

Chapitre 15. L'analyse des données qualitatives : voyage au centre du codage

Sébastien Point

DANS **BUSINESS SCIENCE INSTITUTE** 2018, PAGES 262 À 282

ÉDITIONS **EMS EDITIONS**

ISBN 9782376871798

DOI 10.3917/ems.cheva.2018.01.0262

Article disponible en ligne à l'adresse

<https://www.cairn.info/les-methodes-de-recherche-du-dba--9782376871798-page-262.htm>



CAIRN.INFO
MATIÈRES À RÉFLEXION

Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...

Flashez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



Distribution électronique Cairn.info pour EMS Editions.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

CHAPITRE 15.

L'analyse des données qualitatives : voyage au centre du codage

Sébastien Point

Résumé

Qu'est-ce que le codage ? Comment donner du sens aux données ? Comment lier concepts et données ? Ce chapitre vise à introduire la notion de codage et susciter une véritable réflexion sur ce processus d'analyse des données. Il s'agit donc de lever le voile sur ce processus de transformation. Si, en théorie, la démarche de codage reste simple, en pratique ce n'est pas un processus neutre, puisqu'il mobilise une grande subjectivité du chercheur à travers la recherche de sens. Reste au chercheur de faire les bons choix de codages et d'interprétations.

Mots-clés : codage, données qualitatives, concepts, subjectivité, interprétation.

*"Coding is fun isn't it? You never know what you might discover
in a simple encounter."*

Strauss et Corbin (1990, p. 73)

INTRODUCTION¹

Les travaux en recherche qualitative visent à comprendre et interpréter un phénomène spécifique étudié. Les extrapolations ou généralisations sont quasiment impossibles. Ce n'est d'ailleurs pas le but des recherches qualitatives fortement contextualisées et qui reposent sur des échantillons beaucoup plus restreints que les recherches quantitatives. Pourtant, avoir de petits échantillons et/ou des problématiques de recherche fortement contextualisées ne veut pas dire avoir une petite quantité d'information à devoir gérer : ces dernières années, nous assistons à une véritable mutation des données qualitatives, et qualifier la recherche qualitative de nuisance attractive (Miles, 1979) constitue plus que jamais une tautologie, et ce, pour trois raisons essentielles. En premier lieu, la taille des échantillons tend à largement augmenter, puisque pour la plupart des chercheurs, un échantillon de taille plus importante garantirait la robustesse et la validité de l'analyse – ce qui n'est pas forcément vrai ! Deuxièmement, les données recueillies sont particulièrement variées : outre les discours et toutes les observations qui conduisent le chercheur à manipuler du texte, des images, des fichiers audio ou même vidéo peuvent aujourd'hui venir enrichir la diversité des données ainsi recueillies par le chercheur². Troisièmement, la frontière entre les approches quantitatives et qualitatives devient de plus en plus floue et conduit parfois le chercheur à opérer toutes sortes de comptages qui tendent à quantifier les informations obtenues à partir de données qualitatives. En conséquence, rigueur et flexibilité sont aujourd'hui devenus les maîtres mots de la méthodologie en recherche qualitative (Gioia *et al.*, 2013 ; Saldaña, 2013).

1. Ce chapitre ne se veut pas être un inventaire *in extenso* des techniques de codage existantes. Nous proposons plutôt ici aux lecteurs une réflexion sur le processus de codage et plus généralement de l'analyse des données qualitatives. Les recensements effectués ici ne sont que des résumés de la pensée originelle et nous incitons fortement les chercheurs, intéressés par une méthode particulière, à se référer aux travaux sources : chacun des auteurs cités possède effectivement une interprétation bien particulière des concepts soulignés dans ce chapitre et nous encourageons également les chercheurs non seulement à en prendre connaissance, mais également à se faire leur propre appropriation.

2. Nous nous référons ici au dicton « *All is data* », cher à Glaser (1978).

Au lieu de parler de méthodologie de recherche qualitative, nous devrions plutôt évoquer des méthodologies : en effet, tant au niveau de la recherche qualitative que quantitative, la méthodologie se décline en trois temps distincts : le recueil, l'analyse et la mesure/validité de la recherche. L'objectif de ce chapitre est de se focaliser sur la deuxième étape, celle de l'analyse qui passe par un processus de (dé)codage des données qualitatives. Mais avant d'évoquer les processus de (dé)codage des données qualitatives et de mieux appréhender l'analyse de celles-ci, rappelons l'objectif de toute recherche qualitative.

L'objectif de toute recherche qualitative reste de comprendre un phénomène par la construction d'un modèle : elle assure avant tout une visée compréhensive qui conduit généralement à une construction théorique. La construction théorique peut être à visée descriptive ou explicative (Snow & Thomas, 1994) : la première correspond au « quoi » en identifiant les construits et concepts clés de la recherche ; la seconde répond au « comment » et « pourquoi » en cherchant à établir des relations entre les concepts et les construits de la recherche. La construction théorique correspond alors plutôt à un réseau de relations non hiérarchiques, exprimées avec des énoncés définissant les liens parmi les concepts (Carley, 1991) ou encore à un modèle prédictif d'événements, pouvant être comparé avec ce qui est réellement observé (Yin, 2017). Cette construction théorique est l'aboutissement de l'analyse des données qualitatives opérée *via* le codage. En l'occurrence, ce processus appelle à plusieurs questionnements qui structurent ce chapitre : 1) qu'est-ce que le codage ? 2) Comment lier concepts et données ? 3) Comment donner du sens aux données ? 4) Comment rendre compte du codage ?

1. UN PEU DE THÉORIE : QU'EST-CE QUE LE CODAGE ?

Commençons par démystifier le processus clé au cœur de l'analyse des données qualitatives : le codage. Le codage est le résultat de l'analyse des données et il va même au-delà ; c'est une démarche captivante, mais aussi chronophage (Ayache & Dumez, 2011). Bon nombre d'auteurs s'accordent à dire que l'analyse des données qualitatives commence dès le recueil de celles-ci !

Pourquoi donc coder ? Le codage permet d'identifier les idées issues des données recueillies. Le codage est un instrument (pour étayer la métaphore de Ayache & Dumez, 2011) : tel un musicien, le chercheur fait corps avec son instrument pour lire, déchiffrer et exécuter sa partition. Il s'agit également de relier les idées entre elles afin de leur donner du sens (Coffey & Atkinson, 1996). C'est en effet à partir du sens donné au matériau recueilli que le processus de codage se réalise. Pour cela, le chercheur rassemble entre eux les codes homogènes puis il va progressivement transformer ces codes en catégories par un travail successif de codage (en repensant puis en recodant les codes). « La catégorie est un code travaillé par le chercheur et renvoie ainsi à l'organisation conceptuelle et à la théorisation menée par le chercheur. Elle va alors bien au-delà de la désignation de contenu » et elle reste apparentée à la notion de concept (Point & Voynnet-Fourboul, 2006, p. 64).

Il existe autant de définitions du codage que de manières d'appréhender les données qualitatives. Les définitions s'accordent toutefois toutes sur un point : le codage est un processus de transformation ; il s'agit de transformer les données pour arriver à un modèle, une construction théorique. Certaines définitions mettent en exergue le principe de réduction des données ; le cerveau humain ayant des capacités limitées, il lui est nécessaire de procéder à une simplification des données face au volume qu'il a à traiter. La catégorisation – principe que nous éluderons ultérieurement – appelle justement à une simplification en invitant le chercheur à classer les informations selon des catégories. Dans cette perspective, le codage consiste à « découper les données en unités d'analyse, à définir les catégories qui vont les accueillir, puis à placer les unités dans ces catégories » (Allard-Poesi, 2003, p. 246). Ce procédé est également cher à Miles et Huberman (1994). D'autres définitions mettent en exergue la cognition et la capacité du chercheur à penser ses données. Dans cette perspective, le codage implique « un effort explicite d'identifier des thèmes, de construire des interprétations telles qu'elles émergent des données ainsi que de clarifier le lien entre les données, les thèmes et les interprétations conséquentes » (Tesch, 1990, p. 113, traduction de l'auteur). Ce processus fournit au chercheur des façons de penser les données (un processus d'indexation). Le chercheur se pose lui-même des questions à propos des données (voir Coffey & Aktinson, 1996).

En théorie, la démarche de codage est simple : il suffit de se poser de multiples questions sur le texte lui-même... et au-delà : bon nombre de chercheurs tendent à se poser uniquement des questions sur le sens du texte ou le contenu du fichier audio/vidéo, en oubliant de lire entre les lignes : silences, hésitations, émotions, métaphores, ton de la voix, gestuelle (en cas de visuel) etc. sont autant d'indices permettant au chercheur d'interpréter et de donner du sens à ces données. À l'instar de l'inspecteur Columbo, il faut donc méticuleusement trouver des indices permettant de construire un sens particulier, celui qui amènera progressivement le chercheur à mobiliser des concepts. Ayache et Dumez (2011) évoquent à juste titre la multi-dimensionnalité du codage pour rendre compte : 1) de la multiplicité du sens que le chercheur peut donner au corpus de données ; 2) de la variété des thématiques auxquelles il peut être fait référence, ou encore ; 3) des perspectives multiples (ou angles d'attaques) que chaque chercheur peut effectuer sur le même corpus de données, de manière convergente ou divergente.

En pratique, le codage fait vraisemblablement douter le chercheur : le codage n'est pas un processus neutre, il amène une grande subjectivité du chercheur puisque c'est lui qui donne du sens : mais a-t-il seulement fait les bons choix ? Que faut-il coder ? En commençant par une codification des données, le processus de codage amène à une catégorisation. Il ne faut en effet pas confondre codification des données et catégorisation des données. En l'occurrence, le premier renvoie à la transformation des données brutes (issues de faits observés, de propos et/ou discours recueillis...) en une première formulation qui relève du sens commun : certains chercheurs appellent cela de l'étiquetage, soit le fait de mettre des étiquettes sur des morceaux de textes ou d'images ou encore sur des extraits audios ou vidéos. La seconde est un processus visant à transformer des thèmes et/ou idées en concepts clés destinés à un système plus abstrait de connaissances. La catégorisation « est l'analyse, la conceptualisation, mise en forme, la théorisation en progression » (Paillé & Mucchielli, 2005, p. 147). Elle renvoie à une réorganisation des données et des codes déjà attribués, au moyen de catégories créées antérieurement ou ultérieurement par le chercheur.

Encadré 1. Le codage en pratique : six pièges à éviter

La paraphrase : de simples citations issues du corpus de données ne tiennent pas lieu d'analyse. Rappelons que le codage va au-delà d'une simple description, puisque le chercheur donne du sens à ses données ; c'est cette recherche de sens qui l'amène progressivement à faire le lien entre concepts et données empiriques.

Le réductionnisme : travailler uniquement sur les codes et non sur le corpus de données de base peut amener le chercheur à perdre la globalité de ses données. Trop de réduction nuit à l'interprétation.

La linéarité des opérations : le codage nécessite des « va-et-vient » entre le corpus de données étudié et les catégories créées, entre les concepts identifiés dans la revue de littérature et les données recueillies ; ce n'est pas un processus complètement linéaire qui part soit des concepts pour les relier aux données, soit des données pour faire émerger des concepts. Le processus reste itératif.

La technicisation : il s'agit d'une surexploitation du corpus de données et l'emploi d'outils non nécessaires à l'analyse. Trop d'expertise peut être contreproductive au chercheur puisqu'il peut facilement se noyer dans ses propres connaissances et interprétations.

L'enlissement : un enlissement dans l'analyse ou même le recueil des données peut nuire à la phase d'interprétation. Le processus de codage reste chronophage mais ne doit pas pour autant nuire aux autres étapes de la recherche.

La détemporisation : le codage nécessite beaucoup de temps, non négligeable dans la procédure d'analyse. Attention toutefois à limiter le temps consacré au codage.

Inspiré de Point et Voynnet-Fourboul (2006, p. 70).

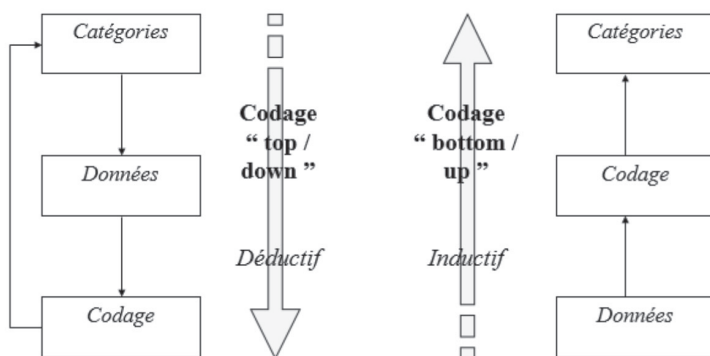
En l'occurrence, le défi pour tout chercheur est justement de passer de la codification à la catégorisation, sinon il risque de rester dans le descriptif, ce qui peut davantage relever d'une démarche journalistique au lieu d'une démarche scientifique et académique. Ce procédé reste complexe mais c'est aussi la qualité de ce passage entre codification et catégorisation qui déterminera la qualité de l'analyse, et donc celle de la thèse. L'étape clé est donc de relier concepts et données.

2. UN PEU DE PRATIQUE : COMMENT LIER CONCEPTS ET DONNÉES ?

Dans sa démarche de recherche, le jeune chercheur se retrouve très vite avec deux types de matériau : d'une part, avec sa revue

de littérature, le chercheur amasse un ensemble de connaissances, de concepts, de modèles, de théories et de réflexions. Il s'agit des données théoriques. D'autre part, les données sont regroupées sous forme de discours, de faits, d'opinions, de blocs, d'attitudes, d'observations..., c'est-à-dire un ensemble de documents regroupés en données primaires (entretiens) et données secondaires (observations, documents, sites internet, etc.) appelées données empiriques. Le défi est de trouver des passerelles ou des liens entre les données empiriques et les données théoriques. Deux visions possibles s'offrent ainsi au chercheur : partir des données empiriques et progressivement arriver à la théorie ou partir des concepts et trouver l'ancrage empirique correspondant à la théorie (Figure 1).

Figure 1. Le passage entre les mondes théorique et empirique



Dans le premier cas, à partir de faits, d'observations et de tous documents dont il dispose, il s'agit de faire émerger des éléments plus conceptuels. Le fait de passer des données empiriques à la théorie (et ses concepts) s'appelle l'abstraction³. L'abstraction est un mécanisme permettant au chercheur de progresser vers une construction théorique ; il s'agit pour le chercheur de se focaliser sur ce qu'il perçoit comme essentiel pour bâtir un modèle

3. Le terme *abstraction* vient du latin *ab trahere* qui veut dire « tirer de ». « La catégorie fait sens, dans la mesure où elle décrit un phénomène d'un certain point de vue, dans la mesure où elle donne lieu à une définition, elle a des propriétés synthétique, dénominateur et explicative d'un concept » (Paillé & Mucchielli, 2005, p. 149).

analytique en relation avec le problème qu'il se pose (Dumez, 2016). Ce principe d'abstraction (ou démarche ascendante, ou *bottom-up*, en anglais) consiste donc à coder les données, formuler des indices (au sens de Lazarsfeld, 1967), pour ensuite progressivement créer des catégories pour faire émerger des éléments plus conceptuels. En d'autres termes, il s'agit de repérer les concepts dissimulés derrière l'ensemble des données que le chercheur a recueilli. Ces éléments empiriques ne représentent jamais directement des concepts et il s'agit justement du travail du chercheur de trouver – par un jeu de codages successifs – la signification des concepts théoriques sous-jacents. Par le jeu de l'induction, le chercheur utilise ses données pour y repérer des concepts clés. Dans l'exemple ci-dessous (Figure 2), le chercheur part de toute la richesse du matériau pour le découper en unités de sens et y attribuer des codes. Aussi, le chercheur commence par faire un repérage avec un codage très proche du texte analysé. Lorsque le label du code est identique au texte, il s'agit d'un codage dit *in vivo*⁴. Puis, progressivement, à la lecture d'autres segments de textes, le chercheur va transformer ses codes en catégories. Ainsi, dans cet exemple, la concurrence pourrait se transformer en catégorie « compétition ». L'ensemble du paragraphe fait aussi référence à la catégorie « ambition ».

Figure 2. Exemple de codification

« Dans un environnement géopolitique incertain et avec une concurrence de plus en plus vive, nous nous sommes donné les moyens d'un développement plus vigoureux nous permettant de prendre notre part de la croissance du transport aérien dans le monde. C'est l'objet de Trust Together. »

Extrait du rapport annuel Air France KLM 2016.

incertitude	
concurrence	ambition
croissance	
projet	

Dans le second cas (c'est-à-dire dans le cas du passage du monde théorique au monde empirique, démarche descendante (*top-down*, en anglais) – généralement plus abordable pour les jeunes chercheurs – ces derniers s'attachent à repérer, au sein même des données recueillies, les éléments empiriques qui illustrent le plus finement possible les concepts identifiés dans la revue de littéra-

4. Le terme *in vivo* vient du latin « au sens du vivant ». En d'autres termes, il s'agit d'utiliser au plus près les verbatim des personnes interrogées pour explorer les différentes dimensions des catégories (voir Saldaña, 2013).

ture. Il s'agit, par une manœuvre de repérage, d'identifier le ou les passages du texte, de la bande-son ou de la vidéo au(x) concept(s) identifié(s) préalablement dans la littérature. Dans l'exemple ci-dessous (Figure 3), les catégories sont données par la théorie et le chercheur tente de les retrouver dans son matériau. En l'occurrence, le chercheur utilise des codes théoriques⁵, c'est-à-dire les codes qu'il a identifiés dans la littérature. S'agissant généralement de concepts, ces codes théoriques constituent de véritables catégories. Dans cet exemple, le codage est clairement à visée théorique (Point & Voynnet-Fourboul, 2006).

Figure 3. Exemple de catégorisation

« Ce que je souhaite pour le Groupe, c'est de l'énergie, de l'audace et de la volonté pour innover. Avec le client au cœur de nos préoccupations, nous travaillons ainsi à une meilleure performance de nos hubs et simplifions nos organisations. »

Extrait du rapport annuel Air France KLM 2016.

innovation
customer-oriented
performance

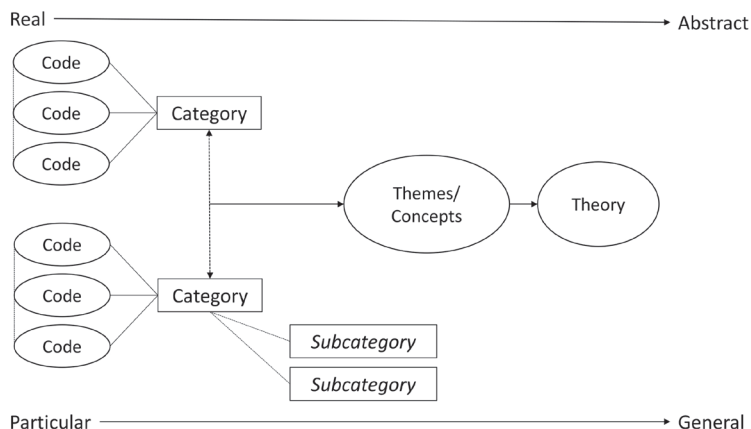
Quelle que soit la démarche privilégiée (inductive ou déductive) le chercheur travaille dans un premier temps sur son corpus de données ou matériau de recherche pour ensuite travailler sur les codes eux-mêmes. En effet n'oublions pas que l'essence du codage reste de travailler sur les ressemblances/différences d'une unité de texte (mot, phrase, paragraphe, etc.) à une autre. Le chercheur va donc adopter le schéma réflexif suivant (Figure 4). Le passage de la codification à la catégorisation va amener le chercheur à passer du concret à l'abstraction et du particulier au général. Le questionnement du chercheur s'articulera alors autour des interrogations suivantes :

- Faut-il créer une catégorie à partir des codes utilisés pour rendre compte de l'idée suscitée à la lecture du texte ?
- Quelle est la place de cette catégorie par rapport aux autres ?

5. Le terme « codage théorique » est perçu comme un oxymore par certains auteurs (Ayache & Dumez, 2011, p. 34). Si le code reste un « label » que le chercheur met sur un mot – ou groupes de mots – issus d'un verbatim, rappelons que cette étiquette peut être soit identique au texte, soit provenir des lectures théoriques du chercheur, ce dernier ayant déjà cheminé sur le sens accordé au code. Nous utilisons le terme « code théorique » pour montrer que le chercheur est déjà dans une démarche d'abstraction au moment où il distribue les étiquettes au sein de ses verbatim.

- Faut-il plutôt aller plus finement dans la catégorisation, en affinant par des sous-catégories ?

Figure 4. Le processus de construction théorique



Source : Saldaña (2013, p. 13)

Dans un premier temps, le chercheur réalisera un codage dit « ouvert », phase durant laquelle il va multiplier les codes à attribuer aux unités de sens. Puis, progressivement, il va chercher à regrouper ces codes en catégories, puis identifier des liens entre codes et catégories ou catégories entre elles. Les catégories peuvent aussi être affinées ; par exemple, la catégorie « performance » peut être affinée en « performance sociale », « performance économique », « efficacité », « efficacité », etc. L'objectif final sera d'identifier les « méta-catégories », c'est-à-dire les trois ou quatre « super catégories » qui permettront de « résumer » l'ensemble des codes et des catégories identifiés. Ces méta-catégories reliées entre elles forment une construction théorique.

Les démarches de codage inductive et déductive ne sont pas incommensurables pour autant. En lisant les détracteurs de l'une ou l'autre de ces deux démarches, on pourrait croire qu'elles s'opposent. La présence de codes théoriques amène le chercheur à avoir de nombreuses « préconceptions » (Glaser, 2012). Si l'absence de préconception reste un des préceptes de la théorie ancrée (*Grounded Theory*, Glaser & Strauss, 1967), il est important

que les codes théoriques ne viennent trop polluer le matériau de recherche (Dumez, 2016). En l'occurrence, le chercheur peut être tenté de « forcer » l'apparition de concepts avec lesquels il est relativement familier au sein de son matériau de recherche. L'affirmation « pas de préconceptions » (Glaser, 2012) doit rester à l'esprit du chercheur pour garantir une certaine ouverture et être au cœur des données, sans véritablement vouloir « trouver » ce quelque chose que le chercheur souhaite absolument trouver ! À l'inverse, le chercheur ne peut que trop rarement se détacher de sa sensibilité théorique, c'est-à-dire de son expérience et de sa revue de littérature ; il reste forcément influencé.

Une des grandes tentations est donc de mêler démarches inductive et déductive, c'est-à-dire laisser libre cours à la créativité à partir du matériau tout en se référant à des codes théoriques par des allers-retours avec la littérature. Cette démarche de va-et-vient entre les modes empiriques et théoriques semble actuellement particulièrement prisée par les chercheurs qui qualifient bien souvent celle-ci de démarche abductive. La notion d'abduction amène effectivement des oscillations inductives et déductives mais elle va bien au-delà. Charles Sanders Peirce (1974) – à qui nous devons l'emploi de ce terme – définit l'abduction comme un processus permettant de former des hypothèses explicatives. Il s'agit d'une opération logique qui introduit une nouvelle idée. L'induction ne fait rien d'autre que de déterminer une valeur, et la déduction extrait les conséquences nécessaires d'une pure hypothèse. En conséquence, le terme « abduction » devra être utilisé avec prudence et parcimonie, puisqu'il renvoie à une véritable démarche de recherche plutôt qu'à une simple oscillation entre les modes théoriques et empiriques. Quelle que soit la démarche adoptée, le défi majeur reste de donner du sens aux données.

3. UN PEU DE RÉFLEXION : COMMENT DONNER DU SENS AUX DONNÉES ?

La littérature propose au chercheur différents principes de « classification » des données ; par exemple, les méthodes de Miles et Huberman (1994) ou encore de Strauss et Corbin (1990) restent particulièrement mécanisées ce qui permet de rassurer le chercheur dans son processus de codage. Ces deux démarches

constituent un bon exemple d'un protocole particulièrement précis et détaillé à suivre pour le codage des données ; un tel protocole pourrait convenir au chercheur soucieux d'exploiter au mieux ses données, mais pourrait aussi devenir particulièrement contraignant pour des chercheurs recherchant davantage de créativité. En effet, ce ne sont pas les seules démarches existantes et Saldaña, (2013) propose pas moins d'une trentaine de méthodes destinées au codage des données qualitatives.

La distinction entre codification et catégorisation nous amène ici à effectuer une différenciation entre les méthodes de codage de premier ordre (ou de premier niveau) et celles de second ordre (ou de second niveau). Les codes de premier ordre renvoient à un codage descriptif ou thématique, ce qui constitue par essence un codage très proche des données. Les méthodes de codage les plus utilisées à ce stade restent les codages *in vivo* (reprise littérale du verbatim) et thématique (identification de grands thèmes). *A contrario*, les codes de second degré sont davantage tournés vers une démarche d'abstraction ; il s'agit de développer une organisation plus conceptuelle et théorique des données à partir des codes de premier ordre (Saldaña, 2013). Pendant cette seconde phase de codage, le chercheur va réorganiser ses codes (les renommer, les fusionner ou encore les abandonner). Saldaña (2013) recense nombre de techniques qui peuvent être regroupées en codage de premier et second ordre (Figure 5).

La figure 5 montre la diversité des codages que le chercheur peut mobiliser pour l'analyse de ses données. Nombre de chercheurs se limitent bien souvent à un codage thématique : même si nous l'avons identifié comme une méthode à part, le codage thématique renvoie à l'ensemble des méthodes de codage de premier ordre. Le codage thématique reste présenté comme un processus intéressant pour les premières phases de codage, mais le chercheur ne saurait se limiter à celui-ci (Boyatsis, 1998). Il est ici intéressant de souligner l'apport de grilles existantes permettant de créer un codage tiré de quelques théories : ainsi les approches émotionnelles, narratives ou dramaturgiques renvoient à des théories existantes permettant au chercheur d'établir une véritable grille de codage en amont de son travail sur le terrain. Néanmoins, la multiplication de ces grilles préexistantes amène

aussi le chercheur à avoir des préconceptions sur son travail de recherche, et parfois même sur ses résultats.

Figure 5. Exemples de codage de premier ordre et de second ordre

Codage de premier ordre	Méthodes usuelles	Codage par attribut	Caractéristiques de la source
		Codage d'ampleur	Focus sur la fréquence ou l'intensité
	Méthodes élémentaires	Codage <i>in vivo</i>	Reprise littérale du verbatim
		Codage processuel	Focus sur les actions
		Codage ouvert	Démultiplication des codes
Codage de premier ordre	Méthodes affectives	Codage des émotions	Focus sur les émotions exprimées
		Codage des valeurs	Focus sur les valeurs, croyances, normes ou visions du monde
		Codage d'évaluation	Focus sur des aspects positifs et/ou négatifs
	Méthodes littéraires	Codage dramaturgique	Grille fondée sur l'approche théâtrale
		Codage mythologique	Grille fondée sur les légendes et les mythes
		Codage narratif	Grille fondée sur la narration d'une histoire
	Méthodes procédurales	Codage culturel	Approche ethnographique
		Codage causal	Approche fondée sur les causes du phénomène étudié
Codage de second ordre	Codage thématique	Approche fondée sur l'identification de thèmes et de descriptions que le chercheur utilise pour créer toutes sortes de taxinomies autour de son corpus de données	
	Codage centré	Identification des catégories les plus saillantes et porteuses de sens	
	Codage axial	Identification des liens entre les codes et/ou catégories	
	Codage théorique	Identification de la ou les catégorie(s) principales	

Source : adapté de Saldaña (2013)

Un véritable travail de codage ne pourra pas non plus prendre place sans la réalisation de mémos. Comparable à un véritable « journal de bord », le mémo est un mini rapport ou une mini analyse fondée sur une critique, une interprétation du chercheur à propos de ce qui émerge de ses données et qui amène à une théorisation (Saldaña, 2013). En effet, la nature même de la recherche qualitative nécessite une véritable réflexion et interprétation du chercheur sur ses données. Si la rédaction de mémos est principalement associée à une démarche de recherche de type *grounded theory* (Glaser & Strauss, 1967 ; Glaser, 1978 ; Strauss & Corbin, 1990), toute approche de codage peut bénéficier de l'utilisation de mémos pour aboutir à une véritable abstraction. L'utilisation de mémos permet en effet d'améliorer la réflexivité du chercheur qu'il n'aurait probablement pas atteint sans la mobilisation de ces derniers.

Le mémo est la véritable « porte d'entrée » entre théorie et données. Les mémos constituent un mode de conservation et de progression des pensées du chercheur ; le chercheur va en écrire à tout moment, du recueil de ses données à la production de son document final. Le fait de prendre du recul sur le codage et de « penser » sur le processus amène le chercheur à avoir une pensée davantage « critique » sur ses données. Les mémos peuvent être de véritables résumés des résultats clés de la recherche ou encore des commentaires ou des réflexions sur des aspects particuliers de la recherche. Glaser (1978, 2014) insiste sur le fait de considérer le mémo comme une priorité pour assurer la rétention de ses idées, réflexions et interprétations ; il considère d'ailleurs le mémo comme une « note conceptuelle ».

Les mémos sont de nature différente et peuvent représenter quelques lignes, voire parfois des pages entières. Nombre d'auteurs proposent des guides de rédaction des mémos (voir par exemple Glaser, 1978, 2014 ; Strauss & Corbin, 1990 ; Charmaz, 2006). Si ces différentes méthodes peuvent effectivement aider le chercheur à mieux comprendre l'intérêt des mémos dans sa recherche et leur construction, nous souscrivons aux propos de Charmaz (2006), conseillant d'écrire des mémos selon la manière qui convient le mieux au chercheur. Strauss et Corbin (1990) distinguent toutefois trois formes essentielles de mémos : opérationnels, de codage et théoriques. Les mémos opérationnels

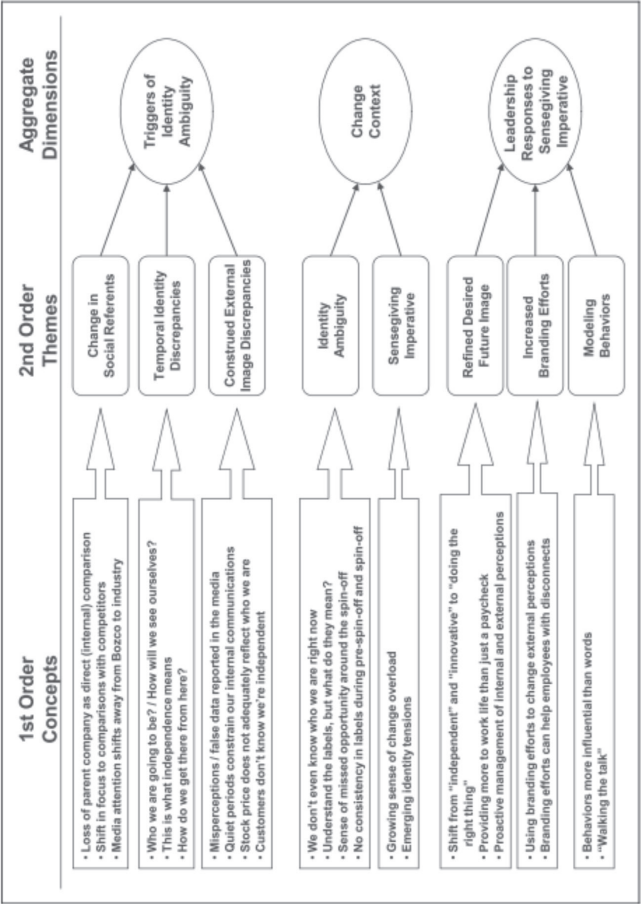
renvoient aux choix du chercheur dans sa recherche. Les mémos liés au codage renvoient à l'extrapolation du chercheur sur ses codes, et la progression vers la production de catégories. Enfin, les codes théoriques – souvent présentés comme une démarche analytique dans la littérature liée à la *grounded theory* – amènent à une plus grande abstraction, en s'interrogeant sur les relations, les explications et les interprétations émergeant des données pour arriver à un niveau conceptuel plus élevé et plus abouti.

4. AVANT DE PASSER À L'ACTION : COMMENT RENDRE COMPTE DU CODAGE ?

La difficulté rencontrée par tout chercheur en recherche qualitative est véritablement de « rendre compte » du processus de codage. Depuis quelques années, un véritable standard est apparu dans les revues académiques de haut niveau. Ce standard, connu sous le nom de « méthodologie à la Gioia ». Gioia, Corley et Hamilton (2013) reprennent le principe du codage de premier ordre et de second ordre pour véritablement montrer au lecteur la progression de la réflexion dans la recherche. Selon les auteurs, l'idée de cette approche est de renforcer la rigueur et la transparence d'une recherche inductive. Si la méthodologie développée par Gioia ces 25 dernières années reste une approche interprétative fondée sur la *grounded theory* (Gehman *et al.*, 2017), la manière de restituer le codage nous semble particulièrement intéressante et à mobiliser dans de nombreuses recherches destinées à la construction théorique. Elle se présente en deux temps.

Dans un premier temps, le chercheur construit sa « structure de données » (*data structure*, en anglais) (Figure 6). En commençant pas être au plus près du terrain, le chercheur travaille sur ces codes de premier ordre. Si Gioia parle de « constellation de codes de premier ordre », il est également fondamental de veiller à garder un nombre gérable de codes selon l'auteur (voir Gehman *et al.*, 2017). Il est vrai que la tentation dans une opération de codage ouvert est de démultiplier les codes *in vivo* et thématiques. Le codage de second ordre permet ensuite de progresser vers des catégories. Il peut s'agir de thèmes, mais ces derniers doivent suggérer des concepts qui expliquent le phénomène observé.

Figure 6. La structure des données ou des codes de premier ordre à l'abstraction

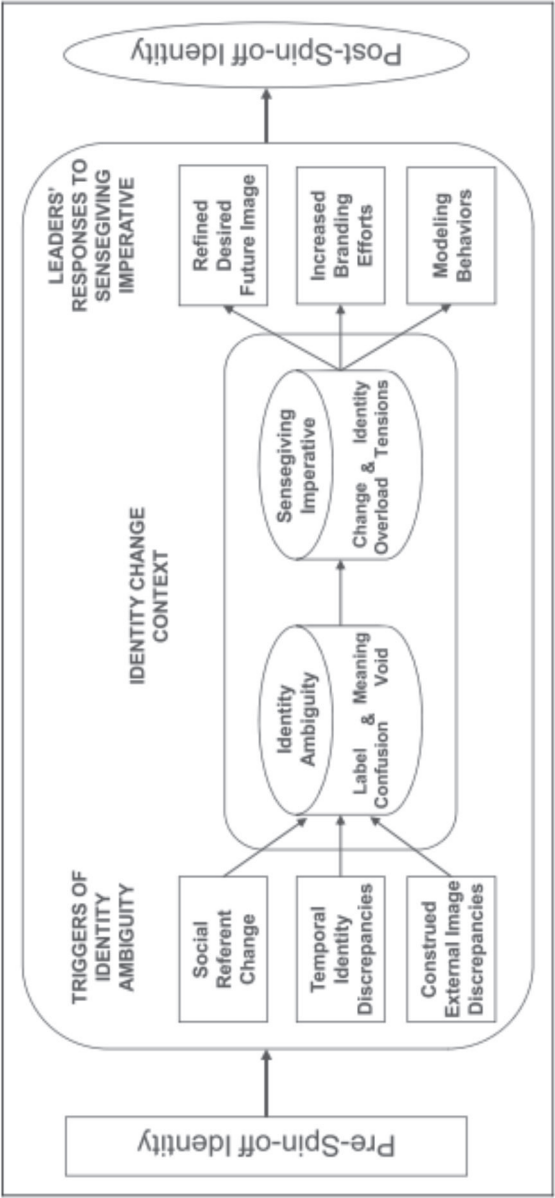


Source : Corley et Gioia (2004, p. 184)

Lorsque les codes de premier et de second ordre ont été assemblés, la base de la structure de données est établie. La structure des données reste indispensable. Cette étape reste le moment pivot de la recherche, car elle montre la progression des données recueillies au codes de premier et de second ordre, ce qui constitue une étape fondamentale dans la rigueur d'une recherche qualitative (Gehman *et al.*, 2017). Cependant, si la figure 6 montre clairement la progression de la recherche dans une démarche d'abstraction, elle ne montre pas encore les liens entre les concepts. Or ces derniers demeurent fondamentaux dans la conceptualisation de la recherche (Dumez, 2016). Aussi importante que soit la structure des données, elle ne constitue qu'une photographie statique éditée à un instant *t*. Aussi, pour comprendre le phénomène étudié, il reste à articuler les liens entre les concepts clés. C'est effectivement la relation entre les concepts émergents qui permettent de générer une théorie (Corley & Gioia, 2011) ; et c'est justement cette articulation entre données et théorie que la « méthode Gioia » s'attache à illustrer. La figure 7 montre le modèle final en montrant les liens entre les catégories et/ou codes de second degré.

L'écriture devient une véritable histoire mettant en exergue le développement de nouveaux concepts (ou de concepts dans un nouveau contexte) avec ce que Gioia *et al.* (2013) appellent « une découverte théorique ». L'objectif reste bien de montrer aux lecteurs l'émergence de ces nouveaux concepts ou de nouvelles perspectives sur les concepts existants.

Figure 7. Modèle émergent à partir de la structure de données



Source : Corley et Gioia (2004, p. 185)

CONCLUSION

La qualité d'une recherche qualitative dépend avant tout de la qualité de son codage. L'analyse des données n'est donc pas à prendre à la légère et il incombe au chercheur de prendre conscience de processus comme la catégorisation, l'abstraction ou encore la production de mémos. Si la phase de codification permet de générer un grand nombre de codes et/ou de thèmes, la phase de catégorisation implique de multiples interprétations de la part du chercheur et favorise inéluctablement l'abstraction. Le travail du chercheur est donc de rentrer dans un processus réflexif – notamment à travers la rédaction de mémos – qui lui permet de prendre progressivement de la distance avec son corpus de données pour arriver à une conceptualisation de sa recherche.

Il existe nombre de méthodes de codage que ce soit de premier ordre ou de second ordre. Chaque auteur a finalement sa propre interprétation, et chaque méthode a parfois un lexique qui lui est propre. L'objectif de ce chapitre n'était pas de proposer une liste exhaustive de ces méthodes, mais plutôt de proposer un début de réflexion et un éclairage sur la diversité des méthodes de codage conduisant le chercheur à exercer sa créativité dans cette étape cruciale de sa recherche. Force est de constater qu'il n'existe pas de stratégie optimale du codage du matériau de recherche : le type de codage ou encore la combinaison des méthodes de codage dépendent avant tout de la nature et de volume de matériau de recherche (Ayache & Dumez, 2011). Il convient donc au chercheur de faire les choix appropriés et surtout... de revenir aux sources.

Références citées

- Allard-Poesi, F. (2003), Coder les données, in Y. Giordano (dir.), *Conduire un projet de recherche : une perspective qualitative*, Caen, Éditions EMS, coll. « Les essentiels de la gestion », p. 245-290.
- Ayache, M. & Dumez, H. (2011), Le codage dans la recherche qualitative une nouvelle perspective ?, *Le Libellio d'Aegis*, vol. 7, n° 2, p. 33-46.
- Boyatzis R. (1998), *Transforming Qualitative Information: Thematic Analysis and Code Development*, Thousand Oaks, CA, Sage Publications.

- Charmaz, K. (2006), *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*, Londres, Royaume-Uni, Sage Publications.
- Coffey, A. & Atkinson, P. (1996), *Making Sense of Qualitative Data: Complementary Research Strategies*, Thousand Oaks, CA, Sage Publications.
- Carley, K. (1991), A theory of group stability, *American Sociological Review*, vol. 56, n° 3, p. 331-354.
- Corley, K. & Gioia, D. (2004), Identity ambiguity and change in the wake of a corporate spin-off, *Administrative Science Quarterly*, vol. 49, n° 2, p. 173-208.
- Corley, K. & Gioia, D. (2011), Building theory about theory building: What constitutes a theoretical contribution ?, *Academy of Management Review*, vol. 36, n° 1, p. 12-32.
- Dumez, H. (2016), *Méthodologie de la recherche qualitative : Les questions clés de la démarche compréhensive*, Paris, Vuibert.
- Gehman, J., Glaser, V.L., Eisenhardt, K.M., Gioia, D., Langley, A. & Corley, K. G. (2017), Finding theory-method fit: A comparison of three qualitative approaches to theory building, *Journal of Management Inquiry*, vol. 27, n° 3, p. 284-300.
- Gioia, D. A., Corley, K. G. & Hamilton, A. L. (2013), Seeking qualitative rigor in inductive research: Notes on the Gioia methodology, *Organizational Research Methods*, vol. 16, n° 1, p. 15-31.
- Glaser, B. (1978), *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory*, Mill Valley, CA, Sociology Press.
- Glaser, B. (2012), No Preconception: The Dictum, *Grounded Theory Review*, vol. 11, n° 2, p. 1-6.
- Glaser, B. (2014), *Memoing: A Vital Grounded Theory Procedure*, Mill Valley, CA, Sociology Press.
- Glaser, B. & Strauss, A. (1967), *The Discovery of Grounded Theory*, Chicago, IL, Adeline.
- Lazarsfeld, P. (1967), *The Uses of Sociology*, New York, Basic Books.
- Miles, M.B. (1979), Qualitative data as an attractive nuisance: The problem of analysis, *Administrative Science Quarterly*, vol. 24, n°4, p. 590-601.
- Miles, M.B. & Huberman, A.M. (1994), *Qualitative Data Analysis: An Expanded Sourcebook*, Thousand Oaks, CA, Sage Publications.
- Paillé, P. & Mucchielli, A. (2005), *L'analyse qualitative en sciences humaines et sociales*, Paris, Armand Colin.
- Peirce, C. S. (1974), *Collected Papers of Charles Sanders Peirce*, vol. 5, Harvard, MA, Harvard University Press.

- Point, S. & Voynnet-Fourboul, C. (2006), Le codage à visée théorique, *Recherche et Applications en Marketing*, vol. 21, n° 4, p. 61-78.
- Saldaña, J. (2013), *The Coding Manual for Qualitative Researchers*, Thousand Oaks, CA, Sage Publications.
- Snow, C. & Thomas, J. (1994), Field research methods in strategic management: Contributions to theory building and testing, *Journal of Management Studies*, vol. 31, n° 4, p. 457-480.
- Strauss, A. & Corbin, J. M. (1990), *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Thousand Oaks, CA, Sage Publications.
- Tesch, R. (1990), *Qualitative Analysis: Analysis Types and Software Tools*, London, Falmer.
- Yin, R. (2017), *Case Study Research and Applications: Design and Methods*, 6th edition, Thousand Oaks, CA, Sage Publications.

Références pour en savoir plus

- Saldaña, J. (2013), *The Coding Manual for Qualitative Researchers*, Thousand Oaks, CA, Sage Publications.

L'ouvrage qui permet d'avoir un aperçu *in extenso* des techniques de codages possibles en recherche qualitative.

- Gehman, J., Glaser, V.L., Eisenhardt, K.M., Gioia, D., Langley, A. & Corley, K. G. (2017), Finding theory-method fit: A comparison of three qualitative approaches to theory building, *Journal of Management Inquiry*, vol. 27, n° 3, p. 284-300.

Un article récent issu d'un séminaire à l'*Academy of Management* où se confrontent les plus grands experts en recherche qualitative.