d'interactions. En effet, dans bien des cas, la

réaction du récepteur ne peut être codée qu'en tenant compte de l'action de l'émetteur. Autrement dit, la réponse de l'informateur ne peut souvent être comprise qu'à partir de la question de l'enquêteur. Dès lors, si l'élément intéressant à coder est réduit aux paroles de l'informateur, le codage doit englober la question de l'intervieweur, car c'est en référence à cette question que la réponse prendra son sens.

Le code attribué à une unité peut être un mot (un concept), une lettre, une couleur (comme sur les cartes géographiques), un trou sur le bord d'une fiche, un chiffre (trop souvent), ou un symbole : tout système de marques qui permet de repérer et de classer. Le code peut être simple, lorsque l'unité est représentée par un seul symbole. Il peut être multiple, lorsque l'unité est représentée par une combinaison de symboles. En effet, on utilise souvent un premier code pour repérer la rubrique (ou la question), un second code pour indiquer la catégorie (le type de réponse) que l'on spécifie parfois par un troisième code (par exemple, pour l'indication du contexte ou d'une pondération). Lorsque l'on travaille en comparant plusieurs textes, on rajoutera encore l'un ou l'autre code pour identifier l'auteur, son affiliation, etc. Mais les codes trop complexes sont lourds à manipuler. Souvent, il faudra réunir plusieurs éléments d'un code multiple sous un supercode lorsque l'on voudra effectuer des calculs. Tous les codes peuvent être combinés et regroupés en super-, en hyper- ou méta-codes.

Pour qu'un codage soit efficace, on s'attend à ce qu'il soit assez discriminant, c'est-à-dire qu'un même extrait ne puisse pas être codé par plusieurs marques concurrentes. Il faut pouvoir trancher; mais, étant donné la polysémie des codes linguistiques, il est rare que les recouvrements soient tout à fait inexistantes. La rigueur exige aussi que le codage soit réalisé d'une manière constante ou standardisée (fidèle), c'est-à-dire que les mêmes codes soient attribués de manière systématique aux mêmes unités de sens. Les exigences de discrimination et de standardisation impliquent la composition d'un lexique des codes et le respect d'une syntaxe de codage.

Le lexique doit dire à quels éléments du code correspondent quelles unités de sens. Chaque entrée du lexique comporte une définition du code, une liste d'exemples qui explicite la définition, et la liste des cas limites pour lesquels une décision non évidente a dû être prise. La notation des cas marginaux dans le lexique est essentielle pour le contrôle du codage. Sans cette notation, aucun autre codeur ne peut comprendre ce qui a été fait et les coefficients de fidélité intercodeur seront faibles.

La syntaxe du codage est l'ensemble des règles de combinaison et de position (par exemple, la suite : prédicat, argument, cas) des éléments du code. Il est impérieux d'en élaborer une et de la respecter lorsqu'on utilise des codes multiples et des codes combinés. Sinon, les inversions risquent d'être nombreuses : prendre tantôt un code prédicat pour un code argument ou l'inverse, annuler un code objet par un code sujet, transformer un code objet en code sujet en passant d'un mode actif à un mode passif, etc. L'amateurisme du codage sans lexique ni syntaxe aboutit à des résultats inexplicables. Après coup, l'analyste ne comprend plus lui-même les conclusions auxquelles il aboutit, et il ne peut plus justifier ses procédures.

2.2 Les éléments du codage

Dans la plupart des cas, le codage comporte au moins deux, sinon trois éléments : les rubriques, les catégories et les valeurs qui leur sont attribuées.

Les *rubriques* correspondent habituellement aux questions de la recherche; ce sont les noms donnés aux cases du modèle de l'objet, cases vides avant l'analyse mais que celle-ci permettra de remplir. Dans la majorité des cas, le chercheur n'entreprend pas sa recherche sans avoir quelques questions : il y a donc au moins une liste provisoire de rubriques.

Les *catégories* correspondent aux regroupements que l'analyste fait des réponses. Comme il s'agit de condenser une information aussi diversifiée dans sa formulation qu'il peut y avoir d'informateurs possibles, l'analyste doit regrouper l'ensemble des réponses possibles sous différentes catégories, de façon à pouvoir coder les réponses semblables sous une même étiquette. Si l'on dispose d'un modèle de l'objet, il est rare qu'on n'ait pas, *a priori*, quelques idées des grands thèmes, des grandes variétés de réponses possibles. Dans ce cas, un index initial des catégories reprenant la liste de ces grandes variétés de réponses peut servir de base au codage. Dans certains cas, les catégories serviront, non pas à regrouper des réponses, mais à distinguer des classes de réponses : on parlera alors de sous-catégories.

Il est cependant des cas où l'on ne sait pas à l'avance quel type de réponses sera fourni par les informateurs. Il est aussi des cas où l'on souhaite expressément que les catégories correspondent à la terminologie utilisée par les informateurs. Dès lors, on peut travailler sans index de catégories *a priori*, et les regroupements se feront progressivement, en cours d'analyse, en fonction des critères que l'analyste choisira. Un premier regroupement s'effectue d'abord sur la base d'une similitude de signification : les

synonymes, les paraphrases, les analogies commençant par se retrouver dans une même classe. Le regroupement des classes entre elles se construit par un processus de catégorisation, première étape du calcul qualitatif. Habituellement le choix des critères de regroupement est commandé par les rubriques, car s'il ne veut pas être inutile, le regroupement se doit d'être pertinent par rapport au problème posé. Cependant, lorsque l'on veut utiliser la terminologie des informateurs, nous recommandons une stratégie de traduction : on commence par utiliser les catégories qui sont suggérées par les rubriques et le modèle de l'objet, puis, au fur et à mesure de l'analyse, on utilise les mots clefs équivalents proposés par les informateurs pour remplacer les catégories initiales (voir la phase 2.5 du codage).

Enfin, les différentes catégories de réponses peuvent apporter une information équivalente par rapport au problème posé. Dans ce cas, elles sont nominales et se suffisent à elles-mêmes. Mais les différentes catégories de réponse peuvent avoir une plus ou moins grande pertinence, une signification relative, ou être ordonnées en fonction d'une référence (p.ex. : l'ampleur d'une perception, la rapidité d'une réaction, etc.); dans ce cas une *valeur* d'ordre peut leur être attribuée.

Deux situations extrêmes se présentent. Ou bien les valeurs attribuables sont déterminées *a priori* par la hiérarchie des catégories associée explicitement au modèle de l'objet ou aux rubriques. On peut, dès lors, les utiliser en cours de codage; le traitement ultérieur des données n'en sera que plus rapide. Par contre, si la hiérarchie n'est pas prédéterminée, l'analyste doit la construire. Mais la hiérarchisation des catégories *a posteriori* constitue un processus de traitement (c.-à-d. une transformation des données) qui, parce qu'il implique une formulation d'hypothèse sur l'objet, ne peut pas se faire en cours de codage, lorsque le cadre conceptuel d'une recherche exploratoire reste encore ouvert. La construction d'une hiérarchie des catégories *a posteriori* ne peut se faire que quand le cadre conceptuel est stabilisé, car on considère comme une faute d'éthique méthodologique d'introduire en cours de processus ce qu'on n'est censé ne trouver qu'à la fin de celui-ci. Cependant, lors du traitement, la recherche d'une hiérarchisation permettant le regroupement le plus pertinent des catégories, est une phase riche pour la formulation et la sélection d'hypothèses. La situation est différente dans le cas d'une recherche vérificative : la hiérarchie des catégories fait partie de l'opérationnalisation du programme, et on en dispose avant de commencer le codage.

2.3 Les types de codage

Formellement, trois types de codage pourraient exister en fonction de la plus ou moins grande prédétermination de la liste des éléments qui constituent le code : le codage peut être fermé, ouvert et mixte. Dans le *codage fermé*, on trouve des rubriques, des catégories et des valeurs dont la liste est définie dès la construction du cadre conceptuel et méthodologique. Le code fait partie des outils, et l'analyste dispose d'un lexique associant unité de sens (entrée lexicale) et code. Il procède en repérant l'unité de sens à coder et en lui accolant la marque qui lui est associée dans le lexique. À l'autre extrême, le *codage ouvert* est sans rubriques ni catégories *a priori* : on ne dispose pas à l'avance d'un lexique indiquant à quel symbole (code) doit correspondre telle unité de sens. Dans ce cas, le lexique se composera en cours de codage; chaque opération de codage produit une nouvelle entrée lexicale lorsqu'elle utilise une nouvelle marque. Enfin, le *codage mixte* comporte un ensemble ouvert de rubriques et de catégories ou un ensemble fermé de rubriques et un ensemble ouvert de catégories, en entendant par «ensemble ouvert» une liste initiale qui peut se modifier, se compléter ou se réduire en cours d'analyse. Semblablement, les listes des valeurs attribuables à chacune des catégories, des sous-catégories ou des super-catégories peuvent être ouvertes, plus ou moins ouvertes, plus ou moins fermées.

L'avantage du codage fermé est son aspect mécanique. Cela permet de le faire effectuer par un non-spécialiste du domaine; cela produit aussi des pourcentages élevés de fidélité inter et intra-codeurs. Mais le codage fermé ne permet pas de tenir compte de la nouveauté que peut apporter le matériel : il convient plus à la vérification d'hypothèses qu'à l'induction d'hypothèses. Le grand désavantage du codage ouvert provient de ce qu'il exige souvent de revenir un nombre élevé de fois sur le codage d'éléments déjà codés. En effet, tant que l'analyse préliminaire du matériel n'a pas conduit à la saturation du codage, le code n'est pas stabilisé, et les codages déjà effectués devraient être corrigés. En fait, l'analyste risque d'être saturé avant que ses catégories ne le soient.

Le codage mixte, souvent plus pratique, répond mieux aux exigences méthodologiques de la recherche exploratoire. En effet, le codage ouvert ne se justifie que dans la mesure où le chercheur n'a aucune idée de ce que les données pourraient être ni du langage avec lequel ses informateurs vont les lui présenter. Or, cela constitue soit un leurre, soit une erreur méthodologique. C'est le plus souvent un leurre, car il est extrêmement rare que le chercheur n'ait pas au moins des préconceptions sur le sujet de sa recherche,

puisqu'il ne pourrait pas l'aborder sans s'en être fait une idée préalable. C'est aussi une erreur méthodologique, car la collecte d'informations significatives et leur interprétation correcte ne peuvent se faire si le chercheur n'est pas familier avec le langage et le contexte de ses informateurs. Outre sa lourdeur, le codage strictement ouvert ne se justifie donc pas. D'où la valeur d'un codage mixte dans lequel l'analyste procède à deux ajustements de la liste préalable : soit il ajoute des éléments au lexique lorsque des nouvelles unités de sens apparaissent, soit il modifie la formulation du code afin de mieux tenir compte du vocabulaire utilisé spontanément par les informateurs, soit encore il fait les deux ajustements.

2.4 Les phases du codage

Un bon codage nécessite un certain nombre de lectures du matériel, puis de segments qui en sont extraits. Le nombre de phases et leur contenu dépendent du type de codage envisagé. Le tableau 19.1 présente les phases et leur contenu en fonction du type de codage. La suite du texte commente ce tableau.

1	Relecture du cadre conceptuel	F	O	M
	liste des rubriques	F		M
	liste des catégories	F		
	avec ou sans valeurs possibles	F		
2	Lecture du sommaire du matériel (matériel invoqué et suscité)	F	O	M
3	Pré-lecture du matériel		O	M
	détermination des unités d'analyse		O	M
	identification des clefs, catégories (C1)		O	M
4	Lecture du matériel	F	O	M
	pointage (découpage) des unités ou	F	O	M
	extraction des passages significatifs	F	O	M
5	Lecture des unités	F	O	M
	codage des unités	F	O	M
	extraction des segments représentatifs	F	O	M
	extension de la liste des catégories (C2)		O	M
6	Reprise de 2-5 sur un second matériel		O	M
	ajustement de la liste de codage (C3)		O	M
7	Vérification par un second codeur de la liste ajustée		O	M
8	Confrontation avec le codeur 2 et nouvel ajustement de la liste (C4)		O	M
9	Reprise des étapes 2 à 5 avec la liste C4 sur le reste du matériel		O	M
10	Correction du codage des 2 premiers matériaux avec la liste C4		O	M
11	Vérification de la fidélité intra- et inter-codeur	F	O	M

Tableau 19.1 : Phases du codage dans le cas fermé (F), ouvert (O) ou mixte (M).

Phase 1. Quel que soit le type de codage envisagé, toute analyse du matériel gagne en pertinence et en efficacité lorsque le chercheur commence par relire le cadre conceptuel et par se remémorer comment le problème a été conceptualisé et modélisé. Lors d'un codage fermé, cette relecture permet de dresser, si cela n'avait pas été fait, la liste des rubriques, celle des catégories et celle des valeurs s'il y a lieu. Lorsque l'on se prépare à un codage mixte, cette relecture permet de construire une liste provisoire de thèmes clefs, ou d'indicateurs, à propos desquels on va chercher des indices dans le matériel : c'est la liste provisoire des rubriques.

Phase 2. Lorsqu'il s'agit de coder des données suscitées, on procède, dans un deuxième temps, à une lecture (ou à une relecture) du sommaire de l'entrevue à coder. Cette relecture du sommaire est importante afin de se remémorer le contexte de la production de l'information. Elle devient essentielle pour le deuxième codeur (ou lors d'un codage aveugle quand le codeur n'a pas fait lui-même l'entrevue) afin qu'il dispose au moins des impressions de celui qui a fait l'entrevue, qui a vu le sujet, qui a dépouillé le matériel pour la première fois. La fidélité d'un second codeur avec le premier est illusoire si celui-là n'a pas pu lire le sommaire de l'entrevue. De même, la validité du codage aveugle manque de pertinence si le codeur ne sait pas dans quel contexte l'informateur a produit le matériel.

Lorsqu'il s'agit d'un matériel invoqué, c'est-à-dire qui existe déjà comme des articles de journaux, des manuels scolaires ou autres documents produits en dehors de la recherche, il est essentiel de faire une première lecture et de rédiger un sommaire de cette lecture. Quelles sont nos impressions ? Qu'est-ce qui nous accroche ? Qu'est-ce que cela peut susciter comme interprétation ? Quels sont les thèmes qu'on a retenus en lisant ? Cette première perception de thèmes va influencer le reste du travail et il est important de les dépister à l'avance et de noter, à l'intention de ceux qui vont valider le codage, les impressions que cette première lecture d'un matériel invoqué a suscitées chez le codeur.

Phase 3. Après avoir lu le sommaire du matériel, une prélecture du matériel permet de déterminer les unités d'analyses et d'identifier les clés (indicateurs) des catégories. Dans un codage fermé, cette étape a été faite lors de l'élaboration du cadre conceptuel et méthodologique. Cette phase concerne surtout les codages ouverts et mixtes.

Dans un codage ouvert, cette phase commence par une étude du texte afin de déterminer ce qui sera important de retenir. Quelle serait la taille optimale des unités d'analyse, quelles seraient les clés (les indicateurs) et quelles seraient les premières catégories ? Une première liste de thèmes est

dressée, un premier codage apparaît. L'identification des questions clés des catégories est essentielle en codage ouvert (C1). Dans le codage mixte, le cadre conceptuel a déjà fourni une liste provisoire de rubriques et donc le principe de sélection des unités. Cette étude du texte permet d'établir la première liste des catégories et de rajuster le principe de sélection des unités, si nécessaire. Par exemple, on peut décider, devant l'ampleur du matériel, de chercher le thème d'un paragraphe plutôt que celui de chacune des phrases.

Phase 4. Ensuite, on effectue une lecture du matériel pour découper, pointer, repérer les unités : il s'agit de trier et d'extraire les passages significatifs qui seront codés plus tard. Un codage immédiat risque de ne pas tenir compte de l'ensemble de l'information : chaque extrait a bien une signification en lui-même, mais il en a aussi une par rapport aux autres. Il est donc préférable de faire un premier repérage des passages significatifs, ceux qui contiennent l'unité de sens. Il est utile d'identifier les passages retenus par un chiffre et une marque (un soulignement), qui ne sont pas le codage proprement dit, mais qui permettent de repérer dans le texte chacun des passages sans avoir à relire le texte. On prendra une autre fiche pour faire le codage afin de ne pas inscrire le code sur le matériel-source et de le laisser vierge pour le deuxième codeur. En effet, si on inscrit le codage sur la feuille originale, le contre-codage sera biaisé, le second codeur voyant ce que le premier a fait.

Cette phase peut poser quelques problèmes : quels sont les passages à extraire du document ? La meilleure procédure consiste d'abord à identifier (par exemple, en les soulignant) les passages du texte selon les rubriques issues de la phase 1 pour le codage fermé, de la phase 3 pour le codage ouvert ou mixte. On recopie sur une fiche (c'est l'extraction) le passage pointé (des logiciels, comme More, sont très utiles à cette étape : on peut y ouvrir des fenêtres de texte sous un titre, et les réorganiser en déplaçant les titres). On ramasse ensuite sous un même titre (une même fiche, ou une même pile de fiches) tous les paragraphes qui correspondent à une rubrique. De cette façon, on élimine ce qui est accessoire ou anecdotique. En ramenant sous un même chapeau les passages qui parlent de la même chose, on identifie ce qui est reprise ou explicitation d'un passage significatif. On élimine ensuite ce qui est redondance, effet de style ou autres amplifications, pour ne garder que les passages centraux, synthétiques, ceux qui correspondent à des rubriques : c'est ce que nous avons appelé les «passages significatifs». Les passages significatifs extraits du texte comportent toujours des phrases telles que fournies par l'informateur, avec un cotexte suffisant

pour éviter toute ambiguïté dans les phases suivantes, soit le codage et l'extraction des segments représentatifs.

L'extraction des passages significatifs peut déjà donner lieu à bien des discussions entre analystes. Certains ont tendance à trop en retenir, d'autres à être trop synthétiques. La procédure est semblable à celle suivie lorsqu'on fait un résumé progressif d'un texte pour n'en garder que les passages essentiels : on retient tout ce qui est important pour recomposer l'information du texte, sa charpente, en éliminant tout ce qui encombrerait inutilement la mémoire. Il est évident que les connaissances que l'analyste possède du domaine dont parle l'informateur et du langage de l'informateur, vont influencer l'ampleur des passages retenus. Aussi on peut s'attendre à ce qu'un analyste formé dans le domaine retienne moins de passages significatifs qu'un naïf; la comparaison de deux analystes ayant des formations différentes produit dès lors des indices de fidélité factices. On peut de même s'attendre à des textes (ou entrevues) comportant des passages en nombre et en niveau de langage très différents, si l'informateur pense s'adresser à un enquêteur expert dans son domaine, ou s'il pense devoir faire de la vulgarisation. Dès lors, si l'on tient compte du rapport entre la longueur des textes parlant d'un thème et le nombre de passages significatifs pour exprimer ce thème, afin d'établir la pondération relative d'un thème chez plusieurs informateurs, il faut aussi tenir compte du contexte de chaque entrevue, sans quoi cette pondération est factice.

Phase 5. Ensuite, on ne relit plus que les unités de sens qui ont été pointées, les passages significatifs, pour extraire de ces unités ce qui apparaît représentatif. Les «segments représentatifs» qui peuvent être des mots clés ou, au plus, des propositions, sont les éléments les plus réduits du texte correspondant à un code. Comme il est certain que l'on aura des doutes pour coder certains passages, il est préférable d'extraire les segments représentatifs et d'en faire le codage ensuite. Les segments représentatifs retenus pour un code seront aussi utiles lors la rédaction du rapport pour illustrer les interprétations.

Si le codage d'un segment représentatif laisse des ambiguïtés, on a intérêt à retourner au passage significatif et à mieux choisir les mots qui composent le segment représentatif. Si le codage d'un segment n'est pas évident et implique qu'une décision soit prise pour trancher entre diverses possibilités, il importe de noter dans le lexique l'entrée codée, son code et le rationnel de la codification. Cela permettra de coder avec constance lorsque des segments similaires apparaîtront ensuite.

Dans le cas d'un codage fermé, les phases 6 à 10 sont inutiles; après le codage d'un premier document, on passe au suivant et le travail se termine par la vérification de la fidélité intra- et inter-codeur (phase 11). Par contre, lors d'un codage ouvert ou mixte, les passages qui apportent une information nouvelle seront codés par nouvelles catégories. À la fin de cette étape, on aboutit à une première extension de la liste des codes (C2). Mais comme celle-ci sera encore ajustée, il est prudent et économique d'indiquer le numéro du passage significatif, dont le segment a été extrait, sur la fiche où l'on inscrit les segments représentatifs et leur code. Cela, afin de ne pas devoir relire tout le texte si l'on a besoin de remettre le code en contexte pour reprendre le codage lorsque la liste sera stabilisée.

Phase 6. Après avoir terminé le codage d'un premier matériel (document), on procède tout de suite au codage d'un deuxième matériel. Cela permet un nouvel ajustement (C3) de la liste des codes. Le codage d'un deuxième document permet d'évaluer la pertinence de la liste de codage issue de l'analyse du premier document (C2). À cette fin, on choisit de préférence un matériel produit par un sujet qui a des conceptions différentes du premier ou qui travaille dans un contexte différent. Des sources trop semblables produisent en effet des saturations artificielles de la liste des codes. La reprise des phases 2 à 5 sur un matériel obtenu dans un contexte différent du premier, provoque un ajustement de la liste de codage, qui devient plus robuste et plus pertinente pour une application aux autres documents du corpus.

Phase 7. Une fois que la liste a été ajustée sur les deux premiers documents, on demande à un second codeur de procéder sur le même matériel, puis on confronte les codages. Un second codeur est important dès cette étape : ce n'est pas la peine de se lancer dans l'analyse d'un matériel plus abondant si la construction ou l'utilisation de la liste de codage comporte un biais. Autant qu'un contrecodage vérifie tout de suite les biais possibles. Pour pouvoir être critique, cet autre codeur devra être suffisamment indépendant du premier.

Le deuxième codeur doit posséder les questions de la recherche, le cadre conceptuel et le sommaire des documents dont il va contrôler le codage. Il doit avoir la liste de codage, le matériel avec les unités (passages significatifs) qui ont été choisies. S'il choisit d'autres unités d'analyse, une erreur de sélection est possible. Une discussion entre les codeurs doit régler ce problème préalable. Pour calculer la fidélité du codage, la comparaison porte sur les mêmes passages significatifs. On n'indique pas

au second codeur les segments représentatifs des passages, car ceux-ci sont normalement trop réduits et laissent trop peu de marge à l'interprétation pour qu'un contre-codage ait de l'utilité.

Phase 8. Ensuite, la confrontation des deux codages provoque souvent un deuxième ajustement de la liste (C4), liste avec laquelle on procédera sur le reste du matériel. En général, il est plus riche de procéder à cet ajustement final par un échange entre les codeurs, afin de comprendre les écarts entre les connotations et les dénnotations subjectives et d'explicitier au mieux les définitions et les listes d'exemples correspondant à chaque entrée du code. Si cette explicitation est bien faite, plusieurs codeurs pourraient travailler et aboutir à des codages assez fidèles.

Avant les corrections et l'ajustement, on devrait, selon Miles et Huberman, obtenir un indice de fidélité d'au moins 70 %. Si on n'a pas 70 % avant l'ajustement, c'est que la liste est mal faite et que les catégories ne sont pas bien définies : le codage est flou. Après l'ajustement, on s'attend idéalement à ce qu'il y ait 90 % de fidélité, tant pour la correspondance intra-codeur (quand on reprend son propre codage sur le premier matériel) que pour la correspondance inter-codeur (avec le deuxième codeur). Chercher 100 % est absurde car deux mesures ne sont jamais identiques. Si c'était le cas, cela voudrait dire que la procédure est faussée de telle manière qu'un des codeurs opère en fonction de l'autre, et non pas sur la base du matériel.

Phase 9. La liste C4 sert au codage de l'ensemble du matériel en effectuant les étapes deux, quatre et cinq, c'est-à-dire la lecture ou la rédaction du sommaire de chaque document (2), son découpage et la sélection des passages significatifs (4) et l'extraction des segments représentatifs et leur codage (5). Normalement, cette liste devrait être suffisante pour coder toute l'information : elle ne devrait plus trop bouger.

Phase 10. Après avoir effectué le codage sur l'ensemble du matériel on corrige, avec la dernière liste, le codage qui avait été effectué sur les premiers matériaux.

Phase 11. Enfin, pour satisfaire aux exigences de la vérification interne et externe des comptes, on procède à un nouveau calcul de la fidélité inter- et intra-codeurs, non pas sur tout le matériel, mais sur quelques échantillons extraits au hasard, et l'on rapporte les indices obtenus quels qu'ils soient. En effet, les pourcentages de fidélité recommandés par les auteurs ne doivent pas être magiques au point de recommencer le codage tant qu'ils ne sont pas atteints. On rencontre des paires de codeurs pour

lesquelles ces pourcentages sont inaccessibles parce que la formation des analystes diffère trop, ou parce qu'ils ont des *a priori* ou un tempérament trop différents. Non seulement un analyste dont la minutie frise l'obsession mettra beaucoup de temps pour effectuer le codage (avec une fidélité intra excessive), mais en outre, s'il est confronté à un analyste synthétique, l'entente risque d'être impossible et la fidélité inter-codeur plafonnera aux alentours des 40-50 %. Dans ce cas, il vaut mieux ne pas s'obstiner; on procède en demandant à un troisième codeur moins minutieux de coder quelques extraits afin de mieux estimer la fidélité et l'on rapporte tout simplement les chiffres obtenus. Ces chiffres ne feront que souligner la relativité inhérente à tout codage, étant donné la nécessaire subjectivité des analystes, la polysémie des langages naturels et la particularité des expériences et des expressions individuelles.

Tel que décrit, le codage avec toutes ses étapes est une opération qui paraît lourde. Plus d'un sautent cette étape et proposent tout de suite de grands thèmes et font du repérage plus ou moins intuitif. Malheureusement l'analyse intuitive, en évitant les étapes 2,4 et 5, conduit à une sélection d'extraits telle que l'information retenue permet plus de reconnaître l'idéologie des analystes que celle des informateurs. Cependant, le travail n'est pas si lourd qu'il apparaît. En effet, sauf pour les deux premiers documents où il y a deux lectures complètes (étapes 3 et 4), l'ensemble du matériel n'est lu au complet qu'une fois, lors de la sélection des passages significatifs. La deuxième lecture est moins aride parce qu'on ne relit que les passages significatifs pour noter sur une fiche les segments représentatifs; et ce n'est que cette fiche qui sera lue pour être codée. Même s'il y a trois lectures, le nombre de pages à lire se réduit considérablement à chacune d'elles. La lecture la plus longue est celle du découpage, parce que ce dernier doit être fait attentivement. De plus, si l'on ne dispose pas d'un sommaire du document, la première lecture est assez rapide, puisque c'est plus une lecture d'ensemble où on essaie d'appréhender la thématique générale et la cohérence du document.

Le codage en lui-même constitue une étape assez technique, bien documentée par ailleurs³. Là où, en éducation, les problèmes réapparaissent, c'est au niveau de l'analyse des données et des calculs sur les codes, que ceux-ci soient qualitatifs ou quantitatifs. À ce moment, on déborde la question de l'analyse de contenu pour se situer dans l'examen et dans le traitement des données codées.

3 Voir, par exemple, Miles M.B. et A.M. Huberman, *Qualitative data analysis*, a Sourcebook of New Methods. Beverly Hills, CA : Sage , 1984, p. 54 - 69.

3 L'examen des données codées

Comme annoncé au début du chapitre 18, une fois que l'on a trié le bruit et l'information, puis qu'on a condensé les passages informatifs jusqu'à les représenter par un code, on poursuit l'analyse par un examen des données codées. Non seulement cet examen pourra peut-être produire quelques hypothèses inattendues et intéressantes, mais il permettra de savoir si les données sont vraisemblables et pertinentes. Pour ce faire, on s'inspire largement des procédures utilisées dans l'analyse descriptive des données quantitatives; cela est d'autant plus justifié que les chiffres utilisés dans ce cas ne sont qu'un code parmi les codes possibles. Ces procédures vont, de manière plus précise, examiner quel usage on a fait de l'ensemble des éléments du code disponible.

Par analogie, si le codage s'était effectué à partir d'un alphabet comportant un nombre limité de possibilités, par exemple les 26 lettres de l'alphabet utilisé en français, quelques questions pourraient se poser :

- Toutes les lettres de l'alphabet ont-elles été utilisées ? Autrement dit, le codage a-t-il utilisé tous les éléments du code ? Certaines lettres (ou éléments du code) sont-elles utilisées plus que d'autres ?
- Si l'on range les discours codés à partir d'un critère pertinent à la recherche, les lettres utilisées pour le codage se placent-elles selon un certain ordre ? L'ordre des codes qui apparaît lorsque l'on range les discours a-t-il un sens (les lettres forment-elles des mots) ? Si l'on change de point de vue pour classer les discours, le codage utilisé forme-t-il de nouveaux sens ?
- Si l'on compare entre les informateurs le codage de certaines rubriques, découvre-t-on des codages semblables chez certains sujets (des mots semblables, en gardant l'analogie avec l'alphabet) ? Si l'on regroupe les sujets chez qui on observe des codages semblables, quelle forme prennent ces regroupements ?
- Les codages utilisés (les mots formés) recouvrent-ils des observations naïves, observables et repérables ou débouchent-ils sur des énoncés inobservables, uniques, non reproductibles ou figés (question de la vraisemblance-fidélité) ?
- Les codages utilisés forment-ils des mots qui ont un sens par rapport à notre cadre conceptuel, renvoient-ils aux concepts, aux idées, aux catégories avec lesquels les problèmes ont été formulés (pertinence-validité) ?

Pour arriver à répondre à ce genre de question, il faut construire des représentations des données qui les présentent de manière condensée. Ces représentations se font en référence à certains critères (les questions de la recherche et les caractéristiques des informateurs et des données) que l'on utilise l'un après l'autre ou en les combinant l'un avec l'autre. La manière la plus efficace de faire consiste à représenter graphiquement les données : construire des tableaux et des matrices simples ou croisées, et dessiner des figures comme celles illustrées dans le tableau 19.2 : des histogrammes, des diagrammes en bâtons, en feuilles, en boîtes, en secteurs, en cible, des diagrammes de dispersion spatiale (nuages de points, circomplex), des courbes et des droites, des organigrammes, des réseaux, des dendogrammes, etc. Ces constructions graphiques vont être plus artisanales que systématiques dans la mesure où les codes qualitatifs n'imposent pas, par les axiomes qui ont présidé à leur élaboration, des règles qui disent comment faire les opérations. Il faudra donc procéder à divers essais, chaque modification apportée à une représentation pouvant la rendre plus ou moins parlante.

Ces représentations, quelle que soit leur forme, ont pour rôle de montrer quelles hypothèses on peut inférer à partir d'une organisation des données. Il n'y a pas une forme de présentation qui prévale : c'est le chercheur qui essaye diverses organisations à partir des questions du cadre conceptuel par lequel le problème a été délimité, ou à partir de questions ou d'intuitions qui se sont présentées au cours de la recherche. Le chercheur a donc intérêt à essayer diverses représentations afin de mieux voir ce qu'il peut mettre en évidence. Or, la crédibilité accordée à la forme qui lui paraît évidente dépend de la visibilité de cette forme aux yeux des lecteurs : l'esthétique de la présentation, le rapport figure-fond, la lisibilité seront donc des qualités que devront avoir ces représentations. À cette fin, il est important de ne pas les surcharger : il vaut mieux ne pas vouloir tout y mettre quand ce n'est pas nécessaire. Quand les données d'un tableau sont nombreuses, on scindera un graphique complexe en deux ou plusieurs sous-graphiques plus lisibles, en indiquant une ligne de référence (la courbe de base) qui se répétera de graphique à graphique pour mieux permettre leur lecture par comparaison.

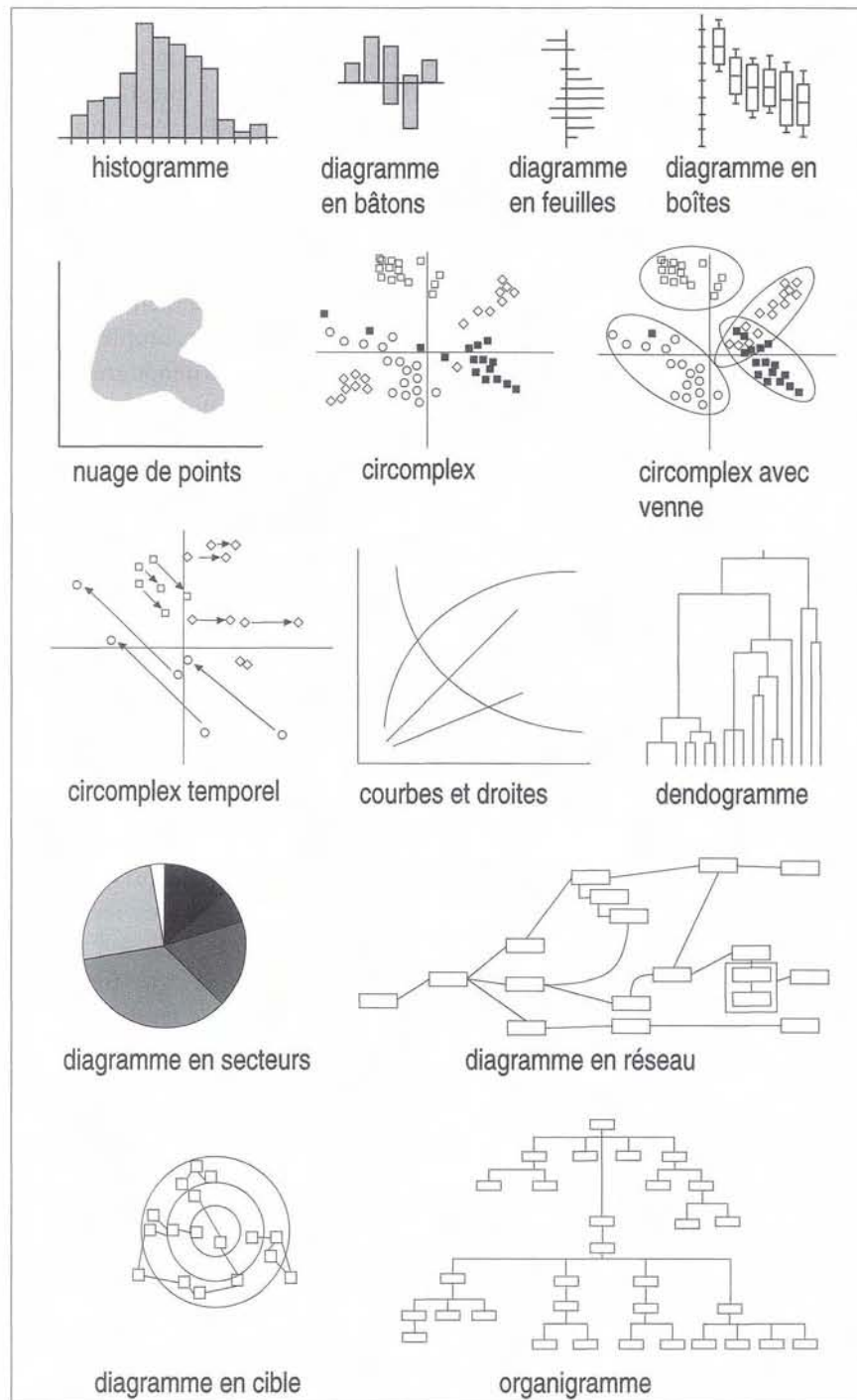


Tableau 19.2 : Exemples des représentations graphiques de résultats.

4 Le traitement des données codées

Une fois que l'analyse descriptive des données est complétée, il faut la poursuivre par un traitement des données qui cherche à mettre en évidence de nouvelles relations entre les premières relations perçues entre les catégories issues du codage. Pour y arriver, il faut transformer les données : remplacer les catégories par les relations qui existent entre elles et effectuer des analyses (ou calculs) sur ces relations.

Lorsqu'on aborde le traitement, une remarque semble essentielle, car une règle y est souvent transgressée. On ne peut travailler sur le code en l'isolant du matériel qu'il représente, car les opérations sur les codes ne sont licites et pertinentes que si on peut les effectuer sur l'objet que le code représente. Par exemple, si on code, en s'inspirant de la géographie, des revenus professionnels par des couleurs différentes, il ne viendrait guère à l'esprit de superposer le code jaune des professions libérales avec le code bleu des chefs d'entreprise pour en conclure que leur moyenne correspond au code vert des enseignants ! Il paraîtrait évident que les opérations qui sont possibles et permises sur les couleurs en tant que telles, ne correspondent pas aux opérations permises sur les revenus ou sur les catégories professionnelles, ou du moins qu'elles ne s'effectuent pas en suivant les mêmes règles. Malheureusement, ce qui paraît évident pour les couleurs semble souvent moins évident lorsque le code est chiffré. Trop de chercheurs considèrent les codes en chiffres comme des nombres, pour effectuer, sur ces codes chiffrés, des opérations arithmétiques qui seraient impossibles à réaliser sur les objets que ces codes représentent. Dans tous ces cas, non seulement le traitement, c.-à-d. les calculs, est illicite, mais en outre il est probable que les résultats ne puissent pas correspondre à un objet identifiable dans la réalité ou dans l'inventaire des objets codés. Si, convaincu de la puissance des calculs numériques afin d'identifier des relations difficiles à percevoir sans eux, l'analyste procède quand même à des opérations illicites, il est essentiel qu'il n'oublie pas la faible légitimité des résultats observés, et qu'il vérifie sur le matériel original, non codé, non transformé, la pertinence et la vraisemblance des relations ainsi mises en évidence.

Effectuer des transformations sur des codes qualitatifs consiste en opérations de catégorisations, de hiérarchisation et de mise en réseaux, de conceptualisation et de modélisation. Il faut, évidemment, disposer d'une collection d'éléments codés recueillis à partir d'une même problématique. On n'effectue pas de traitements, même qualitatifs, sur des éléments provenant d'univers problématiques différents et sans rapports entre eux.

Comme pour l'analyse des données, la présentation graphique et synthétique des relations mises en évidence constitue une partie importante des traitements. À ce niveau, ce seront surtout les matrices et les graphiques permettant d'afficher simultanément plusieurs relations qui seront utilisés. Miles et Huberman⁴ suggèrent plusieurs modèles de matrices et de graphiques. Leur liste n'est pas exhaustive, et les possibilités dépendent de l'imagination des chercheurs et du développement des logiciels avec lesquels ils travaillent.

4.1 Le traitement qualitatif des données qualitatives

4.1.1 La catégorisation

La première phase consiste à réduire, à condenser la collection des éléments. On y parvient en mettant en relief les traits communs et généraux, les éléments qui se répètent, et en dégagant les structures ou les organisations qui peuvent apparaître dans les données. À partir du codage, on crée ainsi des classes d'éléments. Les éléments de la collection pourront être réunis en classes sur la base de leur ressemblance de forme, de la similitude de leur structure ou de leur fonction, ou de tout autre trait qui semble pouvoir être partagé par une partie d'entre eux. Les éléments peuvent aussi être réunis en fonction de leur proximité spatiale ou temporelle (ils apparaissent dans des mêmes lieux, à des moments semblables) ou sur base de relations causales, fonctionnelles (ils jouent un rôle semblables), structurales (ils servent de liens, de passages entre d'autres), etc. Toutes les caractéristiques pouvant être attribuées à plusieurs éléments de la collection peuvent servir de base à une catégorisation en tant que condensation.

4.1.2 La hiérarchisation et la connexion en réseaux

Une fois les classes d'éléments constituées, on poursuit l'analyse en identifiant les caractéristiques typiques des catégories qui recouvrent les classes isolées. En effet, un trait typique à une classe peut ne pas apparaître lorsqu'on examine les individus qui composent cette classe, alors qu'elle est bien identifiable lorsqu'on compare plusieurs classes entre elles. On procède donc à un examen plus attentif des traits non évidents mais qui

4 Huberman A.M et M.B. Miles, *Analyse des données qualitatives*. Bruxelles : De Boeck-Wesmael, 1991. Les chapitres 4 (analyse de cas intra-site, pages 144- 267) et 5 (analyse de cas multiples -inter-sites, pages 271-369) présentent de nombreux exemples de matrices, de tableaux, de circomplex et de diagrammes. D'autres exemples de présentations graphiques sont présentés dans les livrets accompagnant de nombreux logiciels pour micro-ordinateurs.

constituent peut-être des caractéristiques plus fondamentales ou plus générales de chacune des classes. Ces traits moins évidents vont permettre de différencier ou d'associer les classes ou les catégories les unes par rapport aux autres. De plus, certains traits seront jugés essentiels et retenus comme critères d'appartenance qui permettent de rejeter ou d'inclure dans cette catégorie de nouveaux éléments. Ce sont les caractéristiques discriminantes des catégories qui permettent de rejeter de la classe de nouveaux éléments et les caractéristiques associantes qui permettent d'associer à la classe de nouveaux éléments. Enfin, le partage par plusieurs catégories de caractéristiques associantes ou discriminantes, va permettre de réunir entre elles certaines catégories pour composer les super-catégories, soit des regroupements encore plus réduits de classes plus englobantes. Il importe de documenter la manière dont les regroupements sont effectués et d'établir les relations d'inclusion, d'exclusion, de recouvrement entre les différentes catégories et niveaux de catégories, et de préciser sur quel type de traits (proximité, forme, structure, fonction, etc.) les regroupements ont été effectués.

Toutes les relations possibles entre les catégories, entre les catégories et leurs super-catégories, puis entre les super-catégories vont être établies. Cela conduit à la construction de réseaux permettant de formuler des hypothèses. Un réseau sera hiérarchique et ascendant lorsqu'il est construit sur le partage de plus en plus réduit de traits communs; inversement, une hiérarchie descendante est construite sur un partage de plus en plus élevé. Mais tous les réseaux ne sont pas hiérarchiques : les connexions et les différenciations entre catégories et niveaux de catégories peuvent se faire sur des relations fonctionnelles ou structurales. Pour comprendre ces procédures, on peut s'inspirer du travail de classification en zoologie, où les premières classes se construisent sur des traits morphologiques apparents (à l'évidence, beaucoup de chiens se ressemblent), pour ensuite se raffiner et se corriger à partir des caractéristiques anatomiques et physiologiques internes, moins évidentes, réclamant une analyse plus raffinée (mode de reproduction, type d'appareil digestif et respiratoire, sang chaud *versus* sang froid), et encore se perfectionner par des analyses biochimiques ou génétiques (présence d'hormones, de gènes). Si les filiations génétiques conduisent encore à des réseaux hiérarchiques, elles se rapprochent des réseaux non hiérarchiques. Les analyses écosystémiques produisent des réseaux non hiérarchiques, appelés communément réseaux (tout court), lorsqu'elles mettent en évidence que telle famille d'animaux (des puces, par exemple) est un parasite de telle autre (des oiseaux), qui est elle-même un prédateur pour une autre classe (des moustiques).

4.1.3 La conceptualisation et la modélisation

Une fois les catégories, les supercatégories et les réseaux établis, on peut les étiqueter par des abstractions, c'est-à-dire par un terme qui identifie l'élément typique partagé par chacun des regroupements. Cet étiquetage par abstraction produit une définition conceptuelle de l'objet : il est défini par les relations entre les concepts étiquettes de ses éléments. Une fois la définition conceptuelle de l'objet obtenue, on peut la représenter par des matrices, des graphiques ou des cartes conceptuelles, ou dans des formules en sachant que cette représentation est une modélisation induite, une simplification par abstraction. Le modèle induit est une représentation réduite de l'objet : c'est une mise en forme parlante, visible en un coup d'oeil, manipulable, d'un ensemble d'événements ou de faits. Souvent, la représentation à laquelle l'induction conduit s'inspire d'une structure préexistante dans un autre domaine : elle s'appuie sur une analogie. Cela suppose un certain isomorphisme entre l'ensemble modèle et l'ensemble modélisé. Comme indiqué au chapitre 11, les modèles pourront être physiques (des maquettes, des modèles réduits), schématiques (des graphiques, des images, des organigrammes, des plans), ou langagiers comme les analogies et les mythes, les fables, les paraboles, les contes qui sont tous très parlants et qui pourront faire comprendre les explications conceptuelles. Enfin, les modèles formels peuvent parfois interpréter une problématique : programmes, lois, règles, principes, algorithmes, formules.

En fin d'analyse, les hypothèses seront l'explicitation de ces réseaux conceptualisés et modélisés.

Les résultats de ces «traitements qualitatifs» ont une qualité majeure par rapport aux résultats des calculs quantitatifs : les concepts et les modèles auxquels ils aboutissent sont en référence directe avec les éléments qu'ils représentent, alors que les résumés numériques et statistiques ne représentent que l'ensemble du code utilisé. Le concept, la catégorie, le résumé qualitatif ont un statut différent du résumé numérique : ils parlent directement, non pas des individus regroupés dans une classe, mais d'au moins un trait partagé par eux et qui leur est typique. Ce statut différent souligne la richesse qu'une analyse catégorielle apporte en complément de l'analyse numérique habituelle. C'est l'analyse catégorielle qui produira les explications les plus pertinentes aux phénomènes observés à la suite de l'analyse numérique. L'analyse numérique peut mettre en évidence des relations que l'analyse catégorielle permettra d'interpréter.

4.2 Le traitement quantitatif de données qualitatives

Le traitement quantitatif des données qualitatives est important dans deux cas : lorsqu'on a épuisé les ressources du traitement qualitatif sans pouvoir mettre en évidence quelques hypothèses intéressantes, et lorsqu'on veut se donner rapidement un premier aperçu de l'organisation possible des données avant d'effectuer les traitements qualitatifs. En effet, le traitement quantitatif fournit des outils qui permettent d'effectuer, sur base d'un codage numérique et par l'informatique, les mêmes procédures que celles envisagées dans le traitement qualitatif : construire des réseaux, modéliser et simuler sur ordinateur. Le traitement quantitatif des données qualitatives est une stratégie alternative très puissante, à condition de ne pas oublier que la même règle s'applique à tous les traitements : on ne peut effectuer sur les codes numériques que les opérations qui sont légitimes sur les objets que ces codes représentent. De plus, comme les transformations numériques forcent souvent la nature des données qualitatives, il faut ne jamais perdre de vue que les catégories, les réseaux et les modèles qu'elles induisent sont, en grande partie, des artifices. Ces transformations produisent des résultats pour faire voir des choses, mais les résultats ne sont pas les choses. Trois formes principales de traitement sont utilisées en fonction des questions posées : constituer des grappes, établir des corrélations, et factoriser. Mais au préalable, le traitement quantitatif suppose la transformation numérique des données qualitatives.

4.2.1 La transformation numérique des données qualitatives

Avant d'effectuer les traitements quantitatifs, la première étape consiste à superposer un codage numérique aux codes catégoriels déjà utilisés. Pour ce faire, toutes les conventions de quantification sont utilisables (voir chapitre 4). Dans la pratique, deux techniques se rencontrent plus souvent.

4.2.1.1 Le codage binaire

On peut considérer que chaque trait ou chaque catégorie peut être présente ou absente, ce qui conduit à un codage binaire. Certains logiciels d'analyse en grappes (*clustering*), de classification hiérarchique et d'analyse factorielle (comme l'analyse des correspondances) travaillent très bien sur des codes binaires. Le seul inconvénient de ce codage réside dans la multiplication des variables. En effet, imaginons une variable qualitative, par exemple l'absentéisme, qui serait codée selon trois catégories, soit

rarement, moyennement, fréquemment. Le codage binaire transforme cette variable qualitative en trois variables binaires puisque l'on devra coder : absentéisme rare : 1 ou 0, absentéisme moyen : 1 ou 0, absentéisme fréquent : 1 ou 0. En binaire, cette variable prendra trois colonnes d'une matrice.

Le codage binaire a le grand avantage de respecter le caractère catégoriel du codage initial, mais cela se paie par un éclatement des variables. Cette atomisation aboutit parfois à des listes suffisamment longues pour saturer la mémoire de l'ordinateur avec lequel on se propose de faire les analyses. Un tel codage implique souvent le recours aux gros ordinateurs d'un centre de calcul et l'on doit se renseigner au préalable sur les logiciels qui y sont implantés et sur leurs exigences quant au format de saisie des données.

4.2.1.2 Le codage ordinal

Une autre manière de transformer des données catégorielles consiste à ranger les catégories ou les valeurs que peuvent prendre les réponses aux rubriques (aux questions). Il est des cas où cette opération est facile, comme dans l'exemple de l'absentéisme. En effet, certains codages qualitatifs correspondent à des attributions de valeurs qui peuvent facilement être ordonnées. Dans l'exemple de l'absentéisme, on peut attribuer le rang 1 à rare, le rang 2 à moyen et le rang 3 à fréquent, ou prendre le rangement dans l'autre sens et inscrire respectivement 3, 2 et 1. Mais il est des cas où la hiérarchisation des catégories n'est pas évidente. Par exemple, comment ordonner les étiquettes de maladies mentales comme la psychose maniaco-dépressive, la paranoïa, la schizophrénie et la maladie d'Alzheimer ? Il faut inventer un critère qui permet le rangement tout en étant pertinent pour les questions de la recherche. Dans l'exemple des maladies mentales, on peut imaginer plusieurs critères possibles : leur fréquence dans la population, leur résistance aux médicaments ou aux psychothérapies, l'invalidité qu'elles provoquent, leur dangerosité sociale, etc.

Le premier avantage du codage ordinal réside dans un gain de place : pour une variable qualitative, on garde une variable ordinale, une colonne dans la matrice. Cela permet donc de travailler avec un grand nombre de rubriques sans risque de saturer aussi vite la mémoire de l'ordinateur. Le second avantage est qu'il ouvre la porte à des calculs d'indices inexistants pour les codes binaires.

Ce codage pose cependant des problèmes, dont celui de modifier les caractéristiques catégorielles des données. Cela veut dire que cette

transformation dépend du critère de hiérarchisation choisi et que tout changement de critère aboutirait probablement à des classements différents. Autrement dit, la stabilité des résultats entre plusieurs recherches dépend de la concordance (ou du recouvrement) des critères choisis. Un autre problème réside dans le poids que ce classement attribue aux rangs extrêmes. Par exemple, si l'absentéisme se classe en 3 rangs et qu'une autre variable se classe en 10 rangs, et si ce sont les rangs 1 qui sont les plus fréquents pour l'absentéisme mais que ce sont les rangs s'étalant de 5 à 10 qui le sont pour l'autre variable, cette dernière variable va peser bien plus lourd dans les analyses que l'absentéisme. De plus, si le critère retenu produit une échelle et que, une fois les données codées, certaines positions sont inoccupées, les rangs non utilisés auront aussi une influence sur le résultat des traitements. Cela signifie donc que le rangement des catégories de réponse risque de conduire à des résultats artificiels sans rapport avec ce que les données de base présentent effectivement. Le traitement quantitatif des données qualitatives impose donc au chercheur de ne le considérer que comme un outil d'exploration qui peut être utile, mais qui n'est pas fidèle et qui doit toujours être vérifié par un retour aux données originales.

4.2.2 Le traitement conduisant à des classes : l'analyse en grappes

Constituer des grappes c'est faire des regroupements de sujets à partir de la présence d'indices semblables dans leur profil (familles sur portraits similaires), ou regrouper des indices à partir de leur présence chez des individus. Cet objectif n'exige pas nécessairement le recours aux transformations numériques, mais il est bien des cas où elles le rendent plus facile à atteindre.

Lorsqu'on n'a pas trop de sujets et pas trop de variables (ou de thèmes, d'indicateurs) par sujet, on peut noter le profil de chaque sujet sur des feuilles plus ou moins transparentes (papier calque, papier de soie), puis, en superposant les feuilles, on peut voir ceux qui se ressemblent et par quelles variables ils le font. On peut se faciliter la tâche en remplaçant les catégories conceptuelles par des codes symboliques ou par des codes couleurs, comme les géographes le font. On code les concepts avec des couleurs pour voir plus facilement apparaître des similitudes. Sur la base des similitudes observées ou calculées (des indices de similitudes ou, au contraire, de distances), on constitue des grappes, des regroupements. Ces regroupements peuvent donc se faire de manière artisanale, en comparant les profils, et en se fixant des critères de regroupement plus ou moins exigeants (par exemple, tous les traits partagés, ou une majorité de traits). Ils s'obtiennent aussi en

appliquant, sur les fichiers de données, des logiciels de classification automatique (*clusters analysis*) qui permettent, plus ou moins, de choisir les paramètres d'association et de hiérarchisation.

Lorsqu'on travaille sur des fichiers importants, le recours à l'ordinateur est indispensable. Dans ce cas, il est essentiel de revenir aux données pour y confronter les résultats, parce que les logiciels de classification qui exigent un codage numérique (binaire ou par transformation numérique) sont très puissants et qu'ils parviennent toujours à fournir des grappes. En effet, pour ces logiciels, les chiffres sont considérés comme des nombres, sans restriction, et ils en tirent toutes les relations possibles, même les plus infimes. De plus, il arrive qu'on leur donne à traiter des données qu'ils n'auraient pas dû recevoir, soit parce qu'elles ont entre elles des relations parasites qu'on n'avait pas vues au départ et que seuls les résultats aberrants de cette analyse font voir, soit parce que, par inattention, on y a mis des données non pertinentes ou impropres à ce genre de traitement. Parce que les grappes sur des chiffres se font facilement, même si elles n'ont aucune signification, la comparaison des résultats avec les données originales permet de juger si les grappes sont vraisemblables ou artificielles. Or, on ne peut pas se permettre les résultats qui ne peuvent pas s'interpréter, qui sont erronés ou qui ne sont que des fictions techniques.

4.2.3 Le traitement établissant des relations : les corrélations

Les réseaux indiquent des relations de structure hiérarchique ou de fonction. L'étude des événements est aussi intéressée à leur évolution, aux changements qui s'opèrent et à leur concomitance : savoir s'il n'y a pas des choses qui bougent en même temps ou qui bougent à la suite l'une de l'autre. La question est celle de la simultanéité et de la similitude des variations, problème abordé par l'observation d'indices de corrélations, de covariations, de dépendances, d'inter-dépendances. Du point de vue qualitatif, les indicateurs de variation semblable et simultanée seront surtout mis en évidence lors de la construction des matrices et de la comparaison des matrices, bien que quelques relations peuvent apparaître dès l'analyse. Mais, comme pour les classifications en grappes, l'observation artisanale de ces relations est souvent insuffisante : elle a tendance à voir comme évidentes des relations qui ne sont que faibles ou peu régulières (mais c'était celles que l'on attendait), alors qu'elle ne voit pas des relations intéressantes mais peu visibles à l'oeil nu. L'utilisation des calculs de corrélation et de régression, avec leurs coefficients et les diagrammes de dispersion spatiale, permet d'identifier des corrélations importantes bien que cachées par une

relation qui sautait aux yeux. Il est aussi important de rechercher les relations non linéaires, moins visibles mais déterminantes dans les organismes complexes. Comme pour la constitution des grappes, le recours aux logiciels quantitatifs apporte une aide précieuse, à condition de respecter les mêmes règles de prudence dans l'interprétation. En outre, on sera attentif à ne pas interpréter des relations de variations simultanées ou consécutives comme des relations causales. Les deux variables corrélées, celles qui bougent en même temps ou à la suite l'une de l'autre, pourraient ne pas être reliées entre elles mais dépendre d'une troisième variable.

4.2.4 La factorisation et la dispersion dans l'espace des axes principaux

À un niveau plus sophistiqué et recourant nécessairement à des codages chiffrés (au moins des fréquences transformées en proportions), on peut factoriser les données, c'est-à-dire chercher quels pourraient être les facteurs qui, en un nombre réduit, organisent l'ensemble des données et permettent de les représenter. Différents types d'analyses sont maintenant disponibles sur des données qualitatives : les méthodes d'analyse factorielle, d'analyse discriminante, d'analyse en composantes principales et les méthodes d'analyse de correspondance qui permettent de disposer les données dans des espaces géométriques (voir dans le tableau 19.2, les exemples de nuage de points et de circomplex). Certaines de ces méthodes présupposent un mode d'organisation préalable des données (les analyses factorielles classiques), d'autres n'ont pas de tels présupposés; il faudra en tenir compte dans le choix des méthodes.

Mais ici aussi, les facteurs ou les vecteurs auxquels ces analyses conduisent par des artifices mathématiques, doivent être interprétés en retournant aux données pour vérifier l'organisation qu'ils suggèrent. Ces factorisations ne comportent, par elles-mêmes, aucune signification. Elles suggèrent que certaines données ont une relation quantitative avec un facteur X ou Y qui leur est commun, que ces facteurs semblent avoir un effet plus prépondérant sur les variations de certaines variables ou sur la position relative de ces variables les unes par rapport aux autres. La considération des facteurs X et Y, ceux qui ont le plus de poids dans «l'explication» des relations que l'ensemble des variables ont entre elles, permet des regroupements de variables ainsi que la construction de diagrammes de dispersion spatiale. Mais l'identification de ces facteurs et la dénomination des axes utilisés pour construire les diagrammes, reste problématique : ces facteurs sont des abstractions mathématiques sur des transformations de données. C'est l'analyse du regroupement des données originales et de leurs

positions relatives dans les diagrammes, qui fournira les significations en fonction du cadre conceptuel. Malgré la difficulté d'identifier les facteurs et les axes, les possibilités qu'ont ces méthodes d'analyse, grâce à l'informatique, de produire des représentations spatiales, constituent une aide précieuse à l'interprétation des résultats.

Un problème de collaboration se pose alors : dans la majorité des cas, le chercheur ne dispose pas des compétences techniques pour effectuer ces factorisations, alors que le statisticien, qui peut les réaliser, n'a pas les connaissances du champ conceptuel pour les interpréter. De notre point de vue, le chercheur qui a conçu le cadre conceptuel, doit rester le timonier dans cette phase; il passe une commande au spécialiste du traitement informatique qui devrait choisir le modèle de traitement le plus approprié à la demande, sans forcer le chercheur à utiliser tel ou tel procédé, peut-être très esthétique, très puissant, mais ne fournissant pas de résultats interprétables. Par contre, le chercheur doit pouvoir accepter que le spécialiste ne peut pas toujours lui offrir un modèle adéquat, qui effectuerait le traitement rêvé, sans modification de la formulation des questions ou sans transformation des données. Dans ce cas, on se trouve devant un problème d'éthique de la recherche : le chercheur est le seul à pouvoir estimer le risque à courir dans une transformation s'il maintient sa demande, alors que le spécialiste est souvent le seul à pouvoir annoncer (ou cacher) la transformation que le modèle proposé implique. Le spécialiste du traitement ne peut cacher au chercheur l'écart entre ses possibilités de réponse et la demande qui lui a été formulée, alors que le chercheur ne peut faire porter au spécialiste la responsabilité du traitement qu'il a accepté qu'on applique à ses données.

4.3 Remarques finales sur le traitement

1. Tous les traitements envisagés devraient être présentés au lecteur de manière telle qu'ils soient lisibles et contrôlables. Cela signifie que le chercheur doit décrire la démarche suivie, sinon dans le détail des procédures, ce qui est parfois long, au moins dans le principe des procédures. Cela veut aussi dire que le chercheur a tout intérêt à présenter ses résultats sous forme de représentations graphiques. De ce point de vue, les suggestions apportées lors de l'analyse des données peuvent toutes être reprises et appliquées à la présentation des résultats.

2. Contrairement à ce que certains pensent, on peut faire maintenant des analyses sophistiquées sur des données qualitatives qui se rapprochent des traitements quantitatifs classiques. Il est vrai que, technologiquement,

les méthodes sur le matériel qualitatif ont été développées après celles sur le quantitatif; par exemple, les analyses de correspondance sont plus récentes que les analyses factorielles classiques. Les centres de calcul et les informaticiens ont moins développé de logiciels permettant de traiter le matériel qualitatif. Cependant, certains logiciels permettent de faire des transformations qui respectent suffisamment le matériel qualitatif, pour effectuer des calculs numériques, comme des résumés et des coefficients, ou produire des classifications et des analyses en facteurs. On dispose maintenant de techniques qui sont suffisamment intéressantes et abordables pour ne plus avoir l'excuse d'être obligé de se limiter à faire du mauvais journalisme parce que l'on travaille avec des méthodes, des démarches et des matériaux qualitatifs.

L'INTERPRÉTATION DES RÉSULTATS ET SA VALIDATION

- 1 Les règles de l'interprétation**
 - 1.1 Définition**
 - 1.2 Les deux moments de l'interprétation dans la recherche**
 - 1.3 Les quatre analogies et leurs leçons**
 - 1.3.1 La fonction d'interface
 - 1.3.2 La fonction de traduction
 - 1.3.3 L'interprétation comme mise en scène
 - 1.3.4 La fonction de divination
 - 1.4 Les balises de l'interprétation**
- 2 Les techniques d'interprétation des résultats**
- 3 La validation des interprétations**
 - 3.1 La vérification des comptes (la rigueur du processus)**
 - 3.2 L'évaluation de l'effet des biais (les interférences)**
 - 3.3 Le contrôle des hypothèses**
 - 3.3.1 Le contrôle sur le matériel résidu
 - 3.3.2 Le contrôle par d'autres chercheurs
 - 3.3.3 Le contrôle par les acteurs
- 4 Notes sur la rédaction du rapport**

1 Les règles de l'interprétation

Indépendamment des discussions épistémologiques sur le sens de l'interprétation, la méthodologie de la recherche doit examiner les problèmes de la pratique de l'interprétation et de ses règles. Pour ce faire, les définitions du sens commun et les définitions techniques du mot seront d'abord examinées. Ensuite, on envisagera sa place dans le cours de la démarche de recherche, puis quatre analogies à partir desquelles on infère des balises pour la pratique de l'interprétation.

1.1 Définition

L'interprétation a, selon le dictionnaire *Le Petit Robert*¹, au moins trois sens courants.

- a. D'abord *l'explication*, soit l'action d'expliquer, de donner une signification claire à une chose obscure; elle explique un résultat par le dévoilement des raisons, des motifs plus ou moins évidents; elle attribue un sens caché à des événements.
- b. Ensuite, l'interprétation est aussi le processus par lequel s'effectue la *mancie* qui donne une signification aux faits ou aux actes de quelqu'un. La *mancie* renvoie aux sciences divinatoires et à ceux qui les pratiquent, comme la cartomancie, la chiromancie, etc.
- c. Enfin, on parle aussi d'interprétation pour désigner l'*exécution*, soit la manière dont une oeuvre (musicale, poétique, théâtrale) est interprétée; dans ce cas, il s'agit de la mise en vie, perceptible par les sens, d'une trace imprimée et cela par l'intermédiaire d'un interprète.

Les définitions a et b, soit l'explication et la *mancie*, peuvent se rapprocher dans une seule dimension : le dévoilement du caché, de ce qui n'est pas évident, ou de ce qui serait même inconscient. Il s'agit là, aux yeux du profane, d'une sorte de divination, même si des règles sont suivies, puisqu'il n'a pas accès à ce qui est dévoilé sans l'intermédiaire de l'interprète. Nous retiendrons donc pour notre analyse analogique une seule dimension pour ces deux définitions : la *divination*. La définition c renvoie à l'artiste qui joue un morceau de musique, une pièce de théâtre, un scénario; pour les fins de notre analyse nous retiendrons l'image de la *mise en scène*.

¹ Robert P., *Le Petit Robert 1*. Paris : Dictionnaires Le Robert, 1985, édition revue 1986, p. 1023.

Les dictionnaires techniques² nous livrent au premier abord deux sens à l'interprétation.

- d. L'interprétation est d'abord l'*interface*, soit la connexion entre des éléments d'un système. Dans le même domaine, elle est aussi, soit la limite commune entre deux éléments (leur lieu de jonction), soit le dispositif qui permet à un système de communiquer avec un autre, soit l'ensemble des règles et des conventions qui permettent cette communication.
- e. L'interprétation est aussi la *traduction*. La traduction s'effectue par l'intermédiaire d'un interprète qui transpose dans la langue d'un auditeur le message d'un émetteur ne partageant pas le même univers linguistique. C'est donc produire dans une langue un énoncé qui est équivalent à un énoncé formulé dans une autre langue.

1.2 Les deux moments de l'interprétation dans la recherche

Que la recherche soit dénommée quantitative ou qualitative, l'interprétation s'y glisse en deux moments. L'interprétation intervient d'abord *avant le traitement des données*, lorsqu'il faut systématiser les réponses des sujets en un format qui permet de résumer et d'organiser leur production pour y voir clair. En d'autres mots, une première interprétation prend place lors du codage des données, avant le traitement, lorsqu'il s'agit de *formater* des réponses particulières, individuelles, dans un même système commun à l'ensemble des sujets. On peut dire qu'à ce moment l'interprétation consiste en une réduction des particularités à un cadre général et manipulable. Tout système inscripteur effectuée, dans ce sens, une interprétation du matériel original.

Dans le cas des données *invoquées* et *suscitées*, c'est évidemment au chercheur d'effectuer cette interprétation qu'il réalisera en fonction des questions de sa recherche et du cadre conceptuel qu'il a construit pour l'entreprendre. Dans le cas des données *provoquées*, cette interprétation n'est pas moins présente, mais le chercheur la fait effectuer en partie par l'instrumentation, en partie par ses sujets. En effet, une fois l'instrument de collecte des données construit, ce sont les sujets de la recherche qui doivent «comprendre» (interpréter ?) les consignes, les questions formulées par le chercheur, puis qui doivent réduire eux-mêmes les particularités de leur expérience pour la faire coïncider avec les catégories de réponses qui leur

2 Luca J. de, *Dictionnaire de la micro-informatique*. Montréal-Paris : Modulo- Belin, 1984, p. 112-113.

sont imposées. On voit, dès lors, que la question de l'erreur d'interprétation lors de la constitution des données quantitatives est complexe, et que les processus en jeu (comprendre la consigne, d'une part, et formater son expérience, d'autre part) échappent au contrôle et à l'analyse du chercheur. Les mesures de fidélité et de validité n'en donnent qu'un reflet difficile à saisir. De ce point de vue, on peut se demander si une cueillette de données qualitatives accompagnée d'un codage bien documenté, et il peut l'être lorsque effectué par le chercheur lui-même, n'aboutit pas à des données plus pertinentes et plus vraisemblables.

Une deuxième phase interprétative se situe lors de ce que l'on appelle typiquement «l'interprétation» des résultats, soit *après le traitement*, lorsque l'on veut formuler des hypothèses à partir du matériel traité ou lorsque l'on veut décider de l'infirmer ou non des hypothèses déduites des théories. Lors de la première phase, le codage, il s'agit de faire rentrer dans un même format des données exprimées avec plus ou moins de liberté et avec plus ou moins d'idiosyncrasie. Lors de la seconde phase, il s'agit plutôt de sortir du format utilisé dans le traitement pour le déborder et trouver ou retrouver un sens. L'interprétation est donc d'abord réductrice, avant le traitement, pour être, après le traitement, créatrice par l'élaboration d'explications et de nouvelles questions qui transcendent la sécheresse des résultats.

En lisant les recherches publiées, on se pose souvent la question des règles qu'ont pu suivre les chercheurs pour réduire et pour créer. Car ces transformations du matériel, avant et après le traitement des données, sont des opérations dont la richesse de la recherche va dépendre. Le traitement des données est relativement formalisé : que l'on travaille avec des données qualitatives ou quantitatives, après une première interprétation ou une analyse qui permet de catégoriser, il s'agit toujours de condenser ou de résumer, puis d'organiser, de structurer ou de factoriser pour enfin afficher, c'est-à-dire présenter graphiquement aux lecteurs les relations ou les structures ainsi établies. En bref, le traitement consiste surtout en condensations et représentations, opérations qui ne peuvent s'effectuer que selon certains patrons relativement bien codifiés et limités par les capacités des machines qu'on utilise pour le faire.

Or, les opérations de transformation avant et après le traitement ne reçoivent qu'une attention relative des chercheurs, et cela même dans la tradition quantitative³. Pourtant, le traitement des données, si sophistiqué

3 Pourtois J.P. et H. Desmet., *Instrumentation et épistémologie en sciences humaines*. Bruxelles : Mardaga, 1988. Ils y rapportent une analyse de trois revues bien cotées en sciences de l'éducation, dont une américaine. Les résultats sont étonnants .../...

soit-il, ne peut pas leur accorder une qualité qu'elles n'auraient pas acquise au moment de leur constitution.

1.3 Les quatre analogies et leurs leçons

Pour éclairer le problème des règles qui gouvernent l'interprétation et qui témoigneraient de la crédibilité de l'opération effectuée, on peut suivre les quatre analogies qui recourent à leur manière les significations usuelles rappelées au début de ce texte : comment s'effectue l'interprétation à la lumière des fonctions d'interface, de traduction, de mise en scène et de divination ?

1.3.1 La fonction d'interface

Le langage technique définit l'interface comme la limite commune à deux ensembles (ou appareils), et comme la jonction entre deux éléments d'un système (informatique, par exemple). En fait, il s'agit plus souvent du dispositif permettant la connexion entre deux équipements exécutant des fonctions différentes mais enchaînées, et des règles et des conventions qu'il faut respecter pour que ces équipements puissent communiquer. L'interface est omniprésente en micro-informatique. Ce sont les connexions physiques reliant, par exemple, le clavier, le processeur, l'écran et l'imprimante, et les conventions (logiciels systèmes) fonctionnant de manière telle que les symboles inscrits sur les touches du clavier apparaissent à l'écran et que l'image figurant sur l'écran peut être imprimée sur papier. L'interface fait en sorte que le transfert de l'information d'un support à un autre est tel que cette information peut être reconnue quel qu'en soit le support. En tenant compte des modifications imposées par le support ou demandées par l'utilisateur, le principe de ces modifications doit donc être constant. L'information présentée par un équipement devra correspondre à celle présentée par l'autre, quelle que soit la forme utilisée par l'interface pendant le transfert, et quelles que soient les particularités de l'affichage dues à chacun des équipements. L'interface *transfère* et *transforme*, mais en maintenant une *correspondance terme à terme* entre chacune des présentations. Si elle ne maintient pas cette correspondance, on diagnostique un pépin, on se plaint d'un bogue.

.../... à leurs yeux : selon l'analyse des revues, effectuée pour les années 1972-74 et 1982-84, seulement 20,4 % des articles fournissent des données quant à la fidélité des mesures ! Voir aussi Pourtois J-P. et H. Desmet, La recherche en éducation : épistémologie et instrumentation. *Les sciences de l'éducation pour l'ère nouvelle*, 1-2, 1986, p. 19-23.

1.3.2 La fonction de traduction

La traduction consiste en la transposition dans une langue de l'information d'un message exprimée dans une autre langue. Dans la traduction, on rencontre rarement la correspondance terme à terme des éléments du lexique, de la syntaxe, des expressions d'une langue dans l'autre. Ainsi, les formes de genre masculin et féminin ne sont pas toujours les mêmes pour un objet identique («lune», féminin en français, masculin en allemand) et l'on rencontre des langues où il existe un genre neutre, comme en néerlandais ; il en va de même pour les modalités de conjugaison et de déclinaison. Il est aussi des expressions directement intraduisibles d'une langue à l'autre, d'une culture à l'autre. L'on connaît les difficultés qu'il existe à vouloir traduire de l'humour. Bref, la traduction est confrontée avec l'impossibilité d'une correspondance terme à terme, mais elle cherche à maintenir le *parallélisme des significations*.

Cette exigence, le parallélisme des significations, devient évidente lorsqu'on observe les défauts de la traduction. On dira qu'une traduction est une trahison lorsque l'interprète ajoute des significations à celles du message original, lorsqu'il fait glisser le sens du message et ainsi le déforme, lorsqu'il ajoute ou insère, dans la traduction du message, des commentaires qui annulent le message ou des parties du message (font passer du sens au non-sens), ou qui en orientent la compréhension. Autrement dit, la traduction doit rester neutre, objective, fiable : les préjugés, les conceptions, les valeurs du traducteur ne doivent pas interférer avec son travail de traduction. La version traduite doit livrer au lecteur le même message (avec les mêmes valeurs) que celui de la version originale composée par l'auteur. Lorsque l'absence de correspondance terme à terme des lexiques ou des expressions rend la traduction approximative, l'interprète doit signaler aux lecteurs sa difficulté et fournir les diverses versions approchées du message original.

1.3.3 L'interprétation comme mise en scène

En tant que mise en scène, l'interprétation est une *transposition analogique* qui présente en un ensemble, simultanément et en interaction, des éléments présentées ailleurs de manière linéaire, les uns à la suite des autres. Elle est aussi une transposition *dynamique* qui incorpore le mouvement, l'espace et le temps. La mise en scène ne peut pas se réaliser comme une interface, sinon elle paraît insignifiante et sans relief. Elle ne peut pas se limiter à n'être qu'une traduction, car elle resterait alors insaisissable : il est des messages qui ne peuvent se transposer intelligemment que si, tout en respectant leur signification initiale, on indique aussi au récepteur

le ou les sens qu'ils peuvent prendre, ainsi que le sens que leur a attribué l'interprète pour effectuer sa tâche.

Au contraire de la traduction, la mise en scène introduit le commentaire, elle insère l'émotion de l'interprète dans la transmission du message. Trois niveaux de sujets s'y trouvent et communiquent : l'auteur parle à l'auditeur par interprète interposé. Le rôle de l'interprète devient essentiel lorsque l'auteur produit un document qui ne donne que des indications sommaires quant à la dynamique du message. Alors, ces indications doivent être, d'une part, situées en rapport au contexte d'écriture de l'auteur et, d'autre part, élaborées en fonction des conditions de réception de l'auditeur. L'interprète, dans sa lecture du texte de l'auteur, va retrouver l'émotion que l'auteur n'a pu qu'esquisser à cause de linéarité de l'écriture. Ensuite, par sa mise en scène, l'interprète va recréer, pour l'auditeur, cette émotion retrouvée. L'interprète ajoute donc au texte l'émotion ressentie, il indique au spectateur (auditeur) ce qui a pu mouvoir l'auteur dans le texte : l'interprétation suggère donc le motif absent ou implicite du texte.

La mise en scène doit respecter une règle particulière : celle du *parallélisme intersubjectif*. Cette règle veut que le rapport entre l'expression (face externe) et l'émotion ou la pensée (face interne) suive une convention de transposition parallèle chez chacun des partenaires.

Auteur	Interprète	Auditeur
<u>Face externe</u>	<u>Expression</u>	<u>Signifiant</u>
Face interne	Pensée	Signifié

Tableau 20.1 : Le parallélisme intersubjectif dans la mise en scène.

Cela s'observe bien dans la mise en scène théâtrale (tableau 20.1) : il ne peut y avoir de communication entre la pensée d'un auteur, situé dans un contexte donné, et celle de l'auditeur, localisé ailleurs, que si le rapport entre signifiant (face externe) et signifié (face interne) reste semblable alors que les moyens (les mots) et le contexte de ce rapport ont changé. Autrement dit, le parallélisme intersubjectif exige que l'expression, par l'interprète, des émotions attachées au contexte de l'auteur soit ajustée au contexte émotionnel de l'auditeur. Cela présuppose, dit la littérature phénoménologique⁴, une rencontre intentionnelle impliquant une valorisation réciproque des partenaires.

4 Thinès G. et A. Lempereur, *Dictionnaire des sciences humaines*. Paris : Éditions Universitaires, 1975, p. 513-514.

Le non-respect de cette règle se manifeste dans les erreurs de la mise en scène. Parmi ces erreurs, on trouve le jeu sans émotion d'une interprétation «technique», qui se rapproche de l'interface lorsque l'auteur est contemporain, ou de la traduction quand le texte nous vient du passé. Une autre consiste à mettre en scène un excès d'émoi au détriment du texte, car le commentaire et l'expression de l'émotion ne sont qu'un accompagnement du texte, et ils ne doivent pas le cacher. Leur rôle est de mettre le texte en valeur, d'en élucider le sens, et non pas de produire un excès de sens dans lequel l'auditeur perdrait le sens du texte. Enfin, les émois personnels de l'interprète ne peuvent pas rompre leur connexion à l'univers émotionnel de l'auditeur (sinon, la salle se vide avant l'entracte), ni à celui de l'auteur (il s'agit alors d'une trahison).

1.3.4 La fonction de divination

En tant que divination, l'interprétation est *dévoilement du caché*. Elle implique trois conditions. D'abord, que deux lectures soient possibles d'un même message : une première, immédiate, par l'auditeur et une seconde, médiatisée, par l'interprète. Ensuite, que le caché soit latent, c'est-à-dire qu'il soit accessible à la perception, à la conscience, même si cet accès n'est provisoirement pas immédiat. Enfin, qu'une connexion soit possible entre les deux niveaux de lecture, entre le déjà vu (passé ou présent) et le voilé que l'on va découvrir.

Pour que l'auditeur lui accorde quelque crédit, la divination doit respecter la règle de la *superposition des plans* : celui du caché qui sera dévoilé et celui du déjà vu, même si un décalage temporel, spatial ou sémantique existe. L'auditeur attribue un plan, une organisation ou une structure à ce qu'il voit ou a déjà vu, c'est-à-dire aux faits, aux gestes, aux événements. Il attend de l'interprète qu'il lui en dévoile les motifs, les raisons cachées. Si l'interprète d'une divination ne peut faire correspondre le plan des motifs ou des raisons (le caché - dévoilé) avec le plan de ce que l'auditeur a vu, l'interprétation sera reçue comme une mystification. C'est la superposition des plans, perçue par l'auditeur, qui confère à l'interprétation l'évidence attendue et exigée quant à la signification du message.

La mystification par défaut de superposition des plans s'illustre par le sentiment d'avoir été berné chez le lecteur de certains romans policiers. En effet, le défaut de superposition des plans s'observe lorsque, vers la fin d'un roman policier, l'auteur modifie le plan du déjà vu en introduisant des faits nouveaux, inconnus jusque-là du lecteur, au moment même où il dévoile les motifs du crime. C'est pour rendre plausibles les motifs du crime que le

romancier doit ajuster la structure des faits connus à la structure des motifs qu'il dévoile au lecteur.

1.4 *Les balises de l'interprétation*

En confrontant les phases de la recherche et les fonctions de l'interprétation telle qu'éclairée par les analogies que l'on vient d'envisager, les fonctions de l'interprétation se répartissent de part et d'autre du traitement des données, et cela, qu'il s'agisse de données quantitatives ou qualitatives. Ensuite on peut en déduire quelques principes applicables à ces phases.

Avant le traitement, l'interprétation des données consiste essentiellement en un codage. L'interprétation y fonctionnera à la manière d'une *interface* ou d'une *traduction* selon les caractéristiques des données et selon le traitement que l'on souhaite leur appliquer. Mais dans la mesure où le codage avant le traitement ne peut rien ajouter à l'information déjà contenue dans les données, l'interprétation ne pourra jamais y fonctionner à la manière de la mise en scène ou de la divination.

Après le traitement, l'interprétation devra relier les résultats (les données traitées, condensées, organisées, représentées) aux questions et au cadre conceptuel de la recherche. Alors, il est important que l'interprétation fonctionne comme une *mise en scène* (commentaire et mise en action) et comme une *divination* (dévoilement par la dénomination des représentations). Mais dans cette phase, elle ne pourra fonctionner ni comme une interface, ni comme une traduction qui ne fourniraient qu'une lecture banale et redondante des résultats.

De ce qui précède, on peut inférer des *guides* pour la pratique de l'interprétation.

Dans la première phase, *avant* le traitement, quatre principes sont importants :

- 1° L'exigence d'un lexique et d'une documentation précise des conventions et des règles utilisées dans le codage, lorsque des réponses des sujets sont transposées dans le système d'étiquettes utilisé pour les réduire et les analyser;
- 2° Le respect d'une constance de base et d'une cohérence dans l'application des règles de transposition (de codage);
- 3° Le maintien de la correspondance dans l'orientation des plans avant et après le codage, c'est-à-dire le maintien des caractéristiques des données indépendamment de la réduction et du système d'étiquettes qu'on a utilisé pour ce faire;

- 4° Le non-débordement des formats originaux : ne pas ajouter ou attribuer aux codes des valeurs ou des pouvoirs qui ne sont pas présents dans les données.

Les règles 3 et 4 protègent de la prophétie qui se réalise avant coup : elles interdisent au chercheur de projeter ce qu'il anticipe du «dévoilé» dans l'opération même du codage, car cette projection dans le codage forcerait le traitement à ne pouvoir mettre en évidence que ce que le chercheur a anticipé.

Dans la seconde phase, *après* le traitement, deux principes sont essentiels :

- 1° Respecter le parallélisme intersubjectif qui impose au chercheur d'ajuster l'interprétation des résultats à tout ce que permettent et à tout ce qu'impliquent les données en tant que productions de sujets ayant répondu au chercheur. Le chercheur, dans cette intersubjectivité, doit respecter les contraintes et les exigences contenues dans les résultats sans y projeter de désir ni détruire ce qui le générerait, que ce désir soit le sien ou celui de ses commanditaires;
- 2° Maintenir la superposition des plans, soit la correspondance des structures apparentes et dévoilées, en s'interdisant d'introduire un écart ou de forcer l'écart favorisant l'interprète et son pouvoir au détriment de l'évidence des données ou de l'intelligence de ses informateurs et de ses lecteurs.

Ces deux règles de l'interprétation visent à protéger le public, soit les informateurs et les lecteurs, des mystifications ou des délires possibles de la part du chercheur, que ces manipulations soient conscientes ou non, calculées ou non.

Les six principes ainsi dégagés sont des *balises minimales* dans une méthodologie de l'interprétation. Ils correspondent assez bien avec différentes stratégies d'élaboration et de validation des hypothèses telles que proposées par Miles et Huberman⁵ ou Goetz et Lecompte⁶, en ce qui concerne la recherche qualitative. On pense, par exemple, au contrôle interne et externe des comptes, à la triangulation, à l'examen des cas extrêmes et à la rétroaction par les informateurs. En ce qui concerne la tradition quantitative, elle documente assez bien, sans pour autant nécessairement

5 Miles M. B. et A.M. Huberman, *Qualitative Data Analysis, A Sourcebook of New Methods*, Beverly Hills, CA : Sage, 1984.

6 Goetz J. P. et M. D. Lecompte, *Ethnography and Qualitative Design in Educational Research*, Orlando, FL : Academic Press, 1984.

les respecter, les principes se rapportant à ce qui se passe avant le traitement; on les y trouve traités sous les titres de la fidélité et de la validité des observations et des plans de recherche. Cette tradition semble cependant plus pauvre en ce qui a trait à l'interprétation après le traitement, et la prudence y conduit bien des jeunes chercheurs à la sécheresse ou à la simple redondance. Ces six principes éclairent les exigences de l'interprétation et pourraient permettre aux chercheurs plus de souplesse et d'ingéniosité dans le choix de leurs stratégies, que des listes de techniques dont le respect paraît parfois plus rituel que fondé.

2 Les techniques d'interprétation des résultats

Les données ont été constituées, analysées et traitées, ce qui a produit des résultats. Ces résultats (et non pas les données) doivent maintenant être interprétés. Faire l'interprétation des résultats, c'est élaborer des hypothèses, soit énoncer des relations plausibles, probables, vraisemblables et pertinentes qui rendent compte des phénomènes interrogés, qui nous permettent de les comprendre pour les expliquer aux lecteurs.

Comment induire des hypothèses ? On commence par travailler les représentations graphiques des résultats jusqu'au moment où ce qu'elles font voir, permet de les commenter. Un graphique ou une matrice sans commentaire ne vaut pas grand-chose. Le lecteur ne sait pas l'interpréter sans que l'auteur de la recherche ne lui donne des indications. Interpréter consiste donc d'abord à commenter les graphiques et les matrices : on verbalise, on conceptualise les représentations. Analogiquement, l'interprétation est comme un récit qui relie entre eux des personnages-concepts, qui raconte leurs interrelations et leur évolution : comment cela marche, comment cela fonctionne. Ce récit est aussi une mise en scène et un dévoilement de signification : il s'agit de montrer la dynamique des relations, du fonctionnement, et de mettre en évidence les jeux de rapports qui sont parfois implicites aux données.

Une fois que les graphiques ont été commentés, on cherche quelle analogie, ou quelle partie de théorie, on peut superposer aux résultats, afin d'en rendre compte d'une manière synthétique. Quelle formule, quelle image, quelle relation entre concepts, quel principe permettra d'exprimer le commentaire en quelques mots ? N'en existe-t-il pas, déjà formulé par ailleurs, qui pourrait fournir une part de l'explication recherchée ? On doit souvent examiner l'apport possible de concepts et de théories autres que ceux qui nous ont servi pour formuler le cadre conceptuel. En effet, celui-ci est nécessairement limité : il sert à fournir les outils pour obtenir des

données, son rôle n'est pas de les interpréter. Dès lors, si on ne peut pas construire d'interprétations intéressantes des résultats à partir des théories ou des modèles qui ont servi au cadre conceptuel, il faut aller voir ailleurs. Des analogies avec d'autres champs théoriques sont parfois très riches. Les physiciens n'ont pas rejeté le modèle de l'atome parce qu'il était issu d'une analogie avec le modèle de la rotation des planètes autour du soleil. Cette analogie extérieure à la théorie a été acceptée parce qu'elle suggérait des réponses aux problèmes posés. En éducation, on se trouve devant la même situation : toute analogie qui apporte des solutions possibles aux questions posées dans la délimitation du problème doit être considérée. Cependant, il faut examiner la pertinence des solutions fournies par une analogie avant de la rejeter. Le recours aux analogies, aux théories extérieures à la formulation du problème, implique évidemment que le chercheur ait une culture assez large. L'histoire des sciences humaines montre que cette culture se doit non seulement d'être scientifique, mais quelle gagne aussi à être littéraire et artistique. Pour pouvoir emprunter une théorie, l'analyste doit connaître comment se présentent les faits, les données ou les résultats qui vérifient cette théorie ou que cette théorie explique : quelles représentations ont permis la formulation de quels commentaires ? Il y a donc un mouvement de va-et-vient entre l'interprétation, le cadre conceptuel, le traitement des données, l'interprétation des résultats et la culture de l'analyste.

L'élaboration d'hypothèses ne peut se faire en évitant la confrontation avec le cadre conceptuel. Si on a pu obtenir des résultats, c'est parce qu'on a posé des questions à partir d'un cadre conceptuel. La formulation des interprétations ne peut pas se faire sans rapport avec la manière dont le problème a été posé dans le cadre conceptuel. La validité de l'interprétation des résultats, ou celle des données qui ont produit les résultats, sera mise en cause si elle présente un trop grand écart avec la formulation du problème : les règles du parallélisme intersubjectif et de la correspondance des plans servent de critère. Il faut donc retourner au cadre conceptuel et vérifier en quoi et comment il trouve des réponses dans les hypothèses. Si les réponses ne sont pas pertinentes, il faudra reprendre l'analyse des données : examiner s'il n'y a pas autre chose à y voir, si d'autres traitements ne pourraient pas suggérer d'autres structures. Dans certains cas, il suffit d'examiner les données d'un autre point de vue : une matrice se lit dans les deux sens et on oublie souvent de le faire.

Miles et Huberman suggèrent quelques tactiques pour induire des hypothèses. En synthétisant leurs recommandations, on retient quelques suggestions :

- Relever les formes (gestalt), les patterns, les structures, les thèmes et les images qui se répètent;
- Sentir ce qui est plausible, ce qui a de l'allure, du bon sens... tout en ne rejetant pas trop vite les résultats qui ne correspondent pas avec l'intuition, avec les autres données. Les résultats qui ne marchent pas avec l'intuition exigent plutôt que l'on corrige l'intuition !
- Regrouper, composer des ensembles en établissant tous les liens possibles entre individus, processus, etc.;
- Imaginer des métaphores, des analogies, des symbolisations qui expriment en bref la complexité des choses;
- Diviser, analyser les variables (sans aller jusqu'à l'explosion atomique...);
- Prendre du recul, tracer les frontières, les bornes, les grands ensembles (regarder la forêt plutôt que les arbres);
- Induire, généraliser, abstraire : chercher et noter les traits communs aux personnes, aux situations, aux événements pour composer des classes et examiner les relations entre ces classes et les individus composant d'autres classes, plutôt qu'entre individus seulement;
- Essayer de dénommer les facteurs communs, les proximités et les relations entre variables;
- Identifier les variables intermédiaires qui relient deux autres variables covariantes;
- Chercher les variations systématiques autres que linéaires, identifier les points de rupture dans l'orientation des courbes exprimant les variations;
- Construire l'enchaînement logique ou sociopsychologique des évidences : montrer comment X conduit à Y puis à Z;
- Intégrer ce qui précède dans des constructions cohérentes, fonctionnelles, qu'elles soient théoriques ou narratives.

On oublie aussi trop souvent que, dans la plupart des cas, plusieurs hypothèses concurrentes sont possibles : on se précipite à accepter pour évidente l'hypothèse qui nous plaît le plus. Or, les raisons de cette évidence sont souvent personnelles et conduisent à sauter l'étape de la validation, du moins, celle d'une validation non biaisée. Les méthodologies quantitatives (données et traitements quantitatifs) n'échappent pas à ce problème commun à toute recherche exploratoire. Quel que soit le matériel recueilli, il faut faire l'effort de chercher toutes les interprétations possibles, de formuler plusieurs hypothèses, de procéder à la validation de toutes celles-ci et non pas d'une seule d'entre elles.

3 La validation des interprétations

La validation des hypothèses consiste à mettre en place une procédure comparative de la validité et de la pertinence de chacune des hypothèses pour retenir l'hypothèse ou les quelques hypothèses qui fournissent l'explication la plus consistante des résultats. Le procédé, trop souvent rencontré, qui consiste à tester une seule hypothèse contre son ombre, n'est qu'une caricature de la vérification expérimentale. Lorsqu'on ne retient qu'une seule hypothèse, on est tenté de construire un test qui vise sa confirmation et, si elle ne résiste pas au test, on essaie de la justifier *a posteriori* par des explications *ad hoc*. Or, il faut reconnaître que les explications *a posteriori*, dites «*ad hoc post hoc*», ne respectent pas l'exigence d'une phase de validation et violent souvent la règle de la clôture sémantique. L'interprétation des résultats exige l'élaboration de plusieurs hypothèses concurrentes et de procéder à une phase de validation. Cette validation va prendre plusieurs formes s'adressant aux différentes étapes du processus de recherche.

3.1 La vérification des comptes (la rigueur du processus)

L'analyse de la documentation décrivant le déroulement de la recherche garantira sa crédibilité. Elle implique qu'une chronique, quasi quotidienne, soit tenue (un journal de bord ou journal de recherche), couvrant la succession des étapes allant de la collecte des données, y compris les modifications apportées au cadre conceptuel, jusqu'à l'interprétation. Le récit, ou compte rendu, de ce qui a été fait tout au long de la recherche permettra de vérifier la fidélité du codage, la constance dans l'application des règles d'analyse, de traitement, d'interprétation. Si, comme on est censé le faire, des principes ont été suivis lors des analyses et des interprétations, il faut documenter leur application afin que le chercheur et ses contrôleurs puissent évaluer dans quelle mesure le programme de recherche a été respecté dans toutes ses procédures. Le compte rendu de chacune des phases de la recherche doit être suffisant pour qu'on puisse vérifier la fidélité et la constance dans l'application des principes ou, autrement dit, la standardisation des règles d'analyse, de traitement et d'interprétation. En fait, il faut qu'un autre chercheur puisse voir si on a bien appliqué les règles qu'on s'était données. On observe ici un des avantages du travail en équipe : chacun peut contrôler l'autre. C'est plus simple que de devoir engager un expert. Cette lecture critique d'une chronique bien tenue de la recherche est appelée *vérification interne des comptes* (ou comptes rendus) lorsque

ce sont des membres internes à l'équipe de recherche qui l'effectuent, ou *vérification externe des comptes*, lorsque ce sont des experts extérieurs à l'équipe.

3.2 L'évaluation de l'effet des biais (les interférences)

L'estimation des effets dus aux conceptions et aux choix du chercheur s'énonce souvent à la fin du rapport, mais on doit l'effectuer lors de la vérification des comptes. Puisque les conceptions et choix du chercheur conditionnent la qualité des données et des résultats, il faut prendre le temps d'évaluer leur impact sur la sélection du matériel, sur le choix du matériel, sur le traitement, sur les interprétations. L'estimation des effets dus au chercheur peut imposer d'examiner la sélection des sujets, la manière dont le contact a été établi avec les informateurs, le contexte des entrevues et la consigne qui a été transmise.

L'exemple d'une enquête évaluative sur l'introduction d'un matériel informatique illustre ce problème des biais. Le matériel informatique était entré dans une école à la suite de l'implantation d'un projet par lequel quelqu'un, qui n'avait plus de travail, était parvenu à maintenir son poste. Grâce à une subvention du ministère, l'école avait acquis sept micro-ordinateurs et quelques accessoires. Il est évident que l'initiatrice du projet, qui a, de plus, pu maintenir son poste par ce projet, y croit : «Tu as vu comme les enfants sont devenus autonomes ? Si c'est splendide ! Et ils ont appris des choses. Ils ont joué avec la tortue et du même coup ils connaissent la géométrie.» Et d'autres déclarations du genre. La personne initiatrice du programme avait des attentes et elle les a projetées. Mais comment les enseignants qui ont été impliqués dans ce programme, l'ont-ils vécu ? Ont-ils rencontré des difficultés ? Est-ce que cela leur a demandé de réaménager leurs horaires, de consacrer du temps et de l'énergie à une formation, de prendre plus de temps pour préparer leur matériel ? Une entrevue avec les enseignants s'impose donc. À la suite de diverses contraintes, l'entrevue avec les enseignants se fait dans l'école, dans le local de l'initiatrice du projet en sachant qu'elle peut rentrer, sortir, interrompre. On s'étonne, après coup, de ce que le contenu des entrevues correspond exactement à ce que la personne responsable du projet a dit. Ils travaillent dans ce milieu et ne veulent pas trop le bousculer. Après l'analyse, le chercheur se rend à l'évidence que le contexte du contact avec les sujets avait été important, qu'il avait interféré avec le processus de constitution des données, et qu'il fallait en tenir compte dans leur interprétation.

Prenons une autre anecdote, celle d'une recherche en milieu universitaire. Un membre du comité d'évaluation du projet avait estimé qu'il fallait absolument qu'une partie de l'échantillon soit extraite de l'institution des chercheurs. Ce à quoi les chercheurs n'étaient pas favorables. Mais ils s'y soumièrent. Ils constatèrent qu'une partie des informateurs anticipent cette entrevue dans un climat paranoïde lorsque l'enquêteur provient de la même institution : «Est-ce qu'il ne va pas évaluer notre pratique ? Est-ce qu'il ne va pas en parler au doyen ? Est-ce que cela ne va pas aller dans notre dossier de promotion ? Est-ce que nous n'allons pas nous faire déposséder de notre démarche ?» On aurait pu croire que l'évaluateur avait bien raison : commencez donc dans votre institution, ils vous connaissent, c'est tout près; pour l'organisme qui subventionne, c'est moins cher, ça c'est clair. Mais la sélection des sujets et le contact avec eux auront une incidence sur ce qu'ils vont dire et sur les interprétations qu'on en tirera. Il faut évaluer ces effets lorsqu'on ne peut pas les anticiper et aménager le contexte du contact pour les amoindrir. De toute façon, ces biais ne sont jamais totalement éliminés; il faut donc toujours en évaluer l'impact.

L'estimation des biais interprétatifs dus aux orientations théoriques du chercheur est aussi importante. Certaines théories de référence conduisent nécessairement à percevoir certaines choses, soit au niveau de la constitution des données, soit au moment de l'interprétation des résultats. Par exemple, il ne serait pas étonnant qu'une orientation psychanalytique conduise le chercheur à interpréter la dynamique d'une situation en termes de conflits intra-psychiques, de répression des pulsions, alors que la sociologie critique y verrait plutôt un système de répression et d'exploitation sociale, etc. Pour deux raisons, il convient d'examiner les biais que les orientations théoriques induisent. D'abord, pour essayer d'élargir les horizons interprétatifs, alors l'horizon du cadre conceptuel est nécessairement restreint. Ensuite, pour vérifier si, en cours d'interprétation, le chercheur n'a pas tendance à rétrécir son horizon pour n'y retrouver que ses manies.

3.3 Le contrôle des hypothèses

Après avoir validé la démarche qui a conduit aux hypothèses, il faut aussi contrôler leur pertinence et leur vraisemblance.

3.3.1 Le contrôle sur le matériel résidu

De par ses caractéristiques, la vraisemblance et la pertinence d'une induction ne peuvent guère faire l'objet d'un contrôle empirique par un

test expérimental. Dès lors, un premier contrôle doit être effectué sur le matériel résidu. Le matériel résidu est la partie du matériel qui n'a pas été retenue dans la sélection initiale. Le matériel était trop abondant pour l'analyse des données; avant le codage, on a sélectionné, on a trié les segments qui paraissaient parlants, qui permettaient la détection d'indices. Cela veut dire que toutes les analyses et, en fin de compte, les interprétations n'ont pas été faites sur l'ensemble des données, mais sur la partie de données qui était parlante, celle où les indices étaient les plus visibles. Une question se pose alors : est-ce que les hypothèses tiennent sur la partie des données où les indices n'étaient pas aussi évidents ?

Il faut se méfier des interprétations qu'on construit sur une partie du matériel qui est, en fait, un matériel extraordinaire. On ne peut pas construire une théorie du ciel en voyant la comète de Halley; elle passe à peu près à tous les cent ans. Il ne faut pas construire une interprétation sur un matériel fantastique qui ne se produit qu'une fois ou l'autre. Le matériel résidu est un matériel banal que les hypothèses doivent aussi être capables d'expliquer. Si les hypothèses tombent par terre devant le matériel banal, elles ont un sérieux problème de pertinence et il faut les réviser. Il est évident, pour la plupart des gens, que si une hypothèse n'explique pas les cas extraordinaires, c'est qu'elle ne tient pas compte de facteurs qui leur sont spécifiques; dès lors, il faut réincorporer les cas extraordinaires à l'ensemble du matériel pour qu'en modifiant l'hypothèse, celle-ci puisse en rendre compte. Il en va de même pour les cas banals. Dans ce sens, il faut examiner si on peut sélectionner du matériel des extraits banals qui illustrent les hypothèses : les illustrations doivent être représentatives de l'ensemble du matériel et non pas seulement de la part qui était extraordinaire. Lorsqu'on n'est pas capable d'illustrer les hypothèses par des extraits relativement banals, le rapport des hypothèses au matériel retenu lors de la condensation est défectueux.

C'est aussi le moment de contrôler la *communicabilité* et l'*apport conceptuel* des hypothèses. Si les hypothèses ne peuvent s'exprimer que dans un jargon ésotérique, si on ne peut rendre les choses communicables qu'avec un lexique, c'est qu'il y a un problème : il est probable que le jargon cache des constructions conceptuelles qui sont invérifiables (et donc dogmatiques) ! Par ailleurs, si les hypothèses induites n'enrichissent pas les connaissances, c'est qu'elles aussi sont banales. Peut-être le problème était-il mal conceptualisé, les outils mal choisis, les analyses inadéquates et les traitements trop vite arrêtés pour produire de nouvelles idées ?

3.3.2 Le contrôle par d'autres chercheurs

On dit qu'une méthode qualitative bien appliquée est plus forte, produit des hypothèses plus pertinentes qu'une méthode quantitative. Mais pour atteindre cela, il est important que les hypothèses soient contrôlées par d'autres chercheurs. Et si possible, par des chercheurs ennemis! Si un collègue, qui est adepte d'autres théories que celles du chercheur, est d'accord avec ses interprétations, c'est que les interprétations sont fortes. Cet autre chercheur devrait être situé en dehors du milieu immédiat, car il devrait ne pas avoir à ménager ses relations. Idéalement, le contrôle des hypothèses doit s'effectuer par des gens qui ne partagent pas les positions ou les points de vue du chercheur, qui peuvent diagnostiquer si ce dernier a été illuminé dans ce qu'il a perçu et qui ont la liberté de le lui dire sincèrement.

Quand plusieurs enquêteurs ont opéré sur plusieurs sites (plusieurs écoles, par exemple), il peut être intéressant de faire contrôler les hypothèses que l'on fait sur une école, par un chercheur qui a travaillé dans une autre école, c'est-à-dire par quelqu'un qui a l'expérience de la même situation de recherche mais sur un autre site, sur un autre milieu.

Plus on peut obtenir et tenir compte d'une évaluation faite par un critique externe, qui n'est pas partie prenante au projet, plus l'expérience sera enrichissante. C'est parfois enrageant, car il arrive de recevoir des commentaires à ce point négatifs qu'on ne soit plus capable de voir et de reprendre, pendant quelque temps, ce travail «assassiné». Mais il reste que lorsqu'on est capable d'y revenir, ces jugements extérieurs permettent de produire un matériel de qualité, car il est souvent vrai que, si la première version est enivrante par les trouvailles qu'on y formule, elle comporte aussi des choses qui ne sont pas au point, trop sommaires, pas assez réfléchies. Et pour finir, la critique surmontée apporte beaucoup à la qualité et à la crédibilité du travail⁷.

3.3.3 Le contrôle par les acteurs

Enfin, le contrôle des hypothèses doit aussi être effectué par les sujets, c'est-à-dire par les informateurs. On a construit une belle histoire sur leur

7 Il en va de même pour les demandes de subvention. On peut y gagner beaucoup en les faisant critiquer par un collègue avec qui on fait, de temps en temps, des échanges de bons procédés : «Je vais être méchant avec toi, tu vas être méchant avec moi». Si on le faisait plus souvent, avant d'envoyer les demandes de subvention ou de bourse, elles passeraient plus vite, elles seraient moins refusées.

vie; il est dès lors important qu'ils puissent l'examiner et qu'ils puissent dire si elle colle avec ce qu'ils ont vécu. Même s'ils ne s'attendent pas à tout ce qu'on y dévoile.

Il est même essentiel de permettre au milieu sur lequel on parle, d'évaluer si les hypothèses sur ce qui a pu les conduire à faire telle ou telle chose, sont délirantes ou pas, si les hypothèses ont un rapport avec la manière dont les choses se sont déroulées. Parce que le chercheur n'a pas pensé ou pu récolter d'information sur un événement, ou parce qu'il est prisonnier de ses biais théoriques, il lui arrive de construire une interprétation, alors que c'est un tout autre événement qui, pour les acteurs, est la clef interprétative essentielle. La crédibilité des rapports de recherche et leur diffusion tiennent, entre autres choses, au souci des chercheurs de faire contrôler la pertinence des inductions par les acteurs du terrain. Ceux-ci ne sont pas nécessairement les individus qui ont été interrogés : il y a ceux qui ont vécu aussi la situation de près et ceux qui l'ont vécue de loin. Il est étonnant, par exemple, de constater comment certaines informations sur ce qui se passe dans une classe peuvent être contrôlées en interrogeant le concierge qui voit des choses par terre, sur le tableau qu'il vient nettoyer, dans les poubelles, et qui entend des choses pendant qu'il entretient les locaux. Le rôle du concierge n'est peut-être pas important au moment où on fait l'enquête, mais au moment de la validation, il peut être essentiel de lui poser la question : «Est-ce que cela a du sens ce qu'on raconte sur ce qui s'est passé dans l'école ?» Et là, il peut vous dire : «Mais non. Je vais vous raconter moi ce qui s'est passé. On vous a raconté des blagues.» En plus du contrôle par les informateurs qui ont été sélectionnés pour faire partie de l'échantillon, il faut donc prévoir une phase de contrôle par les acteurs, en envisageant autant ceux qui ont une position périphérique que ceux qui ont une position centrale par rapport aux événements qui nous intéressent.

4 Notes sur la rédaction du rapport

Un des premiers problèmes de la rédaction du rapport d'une recherche exploratoire comparativement à celui d'une recherche vérificative, découle de la nécessité, pour le chercheur, de traduire de manière linéaire ce qui a été, dans la plupart des cas, un processus très circulaire. Pour que le lecteur apprécie le rapport, pour qu'il le comprenne sans éprouver une difficulté qu'il attribuerait à la mauvaise qualité de la recherche, il faut que le rapport soit linéaire. Sa rédaction doit donner l'impression que les choses se sont écoulées sans heurts, du moins sous contrôle. Il faudra rationaliser, d'une

certaine façon, les moments d'hésitation, les angoisses, les retours sur soi-même, les retours à zéro, les données qu'on a mises à la poubelle et le fait qu'on a couru derrière le gars qui ramassait la poubelle parce qu'on s'est dit qu'il y avait peut-être quand même quelque chose dedans. Le premier problème de la rédaction du rapport est donc de traduire en linéaire ce qui a été très circulaire, très chaotique parfois.

Le deuxième problème consiste à rapporter la chronique des étapes de la recherche en la condensant et en éliminant l'anecdotique, mais en rapportant tout ce qui doit l'être. Le rapport de recherche ne doit pas comporter tous les éléments qui ont servi à la vérification interne et externe des comptes. Le rapport n'est pas le récit de la recherche : ce n'est pas parce qu'il s'agit de recherche exploratoire qu'il faut abandonner le format du rapport scientifique pour tomber dans la narration romanesque. Certains auteurs noient le lecteur dans l'anecdote parce qu'ils pensent que l'anecdote est particulièrement riche, alors que l'anecdotique ne doit qu'illustrer le conceptuel. Ce dont le lecteur a besoin, ce sont les extraits qui illustrent des concepts, tout en ayant un minimum de sens par rapport à la situation, et non pas des extraits qui sont marginaux. Un extrait illustratif n'est pas un dédale d'anecdotes; la chronique est donc condensée, le rédacteur en enlève ce qui n'est qu'anecdotique.

Dans la recherche exploratoire, la seule hypothèse formulée est une hypothèse méthodologique. C'est-à-dire qu'étant donné le problème auquel on est confronté, on pose que, si telle stratégie de constitution des données est utilisée, alors on obtiendra un matériel dont l'analyse et le traitement permettront de formuler des hypothèses. En conséquence, le rapport porte sur toutes les étapes qui le précèdent. Donc, le cadre conceptuel, y compris la chronique de l'évolution de ce cadre conceptuel, sont présentés; ensuite on explique comment les analyses, les traitements, les interprétations ont été faits, puis comment on les a validés. La phase de validation des interprétations doit être intégrée au rapport. Certains croient que parce qu'ils font de l'exploration, le rapport se termine à l'interprétation : ils vont directement de l'énoncé du problème aux conclusions, et s'y arrêtent. Mais pour comprendre leurs conclusions, il faut au moins savoir quelles procédures ont été suivies pour y aboutir et le soin que le chercheur a accordé à leur validation.

Enfin, il ne faut pas oublier d'apporter divers commentaires sur les conclusions de la recherche, commentaires pour lesquels l'auteur est le mieux placé :

- 1° Les commentaires sur la signification et la pertinence de la recherche;
- 2° Les commentaires sur l'apport conceptuel de l'interprétation des résultats;
- 3° Les commentaires sur les applications possibles, sur les transferts.

Le chercheur est le mieux placé pour expliquer quelles sont les limites dues aux caractéristiques de cet échantillon. Il en est de même pour l'énoncé des applications possibles. Pourquoi serait-ce à l'enseignant, dans sa classe, de deviner ce qu'il peut faire des conceptualisations auxquelles notre recherche aboutit ? Si le chercheur, qui connaît mieux que tout autre comment le problème de base a été conceptualisé, quel a été le contexte et donc les contraintes de sa recherche, n'indique pas à quelles conditions les transferts sont possibles et comment les applications sont permises, personne ne peut raisonnablement le faire à sa place. Si le chercheur ne commente pas les conditions qui permettent le transfert des conclusions de sa recherche, il ne faut pas qu'il soit étonné de la voir rester sur les tablettes. C'est donc aussi, avec la chronique, un élément important dans la rédaction du rapport.