

Der folgende Text wird über DuEPublico, den Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt.

Diese auf DuEPublico veröffentlichte Version der E-Publikation kann von einer eventuell ebenfalls veröffentlichten Verlagsversion abweichen.

Pistor, Petra; Stammen, Karl-Heinz:

Outils et procédures pour le renforcement de la qualité au sein des établissements d'enseignement supérieur | Module 2

In: Collection de livres pour la formation sur l'assurance qualité interne

DOI: <http://dx.doi.org/10.17185/duepublico/43410>

URN: <urn:nbn:de:hbz:464-20170302-164513-4>

Link: <http://duepublico.uni-duisburg-essen.de/servlets/DocumentServlet?id=43410>

Lizenz:



Dieses Werk kann unter einer [Creative Commons Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/) Lizenz genutzt werden.



Petra Pistor et Karl-Heinz Stammen

Outils et procédures pour le renforcement de la qualité au sein des établissements d'enseignement supérieur

Collection de livres pour la formation sur l'assurance
qualité interne | Module 2

Solveig Randhahn et Frank Niedermeier (Éd.)

Avec le soutien financier du



Ministère fédéral de la
Coopération économique
et du Développement

Édition

Cette publication électronique fait partie de la collection de livres pour la formation sur l'assurance qualité interne qui a été également publiée sous forme de livre de poche (ISBN: 978-3-7345-7797-0) et qui est disponible dans des librairies dans le monde entier. Vous trouverez plus d'informations sur <http://www.trainiqa.org>

Auteurs : Petra Pistor and Karl-Heinz Stammen

Éditeurs : Solveig Randhahn and Frank Niedermeier

Relecture : Barbara Michalk, Frank Niedermeier, Solveig Randhahn

Traduction française : Translate 4 U - Alette Chaput - Emmanuel Pons

Édition : Première édition

Mise en page : Nikolaj Sokolowski, Randi Ramme

Maison d'édition : DuEPublico, Duisbourg/Essen, Allemagne

DOI: 10.17185/duepublico/43410



Copyright © 2017 Petra Pistor and Karl-Heinz Stammen

Ce livre est sujet à la licence de Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Merci de citer ce livre dans des présentations, formations, publications, etc. selon les normes scientifiques. Vous pouvez citer ce livre comme suit :

Pistor, P. & Stammen, K.-H. (2017). Outils et procédures pour le renforcement de la qualité au sein des établissements d'enseignement supérieur. Modul 2, Randhahn, S. & Niedermeier, F. (Éd.) *Collection de livres pour la formation sur l'assurance qualité interne*. Duisbourg/Essen : DuEPublico. Récupéré de: <http://dx.doi.org/10.17185/duepublico/43410>

Remerciements

Nos modules et livres de cours ont été préparés et rédigés dans un effort conjoint des universités de Duisburg-Essen et de Potsdam, sous l'égide du programme DIES (Dialogue sur les stratégies innovantes dans l'enseignement supérieur), dirigé par l'Office allemand d'échanges universitaires (DAAD) et la Conférence allemande des recteurs d'universités (HRK), avec le soutien financier du ministère fédéral de la Coopération économique et du Développement (BMZ). Nous en profitons pour remercier le programme DIES et l'ensemble des partenaires d'Afrique, d'Europe et d'Asie du Sud-Est associés au processus de développement et leur adressons notre plus sincère gratitude pour leur aide, sans laquelle ces modules et livres de cours n'auraient pas pu être rédigés.

Nous tenons également à adresser nos plus sincères remerciements aux partenaires associés au présent projet pour leur aide et leurs précieuses contributions.

- L'Autorité Nationale d'Assurance Qualité de l'Enseignement Supérieur (ANAQ-Sup) du Sénégal ;
- Le réseau d'assurance qualité de l'ANASE (AQAN) ;
- Le réseau universitaire de l'ANASE (AUN) ;
- L'Association des universités africaines (AAU) ;
- Le Conseil Africain et Malgache pour l'Enseignement Supérieur (CAMES) ;
- L'Association européenne pour la garantie de la qualité dans l'enseignement supérieur (ENQA) ;
- L'Agence nationale d'agrément (NAB), Ghana ;
- Le Conseil national de l'enseignement supérieur (NCTE), Ghana ;
- La Commission nationale des universités (NUC), Nigeria ;
- Le Centre régional pour l'enseignement supérieur et le développement de l'Organisation des ministres de l'Éducation d'Asie du Sud-Est (SEAMEO RIHED) ;
- L'Organisation des Nations Unies pour l'Éducation, la science et la culture (UNESCO) ;
- L'Université d'études professionnelles (UPSA), Ghana.

et tout particulièrement Prof Dr Shahrir Abdullah, Richard Adjei, Prof Dr Goski Bortiorkor Alabi, Prof Dr Bassey Antia, Prof Dr Arnulfo Azcarraga, Gudrun Chazotte, Assoc Prof Dr Tan Kay Chuan, Kwame Dattey, Prof Dr Ong Duu Sheng, Prof Zita Mohd. Fahmi, Mae Fastner, Assoc Prof Dr Nantana Gajaseni, Robina Geupel, Josep Grifoll, Juliane Hauschulz, Dr Pascal Hoba, Dato' Syed Hussein, Benjamin Jung, Prof Abdel Karim Koumare, Dr Vipat Kuruchittham, Prof Dr Chiedu Mafiana, Prof Dr Duwiejua Mahama, Barbara Michalk, Prof Dr Le Quang Minh, Nguyen My Ngoc, Johnson Ong Chee Bin, Concepcion V. Pijano, Prof Dr Philipp Pohlenz, Sonja Pohlmann, Dr Suleiman Ramon-Yusuf, Dr Sylvia Ruschin, Dr Chantavit Sujatanond, Dr Oliver Vettori et Marc Wilde.

Les auteurs

Petra Pistor

Centre pour le développement universitaire et le renforcement de la qualité (ZfH)

Université de Duisburg-Essen, Allemagne

petra.pistor@uni-due.de

<https://www.uni-due.de/zfh/>

Petra Pistor exerce actuellement le poste de sous-directeur-général du Centre pour le développement universitaire et le renforcement de la qualité (ZfH) à l'Université de Duisburg-Essen (UDE). Sa responsabilité ici incombe aux deux des cinq équipes du Centre : gestion des données et évaluation. Petra Pistor a entre autres travaillé sur un éventail de projets d'évaluation dans le domaine de l'enseignement, de la recherche et de la structure organisationnelle. Ses travaux portent actuellement sur la conception et la mise en œuvre de différents projets d'évaluation et de formations dans le domaine de la gestion de la qualité dans l'enseignement supérieur, en Allemagne mais aussi dans le cadre de la coopération internationale, en particulier en Afrique de l'Est et de l'Ouest. Elle travaille pareillement comme conseillère indépendante et formatrice dans des différents contextes des établissements d'enseignement supérieur.



Karl-Heinz Stammen

Centre pour le développement universitaire et le renforcement de la qualité (ZfH)

Université de Duisburg-Essen, Allemagne

karl-heinz.stammen@uni-due.de

<https://www.uni-due.de/zfh/>

Karl-Heinz Stammen possède un diplôme de deuxième cycle en sciences sociales et a suivi une formation d'enseignant. Depuis 2005, il est associé de recherche au Centre pour le développement de l'enseignement supérieur et le renforcement de la qualité (ZfH) de l'Université de Duisburg-Essen, où il dirige l'équipe en charge de la gestion des données.

Karl-Heinz Stammen coordonne notamment le panel d'étudiants, l'évaluation des cours et le suivi socioprofessionnel des diplômés de l'université. Ses travaux portent sur la conception et la mise en œuvre de différents projets de recherche des étudiants ainsi que des formations dans le domaine de l'assurance et renforcement de la qualité et aussi dans l'enseignement supérieur, particulièrement en Afrique de l'Est.



Glossaire

AQ	Assurance qualité
AAQ	Agents de l'assurance qualité
Cycle PDCA	Cycle PDCA (planifier-développer-contrôler-agir)
DeGEval	Gesellschaft für Evaluation (Société d'évaluation allemande)
EEES	Espace européen de l'enseignement supérieur
EES	Établissement d'enseignement supérieur
JCSEE	Comité conjoint sur les normes pour l'évaluation de l'enseignement
KIT	Karlsruher Institut für Technologie (Institut de Technologie de Karlsruhe)
SGQ	Système de gestion de la qualité
SSC	Science Support Center (Centre de soutien scientifique)
UDE	Université de Duisburg-Essen
UP	Université de Potsdam
ZfH	Centre pour le développement de l'enseignement supérieur et le renforcement de la qualité, Université Duisburg-Essen

Table des matières

Introduction au module	12
Introduction à la théorie et aux concepts de l'évaluation	14
1 Introduction à la théorie et aux concepts de l'évaluation...	15
1.1 Introduction	15
1.2 Évaluation — Définition du concept	16
1.3 Les paradigmes de l'évaluation	16
1.4 Les concepts de l'évaluation	18
1.5 Les normes d'évaluation	19
L'évaluation au sein des établissements d'enseignement supérieur ²⁴	
2 L'évaluation au sein des établissements d'enseignement supérieur	25
2.1 Les différentes formes de qualité	25
2.2 L'évaluation interne et externe	27
2.3 L'évaluation participative	29
2.4 Les phases d'une évaluation	31
2.5 Évaluation d'un établissement — Le cas de l'Université de Duisburg-Essen ...	34
2.6 Zoom sur l'évaluation de l'enseignement et de l'apprentissage	38
2.7 Études de cas comparatives de l'évaluation des cours	40
2.8 Évaluation des cours — Que demander et comment le demander	44
2.9 Les facteurs de succès des évaluations dans les établissements d'enseignement supérieur	46

Les instruments de la recherche empirique en sciences sociales en tant qu'outils pour l'évaluation 50

3	Les instruments de la recherche empirique en sciences sociales en tant qu'outils pour l'évaluation	51
3.1	La pertinence de la recherche empirique en sciences sociales pour l'évaluation	51
3.2	Finalités et approches de la recherche empirique en sciences sociales	53
3.3	Les stratégies de la recherche empirique en sciences sociales	56
3.4	Méthodes de recherche empirique en sciences sociales	59

Le processus de recherche empirique en sciences sociales : De la question de recherche à la collecte de données..... 62

4	Le processus de recherche empirique en sciences sociales : De la question de recherche à la collecte de données	63
4.1	Les phases idéales d'une étude empirique	63
4.2	Zoom sur les méthodologies d'une étude empirique en sciences sociales.	66
4.3	Zoom sur les méthodes de collecte de données.....	70
4.4	Échantillonnage et représentativité	74
4.5	Méthodes de collecte de données : Enjeux et difficultés.....	75
4.6	Lignes directrices pour le choix de la méthode de sondage appropriée	77

Mesure et opérationnalisation 80

5	Mesure et opérationnalisation	81
5.1	La collecte de données est un processus de longue haleine : De la conceptualisation à la mesure.	81
5.2	Exemple : Opérationnalisation du concept « Réussite des étudiants »	82
5.3	Les niveaux de mesure et les types d'échelle	85
5.4	Critères de la recherche en sciences sociales	87
5.5	Types d'erreurs courants en recherche évaluative	88

La formulation de questions et de la méthodologie	90
6 La formulation de questions et de la méthodologie	91
6.1 Fonctions et types de questions	91
6.2 La formulation de bonnes questions	94
6.3 Conception du questionnaire	98
6.4 Mesures pour l'obtention d'un taux élevé de réponse	99
Traitement et analyse des données :	
Des questionnaires aux résultats d'un sondage	102
7 Traitement et analyse des données :	
Des questionnaires aux résultats d'un sondage	103
7.1 Nettoyage et codification des données	103
7.2 Le livre de codes et le rapport d'enquête	104
7.3 Compte rendu des résultats d'un sondage	104
Bibliographie	108
Liste des tableaux	110
Liste des figures	111

Préface

Introduction au module

Prérequis

- Les apprenants connaissent bien les différentes notions de la qualité et les lacunes de la définition de la qualité ;
- ils connaissent le cycle PDCA ainsi que différents modèles pour définir la qualité et l'assurance qualité ;
- ils connaissent le contexte particulier des établissements sans but lucratif et notamment des établissements d'enseignement, contexte qui influence les concepts de la qualité et l'assurance qualité.

Objectifs du module

Le module 2 commence par définir et analyser le concept de « l'évaluation », en tenant compte du contexte particulier des établissements d'enseignement supérieur. S'agissant de la définition de l'évaluation, elle fait partie intégrante de la gestion de la qualité, à savoir la collecte continue d'informations pertinentes pour le contrôle et le renforcement, en quelque sorte le « C » du cycle PDCA (planifier-développer-contrôler-agir). Dans le présent module, les lecteurs prendront connaissance des lacunes éventuelles de l'évaluation dans l'enseignement supérieur et découvriront des astuces pratiques pour améliorer la qualité des processus d'évaluation. Ces points feront l'objet d'une réflexion à la lumière des normes nationales et internationales pour l'évaluation et l'assurance qualité.

Le présent module aborde également les connaissances de base de la méthodologie de recherche empirique en sciences sociales ainsi que les exigences et procédures techniques de certains instruments et processus spécifiques au domaine de la gestion de la qualité. Ces connaissances sont cruciales car les méthodes et outils de la recherche empirique en sciences sociales sont indispensables pour la collecte de données et informations dans le domaine de l'évaluation. Les éléments clés de la matière de ce module sont dès lors la conception précise et la collecte systématique de données qualitatives et quantitatives ainsi que l'analyse et l'interprétation des données.

Les agents de l'assurance qualité (AAQ) ont des tâches diverses à accomplir dans le domaine de l'évaluation: concevoir un processus d'évaluation adapté au cadre organisationnel et culturel d'un établissement d'enseignement supérieur, en tenant compte de la combinaison des évaluations avec d'autres outils d'AQ (comme la gestion des données) et de l'intégration dans un système d'AQ ; concevoir et mettre en œuvre des méthodes et outils en vue d'obtenir des informations pertinentes, etc.

Les AAQ font office de coordinateurs, de facilitateurs, de médiateurs, de spécialistes en sciences sociales (bien qu'ils aient rarement un diplôme dans cette discipline) et parfois même de psychologues (car la gestion de la qualité et l'évaluation sont des processus).



Au terme du présent module, vous devriez pouvoir...

- appliquer les connaissances acquises à propos des prérequis, paradigmes, normes et lacunes typiques des évaluations dans l'enseignement supérieur lorsque vous organiserez un processus d'évaluation au sein de votre propre établissement ;
- tenir compte des principes de base des méthodes empiriques de recherche qualitative et quantitative en sciences sociales lorsque vous concevrez vos propres instruments de collecte de données, comme les sondages et les entretiens et interpréterez vos propres données ;
- savoir quoi tenir en compte lorsque vous interprétez des données ;
- appliquer des principes de base en faisant des rapports de données ;
- appliquer ces instruments en fonction des caractéristiques et du cadre spécifique de votre établissement.

Chapitre 1

Introduction à la théorie et aux concepts de l'évaluation

1	Introduction à la théorie et aux concepts de l'évaluation	15
1.1	Introduction	15
1.2	Évaluation — Définition du concept	16
1.3	Les paradigmes de l'évaluation	16
1.4	Les concepts de l'évaluation	18
1.5	Les normes d'évaluation	19



Au terme de ce chapitre, vous devriez pouvoir...

- définir le terme « évaluation » et son rapport avec l'appréciation, l'analyse et la vérification ;
- décrire les processus d'évaluation en fonction de leurs paradigmes sous-jacents, intentions et perspectives d'analyse ;
- énumérer différentes normes d'évaluation et en tenir compte dans votre travail quotidien en tant qu'agents de l'assurance qualité.

1 Introduction à la théorie et aux concepts de l'évaluation

1.1 Introduction

Le terme « évaluation » est utilisé dans le monde entier, y compris dans le contexte de l'enseignement supérieur. Mais que représente-t-il dans des contextes différents et qu'implique la diversité des significations selon le contexte pour sa mise en œuvre ? Au début, le présent module proposera une définition générale du terme « évaluation » et introduira différents paradigmes, concepts et normes qui déterminent la nature des processus d'évaluation. Seront ensuite débattues les caractéristiques spécifiques et domaines de conflit potentiels au sein des établissements d'enseignement supérieur. Nous en déduisons des recommandations pratiques pour la mise en œuvre des auto-évaluations et des évaluations externes (également mentionnées comme évaluations par des spécialistes) en tant qu'instruments pour l'assurance qualité et le renforcement de la qualité.

À nos yeux, l'évaluation va au-delà d'une simple appréciation. Outre la collecte et l'interprétation de données, que l'on pourrait qualifier d'appréciation, les évaluations incluent avant tout la prise de mesures d'amélioration et donc un processus décisionnel.

Pour mener les évaluations, il est dès lors crucial de disposer d'informations pertinentes car celles-ci aident les facultés, gestionnaires de la qualité et dirigeants de l'enseignement supérieur dans leurs processus décisionnels. C'est pourquoi il est de la plus haute importance que ces informations soient exhaustives et exactes (valides). Pour la collecte d'informations pertinentes et valides, les évaluations ont recours à des méthodes de recherche empirique en sciences sociales, qui garantissent que les informations recueillies sont le reflet juste et exact de la réalité et peuvent dès lors former une base de discussion et favoriser la prise de mesures d'amélioration.

Les cellules d'assurance qualité au sein des établissements d'enseignement supérieur doivent donc idéalement disposer de personnel expérimenté dans le domaine de la théorie de l'évaluation, de la gestion de l'enseignement supérieur et des méthodes de recherche empirique en sciences sociales. Il importe en effet tout autant de connaître l'objectif dominant d'une évaluation et son lien avec les plans stratégiques (évaluation et gestion de l'enseignement supérieur) que les possibilités et les limites des instruments de collecte de données à utiliser à cet effet (recherche empirique en sciences sociales). En réalité toutefois, ces trois domaines de qualification doivent bien souvent être assurés par un seul membre du personnel, qui fait office d'agent de l'assurance qualité. Nous vous invitons donc à utiliser ce module pour vous familiariser avec les fondements de la théorie de l'évaluation et de la méthodologie de recherche empirique en sciences sociales.

1.2 Évaluation – Définition du concept

Dans le langage de tous les jours, nous qualifions d'évaluation (du latin « valere » = être fort, valoir) le processus d'appréciation d'une chose. Cette description ne donne toutefois aucune indication quant aux acteurs concernés, à l'entité à analyser, au degré de fondement scientifique du processus, etc.

L'intitulé « évaluation » peut englober un éventail de définitions et de concepts différents. Dans les pages suivantes sera proposée une classification simple des différentes approches du concept. Pour le présent module, nous conviendrons de la définition suivante :

Définition de l'évaluation et de l'appréciation

L'évaluation est la description et l'appréciation systématiques de phénomènes particuliers sur la base de critères explicites ou implicites. Elle appuie le processus décisionnel quant à la mise en place de mesures qui peuvent elles-mêmes faire l'objet de processus d'évaluation (voir le cycle PDCA dans le module 1).

Les évaluations ont recours à des méthodes de recherche empirique en sciences sociales, comme nous le verrons en détail dans les chapitres suivants. S'agissant du présent module, les évaluations sont entendues comme un processus englobant la collecte de données, l'interprétation de données et la prise de mesures d'amélioration. Bien souvent, les termes « évaluation » et « appréciation » sont utilisés indifféremment. Ce module les distingue néanmoins, à l'instar de Charles Secolsky et Brian Denison:

Appréciation = « ... collecte, analyse et interprétation d'informations liées à une question particulière ou à un résultat intéressant. »

Évaluation = « ... détermine la valeur ou l'efficacité d'une chose, bien souvent une sorte de programme. »

(traduit de Secolsky & Denison 2012, 18)

L'appréciation fait donc partie intégrante de l'évaluation.

1.3 Les paradigmes de l'évaluation

Les évaluations peuvent avoir des fonctions différentes et interconnectées. En d'autres termes, elles peuvent être menées selon différents paradigmes. En principe, elles ne suivent pas l'un de ces paradigmes individuellement, mais adhèrent plutôt à une série d'entre eux en mettant l'accent plutôt sur l'un ou sur l'autre. Les récents débats en la matière distinguent généralement quatre paradigmes de l'évaluation (Stockmann 2004) :

a) le paradigme du gain de connaissances — évaluation en vue d'un apport de connaissances

Les évaluations servent en premier lieu à générer des informations sur des phénomènes, entités ou processus particuliers, par ex. une étude menée dans le cadre d'un projet de recherche. L'évaluation n'est pas une fin en soi mais les informations fournies par ce biais serviront à l'évaluateur, par ex. la personne responsable ou le groupe cible d'un processus. Il pourrait ainsi être intéressant de déterminer si un programme (par ex. un programme d'étude) fonctionne correctement, quelles sont les exigences d'un groupe cible déterminé ou si

les mesures appliquées ont répondu aux exigences d'un groupe cible. Selon le paradigme du gain de connaissances, les évaluations servent à fournir un aperçu de la nature d'une chose. Après cela seulement, elles sont vues comme le point de départ d'une analyse et de la prise de mesures.

b) le paradigme du contrôle — évaluation en vue d'un contrôle

Les évaluations peuvent également fournir des informations sur le succès d'une mesure, d'un programme, etc. Elles peuvent permettre de déterminer si une activité a abouti au résultat escompté et pourquoi ce fut ou ce ne fut pas le cas. Le lien entre l'évaluation et le contrôle est donc simple. Dans ce cadre, nous envisageons le contrôle comme un processus de vérification du degré de réalisation d'un objectif en particulier, et non la vérification des prestations d'individus ou de groupes. Selon nous, une vérification des prestations contribuera difficilement à l'amélioration de la qualité.

c) le paradigme de la légitimation — évaluation en vue de documenter les résultats

Toute mesure ou tout programme poursuit généralement un objectif et devrait être axé sur les exigences et attentes des parties prenantes concernées (par ex. le groupe cible de la mesure ou l'initiateur d'un programme). Les évaluations permettent de générer des informations sur la réalisation d'objectifs et sur l'adéquation de mesures prises. Elles permettent en outre de documenter les ressources nécessaires et le processus dans son ensemble. Cette documentation peut inclure des renseignements sur la productivité et le succès d'une mesure et être transmise comme preuve aux parties prenantes concernées.

d) le paradigme du développement — évaluation comme point de départ du changement

Les évaluations apportent des connaissances (voir a). Mais la collecte d'informations et la génération de connaissances sur un thème ne sont pas des objectifs en soi. Les informations recueillies et les connaissances acquises peuvent en fait être le point de départ d'une analyse et permettre d'approfondir la compréhension (causalités et contexte) d'un sujet et de prendre des mesures en vue d'une amélioration. Les informations et données recueillies lors d'une évaluation peuvent être le fondement d'un dialogue entre différentes parties prenantes et la personne responsable d'une mesure. Ce dialogue donne lieu à une analyse des informations et de la base de données selon différentes perspectives et favorise la mise au point de mesures conjointes en vue d'une amélioration.

« Agendas cachés » — Les évaluations peuvent également être utilisées abusivement comme une manœuvre stratégique. Elles peuvent servir à justifier des décisions politiques déjà prises et il arrive qu'une évaluation doive essuyer ce type d'objection. Ces pseudo-évaluations ne poursuivent pas les objectifs susmentionnés mais sont réalisées en tant qu'outil à la mode pour le développement commercial et la gestion de l'enseignement supérieur. Elles sont donc un outil de marketing idéal pour un établissement ou pour certains acteurs. On peut estimer que ce type d'évaluation suivant le « paradigme stratégique » est un outil légitime de l'élaboration de politiques. Compte tenu des ressources habituellement allouées aux processus d'évaluation, on peut facilement réfuter cet argument. Les évaluations menées uniquement à des fins politiques ne justifient pas l'utilisation de ces ressources. *(Rossi, Lipsey, et Freeman 2004, 37)*

1.4 Les concepts de l'évaluation

Les évaluations
ex-ante,
intermédiaires
et ex-post

Après avoir décrit les paradigmes de l'évaluation au [sous-chapitre 1.3](#), nous allons à présent exposer les intermédiaires et ex- différentes procédures d'évaluation.

Lors de l'appréciation de mesures ou de séries de mesures (programmes), les évaluations peuvent poursuivre différents objectifs. Ceux-ci déterminent leur principe sous-jacent ainsi que le moment de leur mise en œuvre.

Ainsi, les évaluations peuvent appuyer le processus de planification d'une mesure en fournissant des informations sur le groupe cible ou sur les conditions cadres de cette mesure. Lorsqu'elles visent cette finalité, les évaluations sont menées avant le lancement d'une mesure planifiée et sont qualifiées d'évaluations ex-ante (du latin « au préalable »). Ces évaluations sont donc menées avant le lancement du programme ou de l'action à évaluer.

Lorsqu'elles accompagnent et décrivent le processus d'une mesure ou d'un programme, les évaluations sont qualifiées d'intermédiaires. Ces évaluations sont donc menées au cours de la mise en œuvre d'un programme ou d'une action.

Enfin, les évaluations dont le but est de générer des informations sur l'efficacité d'une mesure une fois cette mesure terminée sont qualifiées d'évaluations ex post (du latin « après les faits »). Ces évaluations sont donc menées au terme d'une action ou d'un programme.

Les termes *ex ante*, *intermédiaire* et *ex post* désignent le **moment** ou la **perspective d'analyse** (anticipative, accompagnatrice, rétrospective) d'une évaluation.

Les évaluations
formatives
et sommatives

Les termes *formatif* et *sommatif* désignent au contraire le concept sous-jacent, la nature d'une évaluation : Les Les évaluations formatives appuient le processus de la mesure à évaluer, par ex. par le biais d'appréciations à différents moments en cours de processus : **Les évaluations formatives sont en quelque sorte des « activités d'évaluation menées afin de fournir des informations qui favoriseront l'amélioration du programme »** (traduit de Rossi, Lipsey, et Freeman 2004, 63).

Les évaluations sommatives en revanche insistent sur le résultat d'un processus. Elles sont des

« activités d'évaluation menées afin de rendre un jugement récapitulatif de certains aspects critiques de l'exécution d'un programme, et ainsi, par exemple, définir si certains objectifs et finalités spécifiques ont été atteints. »

(traduit de Rossi, Lipsey, et Freeman 2004, 65)

La matrice suivante détaille les connexions entre ces deux modes de description :

Phase d'une mesure ou d'un programme	Moment / Perspective d'analyse	Type d'analyse	Concept / Nature de l'évaluation
Phase de Planification	Analyse ex-ante / anticipative	Analyse des conditions cadres et du groupe cible pour une action planifiée -> acquisition de connaissances sur les bases et la situation de départ d'une action planifiée	Évaluation pré-formative / formative : axée sur le processus, constitutive, constructive
Phase de mise en œuvre	Analyse intermédiaire / accompagnatrice	Nouvelle analyse des conditions cadres et du groupe cible pour une action planifiée analyse des premiers effets de l'action planifiée -> acquisition de connaissances à propos des éléments à adapter en vue de favoriser le succès de l'action	Évaluation formative / sommative (les deux sont possibles)
Phase d'impact	Analyse ex-post / rétrospective	Analyse des effets de l'action planifiée analyse de la rentabilité -> acquisition de connaissances à propos du succès de l'action	Évaluation sommative : concluante, récapitulative, axée sur les résultats / effets

Tableau 1 Concept et perspective d'analyse des évaluations (adapté de Stockmann 2002, 6)

1.5 Les normes d'évaluation

La professionnalisation actuelle de la recherche dans le domaine de l'évaluation, qui a débuté dans les années 1970 aux États-Unis et en Europe, a fait naître des institutions et des commissions qui s'intéressent aux bonnes pratiques dans les travaux d'évaluation. Ainsi, le Joint Committee on Standards for Educational Evaluation [Comité conjoint sur les normes pour l'évaluation de l'enseignement (JCSEE)¹] a mis au point et révisé continuellement une série de normes d'évaluation pour le domaine de l'enseignement, adaptées par d'autres associations également (voir, par ex., les normes d'évaluation de la DeGEval, la Société allemande de l'évaluation 2001).

Ces normes d'évaluation combinent des principes procéduraux et éthiques et appuient l'assurance qualité des évaluations mais guident également les évaluateurs dans leur travail. Le JSCEE formule les normes et lignes directrices suivantes pour des évaluations labellisées. Celles-ci se divisent en cinq groupes conceptuels (Yarborough et al. 2011) :

1 Mis sur pied en 1975, le Comité conjoint est une coalition de grandes associations professionnelles (comme l'AERA, l'Association américaine de recherche en éducation, ou encore le NCME, le Conseil national américain sur la mesure en éducation) soucieuses de la qualité de l'évaluation. Le Comité conjoint est hébergé au Center for Evaluation and Assessment de l'Université de l'Iowa.

Normes d'utilité		
<i>Les normes d'utilité visent à rendre les processus et les produits d'évaluation plus utiles pour répondre aux besoins des parties prenantes au programme.</i>		
U1	Crédibilité de l'évaluateur	Les évaluations doivent être menées par des personnes qualifiées, qui instaurent et maintiennent une crédibilité certaine dans le cadre de l'évaluation.
U2	Attention portée aux parties prenantes	Les évaluations doivent prêter attention à l'ensemble des individus et groupes associés au programme et concernés par son évaluation.
U3	Finalités négociées	Les finalités de l'évaluation doivent être identifiées et en permanence négociées en fonction des besoins des parties prenantes.
U4	Valeurs explicites	Les évaluations doivent clarifier et spécifier les valeurs individuelles et culturelles sous-tendant les finalités, processus et jugements émis.
U5	Pertinence des informations	Les évaluations doivent clarifier et spécifier les valeurs individuelles et culturelles sous-tendant les finalités, processus et jugements émis.
U6	Processus et produits porteurs de sens	Les évaluations doivent aboutir à des activités, descriptions et jugements qui stimulent les participants à redécouvrir, réinterpréter ou revoir leurs notions et comportements.
U7	Communication et comptes rendus opportuns et adaptés	Les évaluations doivent veiller à répondre aux besoins d'information continus de leurs multiples publics.
U8	Préoccupation pour les conséquences et l'influence	Les évaluations doivent promouvoir une utilisation responsable et adaptative tout en veillant à éviter les conséquences négatives non intentionnelles et les abus.

Tableau 2 Les normes d'utilité du JCSEE (adapté de JCSEE)

Normes de faisabilité		
<i>Les normes de faisabilité visent à renforcer l'efficacité et l'efficience des évaluations.</i>		
F1	Gestion de projet	Les évaluations doivent recourir à des stratégies efficaces de gestion de projet.
F2	Procédures pratiques	Les procédures d'évaluation doivent être pratiques et tenir compte du mode de fonctionnement du programme.
F3	Viabilité du cotexte	Les évaluations doivent reconnaître, suivre et équilibrer les intérêts et besoins culturels et politiques individuels et des groupes.
F4	Utilisation des ressources	Les évaluations doivent utiliser les ressources de manière efficace et efficiente.

Tableau 3 Les normes de faisabilité du JCSEE (adapté de JCSEE)

Normes de déontologie		
<i>Les normes de déontologie appuient les éléments appropriés, justes, légaux, corrects et équitables dans les évaluations</i>		
P1	Orientation sople et inclusive	Les évaluations doivent tenir compte des parties prenantes et de leurs communautés.
P2	Accords formels	Les accords d'évaluation doivent faire l'objet d'une négociation afin de rendre les obligations explicites et de prendre en considération les besoins, les attentes ainsi que le contexte culturel des clients et autres parties prenantes.
P3	Droits de l'homme et respect de ceux-ci	Les évaluations doivent être conçues et menées en vue de protéger les droits humains et juridiques et de préserver la dignité des participants et autres parties prenantes.
P4	Clarté et équité	Les évaluations doivent être compréhensibles et répondre de manière équitable aux besoins et objectifs des parties prenantes.
P5	Transparence et divulgation	Les évaluations doivent fournir à l'ensemble des parties prenantes une description complète des résultats, limites et conclusions, à moins que ce type de divulgation ne soit contraire aux obligations juridiques et déontologiques y afférentes.
P6	Conflits d'intérêts	Les évaluations doivent identifier et résoudre de manière ouverte et honnête les conflits d'intérêts réels ou perçus qui pourraient compromettre leur processus.
P7	Responsabilité fiscale	Les évaluations doivent rendre compte de l'ensemble des dépenses et respecter des procédures et processus fiscaux sains.

Tableau 4 Les normes de déontologie du JCSEE (adapté de JCSEE)

Normes de précision		
<i>Les normes de précision visent à améliorer la fiabilité et la véracité des déclarations, propositions et conclusions de l'évaluation, en particulier celles appuyant les interprétations et les jugements en matière de qualité.</i>		
A1	Conclusions et décisions justifiées	Les conclusions et décisions découlant de l'évaluation doivent être justifiées explicitement dans les cultures et contextes où elles ont des conséquences.
A2	Informations valides	Les informations découlant des évaluations doivent servir les finalités visées et appuyer des interprétations valides.
A3	Informations fiables	Les procédures d'évaluation doivent générer des informations suffisamment fiables et cohérentes pour répondre à leurs finalités.

Normes de précision <i>Les normes de précision visent à améliorer la fiabilité et la véracité des déclarations, propositions et conclusions de l'évaluation, en particulier celles appuyant les interprétations et les jugements en matière de qualité.</i>		
A4	Descriptions explicites du programme et du contexte	Les évaluations doivent documenter les programmes et leurs contextes au moyen de détails appropriés et de portée suffisante au regard de leurs finalités.
A5	Gestion de l'information	Les évaluations doivent recourir à des méthodes systématiques de collecte, d'analyse, de vérification et de stockage des informations.
A6	Méthodologies et analyses judicieuses	Les évaluations doivent recourir à des méthodologies et analyses techniquement adaptées à leurs finalités.
A7	Raisonnement explicite de l'évaluation	Le raisonnement de l'évaluation, qui permet de passer des informations et de l'analyse aux résultats, interprétations, conclusions et jugements doit être documenté de manière claire et exhaustive.
A8	Communication et comptes rendus	Les communications formulées dans le cadre de l'évaluation doivent avoir une portée adaptée et éviter tout malentendu, parti pris, distorsion ou erreur.

Tableau 5 Les normes de précision du JCSEE (adapté de JCSEE)

Les normes de responsabilité de l'évaluation <i>Les normes de responsabilité de l'évaluation encouragent une documentation adéquate des évaluations ainsi qu'une perspective méta-évaluative d'amélioration et de responsabilité pour les processus et produits d'évaluation.</i>		
E1	Documentation de l'évaluation	Les évaluations doivent documenter de manière exhaustive leurs finalités négociées et leurs méthodologies, procédures et données mises en œuvre, ainsi que leurs résultats.
E2	Méta-évaluation interne	Les évaluateurs doivent utiliser ces normes de même que d'autres normes adaptées en vue de mesurer la responsabilité de la conception de l'évaluation, des procédures employées, des informations recueillies et des résultats
E3	Méta-évaluation externe	Les sponsors, clients, évaluateurs et autres parties prenantes de l'évaluation du programme devraient encourager la mise en œuvre de méta-évaluations externes au moyen de ces normes ainsi que d'autres normes applicables

Tableau 6 Les normes de responsabilité de l'évaluation du JCSEE (adapté de JCSEE)

 **Questions et tâches à accomplir**

1. Selon quel paradigme les processus d'évaluation menés au sein de votre établissement d'enseignement supérieur ont-ils été menés ? Quels sont, selon vous, les éventuels atouts et inconvénients des différents paradigmes ?
2. Pensez-vous que les normes d'évaluation du JCSEE soient sensées et utiles ? Pourquoi (pas) ?
3. Quels problèmes la mise en œuvre des normes d'évaluation pourrait-elle entraîner pour les responsables qualité ?

Chapitre 2

L'évaluation au sein des établissements d'enseignement supérieur

2	L'évaluation au sein des établissements d'enseignement supérieur	25
2.1	Les différentes formes de qualité	25
2.2	L'évaluation interne et externe	27
2.3	L'évaluation participative	29
2.4	Les phases d'une évaluation	31
2.5	Évaluation d'un établissement — Le cas de l'Université de Duisburg-Essen	34
2.6	Zoom sur l'évaluation de l'enseignement et de l'apprentissage	38
2.7	Études de cas comparatives de l'évaluation des cours	40
2.8	Évaluation des cours — Que demander et comment le demander	44
2.9	Les facteurs de succès des évaluations dans les établissements d'enseignement supérieur	46



Au terme de ce chapitre, vous devriez pouvoir...

- expliquer les caractéristiques de l'évaluation interne, externe, de l'évaluation par des pairs et de l'auto-évaluation ;
- planifier une évaluation qui corresponde aux parties prenantes concernées, soit conforme aux phases traditionnelles d'un processus d'évaluation et tenant compte des obstacles habituels potentiels ;
- comparer les atouts et inconvénients des différentes études de cas d'évaluations dans le domaine de l'enseignement et de l'apprentissage ;
- décrire les facteurs différents d'évaluations réussies.

2 L'évaluation au sein des établissements d'enseignement supérieur

2.1 Les différentes formes de qualité

Comme nous avons pu le voir dans le premier module, la définition du concept de la qualité n'est pas chose aisée pour les organisations non lucratives, les organisations de service et, surtout, les établissements d'enseignement. Selon Nickel (2010), on observe différents types de qualité dans les établissements d'enseignement supérieur :

- **la qualité des objectifs** dans la recherche, l'étude et l'enseignement ainsi que dans le domaine des activités réalisées au bénéfice de la collectivité ;
- **la qualité des extrants** (performances) des programmes d'étude et des activités de recherche ;
- **la qualité des résultats** des programmes d'étude, activités de recherche et services au bénéfice de la collectivité ;
- **la qualité des processus internes** (pour atteindre les objectifs, extrants et résultats susmentionnés).

En outre, les établissements d'enseignement diffèrent d'autres types d'organisations de par certaines particularités inhérentes qui compliquent la définition et l'évaluation de la qualité (Heinke, 2012 ; Stockmann, 2006) :

- **Des objectifs vagues** : Si les autres types d'organisation poursuivent des objectifs clairs et transparents (les hôpitaux soignent, les entreprises produisent des biens et maximisent leurs profits, etc.), les établissements d'enseignement supérieur ne poursuivent que des objectifs vagues. L'objectif d'un établissement d'enseignement supérieur est-il l'épanouissement personnel des gens, le transfert des derniers résultats de la recherche à la communauté, la préparation de son public au marché de l'emploi, ou existe-il d'autres objectifs qu'un établissement d'enseignement supérieur peut poursuivre ?
- **La variété des parties prenantes** : Dans le domaine de l'enseignement supérieur, une variété de groupes cibles et d'acteurs majeurs ont leurs exigences en matière de qualité des produits et processus de l'enseignement supérieur. Les étudiants, leurs parents, l'État, les enseignants et les chercheurs ont des attentes différentes envers les établissements d'enseignement supérieur. Ils évalueront dès lors l'offre d'un EES à la lumière de critères implicites différents.
- **La participation du client** : La réussite des programmes d'enseignement dépend fortement de leurs destinataires et leur qualité ne peut dès lors être que partiellement maîtrisée par l'établissement d'enseignement.
- **Des technologies problématiques** : Contrairement aux fabricants de produits, les établissements d'enseignement peuvent difficilement, voire pas du tout, normaliser les « étapes de production » de l'enseignement. Ainsi, comment un processus d'enseignement et d'apprentissage fortement dépendant de la personnalité des enseignants et groupes d'apprenants en interaction peut-il être subdivisé en étapes distinctes et être normalisé ?

Différences entre les concepts de la qualité dans l'industrie/ l'économie et dans les établissements d'enseignement supérieur

- **Le lien avec le sujet** : Des experts de tout un éventail de domaines (organisations d'experts) coopèrent au sein des établissements d'enseignement supérieur. La définition de la qualité peut varier d'un sujet à un autre. Ainsi, une leçon qui consiste uniquement en la lecture d'ouvrages sera jugée de faible qualité en sciences naturelles mais de bonne qualité en sciences humaines.

Au vu de ces facteurs, il est extrêmement compliqué de déterminer si une des nombreuses activités d'un établissement d'enseignement supérieur peut être jugée de bonne qualité. L'évaluation d'un point, d'un programme ou d'un produit nécessite invariablement l'existence de critères et de normes en fonction desquelles une évaluation peut être menée. La définition de normes dans l'enseignement supérieur diverge de celle dans l'industrie et l'économie et n'est pas une tâche aisée, comme il ressort des explications suivantes.

Note : Dans ce contexte, un critère est la caractéristique d'un élément, d'une mesure ou d'un programme qui permet de poser un jugement sur sa qualité. L'étanchéité, par exemple, est un critère pour les vestes d'extérieur. Si celui qui la porte reste sec après avoir été arrosé de 100 litres d'eau, la veste est jugée bonne (sur le plan de l'étanchéité). La protection coupe-vent, l'isolation thermique, etc. sont autant d'autres critères de qualité pour les vestes d'extérieur.

Une norme détermine le niveau qu'un critère doit atteindre pour être jugé bon. Ainsi, le critère de l'étanchéité d'une veste d'extérieur pourrait être atteint si l'intérieur de la veste reste sec après 30 minutes de pluie intense. « 30 minutes sous une pluie intense » serait la norme pour le critère de l'étanchéité.

Critères et normes

La définition de normes dans l'enseignement supérieur est très rare (le terme « norme » est parfois utilisé abusivement). Appliquée au contexte de l'enseignement supérieur, la définition ci-avant d'une norme de qualité deviendrait : « La qualité d'un programme d'étude est bonne si le nombre d'étudiants par enseignant ne dépasse pas 30. » En réalité, la qualité dans le secteur de l'enseignement supérieur est plus souvent évaluée selon des critères ouverts. On trouve bien souvent des formulations de type « le rapport entre le nombre d'étudiants et d'enseignants est approprié » ou « les étudiants disposent de ressources documentaires suffisantes ». La définition, ou plutôt la traduction, des termes « approprié » et « suffisant » en entités observables ou mesurables dépend de l'observateur. En d'autres termes, pour pouvoir procéder à des appréciations valides et pertinentes de la qualité, les acteurs concernés doivent convenir de normes qu'ils estiment refléter la bonne qualité d'un élément, d'une mesure ou d'un programme. Il n'existe pas de norme universelle à cet effet. Les établissements d'enseignement supérieur définissent habituellement ces critères et ces normes en s'inspirant de leur utilisation par d'autres acteurs (d'autres établissements d'enseignement supérieur) et les adaptent ensuite à leurs propres conditions-cadres après débat avec les parties prenantes concernées.

Note : Les définitions de la qualité et les exigences des diverses parties prenantes peuvent également s'avérer contradictoires ! Si pour l'établissement, un taux de réussite élevé et des programmes d'étude courts peuvent être des critères de qualité, les étudiants, en revanche, estimeront peut-être de meilleure qualité des programmes d'étude qui prévoient une certaine liberté et un choix d'options (sans mentionner un diplôme à décrocher dans un délai habituel).

Et enfin, la définition de critères et de normes est également fonction de la perspective d'analyse et de la nature d'une appréciation ou d'une évaluation. Une comparaison des objectifs et des performances sur la base d'indices de performance clés est possible (évaluation selon le paradigme du contrôle) s'il s'agit d'une évaluation ex post du degré de réalisation d'un objectif. Si un établissement d'enseignement supérieur part, par exemple, de l'hypothèse que le nombre de participants à un séminaire donné est faible parce que le créneau horaire ne convient pas aux étudiants, déjà actifs sur le marché de l'emploi, on peut mesurer à nouveau le nombre de participants après modification de l'horaire du séminaire. Ce faisant, on peut déterminer si la mesure mise en œuvre (la modification de l'horaire du cours) a porté ses fruits (= plus de participants). Le critère (un parmi d'autres) déterminant la bonne qualité est, dans ce cas précis, le nombre de participants présents au cours.

Les évaluations menées au sein des établissements d'enseignement supérieur doivent généralement être efficaces et appuyer des processus de changement. Elles doivent veiller à ce que les décisions prises par la direction découlent d'informations judicieuses. Les données et renseignements recueillis dans le cadre d'une évaluation doivent être valides et fiables (voir [sous-chapitre 5.4](#)) pour pouvoir servir d'instrument de contrôle. Il est possible d'améliorer la validité et la fiabilité des données en suivant les règles de la recherche empirique en sciences sociales.

2.2 L'évaluation interne et externe

Les évaluations peuvent se subdiviser en évaluations internes et externes. Les évaluations internes sont des processus d'appréciation menés par des acteurs membres de l'établissement à évaluer. Si ces évaluateurs sont également actifs dans le domaine à évaluer, par ex. des membres de la faculté qui évaluent les processus et les produits de leur propre faculté, il s'agit d'auto-évaluations internes. Si l'appréciation est menée, par exemple, par une cellule AQ de l'établissement d'enseignement supérieur, l'évaluation est dite interne, mais n'est pas une auto-évaluation.

Les évaluations externes sont, pour leur part, menées par des spécialistes non issus de l'établissement.

Si les évaluations internes présentent l'avantage de nécessiter moins d'efforts et de fournir aux évaluateurs des connaissances internes plus complètes sur le sujet évalué, elles présentent également le risque d'un lien étroit entre l'évaluateur et l'évalué. Cela signifie que, dans certains cas, des membres du personnel évaluent les réalisations de leurs collègues, un processus délicat puisque tout le monde n'apprécie pas d'être jugé par ses pairs et une tendance à l'atténuation des points problématiques peut ainsi voir le jour. En outre, un engagement personnel dans des processus soumis à évaluation peut entraver le brainstorming en vue de définir des mesures d'amélioration potentielles.

Si l'évaluation est menée par un institut au sein de l'établissement, par ex. une agence AQ, elle peut faire naître des conflits potentiellement durables au sein de l'établissement. L'évaluation externe permet d'externaliser le conflit potentiel entre les évaluateurs et les personnes liées à l'élément ou au programme évalué et de ne pas le pérenniser au sein du cadre organisationnel. Les évaluations externes ont l'avantage d'être généralement menées par des évaluateurs indépendants et peuvent dès lors fournir de nouvelles perspectives

pour une évaluation sans entrave. Les évaluateurs externes ont, en revanche, besoin de beaucoup plus de temps pour dégager une appréciation finale exhaustive puisqu'ils doivent tout d'abord se familiariser avec le contexte et le cadre de l'élément à évaluer (Stockmann 2006, 263). Des acteurs externes n'auront jamais une compréhension aussi approfondie de l'organisation à évaluer qu'une personne active au quotidien au sein de l'organisation. Aussi est-il idéal de toujours combiner la perspective externe avec des informations issues d'acteurs internes à l'organisation (Stockmann / Meyer 2010, 81). La matrice suivante récapitule les atouts et les inconvénients de l'évaluation interne et de l'évaluation externe :

	Auto-évaluation interne	Évaluation interne par une agence d'AQ ou similaire	Évaluation externe par des pairs
Atouts	■ Rapide, moins de travail	■ Du travail, un certaine indépendance	■ Indépendance
	■ Compréhension détaillée	■ Maîtrise élevée de la méthodologie, compréhension relativement détaillée	■ Maîtrise élevée de la méthodologie
	■ Rentabilité	■ Suivi guidé	■ Crédibilité élevée
			■ Plus grand potentiel de réforme
Inconvénients	■ Faible maîtrise de la méthodologie	■ Conflit interne potentiel	■ Résistance possible
	■ Manque de neutralité et de distance par rapport à l'élément évalué		■ Coût plus élevé
	■ Aucune vision organisationnelle		■ Aucun suivi guidé
	■ Faible potentiel de réforme		

Tableau 7 Atouts et inconvénients potentiels des évaluations internes et externes (tableau des auteurs)

Le tableau 7 ne donne qu'un aperçu des atouts et inconvénients potentiels. Les atouts et inconvénients d'un type d'évaluation peuvent valoir pour un autre et vice-versa, en fonction de la conception de l'évaluation et de la structure de l'organisation au sein de laquelle a lieu l'évaluation.

Note : Dans les établissements d'enseignement supérieur, à tous les niveaux et dans tous les domaines (enseignement et apprentissage, recherche, départements ou universités entières), on observe aujourd'hui une combinaison d'auto-évaluation interne et d'appréciation externe. L'approche dite de la révision par des pairs / spécialistes se révèle particulièrement adaptée, surtout pour les établissements d'enseignement supérieur caractérisés par leur lien avec le sujet et souvent qualifiés d'« organisations d'experts » (voir plus haut). Une évaluation éclairée par des pairs prévoit un processus d'auto-évaluation avant que les pairs ne soient invités à apprécier l'élément en question. Les conclusions de l'auto-évaluation sont rassemblées dans un rapport d'auto-évaluation qui sert de base de renseignement aux pairs chargés de l'évaluation. Ces pairs basent leur appréciation sur les informations contenues dans ce rapport et sur les entretiens menés avec les parties prenantes concernées lors de leur visite du site.

Tous les participants à une évaluation ne sont pas familiarisés avec les concepts sous-jacents et processus pratiques des évaluations. Ils doivent parfois être formés afin de pouvoir remplir leur rôle de sondé ou de pair-évaluateur. Pour la plupart des évaluations, les pairs évaluateurs sont formés à la conduite d'entretiens, à la rédaction de rapports, etc.

2.3 L'évaluation participative

Contrairement à la recherche, l'évaluation dans le domaine de la recherche sociale appliquée vise à fournir des informations utiles à des fins de développement. C'est tout particulièrement le cas lorsque les évaluations suivent le paradigme du développement (Stockmann 2006, 262).

Les mesures d'amélioration découlant d'une évaluation touchent parfois différents acteurs d'une organisation. Ces mesures doivent dès lors être d'autant plus porteuses de sens pour eux. La pertinence des évaluations pour les acteurs concernés est un facteur majeur de réussite du processus, comme l'expliquent également les normes d'utilité du [sous-chapitre 1.3](#).

On peut rendre le processus d'évaluation pertinent pour les acteurs en y intégrant leurs questions et conseils et en leur permettant de contribuer activement au processus de planification (voir U2 Attention portée aux parties prenantes). Comme les acteurs internes connaissent davantage les détails et le contexte, leur contribution influe également sur la qualité du processus même en aidant les évaluateurs à encore mieux comprendre l'élément à évaluer.

Enfin, la participation au processus de planification d'une évaluation donne un bon aperçu des processus d'une organisation et permet d'acquérir des connaissances sur la théorie de l'évaluation. Elle peut dès lors être perçue comme une qualification supplémentaire pour les participants.

Il est conseillé d'associer activement les membres concernés du personnel de l'organisation ou de l'unité à évaluer à la phase de planification d'une évaluation et à la réflexion sur les résultats et la planification de mesures d'amélioration. La collecte de données relève toutefois invariablement d'un évaluateur qualifié,

expérimenté en recherche empirique en sciences sociales, et ce afin de garantir la qualité scientifique de la collecte de données (voir U1 Crédibilité de l'évaluateur). Lors de la phase de collecte de données, le rôle des personnes concernées par l'évaluation doit se limiter à l'apport d'informations.

Lors de la planification d'une évaluation soutenue par des spécialistes en évaluation, il est judicieux de mettre sur pied un groupe de travail qui devrait idéalement se composer de membres issus de l'ensemble des groupes de parties prenantes concernés, de sorte que les objectifs de l'évaluation et les questions puissent être débattus et, idéalement, convenus (U3 Finalités négociées).

Dans l'intégralité du processus, nous considérons la transparence comme un des éléments d'une approche participative de l'évaluation. Une communication opportune et exhaustive sur l'évolution du processus et ses conclusions aux parties prenantes non membres du groupe de travail susmentionné nous apparaît dès lors comme essentielle (voir U7 Communication et comptes rendus opportuns et adaptés).

Même si nous préconisons une approche participative de l'évaluation, il convient de bien réfléchir à l'opportunité d'une participation d'acteurs à l'évaluation, au degré de cette participation et aux acteurs à y associer. Le choix des parties prenantes à associer au processus doit tenir compte des aspects suivants :

- La participation n'exclut en rien des responsabilités claires - La gestion de la qualité et l'évaluation resente une responsabilité des supérieurs. Dans le cadre d'un groupe de travail, il est toujours utile de chercher à dégager un consensus. Il est parfois même judicieux d'associer à une évaluation les parties prenantes ouvertement opposées à ce processus. Si aucun consensus n'est possible au sein d'un groupe de travail, la décision peut alors relever de l'initiateur ou du sponsor de l'évaluation, généralement un responsable (vice-recteur adjoint aux affaires académiques, etc.).
- « Trop de cuisiniers gâtent la sauce. » — Les parties prenantes potentielles ne doivent pas TOUTES être associées de la même manière à une évaluation participative. Ainsi, lors de la planification d'une évaluation avec un groupe de travail, il importe que ce groupe puisse fonctionner. D'expérience, nous pensons que ce groupe ne doit donc pas être composé de plus de 6 personnes. Lors de la sélection des membres de ce groupe de travail, il convient aussi de ne pas oublier que l'on choisit des personnes capables de représenter leur unité, groupe, etc., c'est-à-dire des personnes acceptées par les autres membres du groupe et qui connaissent bien l'élément à évaluer.
- Hiérarchies et rôles — Lors de la constitution d'un groupe de travail et la détermination de ses tâches, ne perdez pas de vue que le travail en équipe reste régi par les mêmes mécanismes que d'ordinaire. Ainsi, même si vous souhaitez voir l'ensemble des membres du groupe coopérer sur un pied d'égalité au sein d'un groupe de travail en charge de l'évaluation, les membres du personnel situés plus haut dans la hiérarchie prendront généralement la direction du groupe de travail. En conséquence, les autres membres de l'équipe pourraient ne pas faire part de leur véritable opinion afin d'éviter de critiquer le travail d'un supérieur, etc. Il est crucial d'enrayer immédiatement et ouvertement les conflits naissants au sein du groupe de travail afin de favoriser une coopération basée sur la confiance. Ce point est tout particulièrement important si des étudiants prennent part au groupe.
- La participation d'étudiants et d'anciens — D'expérience, nous pouvons dire qu'il est particulièrement difficile d'associer ces groupes d'acteurs à l'évaluation parce que les étudiants et les diplômés ne passent généralement pas beaucoup de temps au sein de leur (ancien) établissement d'enseignement supérieur en

raison de leurs engagements professionnels ou de leurs études. Le processus d'évaluation ne les intéresse parfois tout simplement pas. Les diplômés sont également parfois plus difficiles à joindre puisqu'ils ne sont plus présents au sein de l'établissement. Permettre un faible niveau de présence aux réunions à ces catégories de participants s'est cependant déjà révélé être un bon compromis, compte tenu de leur profil particulier.

- Les parties prenantes ne sont pas toutes des spécialistes — Comme mentionné ci-avant, il n'est pas avisé de confier à n'importe quel membre du groupe de travail la conception d'instruments de collecte de données et la planification de l'évaluation dans son ensemble car on ne pourra être assuré d'obtenir des résultats utiles que si la personne qui s'acquitte de ces tâches a une solide expérience dans le domaine de la théorie de l'évaluation et de la recherche empirique en sciences sociales.
- Participation réelle ou non-participation — Il est tout à fait possible de mener une évaluation sans la participation des acteurs concernés pour autant que ceux-ci soient tenus informés de tous les aspects pertinents. Parfois, une évaluation prévoit la participation fictive des parties prenantes. Celles-ci sont ainsi invitées à prendre part au processus mais les décisions cruciales sur les points et questions clés de l'évaluation sont prises en haut lieu (voir [sous-chapitre 1.3](#) sur les « agendas cachés »). Cette pratique peut entamer la confiance envers les supérieurs et les structures d'une organisation à long terme. Aussi recommandons-nous de n'associer les acteurs que si leurs contributions sont les bienvenues.

Note : En tant qu'agent de l'assurance qualité, il vous incombe de formuler des recommandations sur les acteurs et les parties prenantes à associer à l'évaluation et de veiller à ce que ceux-ci puissent contribuer avec fruit aux débats menés au sein d'un groupe de travail ainsi qu'au processus d'évaluation dans son ensemble.

2.4 Les phases d'une évaluation

Nous sommes d'avis que tout type d'évaluation peut se subdiviser en quatre phases, identiques aux phases idéales types d'un processus de recherche, comme nous le verrons au [sous-chapitre 4.1](#) : 1. Phase de planification, 2. Phase de collecte de données, 3. Phase de réflexion, 4. Phase de suivi (pour d'autres subdivisions du processus d'évaluation, veuillez consulter Stockmann 2006, 290, ou Røbken 2012, 73). En règle générale, les différentes phases du processus se chevauchent ou ont lieu simultanément. Ainsi, lors de la phase de collecte de données, il conviendra parfois d'interpréter les données déjà disponibles et de restructurer ou adapter les autres phases. Les sections suivantes détailleront les différentes phases de l'évaluation dans le contexte particulier des établissements d'enseignement supérieur.

I. La phase de planification : Au cours de cette première phase d'une évaluation, la tâche de l'agent de l'assurance qualité et d'un groupe de travail sur l'évaluation consiste à planifier le processus d'évaluation en tenant compte des caractéristiques spécifiques d'un établissement d'enseignement supérieur et en veillant à ce que le processus respecte les normes de faisabilité décrites au [sous-chapitre 1.3](#).

Cela signifie que tous les blocs de tâches et toutes les étapes du processus doivent être planifiés de façon systématique et réaliste en termes de délais par un coordinateur (l'AAQ), responsable de l'avancement du pro-

cessus (voir F1 Gestion de projet). La planification de l'évaluation doit être réfléchie et dès lors correspondre aux objectifs de l'évaluation et de l'élément évalué (voir F2 Procédures pratiques). En outre, la planification doit être effectuée d'un façon qui emporte l'adhésion de l'ensemble des parties prenantes en respectant au mieux leurs intérêts et perspectives respectifs en rapport avec ce processus de planification (F3 Viabilité contextuelle).

Les évaluations ne font généralement pas partie du travail quotidien des scientifiques, enseignants, membres du personnel administratif, réviseurs externes et étudiants. Elles constituent dès lors une charge de travail supplémentaire pour ces groupes d'acteurs. Aussi est-il crucial de faciliter l'évaluation en y allouant des ressources humaines et financières suffisantes pour pouvoir atteindre les objectifs visés (F4 Utilisation des ressources), tout en les limitant au minimum afin de garantir l'efficacité du processus.

Durant la phase de planification, il convient de débattre des points suivants, tous interconnectés, et de prendre une série de décisions afin de veiller au bon déroulement du processus d'évaluation:

- **L'élément à évaluer** : Avant de décider de l'élément à évaluer, il convient de se demander « Que va-t-on évaluer ? », et ce d'une manière critique et rigoureuse. On peut être tenté de regrouper au sein d'une même évaluation toute une série de questions à différents niveaux puisque cette évaluation est de toute manière déjà planifiée. Dans un souci d'économie des données et de réduction de la charge de travail supplémentaire, il est vivement recommandé de limiter l'évaluation à un élément bien défini. C'est d'autant plus justifié que cette limitation à un élément à évaluer bien défini permettra une étude plus approfondie.
- **L'objectif de l'évaluation** : Pour être certains de l'objectif réel d'une évaluation et ainsi pouvoir le faire connaître à l'ensemble des parties prenantes, les AAQ doivent veiller à répondre aux questions suivantes avant le début du processus : Quel est l'objectif de l'évaluation ? Doit-il s'agir d'une comparaison des objectifs et des performances au moyen d'une approche sommative ? Doit-on recueillir des informations qui serviront en premier lieu à informer le public des processus / projets en cours ? Ou l'évaluation doit-elle générer des connaissances qui permettront d'aboutir à des mesures d'amélioration ? Comme on peut le voir, on en revient toujours à la question du paradigme de l'évaluation (voir [sous-chapitre 1.1](#)).
- **Les instruments à mettre en œuvre** : Une appréciation sensée et valide nécessite la collecte d'informations aussi exhaustives que possible permettant d'étayer un jugement. En règle générale, il est recommandé de dégager une compréhension détaillée de l'élément à évaluer selon différentes perspectives bien définies (triangulation). En votre qualité d'AAQ, vous devez tout d'abord recueillir des informations sur les données et informations déjà disponibles au sein de votre établissement d'enseignement supérieur afin d'éviter tout travail inutile. Pour choisir les instruments supplémentaires appropriés, posez-vous les questions suivantes : Quelles sont les données et informations (supplémentaires) nécessaires ? Quelles sont les personnes ou les entités à même de fournir ces données et informations ? Comment ces données et informations devraient-elles être obtenues ? Convient-il d'appliquer une stratégie de recherche qualitative ou quantitative (voir [sous-chapitre 3.3](#) pour les atouts et les inconvénients de chaque stratégie de recherche) ? Le choix des instruments à utiliser dépend par ailleurs des conditions-cadres de l'évaluation. Par exemple, il peut être judicieux de se demander si les prérequis techniques permettent un sondage en ligne ou, de manière plus générale, si les ressources financières et le temps sont suffisants pour mener une évaluation et si le personnel en charge de l'évaluation dispose des qualifications requises. Pour le choix des instruments auxquels recourir dans le cadre d'une évaluation, comme lors de la planification de celle-ci,

tâchez de respecter la règle suivante : Autant que nécessaire, aussi peu que possible !

- **Les évaluateurs et autres parties prenantes** : En fonction de l'élément à évaluer, des conditions-cadres structurelles d'un établissement d'enseignement supérieur (L'EES dispose-t-il d'AAQ ? Dispose-t-il d'une cellule en charge de l'assurance qualité ?) et des instruments à mettre en œuvre, il convient de décider de la personne qui contribuera à l'évaluation et de quelle manière cela doit se faire. L'évaluation doit-elle être menée en interne ou des spécialistes externes seront-ils invités (voir [sous-chapitre 2.2](#)) ? Quel sera le niveau de participation pour cette l'évaluation (voir [sous-chapitre 2.3](#)) ?
- **Calendrier** : En fin de phase de planification d'une évaluation, un plan d'action du projet (PAP) doit être mis en place. Ce PAP (voir Module 1) doit mentionner les phases de l'évaluation ainsi que le calendrier de la collecte de données, les acteurs en charge des activités programmées et les autres parties prenantes à associer au projet.

Note : La « triangulation des données, également appelée la triangulation des sources de données, décrit l'utilisation de sources de données multiples au sein d'une même étude, et ce à des fins de validation. Denzin (1978) recense trois types de triangulation des données : le temps, l'espace et la personne. Ces types de triangulation des données partent du principe que les données peuvent avoir un degré de validité variable selon le moment où elles sont recueillies, selon les personnes associées au processus de collecte et selon le cadre dans lequel elles sont recueillies (Begley 1996) » (cité dans Hussein 2009, 3). La triangulation permet de renforcer la validité des données ou informations parce que les faiblesses (angles morts) d'un instrument peuvent être compensées par les points forts d'un autre.

II. La phase de collecte de données : La phase de collecte de données est la partie centrale d'une évaluation. Au cours de cette phase, les données et informations nécessaires pour répondre aux questions de l'évaluation sont recueillies et traitées afin de les rendre compréhensibles pour les parties prenantes non spécialisées dans la recherche empirique en sciences sociales.

La phase de collecte de données peut se subdiviser en quatre étapes :

- a) l'opérationnalisation et la mise sur pied des instruments ;
- b) la collecte de données et
- c) le traitement des données ;
- d) le compte rendu

(voir [sous-chapitre 4.1](#) pour de plus amples détails sur ces étapes).

Cette phase est donc l'étape du processus d'une évaluation, au cours de laquelle se mettent en place les méthodes issues de la recherche empirique en sciences sociales. La conception et le calendrier de ces étapes dépendent largement de la conception de l'évaluation en général.

III. La phase de réflexion : Les données recueillies sont-elles traitées de manière à permettre aux non-initiés de les comprendre ? Il convient d'entamer une réflexion systématique à propos de l'élément à évaluer sur la base de ces données (voir également la phase de compte rendu et de diffusion, [sous-chapitre 4.1](#)). À ce stade, il ne suffit pas, selon nous, de soumettre les données traitées et les résultats aux personnes responsables

ainsi qu'aux décideurs. Il faut également entamer un processus de réflexion, dont vous serez le modérateur, en votre qualité d'AAQ. La réflexion sur ces données sous la forme d'un débat guidé favorise grandement la qualité des appréciations basées sur les données. En général, il est avisé d'étaler cette réflexion sur un ou plusieurs ateliers. Au cours de ce(s) atelier(s), l'AAQ présentera les informations à disposition de manière compréhensible pour les parties prenantes concernées. Il y sera répondu aux questions et les données feront l'objet d'une réflexion et d'un débat. Des déclarations finales et des mesures d'amélioration (c'est particulièrement le cas pour les évaluations suivant le paradigme du développement) seront ensuite formulées et intégrées à un rapport d'évaluation. Afin de pouvoir mener cette réflexion, les parties prenantes concernées doivent recevoir les données brutes bien avant le début de l'atelier.

IV. La phase de suivi : Les évaluations doivent produire des effets, c'est-à-dire produire des résultats directement liés à l'objectif du processus. Dans le cas des évaluations sommatives, souvent utilisées comme justification pour certains groupes d'intérêt, l'objectif de l'évaluation peut être atteint grâce à la publication d'un rapport d'évaluation. Les évaluations suivant le paradigme du développement, les plus courantes au sein des établissements d'enseignement supérieur, doivent stimuler le changement et être efficaces au sein des structures de l'organisation. Dans ce cas-ci, il ne suffit généralement pas de commenter les résultats de l'évaluation et de soumettre les mesures d'amélioration décidées sur cette base. Il importe, selon nous, de rendre ces mesures d'amélioration contraignantes, par exemple au moyen d'accords entre la faculté (ou une autre sous-unité de l'organisation) et le vice-recteur ou, dans le cas de l'évaluation d'un cours, au moyen d'un dialogue entre le professeur et le doyen. Le caractère contraignant du processus de suivi améliore l'efficacité d'une évaluation et permet de mesurer les progrès enregistrés. Le Module 5 approfondira la conception de la phase de suivi d'une évaluation et la manière de clore un cycle de qualité.

2.5 Évaluation d'un établissement — Le cas de l'Université de Duisburg-Essen

Une procédure d'évaluation de l'ensemble des unités de l'organisation est en cours depuis 2006 au sein de l'Université de Duisburg-Essen (UDE), en Allemagne. Elle tient compte des processus et résultats de toutes les sous-unités sur une période de six ans. L'évaluation est donc directement liée aux accords internes conclus entre les sous-unités et la direction de l'université. L'évaluation concerne les unités suivantes :

- les facultés ;
- les cellules de recherche ;
- les unités de service ;
- l'administration centrale.

Note : Étant donné que tous les segments et toute la structure organisationnelle de ces différentes unités (et donc l'établissement dans son ensemble) sont soumis à l'évaluation, l'UDE et d'autres établissements suivant des procédures similaires ont créé le terme « évaluation institutionnelle ». À l'échelon international, ce terme renvoie généralement à l'évaluation des établissements d'enseignement supérieur dans leur intégralité. Dans le présent manuel, ce terme renvoie en revanche à l'évaluation des sous-unités d'une organisation.

Dans le cadre des évaluations institutionnelles menées au sein de l'Université de Duisburg-Essen, tous les domaines d'activité des diverses unités de l'organisation à évaluer sont pris en considération, le cas échéant (enseignement et apprentissage, recherche, services et gestion). Ces domaines sont évalués par les parties prenantes de l'unité en question et les conclusions sont intégrées à un rapport d'auto-évaluation transmis aux spécialistes extérieurs qui, à leur tour, reprendront leur appréciation dans un rapport externe. Bien que les évaluations institutionnelles menées au sein de l'UDE concernent des sous-unités, il reste de la marge pour aborder des questions spécifiques au niveau des programmes d'étude au sein de l'évaluation.

Toute cette procédure au sein de l'UDE, menée tous les six ans, dure neuf mois et prépare chaque second cycle d'accords, ces accords étant conclus tous les trois ans. Le règlement de l'évaluation fixe cette procédure et la rend donc obligatoire pour l'ensemble des unités de l'organisation. Si la procédure d'évaluation des facultés est plus normalisée, celle portant sur les autres unités met davantage l'accent sur des questions spécifiques et est plus flexible.

Ce processus est axé sur les normes internationales de l'évaluation et de gestion de la qualité, comme les références et lignes directrices européennes pour l'assurance qualité au sein de l'espace européen de l'enseignement supérieur (ESG) (ENQA 2009) ou encore les normes de la Société d'évaluation allemande (DeGEval 2001).

La coordination des évaluations institutionnelles au sein de l'UDE relève de la cellule d'assurance qualité de l'université, le ZfH (Centre pour le développement de l'enseignement supérieur et le renforcement de la qualité). La procédure normalisée d'évaluation d'une faculté se subdivise en cinq phases et est détaillée dans les sections suivantes :

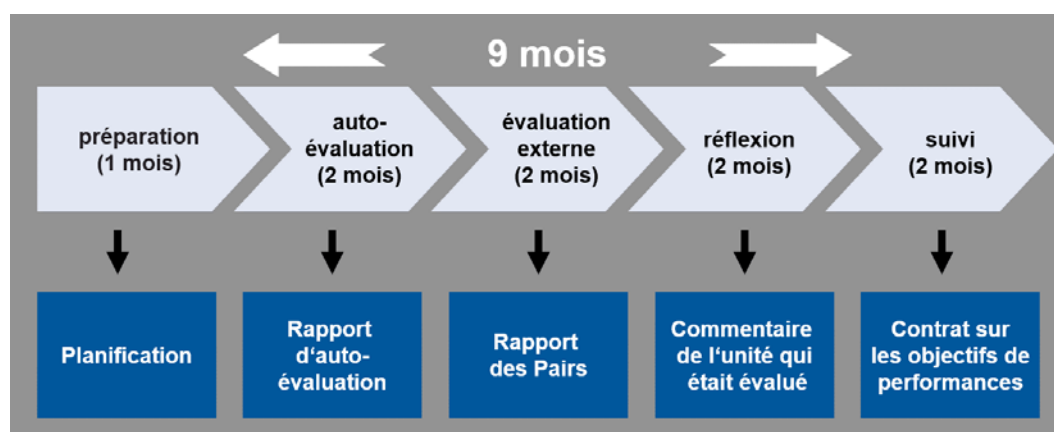


Figure 1 L'évaluation institutionnelle au sein de l'Université de Duisburg-Essen (illustration des auteurs)

I. Phase de préparation (voir Phase de planification) : Comme expliqué au [sous-chapitre 2.4](#), la phase de planification doit éclaircir les points suivants : élément à évaluer, objectifs, instruments, parties prenantes, calendrier. Les évaluations menées au sein de l'UDE suivent le paradigme du développement et la phase de planification doit dès lors impérativement permettre de dégager un consensus entre toutes les parties prenantes et de garantir la transparence du processus. Ceci se fait en trois étapes :

1. Réunion préparatoire	
<i>Participants à la réunion</i>	<ul style="list-style-type: none"> ■ Le vice-recteur à la planification des ressources et au développement ■ L'AAQ du ZfH ■ Les responsables du département de la planification des ressources et du développement ■ Les responsables du Science Support Center (SSC)
<i>Contenu / objectifs</i>	<ul style="list-style-type: none"> ■ Analyse du statu quo de la faculté, à évaluer sur la base des indices clés de performance fournis par la cellule de contrôle de l'UDE. ■ Recueil de questions spécifiques pour l'évaluation, en particulier pour les réviseurs externes
2. Réunion de lancement	
<i>Participants à la réunion</i>	<ul style="list-style-type: none"> ■ Liste identique à celle de la réunion préparatoire ■ Représentants des services du doyen
<i>Contenu / objectifs</i>	<ul style="list-style-type: none"> ■ Présentation du processus par le ZfH ■ Discussion sur les questions spécifiques de l'évaluation ■ Planification détaillée du processus et calendrier ■ Dégagement d'un consensus entre l'ensemble des parties prenantes
3. Information des membres du personnel	
<i>Participants à la réunion</i>	<ul style="list-style-type: none"> ■ Toutes les parties prenantes concernées de l'unité (personnel scientifique et non scientifique, étudiants)
<i>Contenu / objectifs</i>	<ul style="list-style-type: none"> ■ Rendre le processus transparent, minimiser les craintes et les réserves

Tableau 8 Trois étapes pour dégager un consensus et veiller à la transparence dans les évaluations de l'UDE (tableau des auteurs)

II. Auto-évaluation (voir Phase de mise en œuvre) : La faculté à évaluer reçoit des lignes directrices pour le rapport d'auto-évaluation et est invitée à détailler et évaluer les éléments suivants :

- structure organisationnelle ;
- profil dans les domaines de l'étude et de l'enseignement ;
- profil dans le domaine de la recherche ;
- gestion de la qualité ;
- gestion de la diversité.

Pour l'ensemble de ces points, il convient de rédiger une description du statu quo, un plan stratégique à long terme (les six prochaines années) ainsi qu'une planification opérationnelle à court terme (tâches pour les trois prochaines années). Les services du doyen sont chargés de la rédaction du rapport.

La faculté à évaluer procède généralement à un exercice d'auto-évaluation avant de rédiger ce rapport. Il n'existe aucune spécification particulière, comme une liste de contrôle, pour la conception de ce processus. Les AAQ du ZfH proposent toutefois leur aide pour la planification et la coordination de cet exercice, par ex. la coordination d'une analyse SWOT.

Les évaluations institutionnelles menées au sein de l'UDE s'intègrent dans un système holistique de gestion de la qualité (SGQ). Cela signifie que les conclusions des évaluations des cours ou des autres sondages menés parmi les étudiants ou le personnel ainsi que les données issues d'une banque de données contenant des indices clés de performance courants doivent servir de base de réflexion au cours de l'exercice d'auto-évaluation.

III. Évaluation externe (voir Phase de mise en œuvre) : Jusqu'à six experts sont invités à prendre part à l'appréciation externe au sein de l'UDE. Lors de leur sélection, le vice-recteur choisit les personnes à inviter dans une liste fournie par la faculté. L'UDE dispose de lignes directrices et de critères pour la sélection des spécialistes externes (par ex. ceux-ci ne doivent pas entretenir de coopération étroite en matière de recherche avec la faculté à évaluer, etc.) afin de garantir la neutralité et l'impartialité de leur point de vue.

Les spécialistes externes reçoivent le rapport d'auto-évaluation et les questions spécifiques du vice-recteur et de la faculté et sont invités à rédiger un bref commentaire avant la visite du site sur la base de ces documents. Ce commentaire doit mentionner les points forts et faiblesses majeurs identifiés par les spécialistes à la lecture des documents fournis, mais également poser des questions ouvertes et exiger des documents complémentaires. Ceci facilite la préparation de la visite sur site et permet d'en tirer davantage profit.

Lors des deux jours de visite du site, les spécialistes externes s'entretiendront avec des représentants du vice-rectorat, des services du doyen et avec différentes parties prenantes de la faculté (personnel scientifique et non scientifique et étudiants). Tous les débats sont dirigés et animés par le personnel du ZfH. La visite sur site se clôture par un atelier final au cours duquel les spécialistes externes préparent les appréciations pré-finales et leurs recommandations, lesquelles seront consignées dans le rapport externe. L'atelier prévoit également une présentation devant les membres du vice-rectorat et de la faculté ainsi qu'un débat avec ces parties prenantes afin d'éviter tout malentendu, etc.

IV. Évaluation externe (voir Phase de mise en œuvre) : Comme mentionné plus avant, nous sommes d'avis qu'une évaluation ne se limite pas à la collecte de données et d'informations (par le biais de sondages, d'une auto-évaluation ou d'une évaluation externe). Elle doit avant tout permettre de dégager une série de mesures d'amélioration. Avant que ne soient décidées des mesures de ce type, il convient néanmoins de se pencher sur les informations recueillies.

La conception de cette phase de réflexion relève de la faculté. Celle-ci est toutefois tenue de récapituler le processus de réflexion et, plus important, les mesures envisagées dans un commentaire sur le rapport externe.

V. Phase de suivi : À l'Université de Duisburg-Essen, la phase de suivi marque la transition entre l'évaluation d'une part, et les accords pris en interne au sein de l'université et le lancement des mesures convenues d'autre part.

Le processus d'évaluation s'achève officiellement à l'occasion d'un dernier échange entre le vice-recteur à la planification des ressources et au développement stratégique, les représentants des services du doyen de la faculté à évaluer, le Science Support Center et le ZfH. Cet entretien vise avant tout à permettre une analyse conjointe des conclusions de l'auto-évaluation, du rapport externe et du commentaire de la faculté sur le rapport externe.

Enfin, dans le cadre de cet échange sont abordées les mesures recommandées par le vice-recteur, lesquelles seront à nouveau analysées officiellement lors du processus d'accord.

Le processus d'accord au sein de l'UDE sera étudié plus en détail dans le module 5 de la collection de livres pour la formation sur l'assurance qualité interne.

2.6 Zoom sur l'évaluation de l'enseignement et de l'apprentissage

Les évaluations menées au sein des établissements d'enseignement supérieur ont trait à des domaines variés (la recherche, l'enseignement et l'apprentissage, les services, la gestion, etc.) et sont menées à différents l'apprentissage niveaux de l'organisation (facultés, établissements, etc.). Le nombre d'éléments à évaluer au sein des établissements d'enseignement supérieur est quasiment illimité.

Outre la recherche, l'enseignement constitue l'un des processus centraux des établissements d'enseignement supérieur et, en Europe et aux États-Unis, les évaluations de ce processus central existent depuis les années 1970. De nos jours, l'évaluation régulière de l'étude et de l'enseignement se reflète dans la législation relative à l'enseignement supérieur de certains pays ou organes et vise à garantir la qualité de l'enseignement supérieur en général, qui comprend naturellement l'étude et l'enseignement.

Le domaine de l'étude et de l'enseignement dispose effectivement des méthodes et instruments éprouvés parmi les meilleurs et les plus avancés en matière d'assurance qualité, appliqués avec une variété d'axes prioritaires et de niveaux d'analyse. Nous présenterons brièvement certains d'entre eux dans les paragraphes suivants :

a) L'évaluation d'un cours : Presque tous les établissements d'enseignement supérieur dans le monde évaluent leurs cours. Cette évaluation peut se faire à l'aide d'une série de méthodes, allant de l'observation en situation de travail par des collègues aux discussions de groupe avec les étudiants. La méthode de loin la plus courante est cependant celle des questionnaires standardisés, au format papier ou en ligne. Il est demandé aux étudiants d'évaluer leur satisfaction générale par rapport à un cours en particulier, l'approche pédagogique du cours et la prestation de l'enseignant.

b) Les études de suivi périodique des diplômés : Les études de suivi des diplômés occupent désormais une place de plus en plus importante dans l'évaluation de l'étude et de l'enseignement au sein des établissements d'enseignement supérieur. Les diplômés sont en mesure de déterminer rétrospectivement la mesure dans laquelle les contenus d'un cours et l'approche pédagogique ont renforcé leurs compétences, en particulier celles qui les ont préparés au marché de l'emploi.

Ce type d'étude se heurte à des difficultés spécifiques liées au fait que ce groupe cible n'est généralement pas aussi facilement joignable que les étudiants, encore physiquement présents au sein de l'établissement d'enseignement supérieur. Dans la plupart des cas, les diplômés sont sondés au cours d'enquêtes normalisées en ligne. Dans les pays où n'existe aucun système efficace de tenue à jour de fichiers d'adresses qui permette de retrouver les diplômés après leur déménagement, les études de suivi périodique des diplômés privilégient d'autres méthodes, comme les entretiens téléphoniques. Cela vaut également pour les pays où les infrastructures techniques pour les sondages en ligne sont rares.

c) L'évaluation et l'agrément d'un programme d'étude : Pour les programmes d'étude, les évaluations analysent par exemple la structure du programme ou son orientation sur le marché de l'emploi. À l'instar de l'évaluation interne mais aussi de l'évaluation externe des programmes d'étude, dans la plupart des pays, l'agrément des programmes d'étude est juridiquement contraignant. L'agrément veille à ce qu'un programme d'étude respecte des cadres juridiques et des politiques liées à l'enseignement supérieur. Le processus d'agrément prévoit l'analyse d'un programme d'étude à la lumière de critères juridiques et, en cas d'avis positif, l'octroi d'un certificat d'agrément par un organe juridique chargé d'agréer les programmes d'étude. Dans les chapitres 4 et 5 du Module 3, vous pourrez découvrir plus en détail l'évaluation interne et externe des programmes d'étude, leur agrément et les liens entre évaluation et agrément.

d) L'évaluation des conditions-cadres : La qualité de l'étude et de l'enseignement ne se limite pas au contenu ou à la structure des cours ou des programmes d'étude. Elle dépend largement d'une variété de conditions-cadres qui influencent les processus d'enseignement et d'apprentissage. Ainsi, les locaux, l'infrastructure technique, l'équipement de la bibliothèque et l'internet influencent la qualité de l'enseignement et de l'apprentissage. Une série de méthodes et d'instruments peuvent être utilisés pour recueillir des informations à propos de ces différents facteurs, notamment les sondages parmi les étudiants et les enseignants, etc.

e) Les indices clés de performance pour l'étude et l'enseignement : Les indices clés de performance dans le domaine de l'étude et de l'enseignement constituent un bon point de départ quantitatif pour les évaluations. Par exemple, les données relatives au ratio étudiants-enseignant ou le taux d'abandon (nombre d'étudiants d'une promotion qui ne finissent pas leurs études), ou encore les notes finales des étudiants offrent un premier aperçu des éventuels problèmes ou améliorations à apporter (grâce notamment à une comparaison

dans le temps ou à une analyse comparative avec d'autres disciplines ou d'autres établissements). Ceci peut être un point de départ idéal pour une analyse approfondie de ces problèmes, etc. Le chapitre 4 du Module 4 détaille plus avant le concept et l'utilisation des indices clés de performance.

Il existe bien évidemment une infinité d'autres méthodes et outils pour l'évaluation des diverses facettes de l'étude et de l'enseignement, comme les sondages sur l'abandon, les études de suivi des diplômés, l'analyse des documents de descriptif des programmes d'étude, les enquêtes de satisfaction des employeurs, l'analyse secondaire des classements, etc. Dans le cadre limité du présent module, nous ne pouvons toutefois pas tous les aborder.

Nous souhaitons en effet nous concentrer sur la description et sur l'analyse critique des pratiques en matière d'évaluation des cours et ce pour les raisons suivantes : a) l'évaluation des cours est souvent le premier instrument d'assurance qualité de l'enseignement et de l'apprentissage mis en place dans les établissements d'enseignement supérieur et b) l'analyse critique des procédures ci-après peut facilement s'appliquer à d'autres processus d'assurance qualité, plus complexes.

2.7 Études de cas comparatives de l'évaluation des cours

Depuis de nombreuses années, les établissements d'enseignement supérieur dans le monde procèdent à une forme d'évaluation des cours au moyen de sondages menés parmi les étudiants. Les procédures mises en place diffèrent par leur mise en œuvre (par ex. le mode de diffusion des formulaires (en ligne ou en version papier), les instruments utilisés (contenu des questionnaires) et par l'utilisation des résultats. Dans les pages suivantes, deux études de cas seront présentées et analysées selon leurs similitudes et leurs différences dans ces domaines. L'Institut de technologie de Karlsruhe (KIT) et l'Université de Potsdam (UP) procèdent tous deux à l'évaluation de leurs cours depuis de nombreuses années déjà, mais leurs approches et pratiques diffèrent.

Évaluation des cours à l'Institut de technologie de Karlsruhe (KIT)

1. Contexte

Le KIT a été fondé en 2009, à la suite à la fusion du Centre de recherche de Karlsruhe et de l'Université de Karlsruhe. Au premier semestre de l'année académique 2013/2014, 24.530 étudiants étaient inscrits au KIT.

Les activités de renforcement de la qualité au sein du KIT relèvent de la responsabilité d'un membre du personnel directement associé à la direction de l'établissement. Un manuel de la qualité ainsi qu'une convention sur le renforcement de la qualité de l'université détaillent tous les processus d'assurance et de renforcement de la qualité. En mai 2014, le Conseil d'agrément allemand a octroyé un agrément au système d'AQ du KIT. (www.kit.edu/index.php, consulté le 1 septembre 2014)

2. Finalité de l'évaluation

Les évaluations des cours au KIT visent non seulement à offrir aux enseignants un feed-back de leurs étudiants mais doivent également faire office de reddition de comptes vis-à-vis de l'État et du grand public. Les évaluations de cours doivent servir de base empirique à l'amélioration des programmes d'étude. (<http://www.pst.kit.edu/lv-eva.php>, consulté le 1 septembre 2014).

3. Le processus de l'évaluation

Chaque semestre, les étudiants évaluent 35 à 50% des cours dispensés au KIT. Le choix de ces cours relève du doyen des études de chaque faculté. À mi-semestre, les enseignants des cours choisis reçoivent de la personne responsable (voir supra) les questionnaires en version papier, qui seront distribués à leurs étudiants. Ces questionnaires sont complétés par les étudiants lors d'une séance de cours, ramassés par un des étudiants et transmis au responsable pour traitement des données. (<http://www.pst.kit.edu/211.php>, consulté le 1 septembre 2014) et http://www.pst.kit.edu/downloads/6_Evaluationsordnung_2013-05-31.pdf).

4. Contenu des questionnaires

Tous les questionnaires du KIT prévoient une série de questions obligatoires à propos de la qualité du contenu du cours et de l'infrastructure, de la performance de l'enseignant, de l'orientation des étudiants, de l'organisation du cours et de la participation active des étudiants présents ainsi qu'une question relative à la satisfaction générale des étudiants. Toutes les facultés peuvent y ajouter des questions propres à leur contexte. (http://www.pst.kit.edu/downloads/6_Evaluationsordnung_2013-05-31.pdf)

5. Traitement et utilisation des données

Les moyennes arithmétiques des réponses à l'ensemble des questions sont compilées dans un rapport. Une logique de type « feux tricolores » situe, pour 5 questions de base, l'ensemble des moyennes arithmétiques sur une échelle de 5 points : de 1 à 2,5 les moyennes sont jugées bonnes (feu vert), de 2,5 à 3,5, elles sont jugées moyennes (feu orange), et au-delà de 3,5 elles sont jugées faibles (feu rouge). Un indice est également calculé pour les questions de base et pondère différemment ces questions. L'indice de qualité de l'enseignement décrit la qualité globale de l'enseignement pour un cours spécifique. Les cours peuvent donc aussi être jugés bons, moyens ou faibles.

Les conclusions des évaluations de cours sont ventilées comme suit :

- a. Enseignant — reçoit les résultats de tous les cours évalués et est tenu par la réglementation des évaluations du KIT d'analyser ces conclusions avec les étudiants des cours concernés.
- b. Conseil de la faculté, Commission des études et doyen des études — reçoivent les résultats anonymisés globaux, associés à des indices (cours bons, moyens et faibles) et sont tenus d'en tirer des mesures d'amélioration (par ex. un dialogue avec l'enseignant, une recommandation à l'enseignant de prendre part à un cours de pédagogie de l'enseignement supérieur, etc.)
- c. Les vice-recteurs adjoints aux études et la Commission des études et de l'enseignement du sénat reçoivent les résultats globaux de toutes les facultés, associés à des indices (cours bons, moyens et faibles) et analysent les mesures d'amélioration envisagées (voir plus haut) avec les doyens des études.

Question de base	Pondération de l'indice de qualité de l'enseignement
Veillez évaluer votre satisfaction générale par rapport au cours (de pleinement satisfait à pas satisfait du tout)	50%
La charge de travail nécessaire pour ce cours est... (de adaptée à pas adaptée)	12.5%
Comment jugez-vous la structure du cours ? (de très bonne à très faible)	12.5%
Dans quelle mesure l'enseignant est-il motivant ? (de très motivant à pas motivant du tout)	12.5%
Dans quelle mesure l'enseignant répond-il aux questions et intrants des étudiants ? (de beaucoup à très peu)	12.5%

Tableau 9 Indice de qualité de l'enseignement de KIT (<http://www.pst.kit.edu/230.php>)

Évaluation des cours à l'Université de Potsdam (UP)

1. Contexte

L'Université de Potsdam a été fondée en 1991 et comptait près de 20.000 étudiants au premier trimestre de l'année académique 2013/2014.

Outre d'autres tâches, comme la promotion des activités d'apprentissage en ligne, l'évolution de la carrière des enseignants et la pédagogie de l'enseignement supérieur, le Centre pour l'amélioration de la qualité dans l'enseignement et l'apprentissage est également responsable des activités de renforcement de la qualité au sein de l'UP. Un manuel de la qualité ainsi qu'une convention sur le renforcement de la qualité de l'université détaillent tous les processus d'assurance et de renforcement de la qualité. Au cours de l'année 2013, le système d'AQ de l'UP a reçu l'agrément du Conseil allemand de l'agrément.

2. Finalité de l'évaluation

L'évaluation du cours vise avant tout à favoriser le dialogue entre étudiants et enseignants et à mettre en place à long terme une culture de feed-back vital entre les étudiants et leurs enseignants. Ces deux groupes déterminent de manière discursive des mesures d'amélioration sur la base des conclusions de l'évaluation. Les résultats globaux et anonymisés sont toutefois transmis une fois par an aux services

du doyen et sur demande. Les doyens peuvent également obtenir les résultats par enseignant pour qu'ils puissent prendre des mesures, si nécessaire.

3. Processus de l'évaluation

À l'UP, les évaluations de cours se font essentiellement au moyen de questionnaires au format papier ou sur la nouvelle plate-forme d'évaluation en ligne. Cette plate-forme offre aux enseignants un accès à une liste recensant l'ensemble de leurs cours. Ils choisissent dans cette liste les cours à évaluer (chaque faculté dispose de ses propres règles pour le choix des cours à évaluer, le cas échéant) et reçoivent les questionnaires au format papier ou les données d'accès dans le cas d'un questionnaire en ligne. (<https://pep.uni-potsdam.de/articles/kurse-bewerten/Kurse-bewerten.html>, dernière consultation le 5 janvier 2015)

4. Contenu des questionnaires

Les enseignants ont la possibilité de choisir entre a) les questionnaires spécifiques à leur propre faculté ou b) une nouvelle enquête individuelle basée sur les acquis d'apprentissage. Chaque faculté dispose de ses propres questionnaires, un questionnaire générique et des formats spécifiques à certains cours (par ex. pour les travaux pratiques / les laboratoires). Les questionnaires sont élaborés en étroite coopération entre le Centre pour le développement de la qualité et les facultés et constituent dès lors des instruments adaptés aux différentes facultés. Tous prévoient toutefois des questions sur la performance pédagogique, la participation active des étudiants, les infrastructures et les conditions-cadres, ainsi que sur le niveau de satisfaction générale des étudiants.

Le nouvel outil d'enquête basé sur les acquis d'apprentissage a fait son apparition en 2014 ; il permet à chaque professeur de choisir les acquis d'apprentissage du cours à évaluer dans une liste prédéfinie et d'ajouter des questions individuelles. Cet instrument donne lieu à un sondage individuel sur les acquis d'apprentissage du cours. L'idée est de disposer d'un feed-back individualisé plus complet sur les cours. Des questions générales subsistent néanmoins, notamment concernant le niveau de satisfaction des étudiants ou la motivation, et ce afin de pouvoir procéder à une comparaison entre tous les cours.

5. Traitement et utilisation des données

L'Université de Potsdam ne pratique aucun calcul d'indice de qualité. Les moyennes arithmétiques des réponses à chaque question sont rassemblées dans un rapport. Les conclusions des évaluations de cours menées au sein de l'UP sont traitées et communiquées comme suit :

- a. Enseignant — reçoit les résultats de tous les cours évalués et est tenu par la réglementation des évaluations de l'UP d'analyser ces conclusions avec les étudiants des cours concernés.
- b. Doyen / doyen des études — reçoit les résultats globaux anonymisés afin de pouvoir en tirer des mesures générales d'amélioration pour l'ensemble des cours. Les doyens peuvent également demander au Centre pour le développement de la qualité les résultats d'un enseignant en particulier. Les enseignants sont tenus informés de toute demande éventuelle.
- c. Direction de l'université — ne reçoit pas les résultats de l'évaluation des cours.

 **Questions et tâches à accomplir**

1. Quelles différences majeures observez-vous entre les procédures d'évaluation des cours du KIT et de l'Université de Potsdam ? Selon quel paradigme d'évaluation ces procédures reposent-elles, selon vous ?
2. Quels sont les avantages et les inconvénients de chaque procédure décrite ci-avant ?

2.8 Évaluation des cours — Que demander et comment le demander

Comme nous l'avons vu au point 2.1, la question de la définition de la qualité est plus complexe pour les établissements d'enseignement qu'elle ne l'est pour l'économie ou l'industrie. Cette complexité tient au fait que les processus et produits des établissements d'enseignement supérieur sont largement déterminés par les attentes et comportements de différents acteurs. Par conséquent, l'évaluation de la qualité doit tenir compte de ces acteurs.

Avant de pouvoir apprécier des processus d'enseignement et d'apprentissage, nous devons convenir de la définition de la qualité dans ce contexte spécifique. En d'autres termes, comment déterminer si la qualité de nos processus et produits est bonne ?

La réponse à cette question n'est pas simple. Si l'on demandait à dix enseignants ce qu'est pour eux un enseignement et un apprentissage de haute qualité, on obtiendrait plus que probablement dix réponses différentes. La notion de qualité peut grandement varier d'une personne à l'autre. Même si nous examinons un petit élément d'un vaste domaine d'étude et d'enseignement, comme un cours spécifique, il serait très difficile de dégager une idée commune de ce qu'est un niveau de qualité élevé car, dans ce cas également, différents acteurs (à savoir les étudiants et les enseignants), avec des attentes différentes, entrent en ligne de compte et le cours est par ailleurs dispensé dans des conditions-cadres spécifiques.

La qualité d'un cours est dès lors une combinaison de la qualité de différents aspects d'un cours. Il est pour ainsi dire impossible de sonder les participants de façon suffisamment approfondie sur l'ensemble de ces aspects au moyen d'un seul questionnaire, comme il en est utilisé habituellement pour l'évaluation d'un cours. Par exemple, le moment où le cours se donne, l'élocution de l'enseignant, les contributions des étudiants sont autant d'éléments qui ont une influence, positive ou négative, sur la qualité d'un cours.

La recherche empirique en sciences sociales peut nous aider à dégager une notion commune d'un concept (par ex. la qualité d'un cours) et rendre celui-ci empiriquement accessible, c'est-à-dire mesurable.

La difficulté consiste à présent à traduire le concept de la qualité d'un cours dans un questionnaire. Ce processus de traduction est qualifié d'opérationnalisation en recherche empirique en sciences sociales et sera expliqué plus en détail au [sous-chapitre 5.5](#) selon le modèle multifactoriel de Munich pour la qualité des cours de Heiner Rindermann (1998). À travers ce modèle, Rindermann expose un concept de la qualité des cours qui

tente de systématiser les aspects ou les dimensions pouvant influencer la qualité d'un cours : Rindermann distingue les dimensions de l'enseignant, des étudiants et des conditions-cadres en tant que parties du concept de la qualité des cours. Pour en savoir davantage sur ces dimensions (par ex. le comportement de l'enseignant ou les connaissances antérieures des étudiants), nous avons besoin d'indicateurs qui nous informent sur le comportement de l'enseignant ou sur les connaissances antérieures des étudiants.

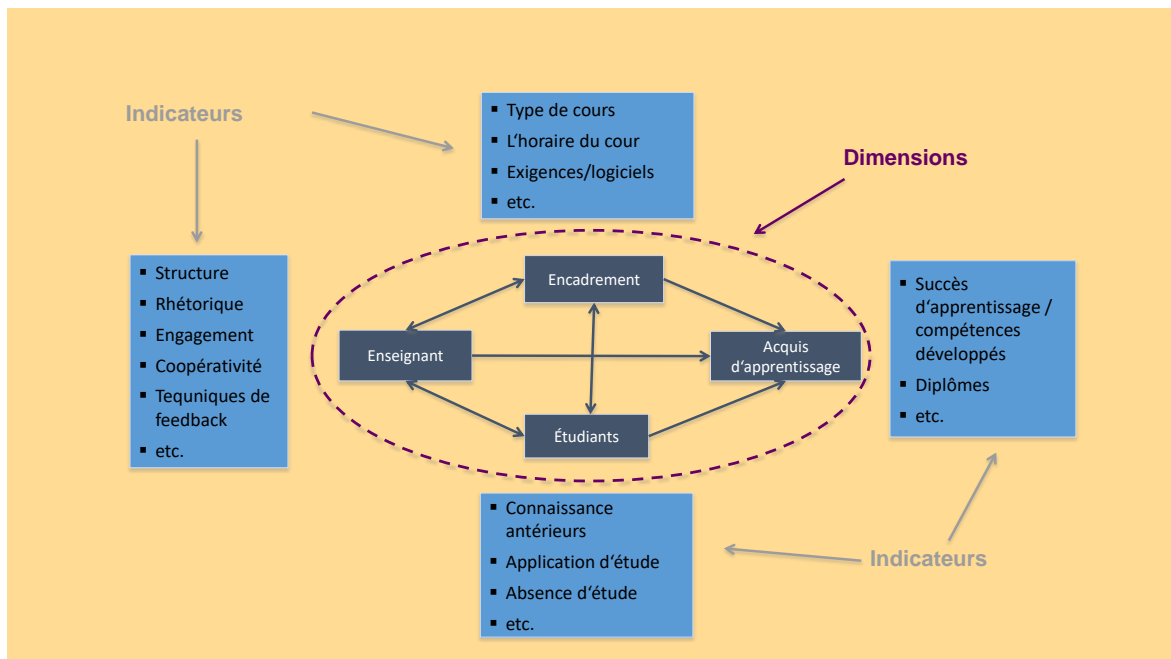


Figure 2 Adaptation du modèle multifactoriel de Munich pour la qualité des cours (adapté de Rindermann 2001)

Comme on peut le voir, la qualité d'un cours est une construction d'une grande complexité. L'influence de chaque aspect ne pouvant être étudiée dans un questionnaire, les conclusions d'une évaluation de cours par les étudiants ne livrent pas une vérité absolue. Ces résultats doivent être interprétés avec prudence, et les renseignements glanés lors des entretiens avec les étudiants doivent également être pris en compte.

Mais comment une évaluation de cours au moyen d'enquêtes auprès d'étudiants peut-elle appuyer le renforcement de la qualité dans l'enseignement et l'apprentissage ? Quelles fonctions peut-elle revêtir et quelles attentes ne peut-elle combler ? Dans les pages suivantes, nous réaliserons une analyse critique de certains aspects d'évaluations de cours usuelles :

■ **Utilisation des résultats** – Parfois, les résultats d'une évaluation de cours sont liés à des récompenses ou à des sanctions financières ou autres. Cette pratique semble reposer sur l'idée que les étudiants sont des clients qui consomment passivement le produit, à savoir la performance de l'enseignant. Nous ne partageons pas cette définition de l'enseignement comme étant un bien fourni à des clients passifs. Nous pensons au contraire qu'un cours est donné en interaction, avec la contribution de l'enseignant et des étudiants, et que le succès de ce cours dépend également des conditions-cadres dans lesquelles il est donné. Il convient de rappeler que l'enseignant a peu d'influence sur la plupart de ces conditions-cadres, comme les locaux, etc. Il n'est pas judicieux d'utiliser les résultats des évaluations de cours pour vérifier et récompenser ou sanctionner des enseignants, qui ne constituent qu'un facteur parmi tant d'autres dont peut

dépendre la qualité d'un cours. L'évaluation d'un cours par les étudiants ne convient dès lors pas pour réaliser des évaluations selon le paradigme du contrôle. Les résultats des évaluations des cours peuvent servir plutôt à encourager le dialogue entre enseignants et étudiants à propos du cours et à dégager des mesures d'amélioration appropriées. Les évaluations de cours par les étudiants conviennent donc parfaitement aux évaluations effectuées selon le paradigme du développement.

- **Portée** — Dans de nombreux établissements d'enseignement supérieur, les résultats des évaluations de cours sont groupés et le résumé de ces résultats sert à apprécier la qualité de l'enseignement dans un programme d'étude. Cette pratique doit être envisagée avec prudence et faire l'objet d'une analyse critique parce que les instruments utilisés pour l'évaluation d'un cours sont, évidemment, conçus pour des cours. Aucune mesure d'amélioration pour un programme d'étude ne peut dès lors être dégagée sur cette base.
- **Intégration du processus** — Un bon enseignement se manifeste sous la forme de cours pédagogiquement bien planifiés, mais peut également être atteint si le cursus est bien structuré et si les conditions-cadres stimulent l'enseignement et l'apprentissage. Aussi l'appréciation de cours spécifiques par les étudiants doit-elle absolument faire partie d'une vision plus globale du très vaste domaine que constituent l'enseignement et l'apprentissage. Cette vision plus globale est possible si l'évaluation du cours est combinée à d'autres instruments (évaluation du programme d'étude etc.), comme nous le verrons en détail au chapitre 5 du module 3 de la collection de livres pour la formation sur l'assurance qualité interne.
- **Suivi** — La communication des conclusions d'une évaluation de cours à l'enseignant et le dialogue avec les étudiants à propos de ces conclusions ne permettront pas d'améliorer la performance de l'enseignant, mais le renforcement des capacités dans la pédagogie de l'enseignement supérieur et l'offre de conseils peuvent aider les enseignants à modifier leurs habitudes d'enseignement.

En votre qualité d'AAQ, il vous incombe de veiller aux bonnes pratiques lors des évaluations de cours et à inclure les recommandations ci-dessus dans votre travail quotidien.

2.9 Les facteurs de succès des évaluations dans les établissements d'enseignement supérieur

Dans les chapitres précédents, nous nous sommes penchés sur des études de cas d'évaluation de cours et de facultés. Il existe cependant de nombreux autres types d'évaluation dans les établissements d'enseignement supérieur. Dans le cadre du présent module, il nous est toutefois impossible d'en aborder plus que les trois évoquées ci-avant.

Dans les pages suivantes, nous présenterons brièvement certains aspects que vous devriez prendre en compte en tant qu'AAQ lors de la planification et du déroulement d'une évaluation selon le paradigme du développement, et ce afin d'améliorer la qualité des résultats et le succès de l'ensemble du processus. Veuillez toutefois noter que nous devons nous limiter à formuler des conseils d'ordre général. Certaines de ces recommandations peuvent ne pas convenir à toutes les formes d'évaluation.

- **Pertinence** — Les évaluations menées au sein des établissements d'enseignement supérieur et en général ne doivent être menées que si elles sont pertinentes pour les parties prenantes. Cela signifie qu'elles doivent générer des résultats qui permettent aux parties prenantes concernées d'améliorer leurs

processus, structures ou produits. Sans cela, l'utilisation de ressources financières et humaines nécessaires à la planification et au déroulement d'une évaluation ne peut se justifier. Il est possible de garantir la pertinence d'une évaluation en :

- menant l'évaluation sur la base de problèmes concrets ou de questions réelles au sein de l'établissement, c'est-à-dire en axant l'essentiel de l'analyse sur la planification stratégique ;
- incluant les attentes et points de vue des différents groupes d'acteurs concernés ;
- traduisant les résultats en mesures d'amélioration selon des processus bien définis, à savoir en lien avec la gestion de l'enseignement supérieur.

■ **Communication et participation** — Les évaluations ont toujours une influence sur différents groupes d'acteurs au sein d'un établissement d'enseignement supérieur et elles suscitent souvent angoisses et ressentiment car les gens s'attendent à des jugements erronés ou négatifs. Ces angoisses et ce ressentiment peuvent à leur tour entraver un déroulement sensé et fluide de l'évaluation car les acteurs impliqués peuvent se montrer très critiques vis-à-vis de l'évaluation en général, être tentés de fournir de mauvaises informations, voire de s'abstenir de prendre part au processus. Aussi importe-t-il de communiquer la finalité, l'avancement et les conclusions de l'évaluation et d'associer les parties prenantes à son processus. Il est possible d'y parvenir en :

- mettant sur pied un groupe de travail pour l'évaluation composé de représentants de tous les groupes d'acteurs pertinents et chargé d'accompagner le processus tout au long de l'évaluation ;
- assurant l'information continue de toutes les parties prenantes concernées par l'évaluation, par le biais de réunions de sensibilisation, d'un site web consacré au projet, d'un bulletin d'information ou d'ateliers qui présentent les objectifs, les méthodes et l'avancement du processus ainsi que les conclusions de l'évaluation.

■ **Organisation du processus** — En général, il conviendrait de prévoir un accompagnement par un spécialiste du domaine pour toutes les évaluations, même s'il s'agit d'auto-évaluations internes. En tant qu'AAQ, vous êtes un expert au sein de votre établissement et il est de votre devoir de :

- planifier l'évaluation, dans le respect des normes communes pour les évaluations et la recherche empirique en sciences sociales ;
- trouver un équilibre entre les normes scientifiques et une approche réalisable, qui peut être nécessaire dans le contexte spécifique d'un établissement d'enseignement supérieur ;
- recourir aux techniques de gestion de projet pour la planification et le déroulement de l'évaluation ;
- prévoir un calendrier réaliste, laissant suffisamment de marge pour les éventuels imprévus ;
- garantir la disponibilité d'un responsable en charge de la coordination du processus ;
- diriger le processus d'une manière aussi intense que possible, mais avec le moins d'effort possible.

■ **Affectation des ressources** — Les évaluations génèrent une charge de travail supplémentaire. Nombre d'évaluations restent décevantes car les ressources financières destinées aux coordinateurs, logiciens, etc. sont insuffisantes. Vous devez prendre en compte les postes budgétaires suivants pour une évaluation :

- salaires / honoraires d'un spécialiste en évaluation, qualifié dans les méthodes et normes de recherche empirique en sciences sociales ;
- charge de travail / heures de travail du personnel participant au groupe de travail de l'évaluation, le cas échéant ;
- honoraires, frais de déplacement et de logement des spécialistes externes ;
- coûts du logiciel d'évaluation (par ex. pour le déroulement des sondages) ;
- coûts d'impression des documents, comme vos propres rapports, etc ;
- coûts liés à l'organisation d'ateliers, etc. (repas, matériel logistique, etc.).

En ce qui concerne les ressources humaines, vous devez veiller à ce que le personnel associé au processus soit, du moins partiellement, déchargé de ses tâches quotidiennes. De manière générale, l'évaluation doit se limiter aux ressources (humaines) strictement nécessaires pour atteindre les objectifs de l'évaluation. C'est particulièrement vrai pour le personnel, car une charge de travail trop élevée peut avoir un effet négatif sur l'ensemble de l'évaluation.

- **Fiabilité** — Les évaluations donnent toujours lieu à une appréciation et donc, indirectement, portent à juger le travail du personnel d'une organisation. Même les évaluations selon le paradigme du développement doivent encourager le changement au sein d'une organisation ou d'un processus. La modification et l'appréciation de son propre travail sont des facteurs d'angoisse et de ressentiment et peuvent donc avoir un effet négatif sur une évaluation. Nous sommes d'avis qu'il est crucial de ne pas perdre de vue les éléments suivants lors de la planification et du déroulement d'une évaluation :
 - En tant qu'AAQ, coordinateur d'une évaluation, vous devez faire preuve de neutralité, d'intégrité, et vous devez comprendre l'évaluation sans préjugé et sans nourrir d'attentes préalables. Vous ne pouvez pas prendre parti. Cela vaut également pour les parties prenantes associées à un groupe de travail sur l'évaluation.
 - Les objectifs d'une évaluation doivent être stipulés clairement, même si cela risque de susciter un certain ressentiment et le mécontentement des parties prenantes concernées. Il faut à tout prix éviter les « agendas cachés » (voir [sous-chapitre 1.3](#)). En tant qu'AAQ, vous devez aborder ouvertement ces sources d'angoisse et de ressentiment ainsi que toute forme de résistance dans le chef des parties prenantes. À cette fin, vous pouvez clarifier les tenants et aboutissants des décisions prises et des débats qui ont précédé l'évaluation.
 - La participation des parties prenantes ne doit être autorisée que si elle est réellement souhaitée (voir [sous-chapitre 2.3](#)). Cela signifie que les avis et recommandations de ces acteurs ne doivent être demandés que si le processus d'évaluation est suffisamment flexible pour les intégrer. Une pseudo-participation remettrait en question la crédibilité de l'évaluation et la fiabilité de l'AAQ.

Chapitre 3

Les instruments de la recherche empirique en sciences sociales en tant qu'outils pour l'évaluation

3	Les instruments de la recherche empirique en sciences sociales en tant qu'outils pour l'évaluation	51
3.1	La pertinence de la recherche empirique en sciences sociales pour l'évaluation	51
3.2	Finalités et approches de la recherche empirique en sciences sociales	53
3.3	Les stratégies de la recherche empirique en sciences sociales	56
3.4	Méthodes de recherche empirique en sciences sociales	60



Au terme de ce chapitre, vous devriez pouvoir...

- reconnaître la pertinence de la recherche empirique en sciences sociales pour l'évaluation ;
- distinguer les techniques des méthodes de la recherche empirique en sciences sociales ;
- décrire et différencier les finalités et les approches principales de la recherche empirique en sciences sociales ;
- énumérer les trois questions fondamentales de la recherche empirique en sciences sociales ;
- faire la différence entre les stratégies quantitatives et qualitatives de la recherche empirique en sciences sociales et décider de l'approche la plus adaptée à une situation donnée ;
- décrire et différencier les données quantitatives et qualitatives ;
- connaître la finalité des méthodes de la recherche sociale.

3 Les instruments de la recherche empirique en sciences sociales en tant qu'outils pour l'évaluation

3.1 La pertinence de la recherche empirique en sciences sociales pour l'évaluation

Comme nous l'avons vu au [chapitre 2](#), dans l'enseignement supérieur, nous avons recours à l'évaluation pour rassembler et réviser des processus et leurs effets sur la recherche, l'enseignement et l'administration. La révision repose dès lors sur des informations relatives à ces processus et leurs effets. En conséquence, l'accès à ces informations et leur collecte sont essentiels à tout processus de planification, développement, contrôle et action dans les processus d'évaluation (cycle PDCA, voir Module 1). Pour que ces informations soient exhaustives et utiles, leur qualité doit répondre à des normes scientifiques. Pour garantir cette qualité scientifique, nous avons recours à des méthodes et techniques issues de la recherche empirique en sciences sociales qui permettent de recueillir et d'analyser des données.

Nous définissons les données comme étant des informations chiffrées, rendues abstraites au moyen de différence entre méthodes et techniques issues des sciences sociales (voir Module 4, sous-chapitre 1.1).

Différence
entre
informations
et données

Ces méthodes et techniques (voir [sous-chapitre 4.3](#)) revêtent différentes facettes et peuvent servir dans le cadre de différents processus d'évaluation (voir [chapitre 2](#)).

Les gestionnaires de la qualité ne doivent pas être des spécialistes de toutes les méthodes en vigueur, mais ils doivent connaître la grande variété de méthodes et techniques de collecte et d'analyse de données en usage en sciences sociales. Ces connaissances sont fondamentales pour leur permettre d'utiliser eux-mêmes des méthodes et techniques de la recherche empirique en sciences sociales dans le cadre de processus d'évaluation. Les gestionnaires de la qualité doivent par ailleurs être à même d'interpréter les données (recueillies au cours des sondages, par ex.). Enfin, une bonne compréhension de la méthodologie est essentielle pour pouvoir interpréter la qualité des données non recueillies par le gestionnaire de la qualité lui-même mais provenant de sources externes. Neumann (2000) décrit comme suit la recherche sociale à envisager par les gestionnaires de la qualité :

Importance
de la recherche
empirique en
sciences sociales

« La recherche sociale est un recueil de méthodes utilisées systématiquement dans le but de générer des connaissances. Il s'agit d'un processus de découverte passionnant, mais il nécessite de la persévérance, de l'intégrité, de la tolérance face à l'ambiguïté, une interaction avec les autres et la fierté d'exécuter un travail de qualité. »

(traduit de Neuman 2000, 2)

Nous pourrions conclure brièvement comme suit :

Les méthodes et techniques de la recherche empirique en sciences sociales permettent de

- recueillir et collecter,
- analyser, et enfin
- interpréter les données et informations afin de pouvoir
- dégager des activités de suivi appropriées.

Différence entre
les techniques
et les méthodes

Nous avons jusqu'ici abordé les méthodes et techniques utilisées en sciences sociales, mais nous n'avons pas encore expliqué la différence entre ces deux notions. Nous utilisons le terme « techniques » pour désigner les instruments et outils nécessaires à la collecte appropriée de données. Les méthodes décrivent les procédures dans lesquelles ces techniques et outils doivent être utilisés. Ainsi, un questionnaire est un outil (technique) utilisé pour mener un sondage (méthode) (Kumar 2005, 126).

Selon le cycle PDCA, les méthodes et techniques de recherche empirique en sciences sociales sont utilisées à des fins de révision (« contrôler »), mais aussi pour dégager et planifier des activités (« agir » ; « planifier »). Travailler à l'aide de ces méthodes et techniques fait généralement apparaître une différence entre les procédures recommandées dans les ouvrages et la réalité. Ainsi, les ressources financières ou le temps recommandés dans les ouvrages scientifiques sont insuffisants. Rossi, Lipsey, et Freeman 2004 résument cette différence comme suit :

« (...) les évaluations doivent répondre à des normes élevées de la recherche scientifique et en même temps répondre aux besoins d'information des décideurs du programme. Or ces deux objectifs sont, dans la pratique, bien souvent incompatibles. Mener des recherches sociales selon une norme scientifique élevée nécessite généralement des ressources supérieures à celles disponibles pour les projets d'évaluation. »

(traduit de Rossi, Lipsey, et Freeman 2004, 23 et seq.)

Face à ces obstacles (par ex. le temps ou les moyens financiers) rencontrés au cours des processus d'évaluation, nous devons toujours tenter de ne pas négliger ou utiliser à mauvais escient les règles de la recherche empirique en sciences sociales. Dans le cas contraire, les résultats pourraient être arbitraires et, dès lors, peu fiables et erronés. Des résultats sans réel fondement scientifique ne résisteront par ailleurs pas à une analyse critique et ne seront donc pas acceptés. Une approche pragmatique des règles de la recherche empirique en sciences sociales permettra aux gestionnaires de la qualité de mener à bien un processus d'évaluation. Rossi, Lipsey, et Freeman (2004) déclarent à ce propos :

« Dans la pratique, l'évaluateur doit par conséquent trouver un compromis raisonnable entre une approche qui privilégie les procédures qui permettent de garantir la validité des résultats et une approche qui privilégie l'obtention de conclusions opportunes, porteuses de sens et utiles pour les consommateurs. (...) Dans de nombreux cas, des évaluations « suffisantes » pour répondre aux questions pertinentes en matière de politique et de programme seront lancées à juste titre, bien que les conditions régissant ces programmes ou les ressources

disponibles empêchent une conception de ces évaluations qui soit optimale d'un point de vue scientifique. »

(traduit de Rossi, Lipsey, et Freeman 2004, 25)

Questions et tâches à accomplir

1. Quelle est la différence entre les informations et les données ?
2. Quelle est la différence entre les techniques et les méthodes de la recherche empirique en sciences sociales ?
3. Quels types de méthodes et techniques de la recherche empirique en sciences sociales devraient être utilisés au cours des processus d'évaluation ? Pourquoi ?
4. Quelles difficultés peut soulever le recours aux méthodes et techniques de la recherche empirique en sciences sociales dans les établissements d'enseignement supérieur ?

Lectures complémentaires

- Babbie, E. R. (2004). *The practice of social research* (10ème éd.). Southbank : Wadsworth, 342-365.
- Neuman, W. L. (2000). *Social research methods : Qualitative and quantitative approaches* (4ème éd.). Boston : Allyn and Bacon, 1-19.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation : A systematic approach* (7ème éd.). Thousand Oaks : SAGE, 21-30.

3.2 Finalités et approches de la recherche empirique en sciences sociales

Comme nous l'avons déjà expliqué, les méthodes issues de la recherche empirique en sciences sociales servent à recueillir et analyser les données de manière systématique. C'est plus difficile qu'il n'y paraît, surtout au vu de la principale finalité de la recherche empirique en sciences sociales, à savoir la collecte et l'analyse d'aspects ou de phénomènes sociaux de la réalité, lesquels peuvent être relativement complexes et impossibles à observer de manière directe (par ex. l'interaction entre enseignants et étudiants, les processus et résultats de l'enseignement et de l'apprentissage, etc.).

Finalités de la recherche empirique en sciences sociales

Au sein des établissements d'enseignement supérieur, nous avons recours à des méthodes des sciences sociales au cours des processus d'évaluation, pour, par exemple, définir la valeur et l'avantage d'un programme d'étude déterminé. Citons encore les processus d'évaluation suivants, basés sur les méthodes et techniques de la recherche empirique en sciences sociales :

- la description ou l'explication de phénomènes ou incidences (par ex. les raisons des abandons).
- la prédiction des incidences à venir (par ex. la demande de formation supplémentaire pour les bacheliers dans les programmes de master).

- le choix et l'évaluation d'interventions ou activités pour atteindre des objectifs définis (par ex. des activités visant à réduire de 5% le taux d'abandon des étudiants).

Le recours à des méthodes et techniques de la recherche empirique en sciences sociales dans les établissements d'enseignement supérieur découle généralement d'une demande concrète. Selon la demande formulée, il est possible d'y associer des sondages basés sur une série d'approches dans le domaine de l'exploration, de la description et de l'explication (Babbie 2004, 87 et seqq.).

Approches de
la recherche
empirique en
sciences sociales

Les analyses exploratoires de données donnent un aperçu de domaines de recherche encore inconnus. Ce type d'enquête n'a donc pas pour but premier de déterminer les raisons de certaines incidences. Par exemple, lorsqu'un établissement d'enseignement supérieur souhaite lancer un nouveau système informatique, cela peut avoir un impact sur l'espace de travail, en termes de structure mais aussi de contenu. Les espaces de travail peuvent être réduits, ils peuvent être revus à la hausse ou à la baisse, ou de nouveaux espaces peuvent être créés. Une analyse exploratoire de données sur, p. ex., un sondage auprès du personnel d'un établissement d'enseignement supérieur peut recueillir des données sur les attentes, les obstacles ou les craintes de ces personnes suscités par le passage à un nouveau système informatique.

Cette analyse exploratoire des données peut être le point de départ d'une **analyse descriptive des données**. Ces analyses visent notamment à décrire au mieux les avis d'un groupe de parties prenantes déterminé, par ex. au moyen de données statistiques. Par exemple, une analyse exploratoire consacrée à l'introduction d'un nouveau système informatique dans un établissement d'enseignement supérieur peut révéler que, d'après le personnel, les étudiants éprouvent des difficultés à concevoir leurs horaires de cours. Si le nouveau système informatique prévoit la possibilité d'ajouter un module sur la conception d'horaires en ligne, vous pouvez mener un sondage auprès des étudiants afin de découvrir quelles fonctions ils jugeraient utiles pour concevoir leurs horaires.

Si vous souhaitez découvrir les liens entre le motif et le résultat, vous devez plutôt procéder à une **analyse explicative des données**. Par exemple, un établissement d'enseignement supérieur part du principe que les étudiants contraints de travailler pour financer leurs études ont besoin de plus de temps pour décrocher leur diplôme et obtiennent de moins bonnes notes que les autres étudiants. Ce type d'hypothèse peut être évalué au moyen d'une analyse explicative. De manière générale, les hypothèses décrivent assez clairement ce que vous attendez de votre étude. Cet exemple illustre aussi parfaitement la raison pour laquelle nous parlons de recherche empirique en sciences sociales : les hypothèses reposant sur des considérations théoriques sont révisées au regard de leurs réalités sociales respectives (Kumar 2005, 74 et seqq).

Cependant, toutes les analyses de données ne reposent pas sur des hypothèses. Parfois, une analyse exploratoire ne part pas d'une hypothèse formelle et la finalité de cette analyse est peut-être d'élaborer une hypothèse ou une prévision bien déterminée qui pourra être testée au cours de recherches à venir (Trochim 2005, 8). Cela signifie que les processus d'évaluation menés au sein des établissements d'enseignement supérieur au moyen de méthodes et techniques issues de la recherche empirique en sciences sociales ne reposent pas systématiquement sur des approches théoriques. Dans ce cas, ces processus se basent sur des critères d'évaluation qui font office d'hypothèses. Ces critères d'évaluation peuvent être définis par les autorités qui demandent l'évaluation (par ex. le vice-rectorat), par le groupe cible (par ex. les membres d'une faculté), par

les parties prenantes associées au processus (par ex. le Conseil d'administration), par un gestionnaire de la qualité ou par toutes ces parties réunies. Dans ce cas, la recherche empirique en sciences sociales a pour finalité la collecte et l'analyse systématiques de données selon les critères d'évaluation définis (Stockmann & Meyer 2010, 77 et seqq).

Nous pouvons dès lors en conclure que les processus d'évaluation au sein des établissements d'enseignement supérieur ne peuvent pas toujours être décrits au sens étroit des projets de recherche en sciences sociales. Ils incluent plutôt des méthodes et techniques de sciences sociales pour, par exemple, définir la valeur et l'avantage d'un programme d'étude déterminé. Dans ce cas, les processus d'évaluation restent factuels vu qu'ils renvoient à des conclusions scientifiques, recueillies de manière empirique, qui contribuent au processus décisionnel (McMillan & Schumacher 2010, 430).

Pour combiner de manière systématique les hypothèses ou les critères d'évaluation, la collecte et l'analyse de données, vous devez vous pencher sur certaines questions fondamentales au début de chaque processus utilisant des méthodes et techniques issues de la recherche empirique en sciences sociales (Atteslander 2010, 4). Au cours des processus d'évaluation, il se peut qu'il incombe au gestionnaire de la qualité de poser ces questions. Si cela ne relève pas de la responsabilité du gestionnaire de la qualité, celui-ci doit veiller à ce qu'une autre personne réponde aux trois questions fondamentales suivantes, étroitement liées et interdépendantes :

Trois questions fondamentales de la recherche empirique en sciences sociales

1. *Pour quelles raisons les données doivent-elles être recueillies ?* Cette question renvoie explicitement à la finalité de l'utilisation des données : Pour quelle raison doit-on procéder à une évaluation de ce type ? Quelles sont les hypothèses de cette évaluation ?
2. *Quelles données doit-on recueillir ?* Cette question renvoie explicitement à la question de la recherche : Que souhaite-t-on exactement découvrir avec cette évaluation ?
3. *Comment les données doivent-elles être recueillies ?* Cette question renvoie explicitement aux méthodes de collecte et d'analyse des données qui seront utilisées pour l'évaluation.

La réponse à ces questions détermine le choix des méthodes et techniques de la recherche empirique en sciences sociales adaptées pour répondre à notre question (voir [sous-chapitre 4.3](#)).

Questions et tâches à accomplir

1. Citez deux ou trois exemples de méthodes et techniques issues de la recherche empirique en sciences sociales appliquées dans votre établissement d'enseignement supérieur.
2. Expliquez les différences entre les analyses exploratoires, descriptives et explicatives de données.
3. Citez des exemples de critères d'évaluation envisageables pour votre établissement d'enseignement supérieur.
4. Citez les trois questions fondamentales de la recherche empirique en sciences sociales et expliquez leur importance du point de vue des critères d'évaluation, de la collecte de données et de la finalité de l'évaluation.

3.3 Les stratégies de la recherche empirique en sciences sociales

En plus des approches décrites au [sous-chapitre 3.2](#), nous pouvons également avoir recours à différentes stratégies de recherche. On distingue les stratégies quantitatives des stratégies qualitatives. Elles se distinguent essentiellement par les méthodes de collecte de données : Les stratégies quantitatives aboutissent à des données numériques, tandis que les stratégies qualitatives aboutissent à des données non numériques (Babbie 2004, 26 et seqq.).

Données et stratégies quantitatives et normalisées

Les stratégies de recherche quantitatives sont des procédures structurées, dans lesquelles le processus de recherche et les données attendues sont déjà prédéfinis (Kumar 2005, 12) et les incidences respectives sont chiffrées. Par exemple, si vous pouvez décrire le niveau de satisfaction des étudiants par rapport à leur programme d'étude à l'aide de chiffres de un à cinq, il s'agit alors de données quantitatives dénombrables et chiffrables. Puisque les étudiants peuvent définir leur niveau de satisfaction de la même manière (à l'aide de chiffres de un à cinq), nous appelons cela une collecte de données normalisées qui fournit à son tour des données normalisées.

Les stratégies quantitatives ont pour grand avantage que la normalisation permet d'analyser les données recueillies relativement facilement et en grandes quantités au moyen de logiciels d'analyse statistique, et ce grâce à la base de données numérique.

Données et stratégies qualitatives et non normalisées

Contrairement aux stratégies quantitatives, les stratégies de recherche qualitatives mettent davantage l'accent sur les mots ou le texte que sur les chiffres. Elles privilégient les relations et décrivent le monde social du point de vue des sujets afin de découvrir son mode de construction. Les stratégies de recherche qualitatives recueillent généralement des données de manière non normalisée. Par exemple, une forme de collecte de données non normalisée consiste à mesurer la satisfaction des étudiants (voir [sous-chapitre 4.3](#)) au cours d'un entretien oral (par ex. au moyen d'une question pour mesurer le niveau de satisfaction par rapport à un programme d'étude). Le texte qui en résulte est une donnée qualitative puisque les propos des étudiants sont un texte non numérique. Les stratégies qualitatives présentent le grand avantage que le niveau de signification des données ainsi recueillies est élevé, en raison de l'absence de normalisation. Cela signifie que les avis formulés lors des entretiens peuvent être recueillis sans restriction. Le tableau suivant offre un aperçu des différences entre les données quantitatives et qualitatives (Babbie 2004, 26 et seqq. ; Trochim 2005, 121 et seqq.).

Différences entre les données quantitatives et qualitatives

Données Quantitatives	Données Qualitatives
Ont trait aux chiffres, plus explicites	Ont trait aux descriptions, peuvent être davantage porteuses de sens
Données mesurables	Données observables (pas mesurables)
Plus faciles à grouper, comparer et résumer	Peuvent être codées quantitativement
Offrent la possibilité d'une analyse statistique	

Tableau 10 Différences entre les données quantitatives et qualitatives (tableau des auteurs)

Jusqu'à présent, nous avons vu que les données qualitatives n'existent pas sous forme numérique (contrairement aux données quantitatives), mais sous forme de mots ou de texte. Ce n'est pas tout à fait exact car les données qualitatives peuvent être traduites en données quantitatives, lorsque les mots ou le texte sont définis au moyen d'une valeur numérique. Ce faisant, les réponses ou commentaires textuels aux questions ou invitations telles que « Veuillez ajouter un commentaire supplémentaire » peuvent être catégorisés. À chacune de ces catégories est associée une valeur numérique (également appelée « code »), laquelle peut faire l'objet d'une analyse statistique. Grâce à cette codification, les données qualitatives sont traduites en données quantitatives (Trochim 2005, 121 et seqq.).

Exemple : Codification de données qualitatives

Au cours d'un sondage mené auprès de dix étudiants, ceux-ci doivent mentionner leur sexe et leur origine. En l'absence de réponse prédéfinie, les étudiants rédigent leur réponse sous forme de texte. Nous recevons donc des données qualitatives, non numériques. Pour pouvoir calculer le nombre de participants masculins et féminins et leur origine, nous avons besoin d'informations quantitatives, numériques.

Aussi associons-nous à la réponse « femme » le code 1, et à la réponse « homme » le code 2. Nous constatons ensuite que cinq étudiants et cinq étudiantes ont pris part au sondage.

Un code est également associé aux réponses concernant l'origine. On constate ainsi que sept étudiants viennent de Bamako, deux de Ségou et un de Tomboctou.

Répondant	Sexe (donnée qualitative)	Sexe (donnée quantitative)	Origine (donnée qualitative)	Origine (donnée quantitative)
Étudiant 01	Homme	2	Bamako	2
Étudiant 02	Femme	1	Bamako	2
Étudiant 03	Homme	2	Ségou	1
Étudiant 04	Femme	1	Ségou	2
Étudiant 05	Homme	2	Tomboctou	1
Étudiant 06	Homme	2	Tomboctou	3
Étudiant 07	Femme	1	Bamako	2
Étudiant 08	Femme	1	Bamako	2
Étudiant 09	Homme	2	Bamako	2
Étudiant 10	Femme	1	Bamako	2

Cet exemple montre que le choix de la stratégie de recherche ne détermine pas nécessairement exclusivement le format des données à analyser, mais que c'est souvent le cas. Le choix d'une stratégie quantitative vous livrera essentiellement des données numériques (Babbie 2004, 26). Les gestionnaires de la qualité doivent savoir que les processus d'évaluation se composent généralement d'éléments qualitatifs et quantitatifs. Il est donc utile et souhaitable qu'ils maîtrisent bien les stratégies de recherche quantitatives et qualitatives.

Le tableau suivant offre un aperçu des différents aspects des stratégies de recherche quantitatives et qualitatives à envisager lors du choix d'une stratégie appropriée :

Aspect	Stratégie de recherche	
	Quantitative	Qualitative
Approche	Structurée, rigide, prédéterminée	Non structurée, flexible, ouverte
Principale finalité	Mesurer des faits objectifs	Élaborer une réalité sociale, une signification culturelle
	Quantifier l'étendue de la variation d'un phénomène	Décrire la variation d'un phénomène
Taille de l'échantillon	Préférence pour les échantillons de plus grande taille	Cas moins nombreux
Thème de recherche dominant	Explique la prévalence, l'incidence, l'étendue, la nature des sujets, avis et comportements ; met en évidence les régularités	Explore les expériences, les significations, les perceptions et les sensations
Analyse des données	Des données numériques aux distributions des fréquences, tableaux croisés et autres procédures statistiques	Des réponses, récits et données d'observation à l'identification des thèmes et leur description

Tableau 11 Différences entre les stratégies de recherche quantitatives et qualitatives (tableau des auteurs, basé sur Kumar 2005, 17 et seq.)

Stratégie quantitative ou qualitative : laquelle choisir ?

En plus des questions fondamentales mentionnées au [sous-chapitre 3.2](#), la réponse à ces dernières questions permet également de trouver une stratégie de recherche appropriée. Demander « pourquoi » permet de répondre à la question de la finalité : Sur qui ou quoi l'analyse va-t-elle se concentrer ? Sur les personnes (par ex. la réussite des étudiants en termes d'apprentissage) ou l'interaction sociale ou encore les processus (par ex. le résultat d'un programme didacticiel) ? Cela signifie que l'analyse et la finalité déterminent les valeurs quantitatives ou qualitatives à étudier (Trochim 2005, 11). Ensuite seulement vient la question de savoir « quelles données » analyser. Quelle est la question concrète de la recherche ? Qui peut aider à répondre à cette question de recherche ? Nous pouvons ensuite répondre à la troisième question, à savoir « comment » recueillir et analyser les données. Après avoir répondu à ces questions, nous pouvons enfin déterminer si les stratégies qualitatives ou quantitatives sont plus appropriées pour répondre aux questions du processus d'évaluation.

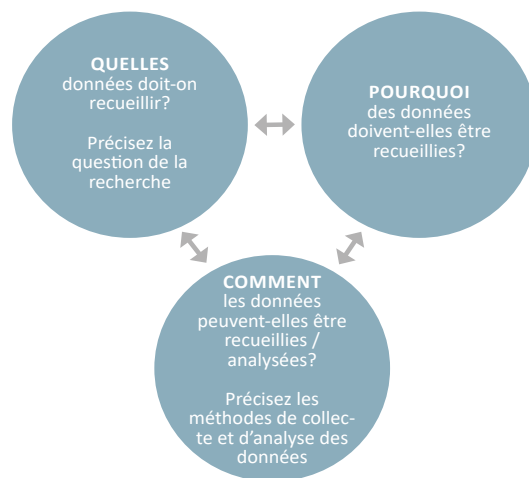


Figure 3 Questions à poser pour le choix d'une stratégie de recherche approprié (illustration des auteurs)

Questions et tâches à accomplir

1. Quelle est la différence entre les stratégies de recherche quantitatives et qualitatives ? Énumérez les principaux avantages de chaque stratégie.
2. Vous avez découvert un exemple de codification des données qualitatives. Veuillez citer un autre exemple de collecte de données, mené au sein de votre établissement d'enseignement supérieur dans le cadre d'un processus d'évaluation.
3. Décrivez une approche possible pour le choix d'une stratégie de recherche appropriée.

3.4 Méthodes de recherche empirique en sciences sociales

Après avoir abordé les stratégies quantitatives et qualitatives de la recherche empirique en sciences sociales et appris leurs différences, nous allons à présent découvrir ce que sont exactement les méthodes de recherche empirique en sciences sociales. Au [sous-chapitre 3.1](#), nous avons déjà pu constater que les méthodes sont des procédures, au cours desquelles sont utilisés des techniques et outils pour la collecte de données. En sciences sociales, il existe une multitude de méthodes reconnues pour recueillir des données. Il peut s'agir de méthodes comme les observations, les sondages, les expériences ou encore l'analyse de contenu (pour de plus amples détails, voir [sous-chapitre 4.3](#)).

L'utilisation de ces méthodes et techniques aide à planifier et structurer la collecte de données de façon à ce qu'elle génère des données utiles, valides et compréhensibles. À l'instar de tout autre type de recherche, la recherche en sciences sociales doit elle aussi répondre aux caractéristiques suivantes : « Elle doit, autant que possible, être contrôlée, rigoureuse, systématique, valide et vérifiable, empirique et critique. » (traduit de Kumar 2005, 7). Au vu de l'écart déjà mentionné entre la reconnaissance scientifique et l'application des

processus de l'évaluation axé sur leur utilisation, une utilisation responsable des techniques et méthodes de la recherche empirique en sciences sociales permettra de réduire les erreurs dans la collecte de données et de garantir l'utilité et l'acceptabilité de ses résultats.

Le tableau suivant illustre dans quel cadre utiliser quelle méthode :

Que faut-il examiner ?	Méthode
Avis formulés par des personnes	Sondage
Comportement humain (en situations naturelles)	Sondage, observation
Comportement humain (en situations expérimentales)	Expérience
Produits de personnes (par ex. un texte)	Analyse de contenu

Tableau 12 Méthodes de recherche basées sur les résultats (tableau des auteurs)

Nous allons à présent analyser plus en détail la méthode du sondage, car elle constitue probablement une des principales méthodes de collecte de données pour les questionnaires de la qualité.

Questions et tâches à accomplir

1. Citez différentes méthodes de recherche sociale.
2. Quelles sont les finalités des méthodes de recherche sociale dans les processus d'évaluation ?

Chapitre 4

Le processus de recherche empirique en sciences sociales : De la question de recherche à la collecte de données

4	Le processus de recherche empirique en sciences sociales : De la question de recherche à la collecte de données.....	63
4.1	Les phases idéales d'une étude empirique	63
4.2	Zoom sur les méthodologies d'une étude empirique en sciences sociales.....	66
4.3	Zoom sur les méthodes de collecte de données.....	70
4.4	Échantillonnage et représentativité	74
4.5	Méthodes de collecte de données : Enjeux et difficultés	75
4.6	Lignes directrices pour le choix de la méthode de sondage appropriée	77



Au terme de ce chapitre, vous devriez pouvoir...

- suivre les principales phases d'une étude empirique au cours de vos propres projets d'évaluation ;
- décrire et expliquer les différences entre les conceptions d'études empiriques ;
- distinguer plusieurs méthodes d'acquisition de données ;
- discuter des avantages et des inconvénients des méthodes de recherche utilisées dans le cadre des sondages ;
- juger du niveau de représentativité d'un sondage ;
- choisir les méthodes de collecte de données appropriées, en fonction des conditions-cadres données.

4 Le processus de recherche empirique en sciences sociales : De la question de recherche à la collecte de données

4.1 Les phases idéales d'une étude empirique

Avant d'approfondir nos connaissances sur les méthodes et techniques de la recherche empirique en sciences sociales, nous nous pencherons d'abord sur les phases idéales de l'analyse socio-scientifique. Ces phases présentent toutes une structure identique à celles d'une évaluation et leur contenu est similaire à celui d'une phase complexe de collecte de données (voir [sous-chapitre 2.4](#)).

Une analyse socio-scientifique se structure en cinq phases idéales. Une phase est idéale en ce sens qu'elle est pertinente et devrait être prise en considération dans l'analyse, tandis qu'au sein d'une phase, certains aspects de moindre importance peuvent être laissés à une procédure pragmatique (voir [sous-chapitre 3.1](#)). Au cours de chaque phase sont prises des décisions et des mesures supplémentaires doivent également être lancées, toutes interdépendantes et en interaction. Aussi n'est-il pas toujours possible de différencier ces phases. Il existe toutefois des passerelles fluides entre ces phases.

Les cinq phases idéales d'une étude empirique

Au cours de **la phase d'orientation et de définition (phase I)**, nous définissons le sujet concret à analyser. Au cours des processus d'évaluation, le sujet est bien souvent déterminé par l'autorité principale, comme le vice-recteur d'un établissement d'enseignement supérieur. Parfois, ces demandes peuvent être vagues, de sorte qu'une ou plusieurs questions de recherche doivent être formulées plus spécifiquement. Une analyse documentaire peut dès lors se révéler utile. Si cela ne suffit pas, nous pouvons également décrire le type d'énoncés à formuler au terme de l'analyse. Par exemple : X% des étudiants contraints de travailler pendant leurs études ont besoin de plus de temps pour décrocher leur diplôme que la moyenne. La description de ces énoncés permet de mieux formuler les hypothèses ou critères d'évaluation nécessaires (voir [sous-chapitre 3.2](#)).

La phase préparatoire (phase II) mène à la planification et à la préparation méthodologiques de la collecte de données. Au cours de cette phase, nous clarifions et conceptualisons la terminologie (voir [sous-chapitre 5.1](#)). Nous devons ensuite choisir le mode d'analyse de ces concepts, et donc choisir la méthode à appliquer. Nous devons parallèlement décider sur qui ou sur quoi nous voulons formuler des conclusions (par ex. des personnes, des groupes, des documents) et, en fonction de cela, déterminer les personnes à interroger au cours des sondages ou les éléments à étudier lors de l'analyse du contenu (par ex. les descriptions de modules). Nous devons ensuite définir un calendrier précis ainsi qu'un budget afin de faciliter la gestion académique.

Au cours de **la phase de collecte de données (phase III)**, nous procédons à la collecte technique de données. Selon la planification technologique choisie en phase II, nous achevons l'opérationnalisation (voir [sous-chapitre 5.2](#)) de l'étude empirique, par ex. en concevant et testant un questionnaire de sondage (voir [chapitres 5 et 6](#)).

Au début de **la phase d'analyse des données (phase IV)**, nous composons un ensemble de données à des fins d'analyse. Cette étape inclut notamment la codification des données (voir [sous-chapitre 3.3](#)) et la recherche d'erreurs (voir [sous-chapitre 4.5](#)).

Après toutes ces étapes seulement, nous pouvons procéder à l'analyse statistique des données (voir [chapitre 7](#)). Enfin, **une étude empirique s'achève par la phase de compte rendu et de diffusion (phase V)**. Au cours de cette dernière phase, nous interprétons et diffusons les résultats, par ex. en rédigeant un rapport de recherche. Les résultats obtenus nous permettent de définir une série d'activités et de les mettre en œuvre.

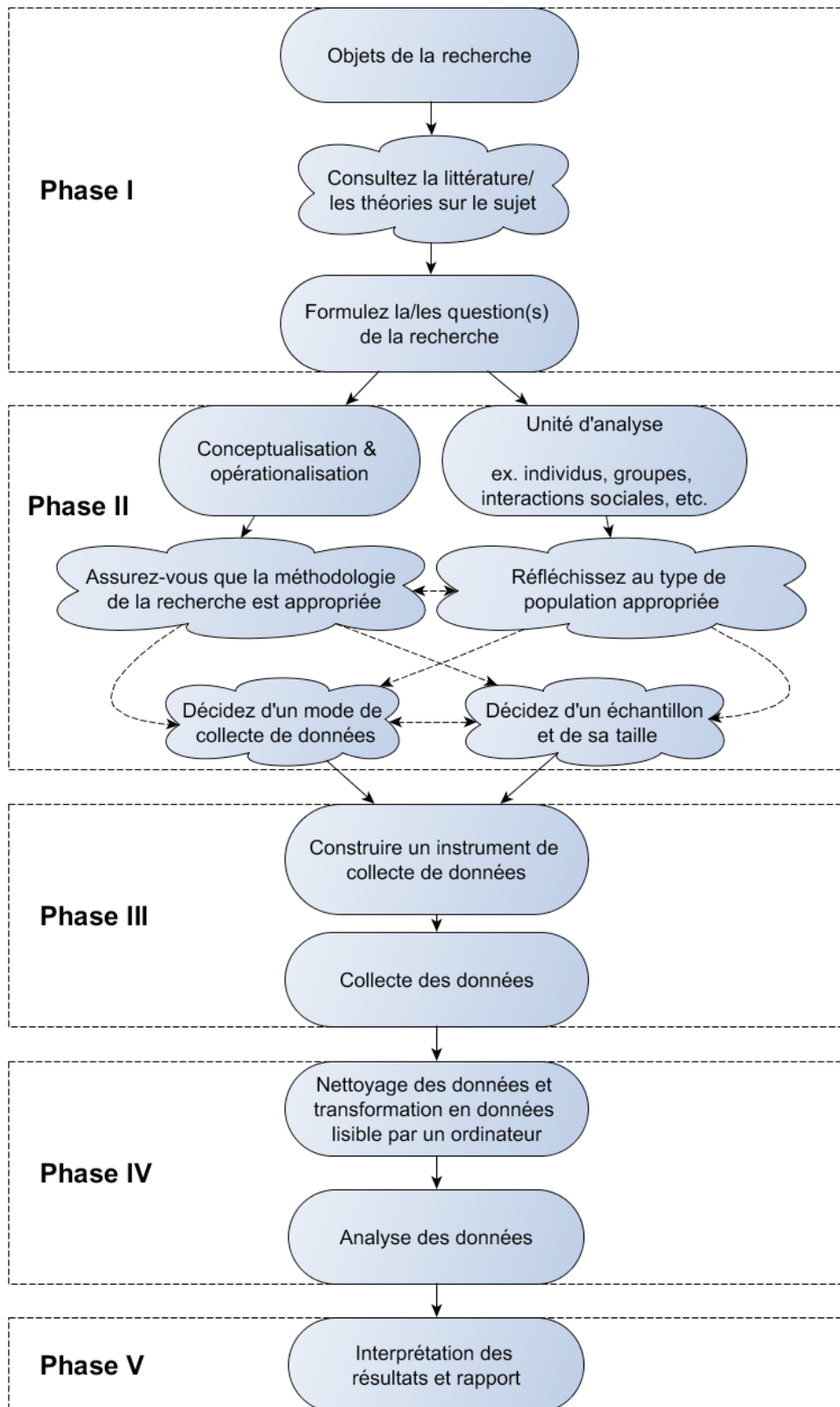


Figure 4 Les cinq phases idéales d'une étude empirique (tableau des auteurs)

Questions et tâches à accomplir

1. Veuillez choisir un exemple de processus d'évaluation mené au sein de votre établissement d'enseignement supérieur et faire une séance de recherche d'idées ou dresser un plan de structuration des phases I à III.

Lectures complémentaires

- Kumar, R. (2005). *Research methodology : A step-by-step guide for beginners* (2ème édition). London : SAGE, 15-25.

4.2 Zoom sur les méthodologies d'une étude empirique en sciences sociales

Nous avons pris connaissance des phases idéales des études empiriques (voir [sous-chapitre 4.1](#)). Voyons à présent plus en détail **la phase préparatoire (phase II) : Selon la phase d'orientation et de détermination (phase I)**, nous devons à présent décider de la manière de procéder pour la collecte de données. Nous devons répondre aux questions suivantes : Quelles méthodes permettront de trouver des réponses à nos questions de recherche ? Comment concevoir les différentes phases du processus pour recueillir les données nécessaires ? Les sciences sociales nous proposent heureusement de nombreuses méthodologies pour nous aider à poser notre choix.

Deux fonctions majeures de la méthodologie de recherche

Les méthodologies de recherche revêtent deux fonctions majeures : Premièrement, elles permettent d'identifier les méthodes et techniques envisageables pour la collecte de données et adaptées à la finalité visée. Deuxièmement, elles sont un meilleur garant de la qualité des processus de collecte et d'analyse des données (Kumar 2005, 84). Les méthodologies de recherche fournissent un cadre pour la collecte et l'analyse des données. Le choix de la méthodologie reflète la priorité accordée à une série de dimensions du processus de recherche (Bryman 2004, 27).

Nous présenterons dans les paragraphes suivants certaines méthodologies de recherche, également appelées méthodologies d'étude, couramment utilisées dans les processus d'évaluation.

Méthodologie d'étude transversale

La méthodologie d'étude transversale est la plus courante en sciences sociales. Dans le cadre d'une étude transversale, les données sont recueillies à un moment déterminé, en fonction d'une observation ou d'un choix particulier. L'objectif n'est pas de recueillir des cas uniques, mais de créer une large base de données sur un thème déterminé. Les études exploratoires et descriptives sont souvent transversales (Babbie 2004, 101 et seq.).

Parmi les exemples de méthodologies transversales, citons celui d'un établissement d'enseignement supérieur qui effectue un sondage auprès des diplômés d'une période spécifique. Ce type de sondage peut aboutir à une sorte de cliché instantané, par exemple pour la révision rétrospective des conditions d'étude. Une seule et unique application d'une méthodologie transversale ne permet toutefois pas de déterminer si les conditions d'étude ont changé.

Contrairement aux méthodologies transversales, les méthodologies longitudinales étudient au moins deux collectes de données réalisées à des moments différents. Voilà pourquoi elles sont utiles pour analyser les changements de manière chronologique. Au total trois types d'études longitudinales seront décrits dans les paragraphes qui suivent : Les études des tendances, les études des cohortes et les études par panel (Babbie 2004, 102 et seqq.).

Méthodologie longitudinale

Les études des tendances analysent les changements de manière chronologique au sein d'une population (voir [sous-chapitre 4.4](#)). Les études des tendances sont donc une répétition d'études transversales sur un même thème. **Les études des cohortes** prévoient également une collecte de données à des moments différents, mais elles se concentrent sur les membres d'un groupe spécifique, dont la composition varie au fil du temps. Ces groupes spécifiques sont également appelés des cohortes. Une cohorte a pour particularité que ses membres partagent une caractéristique commune (par ex. 1986 pour année de naissance ; mariage en 2005 ; études entamées au premier semestre 2011/12), qui influence dans une plus ou moins grande mesure le cours de leur vie. **Les études par panel** observent un phénomène à des moments différents mais ont toujours trait à un même groupe de personnes. Dans ce cas, la composition du groupe ne varie pas.

Études des tendances, des cohortes et par panel

Les études par panel présentent pour inconvénient ce que l'on appelle l'érosion de l'échantillon : Les érosion naturelle de participants au premier sondage (vague) de l'étude par panel ne prennent plus part à la seconde vague ou l'échantillon aux vagues ultérieures. C'est là un point délicat parce que nous ignorons si ces personnes sont essentielles à l'étude et si elles abandonneront leur participation, ce qui peut modifier les résultats.

Érosion naturelle de l'échantillon

Le tableau suivant vous aidera à faire la différence entre les études des tendances, des cohortes et par panel :

Méthodologie	Objectif	Thème	Sondages et populations	Valeur de l'énoncé
Tendances	Changements au sein d'une population au fil du temps	Conditions d'étude	Sondage 1 en 2015: étudiants Sondage 2 en 2017 : étudiants de première année de 2017 : de première année de 2015	Observer si les étudiants de première année en 2017 émettent un avis différent à propos des conditions d'étude par rapport à celui des étudiants de première année en 2015.

Méthodologie	Objectif	Thème	Sondages et populations	Valeur de l'énoncé
Cohortes	Changements au sein d'une cohorte au fil du temps	Conditions d'étude	Sondage 1 en 2015 : étudiants de première année de 2015 Sondage 2 en 2017 : étudiants de première année de 2015	Observer si les membres de la cohorte « étudiants de première année en 2015 » modifient leur avis à propos des conditions d'étude en 2017 par rapport à 2015.
Panel	Changements au sein d'un même groupe de personnes au fil du temps	Réussite professionnelle	Sondage 1 en 2015 : diplômés de 2015 Sondage 2 en 2017 : répondants du sondage 1 Sondage 3 en 2019 : répondants des sondages 1 & 2	Observer l'évolution de la réussite professionnelle individuelle des diplômés de l'année 2015 sur la période allant de 2015 à 2019.

Tableau 13 Différence entre les études des tendances, des cohortes et par panel (tableau des auteurs)

Nous savons désormais que les études longitudinales portent sur des changements dans le temps. Penchons-nous à présent sur les méthodologies qui observent le résultat d'une intervention (par ex. une mesure visant à améliorer les conditions d'étude au sein d'un établissement d'enseignement supérieur). On peut distinguer les études « avant-après » et les études « après-uniquement ».

Méthodologie « avant-après »

Une **méthodologie « avant-après »** repose sur deux ensembles de données transversales collectées à deux moments différents auprès de la même population en vue de découvrir le changement d'un phénomène entre deux points dans le temps. Ce changement se mesure en comparant la différence entre la situation avant et après l'intervention (Kumar 2005, 95). Cette méthodologie est couramment utilisée dans les processus d'évaluation. Nous pouvons ainsi voir, par exemple, si l'évaluation des conditions d'étude à un moment 1 dans le temps change après une intervention découlant de l'évaluation de ces mêmes conditions d'étude à un moment 2 (voir Figure 4).

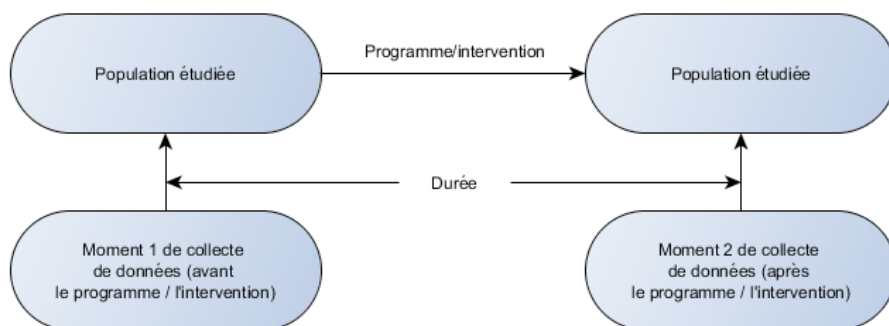


Figure 5 Méthodologie « avant-après » (Kumar 2005, 95)

Nous constatons que les établissements d'enseignement supérieur préfèrent très souvent les **méthodologies « après-uniquement »** aux méthodologies avant-après pour les processus d'évaluation. Dans ce cas, nous mesurons le résultat d'une intervention après sa réalisation. On ne détermine généralement pas une valeur de départ représentant la situation avant l'intervention (voir Figure 4, collecte de données au moment 1) et pouvant servir de point de comparaison. Cela signifie que l'efficacité d'une intervention ne peut être déterminée que sur la base de données recueillies après l'intervention (voir Figure 4, collecte de données au moment 2), par ex. sur la base du souvenir du répondant quant à la situation avant l'intervention ou des données disponibles (Kumar 2005, 102 et seqq.). Le choix de méthodologies après-uniquement est généralement pragmatique : ce type de méthodologie ne prévoit qu'une seule vague de collecte de données et réduit dès lors le temps et les ressources financières nécessaires et bien souvent non disponibles.

Méthodologie
« après-
uniquement »

Les études susmentionnées décrivent des méthodologies basées sur de vastes bases de données et incluant de nombreux cas (par ex. les répondants à un sondage). Une étude de cas, en revanche, porte sur un nombre inférieur de cas, parfois même un seul. Ce type d'étude prévoit l'analyse détaillée et approfondie d'un seul cas, comme un programme, un événement, une activité ou un ensemble d'individus, limité dans le temps et dans l'espace (McMillan & Schumacher 2010, 24). Il permet une analyse plutôt holistique, d'un programme de prêt, par ex. il est également utilisé en cas de contraintes spatio-temporelles (par ex. un incident unique) ou liées au personnel (par ex. très peu de personnes à un certain poste). Ce type d'étude exploratoire (voir [sous-chapitre 3.2](#)) présente l'avantage de fournir des conclusions détaillées sur les opinions et les relations interpersonnelles dans des situations assez spécifiques. Ses résultats sont par contre difficiles à transférer à d'autres cas.

Étude
de cas

Questions et tâches à accomplir

1. Décrivez les deux fonctions principales des méthodologies.
2. Quel avantage présentent les études longitudinales par rapport aux études transversales ?
3. Quelles sont les différences entre les études des tendances, des cohortes et par panel ?
4. Quelles méthodologies observent les résultats ? Veuillez les décrire.

Lectures complémentaires

- Babbie, E. R. (2004). *The practice of social research* (10ème édition). Southbank : Wadsworth, 101-106.
- Bryman, A. (2004). *Social research methods* (2ème édition). Oxford : University Press, 33-54.
- Kumar, R. (2005). *Research methodology: A step-by-step guide for beginners* (2ème édition). London : SAGE, 83-113.

4.3 Zoom sur les méthodes de collecte de données

Nous avons abordé les stratégies de recherche qualitatives et quantitatives au [sous-chapitre 3.3](#). Nous savons donc que le choix de l'une ou l'autre stratégie de recherche prédétermine en grande partie le type de données attendues. Au [sous-chapitre 3.4](#), nous avons appris que les méthodes de recherche empirique en sciences sociales comme l'analyse de contenu, les observations et les sondages favorisent la planification, la structuration et la mise en place efficaces de la collecte de données. Au [sous-chapitre 4.2](#), nous avons pris connaissance de la multitude d'études réalisables sur la base de sondages et nous en avons analysé plusieurs exemples spécifiques aux établissements d'enseignement supérieur. Étant donné que les gestionnaires de la qualité au sein des établissements d'enseignement supérieur travaillent régulièrement avec des méthodes et techniques de sondage, nous allons à présent nous concentrer sur les méthodes de recherche les plus courantes, généralement basées sur des données quantitatives. Avant cela, nous devons aborder brièvement les méthodes d'analyse de contenus et d'observation. Elles sont effectivement couramment utilisées dans les processus d'évaluation menés dans les établissements d'enseignement supérieur.² Ces méthodes reposent généralement sur des données qualitatives.

Analyse de
contenus

Dans une **analyse de contenus**, nous recueillons et analysons des données issues de textes variés, comme les commentaires radio ou télévisés, les images, les films, etc. L'analyse peut être de type quantitative, qualitative ou les deux (Trochim 2005, 127). Lorsque les processus d'évaluation dans les établissements d'enseignement supérieur prévoient une analyse de contenus, ils portent sur un texte tel que les rapports de développement stratégiques, les descriptifs de module / cours, les réglementations des appréciations, mais aussi les registres de publications, etc. Par exemple, si un établissement d'enseignement supérieur souhaite savoir dans quelle mesure les programmes d'étude ont un profil international, il peut analyser les réglementations du programme au moyen d'une analyse de contenus, laquelle cherchera à déterminer si ce programme prévoit des semestres obligatoires de cours à l'étranger ou s'il permet des séjours d'étude à l'étranger à titre volontaire (pour de plus amples détails sur l'analyse de contenus, veuillez consulter Babbie 2004, 314 et seqq.).

Observations

Les observations jouent un rôle relativement secondaire dans le travail des gestionnaires de la qualité au sein des établissements d'enseignement supérieur. Une observation est une manière ciblée, systématique et sélective d'observer et écouter une interaction au moment où elle se produit. Nous parlons d'**observation participante** lorsqu'un chercheur prend part aux activités du groupe observé au même titre que ses membres, que ceux-ci sachent ou non qu'ils sont observés. Une **observation non participante** n'associe pas le chercheur aux activités du groupe observé. Celui-ci reste un observateur passif, qui regarde et écoute les activités et en tire des conclusions (Kumar 2005, 119 et seqq.). Les établissements d'enseignement supérieur ont recours aux observations lorsque, par exemple, ils évaluent un concept pédagogique, observent l'interaction entre professeurs et étudiants (pour de plus amples détails sur l'**observation**, veuillez consulter Bryman 2004, 164 et seqq.).

Les sondages,
méthode de
recherche
la plus
courante

Nous abordons enfin les **méthodes de recherche par sondage**, probablement la forme la plus courante d'analyse de données en sciences sociales. De nombreuses questions de recherche portant, par ex., sur les avis ou

² Nous n'aborderons pas ici les méthodes expérimentales car elles sont rarement utilisées dans les processus d'évaluation au sein des établissements d'enseignement supérieur (pour de plus amples informations sur les expériences, veuillez consulter Bryman 2004, 34 et seqq.).

comportements des êtres humains, peuvent faire l'objet d'une analyse efficace au moyen d'un sondage.

Nous pouvons différencier les sondages sur la base de trois grandes caractéristiques :

Tout d'abord, le niveau de structuration (ou de normalisation) d'un sondage peut varier. Un sondage peut être structuré (ou normalisé), partiellement structuré (ou partiellement normalisé) ou non structuré (non normalisé).

Ensuite, les questionnaires peuvent être complétés par le répondant avec ou sans l'aide d'un sondeur. Dans le second cas, nous parlons d'un sondage auto-administré (par ex. un sondage papier). Dans le premier cas, il s'agit d'un sondage administré (par ex. un entretien téléphonique, au cours duquel le sondeur peut aider le répondant à compléter le questionnaire ou commenter les questions). Nous verrons que les ordinateurs peuvent eux aussi jouer le rôle du sondeur et administrer un sondage (par ex. un sondage en ligne). Fondamentalement, une collecte de données est un entretien si une personne (un sondeur) pose des questions à une autre (un répondant) et s'il y a une interaction entre le sondeur et le répondant (Babbie 2004, 263 et seq.).

Enfin, les sondages peuvent être effectués oralement (par ex. un entretien téléphonique) ou par écrit (par ex. un questionnaire papier).

Penchons-nous à présent de plus près sur les sondages structurés. Dans les **sondages structurés**, toutes les questions du questionnaire sont formulées en détail et l'ordre et les réponses possibles sont prédéterminés. Cela signifie que tout le sondage est effectué selon la même norme. Le niveau de normalisation est généralement très élevé dans le cas des sondages écrits. Les sondages structurés visent à recueillir des données (en grand nombre). En raison de cette normalisation, ces données peuvent être agrégées (voir [sous-chapitre 7.1](#)) et évaluées facilement puisque les commentaires des répondants ont tous la même forme.

Les sondages structurés

Note: Agréger des données, c'est les réunir en un ensemble plus large. Par exemple, les étudiants de physique, chimie et biologie peuvent être rassemblés et former le groupe des étudiants en sciences naturelles.

Agrégation de données

Un **sondage semi-structuré** est une simple ligne directrice du sondage qui permet de structurer celui-ci. Si nécessaire, il est possible de sortir de ces lignes directrices, par exemple pour mieux comprendre le sujet et ainsi poser des questions supplémentaires.

Les sondages semi-structurés

Dans un **sondage non structuré**, il n'est pas du tout fait recours à un questionnaire ou à une ligne directrice de sondage. Seuls des sujet ou mots clefs prédéterminés peuvent être abordés lors du sondage. Les méthodologies qualitatives ont souvent recours à des sondages semi-structurés ou non structurés, et les sondages structurés sont plus courants dans les méthodologies quantitatives.

Les sondages non structurés

Nous avons abordé le niveau de structuration des sondages. Voyons à présent les différents modes d'administration des sondages et la question de savoir si les répondants sont aidés au moment de répondre aux questions ou non.

Les sondages administrés et auto-administrés

Les sondages papier sont **des sondages administrés** si plusieurs répondants complètent un questionnaire au même moment (par ex. l'évaluation d'un cours) et si un sondeur est présent pour répondre aux questions éventuelles.

Sondages papier et parcourrier électronique **Les sondages auto-administrés** sont ceux pour lesquels un groupe de répondants reçoit un questionnaire par courrier électronique, à compléter pour une date définie et à renvoyer au sondeur (par ex. un sondage par courrier électronique dans le cas d'une étude de suivi des diplômés).

Les sondages assistés par ordinateur Les sondages papier peuvent donc être administrés ou auto-administrés. Les sondages électroniques sont généralement auto-administrés, même si l'ordinateur peut aider les répondants (par ex. grâce à des codes programmés définis pour utiliser le modèle de questionnaire normalisé). Les sondages électroniques peuvent être effectués par courrier électronique ou en ligne. Dans le premier cas, les questionnaires sont envoyés par courrier électronique (par ex. en pièce jointe). Le répondant doit répondre au sondage et le renvoyer à l'expéditeur. Les sondages en ligne (ou les sondages web) sont mis en ligne sur un site web spécifique et complétés au moyen d'un navigateur web. Les données sont habituellement sauvegardées directement dans une base de données en ligne. Il existe déjà une variété de logiciels bon marché ou gratuits pour les sondages web. Bien souvent, ils prévoient déjà une fonction d'analyse. Les sondages par courrier électronique sont donc de moins en moins courants. Il reste toutefois utile d'inviter les répondants à y prendre part et de leur envoyer par courrier électronique le lien vers le sondage en ligne (si les coordonnées de courriel sont disponibles).³

Les sondages téléphoniques **Les sondages téléphoniques** sont une forme relativement courante de sondage administré. Il s'agit de sondages structurés et écrits. Contrairement aux sondages papier, les questions sont lues au répondant au cours d'un appel téléphonique. Les réponses sont en général directement transférées dans une base de données électronique par le sondeur.

Les sondages en face à face Les sondages téléphoniques sont également une forme spécifique d'**entretien en face à face**, au cours duquel un sondeur lit les questions à un répondant et transcrit les réponses sur papier ou sur ordinateur (au format électronique).

Les groupes témoins et les discussions de groupe Si plusieurs répondants prennent part à un sondage, on parle d'entretien de groupe. Ces entretiens de groupe sont à ne pas confondre avec les groupes témoins, à savoir des discussions de groupe guidées. Le sondeur fait davantage office de modérateur : il se limite à présenter un sujet et veille à la participation de l'ensemble des répondants à la discussion. Les groupes témoins ont pour objectif de recueillir l'avis des participants sur un sujet particulier. L'avantage est que l'interaction et la dynamique de groupe permettent de recueillir des informations plus détaillées que lors d'un entretien avec une seule personne.

Le tableau suivant illustre le niveau de structuration, le mode et les caractéristiques des différentes formes de sondage.

³ Pour de plus amples informations sur les sondages en ligne, veuillez consulter Couper 2008.

Sondage	Mode	Exemple	Caractéristiques
Structuré	Oral	Entretien individuel ou de groupe, sondage téléphonique (administré)	Structure rigide, contenu rigide, questions et formulation rigides recueillir des données et éléments quantitatifs « mesurer »
	Écrit	Sondage par courrier électronique, sondage en ligne (auto-administré)	
Semi-structuré	Oral	Discussion guidée, entretien de groupe (administré)	Différents niveaux de flexibilité et de spécificité
	Écrit	Sondage de spécialistes (auto-administré)	
Non structuré	Oral	Sondage de spécialistes, discussion de groupe (administré)	Structure flexible, contenu flexible, questions et formulation flexibles
	Écrit	Sondage informel de spécialistes (administré)	

Tableau 14 Caractéristiques des sondages (tableau des auteurs)

Questions et tâches à accomplir

1. Pourquoi un gestionnaire de la qualité privilégierait-il une méthode de sondage structurée à une méthode non structurée pour recueillir des données ?
2. Expliquez dans quelle mesure les méthodes de sondage ne peuvent s'appliquer à votre établissement d'enseignement supérieur.

Lectures complémentaires

- Babbie, E. R. (2004). *The practice of social research* (10ème édition). Southbank : Wadsworth, 256-280.
- Neuman, W. L. (2000). *Social research methods: Qualitative and quantitative approaches* (4ème édition). Boston : Allyn and Bacon, 271-285.
- Salant, P & Dillman, D. A. (1994). *How to conduct your own survey*. New York : Wiley, 101-136.

4.4 Échantillonnage et représentativité

Recensement complet et sondages par échantillon

Au cours d'un sondage, il n'est parfois pas possible de procéder à un **recensement complet**. La « population d'une étude » peut ne pas être sondée dans son ensemble car cette **population** est trop importante et en interroger tous les membres prendrait trop de temps. Par exemple, un sondage mené auprès de tous les étudiants d'un établissement d'enseignement supérieur est parfois impossible en raison de contraintes financières et / ou des délais impartis. Aussi a-t-on souvent recours aux « **sondages par échantillon** ». Les sondages par échantillon permettent de recueillir des informations auprès d'un nombre relativement restreint de répondants pour déterminer les caractéristiques de l'ensemble d'une population. Ce type de sondage nécessite moins de temps et de ressources financières que ceux qui couvrent de nombreux répondants (Salant & Dillman 1994, 53). Pour créer des **échantillons**, nous sélectionnons certaines unités (un échantillon) d'un groupe plus large d'unités (la **population échantillonnée**).

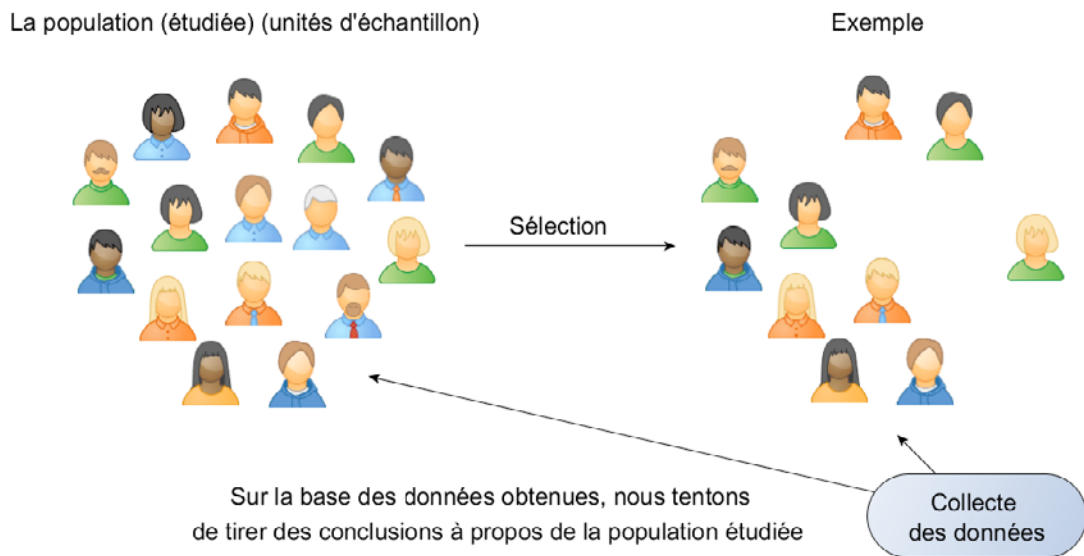


Figure 6 Population (étudiée) et échantillon (Kumar 2005, 164)

Représentativité et sondages par échantillon

Pour pouvoir tirer des conclusions à propos d'une population sans procéder à un recensement complet, le groupe à interroger (l'échantillon) doit être semblable à la population. Cela signifie que la distribution des caractéristiques dans la population étudiée et dans l'échantillon doit être presque identique. Dans le cas contraire, il s'agit d'un biais d'échantillonnage. Les personnes sélectionnées ne sont pas représentatives de la population étudiée dont elles sont issues. Pour citer un exemple : Une population étudiée compte 45% de femmes et 55% d'hommes. L'échantillon doit présenter les mêmes proportions. Cela signifie qu'un échantillon est représentatif lorsqu'il reflète avec précision la population étudiée et en est donc un microcosme (Babbie 2004, 178 et seqq.).

Pour créer des échantillons représentatifs, nous pouvons recourir à différentes méthodes (Kumar 2005, 171 et seqq.; Babbie 2004, 186 et seqq.). Par exemple, un échantillon peut reposer sur une sélection aléatoire. Nous parlons de sélection aléatoire lorsque chaque unité de la population présente une probabilité égale ou indépendante d'être sélectionnée. Une probabilité « égale » signifie, dans ce cas-ci, que la probabilité d'être sélectionné est la même pour tous, et qu'aucune unité ne peut être sélectionnée plusieurs fois. Une probabilité « indépendante » signifie que la sélection est purement aléatoire et qu'aucun facteur n'intervient (par ex. les préférences de la per-

sonne responsable de la sélection). Si les échantillons sont sélectionnés au moyen de la méthode ci-dessus, la probabilité d'obtenir un échantillon qui présente les caractéristiques spécifiques et mesurables de la population étudiée est relativement élevée.

Questions et tâches à accomplir

1. Qu'entendez-vous par le terme « représentativité » ?
2. Quelles caractéristiques un échantillon doit-il présenter pour être le plus représentatif possible ?

Lectures complémentaires

- Babbie, E. R. (2004). *The practice of social research* (10ème édition). Southbank : Wadsworth, 178-217.
- Bryman, A. (2004). *Social research methods* (2ème édition). Oxford : University Press, 83-106.
- Groves, R. M., Fowler, F. J., Couper, M., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2009). *Survey methodology* (2ème édition). Hoboken, NJ : Wiley, 69-139.
- Salant, P & Dillman, D. A. (1994). *How to conduct your own survey*. New York : Wiley, 53-71.

4.5 Méthodes de collecte de données : Enjeux et difficultés

Lors de la collecte de données au moyen de méthodes et techniques socio-scientifiques il existe toujours diverses sources d'erreurs qui influencent négativement la collecte de données et, par conséquent, les résultats à obtenir. Pour en revenir aux considérations sur l'échantillonnage et la représentativité (voir [sous-chapitre 4.4](#)), il est clair que nos résultats peuvent être biaisés et que nous ne devons pas généraliser ces résultats lorsque l'échantillon n'est pas représentatif.

Outre les enjeux posés par l'échantillonnage, il convient également d'ajouter que d'autres effets peuvent entraîner des distorsions ou des erreurs lors d'un sondage, en particulier dans le cas d'entretiens (voir [sous-chapitre 4.3](#)).

*« Dans ce contexte, le terme « erreur » ne revêt pas son sens le plus courant mais fait référence à un écart entre ce à quoi le processus de sondage est supposé aboutir et ce à quoi il aboutit »
(traduit de Groves et al. 2009, 40)*

Les catalyseurs ou sources d'erreurs peuvent être des sondeurs ou des répondants, un questionnaire ou certaines questions du questionnaire, mais aussi les moyens techniques utilisés pour effectuer le sondage (par ex. l'instabilité de l'accès à Internet). Les sources d'erreurs sont variées, comme témoigne le tableau (Bryman 2004, 105 et seqq.).

Les diverses sources d'erreurs lors de la collecte de données suivant

Le choix des sources et des exemples mentionnés ci-dessus démontre que le risque d'erreur est présent à presque tous les stades d'un processus de collecte de données. Parfois, les sources d'erreurs ne peuvent pas être totalement exclues (par ex. on ne peut pas forcer les répondants à prendre part à un sondage car cela engendrerait aussi des erreurs dans la mesure où ils fourniraient de fausses réponses sous la contrainte). Nous savons toutefois qu'il existe diverses méthodes et techniques (par ex. en matière d'échantillonnage) pour réduire ces sources d'erreurs (voir [sous-chapitre 4.4](#)). De plus amples détails sur la façon de réduire les erreurs dans le processus de collecte de données (voir [chapitre 6](#)) et le traitement des données (voir [chapitre 7](#)) seront données dans les chapitres subséquents mentionnés.

Source	Exemples
Échantillonnage et lien avec l'échantillon	<p>Différence entre un échantillon et la population étudiée dont il est issu, malgré le choix d'un échantillon aléatoire.</p> <p>Les répondants refusent catégoriquement de prendre part à un entretien (auto-sélection, non-réponse de l'unité).</p> <p>Les répondants ne répondent pas à certaines questions du questionnaire (non-réponse à un point).</p>
Collecte de données	<p>Mauvaise formulation de la question ou question mal posée par le sondeur.</p> <p>Mauvaise compréhension ou problèmes de mémoire du répondant.</p> <p>Désirabilité sociale : les répondants peuvent avoir tendance à répondre aux questions d'une manière jugée favorable par d'autres.</p>
Traitement des données	<p>Gestion erronée des données.</p> <p>Erreurs d'encodage des réponses.</p>

Tableau 15 Sources d'erreurs dans les sondages (tableau des auteurs)

Questions et tâches à accomplir

1. Expliquez pourquoi et quand des erreurs peuvent apparaître dans un sondage.
2. Avez-vous déjà pris part à un sondage dans lequel vous avez identifié des erreurs du type de celles mentionnées ci-avant ? Veuillez décrire ces erreurs et expliquer en quoi il s'agit d'erreurs.

Lectures complémentaires

- Groves, R. M., Fowler, F. J., Couper, M., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2009). *Survey methodology* (2ème édition). Hoboken, NJ: Wiley, 39-211.

4.6 Lignes directrices pour le choix de la méthode de sondage appropriée

Le choix d'une méthode de sondage appropriée est probablement l'une des décisions les plus délicates au cours d'une étude. Pour pouvoir choisir ou écarter une méthode, il peut être utile de répondre aux questions suivantes, basées sur les trois questions fondamentales de la recherche empirique en sciences sociales (voir [sous-chapitre 3.2](#)) :

- Que souhaitez-vous savoir ?
- Que souhaitez-vous faire des résultats du sondage ?
- Après de qui souhaitez-vous effectuer ce sondage ? Comment pouvez-vous joindre ces personnes ?
- Comment souhaitez-vous poser votre question ? Comment peuvent-ils répondre à votre question ?

Par exemple, la population étudiée influence largement le choix de la méthode de sondage. Les sondages en ligne ne sont pas envisageables si les répondants ne disposent pas d'un ordinateur et d'une connexion à Internet, même s'ils disposent d'une adresse électronique à utiliser pour l'administration du sondage. Les méthodes de sondage qui prévoient la présence de sondeurs sont impossibles si les ressources humaines et financières nécessaires ne sont pas suffisantes. Si les résultats doivent être disponibles rapidement, les méthodes de sondage semi- ou non-structurées ne conviennent pas non plus.

Le tableau 16 suivant présente davantage de questions à poser à propos des avantages et limites des méthodologies présentées dans le présent module. Note : Il doit être répondu à toutes les questions dans le respect des conditions préalables en vigueur au sein de votre établissement d'enseignement supérieur. Certaines réponses aux questions mentionnées dans ce tableau peuvent varier en raison de leur interdépendance (par ex. un grand échantillon augmente le prix d'un sondage en face à face).

Méthode de sondage appropriée				
Question à poser	Courrier électronique	En ligne	Téléphonique	En face à face
Coordonnées : De quels éléments ai-je besoin pour prendre contact avec le répondant ?	Adresse postale (si nécessaire, autres renseignements pour trouver l'adresse)	Adresse électronique (si nécessaire, autres renseignements pour trouver l'adresse)	Numéro de téléphone (si nécessaire, autres renseignements pour trouver le numéro)	Adresse postale ou électronique ou numéro de téléphone

Question à poser	Méthode de sondage appropriée			
	Courrier électronique	En ligne	Téléphonique	En face à face
Coûts : Quelles sont les ressources financières disponibles pour ce sondage ?	Moyenne (par ex. impression des questionnaires, distribution et emballage, port de retour, collecte de données, etc.)	Faibles (aucun coût d'impression de questionnaires, distribution et emballage, port de retour, collecte de données, etc., peut-être des coûts pour le logiciel de sondage en ligne)	Élevés ou très élevés (par ex. coûts pour un logiciel de sondage téléphonique assisté par ordinateur, coûts téléphoniques, salaire du sondeur, collecte de données, etc.)	Très élevés (par ex. coûts pour le personnel assisté par ordinateur, logiciel de sondage, coûts téléphoniques, applications mobiles, salaire du sondeur, frais de transport et de logement, collecte de données, etc.)
Appui technique : Combien de personnes sont nécessaires à la réalisation du sondage ? Quelles sont les infrastructures déjà disponibles (logiciel, matériel) ?	Léger (sans logiciel pour collecter les données), minime (avec logiciel pour collecter les données)	Minime	Important ou très important	Important ou très important
Échantillon : Combien d'entretiens doivent être menés ? Quelle est la taille de l'échantillon ?	Nombre / taille moyen(ne) à grand(e)	Très élevés	Nombre / taille grand(e)	Nombre / taille faible à moyen(ne)
Complexité : Quel doit être le niveau de complexité du contenu du sondage ?	Nombre / taille faible à moyen(ne)	Élevé à très élevé	Faible à moyen	Élevé à très élevé
Longueur : Quelle doit être la longueur du questionnaire ? Les répondants pourront-ils répondre aux questions ?	Moyens	Moyenne à longue	Moyenne à longue	Très longue

Question à poser	Méthode de sondage appropriée			
	Courrier électronique	En ligne	Téléphonique	En face à face
Durée : Combien de temps faudra-t-il pour collecter les données ?	Beaucoup (sans logiciel pour la collecte de données) peu (avec logiciel pour la collecte de données)	Très peu	Très peu	De beaucoup à un laps de temps moyen
Taux de réponse (voir sous-chapitre 6.4) Quel est le taux de réponse attendu ?	Moyens	Faibles	Faible à moyen	Nombre / taille grand(e)

Tableau 16 Questions pour le choix de la méthode de sondage appropriée (tableau des auteurs)

Questions et tâches à accomplir

1. Si vous vouliez savoir de quels services complémentaires les étudiants de première année ont besoin au cours du premier semestre au sein de votre établissement d'enseignement supérieur, quelle méthodologie choisiriez-vous et pourquoi ?

Lectures complémentaires

- Salant, P & Dillman, D. A. (1994). *How to conduct your own survey*. New York : Wiley, 33-52.
- Trochim, W. M. K. (2005). *Research methods: The concise knowledge base*. Mason, Ohio : Cengage Learning, 94-97.

Chapitre 5

Mesure et opérationnalisation

5	Mesure et opérationnalisation	81
5.1	La collecte de données est un processus de longue haleine : De la conceptualisation à la mesure.	81
5.2	Exemple : Opérationnalisation du concept « Réussite des étudiants »	82
5.3	Les niveaux de mesure et les types d'échelle	85
5.4	Critères de la recherche en sciences sociales	87
5.5	Types d'erreurs courants en recherche évaluative	88



Au terme de ce chapitre, vous devriez pouvoir...

- citer et décrire les différentes étapes depuis les concepts jusqu'à la mesure ;
- opérationnaliser vos propres concepts ;
- citer et décrire les niveaux de mesure et les types d'échelle ;
- expliquer l'importance des critères de qualité des méthodes de recherche en sciences sociales en termes d'objectivité, de validité et de fiabilité ;
- citer et décrire les types d'erreurs courants en recherche évaluative.

5 Mesure et opérationnalisation

5.1 La collecte de données est un processus de longue haleine : De la conceptualisation à la mesure

Un sondage est souvent également appelé **une mesure** car il fait partie d'un processus scientifique de collecte de données (Trochim 2005, 49). Les mesures sont un processus courant dans la vie de tous les jours : En réalité, nous mesurons tous les jours, des précipitations moyennes en litres par mètre cube à la portion de riz nécessaire pour notre plat préféré, en passant par la taille des enfants en centimètres. Ces mesures sont différentes de la collecte de données parce qu'elles reposent sur un concept structuré et sur une échelle structurée. Par exemple, la taille d'une personne, estimée à 1m75, peut être exprimée en centimètres (175) et cette mesure repose sur une échelle métrique. Mais comment faire si nous souhaitons mesurer la qualité de l'enseignement au sein de notre établissement d'enseignement supérieur ?

Les mesures dans la recherche en sciences sociales

Pouvons-nous mesurer **des variables** aussi facilement ? La réponse est oui, car nous pouvons tout mesurer (Babbie 2004, 119) ! Cela a l'air plus facile que cela ne l'est en réalité. Bien souvent, nous ne savons pas exactement ce qui doit être mesuré. Chacun associe le concept de « qualité de l'enseignement », à différentes perceptions et images mentales. En outre, nous ne pouvons pas toucher ni voir cette image mentale. Nous ne pouvons dès lors pas directement l'observer ni la mesurer.

Les concepts nécessaires au processus de mesure

Comme nous l'avons vu au [sous-chapitre 2.8](#), un **concept** est à la base de tout processus de collecte de données (et, partant, des mesures). La première étape d'une mesure est donc la création d'une image mentale (ou concept) qui fasse l'unanimité. Le processus visant à dégager un consensus autour de la signification des termes s'appelle la **conceptualisation**, et son résultat est un concept (Babbie 2004, 120). La deuxième étape est la traduction de cette image mentale afin que ce concept non observable devienne mesurable. Pour des raisons de complexité, ce processus peut nécessiter de diviser le concept en différentes composantes ou dimensions. Comme nous l'avons vu au [sous-chapitre 2.8](#), cette « traduction » est appelée **opérationnalisation**. Au cours de cette opérationnalisation, nous cherchons des critères observables ou mesurables, et donc quantifiables. Ces critères sont appelés **indicateurs** car ils fournissent une indication à propos d'un concept et de ses dimensions, et ils ont donc un lien logique avec ce concept. Les indicateurs représentent le concept et nous les utilisons pour repérer les concepts qui ne sont pas directement quantifiables.

Trois étapes, des concepts aux mesures

Pour rendre les indicateurs quantifiables ou mesurables, nous devons les convertir, au cours d'une troisième étape, en **variables** (Bryman 2004, 66 et seq.; Kumar 2005, 54 et seq.). Ces variables peuvent avoir différentes valeurs connectées de manière logique. Par exemple, la variable « genre » peut avoir pour valeurs « homme » ou « femme » (voir [sous-chapitre 5.3](#)).

Que sont les variables ?



Figure 7 Étapes de l'opérationnalisation à la mesure (illustration des auteurs)

Les résultats de ces étapes constituent le point de départ pour la conceptualisation des questions, des lignes directrices associées aux questions et des questionnaires des sondages. Kumar (2005) insiste sur la réussite de l'opérationnalisation et formule à cet effet les recommandations suivantes :

« Si vous utilisez un concept au cours de votre étude, veillez à bien définir ses indicateurs et leur mesure. (...) N'intégrez pas dans votre problème de recherche des concepts si vous n'êtes pas sûr de la manière de les mesurer. Cela ne vous empêche toutefois pas de mettre au point une procédure de mesure en cours d'étude. La majeure partie du travail d'élaboration aura lieu au cours de l'étude. Il est toutefois impératif d'avoir une idée déjà assez claire du mode de mesure des concepts à ce stade. »

(traduit de Kumar 2005, 43)

Au [sous-chapitre 5.2.](#), nous présenterons plus clairement les étapes de l'opérationnalisation à l'aide d'un cas-précis : l'opérationnalisation du concept « Réussite des étudiants ».

5.2 Exemple : Opérationnalisation du concept « Réussite des étudiants »

Le modèle multifactoriel de Munich pour la qualité des cours, créé par le psychologue Heiner Rindermann, tente de systématiser les facteurs qui influencent la réussite, en d'autres termes la qualité.

À ce jour, il n'existe pas de modèle scientifique pour la qualité de l'enseignement. Pas exemple, selon le modèle de Rindermann, un indicateur de bon enseignement est l'assimilation d'acquis d'apprentissage d'un cours par les étudiants. La réussite d'un cours dépend dès lors d'une influence positive sur les acquis d'apprentissage. Cette influence est la première dimension de notre concept.

Dimensions et indicateurs

En somme, Rindermann part du principe que trois dimensions contribuent à l'assimilation d'acquis d'apprentissage et donc à la réussite de l'enseignement et de l'apprentissage : les capacités et le comportement de l'enseignant (par ex. une élocution claire et intelligible), les capacités et le comportement des étudiants (par ex. une contribution active aux débats en classe) et les conditions-cadres au sein desquelles le cours se donne (par ex. le cours compte 10 étudiants). Ces dimensions peuvent à leur tour être opérationnalisées et transférées à différents indicateurs et donc être utilisées dans un sondage (voir [sous-chapitre 5.1](#)). Il nous reste donc encore à différencier les indicateurs, afin que ceux-ci produisent des résultats concrets et mesurables. Le niveau de différenciation peut porter à conclure qu'un aspect particulier existe ou non (par ex. question: « Des outils pour la préparation et la relecture (comme des transcriptions, des transparents, des services

d'apprentissage en ligne) ont-ils été fournis ? Réponses possibles : « oui » ou « non »). Il peut aussi amener à déterminer le niveau ou la portée d'un aspect déterminé (par ex. question : « Combien de temps avez-vous consacré à la préparation ou aux tâches assignées pour le cours la semaine dernière ? » Réponse possible : « Veuillez compléter par un chiffre ») (voir [sous-chapitre 5.3](#)).

Le tableau 17 suivant illustre la manière d'opérationnaliser les dimensions, selon Rindermann, et de les traduire en indicateurs. Précisons qu'un indicateur peut être décrit au moyen de différentes questions. Par exemple, pour en savoir plus sur l'investissement des étudiants dans un cours déterminé, vous pouvez demander comment les présentations des étudiants sont perçues par leurs camarades, ou comment les étudiants jugent leur participation active au cours, ou si le travail de groupe est jugé utile, etc.

Des concepts
aux variables

Dimension	Indicateur	Variables (exemples de questions, énoncés dans un sondage d'étudiants)
Prestation de l'enseignant	Caractère structurel	Les objectifs d'apprentissage sont clarifiés. Les exigences du cours sont clairement annoncées. Le contenu d'une séance de cours est cohérent et intégré dans le contexte plus large de la matière. ...
	Rhétorique	Le professeur parle clairement et de manière intelligible. Le contenu du cours est présenté d'une manière compréhensible. L'utilisation des méthodes et techniques d'apprentissage est appropriée. Je parviens à suivre la présentation du contenu du cours et des tâches y afférentes. ...
	Investissement	Le professeur... ... semble bien préparé aux différentes séances de cours. ... montre un intérêt réel pour la réussite de l'apprentissage des étudiants. ... facilite les questions des étudiants et leur participation active. ... apporte une réponse appropriée aux questions ou remarques des étudiants. ... crée un environnement d'apprentissage constructif. ... est suffisamment disponible pour les requêtes ou conseils supplémentaires éventuels. ...
...

Dimension	Indicateur	Variables (exemples de questions, énoncés dans un sondage d'étudiants)
Étudiants	Connaissances préalables	Le contenu du cours est en lien avec mes connaissances préalables. ...
	Investissement des étudiants	Je pense que la plupart des étudiants de ce cours... ... assistent régulièrement aux différentes séances de cours ... sont correctement préparés aux différentes séances de cours. ... participent activement dans la mesure du possible. ... suivent le cours avec intérêt et attention. ... contribuent suffisamment à la réussite du cours. ...

Conditions-cadres	Locaux	Le nombre d'étudiants par rapport à la taille du local est acceptable. Les équipements prévus dans l'espace d'apprentissage (meubles, technologies médias, conception d'une plateforme d'apprentissage dans le cadre de l'apprentissage en ligne, etc.) semblent utiles.
	Exigences du cours	Je peux remplir les exigences du cours (préparation et post-traitement, participation active). De manière générale, je peux assurer la préparation requise et effectuer les tâches assignées pour le cours dans les délais que je me suis fixés.

Acquis d'apprentissage	Renforcement des compétences	Le cours m'aide à... ... présenter les faits et concepts de base de la matière du cours. ... être capable de travailler sur les questions ou problèmes typiques de la matière du cours. ... débattre de manière critique des limites et possibilités de la matière. ... améliorer mes méthodes d'apprentissage. ... améliorer ma capacité d'étude/travail autodidacte. ... éveiller mon intérêt pour la matière. Le contenu du cours est en lien avec le domaine professionnel et l'application des connaissances. Le mode d'apprentissage me permet d'approfondir ma compréhension du contenu du cours à titre individuel.

Tableau 17 *Opérationnalisation de la qualité du cours, (adapté de Rindermann 1998)*

 **Questions et tâches à accomplir**

1. Pourquoi avons-nous besoin de concepts pour les mesures ?
2. Décrivez le processus d'opérationnalisation.
3. Imaginez un concept, ses dimensions et indicateurs possibles.

 **Lectures complémentaires**

- Babbie, E. R. (2004). *The practice of social research* (10ème édition). Southbank : Wadsworth, 119-140.
- Bryman, A. (2004). *Social research methods* (2ème édition). Oxford : University Press, 65-69.

5.3 Les niveaux de mesure et les types d'échelle

La définition d'indicateurs ou de questions pour la mesure de nos concepts peut se révéler une tâche difficile. La formulation des questions détermine le type de variables et la classification de notre mesure en fonction d'une **échelle de mesure** déterminée.

Nous pouvons classer les différents types de variables pour notre mesure : Les variables peuvent être **constantes** et ne définir qu'une seule **valeur** ou une seule **catégorie** (par ex. un étudiant). Elles peuvent au mieux définir deux valeurs (par ex. oui/non). Dans ce cas, elles sont **dichotomiques**. Elles peuvent au contraire décrire plus de deux catégories, et sont dès lors **polytomiques** (par ex. chrétien, musulman, hindou, et juif). Ces **variables nominales** se distinguent des **variables continues**, **également appelées variables métriques**, comme les revenus ou l'âge. Les variables métriques peuvent renvoyer à n'importe quelle valeur sur l'échelle de mesure (par ex. revenus : dollars et centimes ; âge : année, mois et jour). Les variables peuvent par ailleurs être qualitatives ou quantitatives.

Types de variables

Variables nominales					
Con-stantes	Dichotomi-ques	Polytomiques	Continues	Qualitatives	Quantitatives
Étudiant	Oui / non	<i>Comportements</i>	<i>Revenus (\$)</i>	<i>Sexe</i>	<i>revenus</i>
Professeur	Homme / femme	fortement favorable favorable incertain fortement défavorable	<i>Âge (an-nées)</i>	homme femme <i>Âge</i> jeune enfant	\$ par an <i>âge</i> année / mois

Tableau 18 Catégorisation des variables (tableau des auteurs, basé sur Kumar 2005, 6)

Différencier les variables quantitatives et qualitatives signifie déterminer si une variable mesure une certaine caractéristique de manière qualitative ou quantitative. Ce faisant, nous définissons l'échelle de mesure d'une variable. Il est essentiel de connaître cette différence car elle aura un effet sur l'analyse de données ultérieure, sur l'interprétation et sur la communication des résultats (voir [chapitre 7](#)).

Types d'échelles de mesure

En règle générale, on distingue quatre types d'échelles de mesure : **l'échelle nominale** (ou de classification), **l'échelle ordinale** (ou de classement), **l'échelle d'intervalle** et **l'échelle de proportion**. Le choix entre l'égalité et l'inégalité est une échelle nominale. Les échelles nominales nomment, classent ou chiffrent. Mais les valeurs de ces variables ne peuvent être classées selon leur taille (par ex. le sexe). Chaque sous-groupe présente une caractéristique spécifique commune à tous au sein de ce sous-groupe (par ex. homme). Une **échelle ordinale** présente les caractéristiques d'une échelle nominale. En outre, les sous-groupes sont en relation. Ils peuvent être classés par ordre croissant ou décroissant. On ignore dès lors les intervalles entre les valeurs individuelles (par ex. statut socio-économique : supérieur / moyen / inférieur). Si nous pouvons définir les intervalles entre les valeurs d'une variable, nous parlons d'une **échelle d'intervalle** (par ex. la température). Cette échelle présente toutes les caractéristiques des échelles nominale et ordinale. Elle a par ailleurs une unité de mesure dotée d'un point de départ et d'un point d'arrivée arbitraires. Nous pouvons ainsi également calculer des valeurs moyennes (voir [sous-chapitre 7.3](#)). L'échelle de proportion présente toutes les caractéristiques d'une échelle d'intervalle. Elle dispose en outre d'un point de départ fixe (par ex. revenus en \$) (Kumar 2005, 69).

Techniques et éléments d'échelonnement

Les sondages réalisés dans le cadre des processus d'évaluation ont souvent recours à des **techniques** et éléments **d'échelonnement** pour mesurer les comportements ou les croyances. **Ces techniques** peuvent récapituler d'échelonnement plusieurs éléments liés sur la base d'une structure logique ou empirique. **Ces éléments** sont les plus petites parties d'un outil de recherche, comme les questions ou énoncés auxquels les répondants peuvent attribuer leur degré d'accord ou de désaccord. Par exemple, au lieu de ne poser qu'une seule question sur l'amélioration des compétences, nous pouvons poser une série de questions, comme l'illustre le tableau suivant :

Veillez indiquer votre degré d'accord ou de désaccord avec chaque énoncé (veillez cocher la réponse appropriée).

Énoncés	tout à fait d'accord	plutôt d'accord	indécis ou pas sûr	plutôt pas d'accord	pas du tout d'accord
... présenter les faits et concepts de base de la matière du cours.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... être capable de travailler sur les questions ou problèmes typiques de la matière du cours.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... débattre de manière critique des limites et possibilités de la matière.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Veuillez indiquer votre degré d'accord ou de désaccord avec chaque énoncé (veuillez cocher la réponse appropriée).

Énoncés	tout à fait d'accord	plutôt d'accord	indécis ou pas sûr	plutôt pas d'accord	pas du tout d'accord
... améliorer mes méthodes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... améliorer ma capacité d'étude / travail autodidacte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... éveiller mon intérêt pour la matière	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Tableau 19 Recours à une série de questions pour l'analyse d'un seul thème (tableau des auteurs)

L'avantage du recours à une série de questions est que l'on évite de trop simplifier les questions complexes. La technique d'échelonnage nous permet donc de résumer les réponses à toute la série de questions en un seul indicateur (Salant & Dillman 1994, 87 ; Babbie 2004, 150 et seqq.)

Questions et tâches à accomplir

1. Pourquoi est-il essentiel de faire la différence entre les différents types de variables ?
2. Assurez-vous de bien connaître les différences entre les types d'échelle de mesure présentés.
2. Décrivez la finalité des éléments liés aux techniques d'échelonnage.

Lectures complémentaires

- Babbie, E. R. (2004). *The practice of social research* (10ème éd). Southbank : Wadsworth, 132-177.

5.4 Critères de la recherche en sciences sociales

Nous avons déjà beaucoup appris sur les méthodes de collecte de données et de mesure. Nous savons que le processus de collecte de données comprend différentes phases et étapes (voir [sous-chapitre 4.1](#)), qui ont des effets sur la précision et sur la qualité de nos conclusions. Heureusement, il existe des **critères de qualité** qui permettent de juger la qualité de la recherche en sciences sociales. Ces critères sont cruciaux pour pouvoir faire confiance aux données recueillies et les injecter dans notre travail. Ces trois principaux critères sont **l'objectivité, la fiabilité et la validité**. Ils seront présentés dans les points suivants.

Principaux critères de la recherche en sciences sociales : objectivité, fiabilité et validité

L'objectivité est assurée lorsque les résultats des données recueillies sont indépendants de la personne responsable. Cela signifie qu'un questionnaire normalisé ne doit pas permettre aux sondeurs de formuler les

Objectivité

questions de différentes manières. L'analyse des données doit par ailleurs être indépendante des personnes responsables de cette étape. C'est la raison pour laquelle nous avons recours à des procédures statistiques éprouvées, qui permettent de vérifier l'analyse. Enfin, l'interprétation des résultats doit elle aussi être indépendante des personnes en charge de cette étape. Cela signifie que les résultats de l'interprétation doivent reposer sur une norme unitaire (par ex. une norme).

Fiabilité La **fiabilité** est la précision formelle de nos résultats. En d'autres termes, les résultats doivent être dénués de toute erreur (voir [sous-chapitre 4.5](#)), de sorte que chaque mesure répétée dans des conditions identiques aboutisse au même résultat.

Validité La **validité** d'une mesure est respectée lorsque les résultats d'une mesure reflètent ce qui doit être mesuré. Par exemple, un sondage sur « la réussite professionnelle » ne serait pas valide s'il portait uniquement sur la réussite des études et non sur la réussite professionnelle. Dans ce cas, toutes les dimensions de notre concept n'auraient pas été prises en considération lors de notre mesure.

Une mesure valide doit également être objective et fiable. Cela signifie que nos opérationnalisations, mesures et conclusions ne sont fiables que si nous prenons les critères de qualité pour normes dans le cadre de nos propres travaux. Nous devons procéder de la sorte dans notre propre intérêt.

Questions et tâches à accomplir

1. Citez un exemple de mesure non objective.
2. Citez un exemple de mesure non fiable.
3. Citez un exemple de mesure non valide.

Lectures complémentaires

- Babbie, E. R. (2004). *The practice of social research* (10ème éd). Southbank : Wadsworth, 140-146.

5.5 Types d'erreurs courants en recherche évaluative

Nous avons découvert les principaux critères de qualité de la recherche empirique en sciences sociales et sommes conscients de leur pertinence. L'expérience a démontré que les principes et normes méthodologiques de la recherche empirique en sciences sociales ainsi que les critères de qualité mentionnés ci-avant ne peuvent être atteints, en particulier au cours des processus d'évaluation. Par exemple, les participants à une évaluation ne peuvent travailler en toute indépendance en raison d'un manque de ressources (voir [sous-chapitre 3.1](#)). Malgré la présence d'obstacles de ce type, il convient de toujours éviter **les types d'erreur** suivants en recherche évaluative.

Type d'erreur	Raison / exemple
Poser de « mauvaises » questions de recherche	Les processus d'évaluation sommatifs ou formatifs reposent sur des critères de qualité ou hypothèses qui ne correspondent pas à la finalité de la recherche.
Demander une étude évaluative après qu'une décision a été prise sur un programme	Se servir d'une étude uniquement pour retarder ou justifier une décision déjà prise
Exiger l'utilisation d'une méthodologie de recherche / d'une technique de collecte de données qui ne correspond pas à la finalité de l'évaluation du programme	Les résultats peuvent ne pas être valides, éthiques ni utiles ; l'évaluation serait inutile et une perte de temps et d'argent
Influencer la méthodologie de recherche ou le processus de collecte de données pour s'assurer qu'il produise les résultats attendus	Les résultats peuvent ne pas être valides, éthiques ni utiles ; l'évaluation serait inutile et une perte de temps et d'argent
Poursuivre un programme même si les résultats de l'évaluation montrent sans équivoque son inefficacité ; ou fermer un programme même si les résultats montrent sans équivoque sa grande efficacité	L'évaluation serait inutile et une perte de temps et d'argent
Effacer des résultats positifs s'ils s'opposent à la fermeture d'un programme ; ou effacer des résultats négatifs s'ils s'opposent à la poursuite / l'extension d'un programme	Les résultats peuvent ne pas être valides, éthiques ni utiles ; l'évaluation serait inutile et une perte de temps et d'argent

Tableau 20 Types d'erreur en recherche évaluative (tableau des auteurs, basé sur Neumann 2000, 105)

Les types d'erreur mentionnés dans le tableau ci-avant ont trait à des utilisations abusives volontaires des méthodes et techniques en recherche empirique en sciences sociales. Cela va à l'encontre du paradigme de « bonne pratique scientifique » et doit être évité.



Lectures complémentaires

- Babbie, E. R. (2004). *The practice of social research* (10ème édition). Southbank : Wadsworth, 61-82.
- Neuman, W. L. (2000). *Social research methods : Qualitative and quantitative approaches* (4ème édition). Boston : Allyn and Bacon, 89-119.

Chapitre 6

La formulation de questions et de la méthodologie

6	La formulation de questions et de la méthodologie .	91
6.1	Fonctions et types de questions	91
6.2	La formulation de bonnes questions	94
6.3	Conception du questionnaire	98
6.4	Mesures pour l'obtention d'un taux élevé de réponse	99



Au terme de ce chapitre, vous devriez pouvoir...

- distinguer les différents types de questions et leurs fonctions ;
- formuler des recommandations pour la formulation des questions ;
- appliquer les règles de la conception de questionnaires à vos propres projets de sondage ;
- appliquer les mesures appropriées pour obtenir un taux élevé de réponse lors de la réalisation de vos propres sondages.

6 La formulation de questions et de la méthodologie

6.1 Fonctions et types de questions

Le choix et la formulation de types de questions adaptés à notre méthodologie sont essentiels parce qu'ils influencent grandement la qualité des données recueillies. Comme nous l'avons vu au [sous-chapitre 5.2](#), les dimensions peuvent être opérationnalisées et transférées à différents indicateurs et donc utilisées dans un sondage. Dans les paragraphes suivants, nous apprendrons comment traduire ces indicateurs en questions adaptées. Tout d'abord, il convient de bien comprendre que les questions peuvent avoir des finalités diverses. Soit nous souhaitons en savoir davantage sur les faits et les connaissances des répondants, soit nous interrogeons ceux-ci à propos de leur comportement et de leurs avis ou d'autres aspects intéressants.

On peut distinguer les questions selon leur fonction dans un questionnaire : **Les questions de départ ou de prise de contact** servent à introduire le sondage et à créer une atmosphère favorable au sondage. Ces questions sont essentielles pour assurer la participation des répondants à un sondage. Si les premières questions sont ennuyantes ou incompréhensibles, de nombreux répondants abandonneront le sondage. Les questionnaires deviennent également ennuyeux pour les répondants si ceux-ci jugent les questions non pertinentes pour eux. Par exemple, dans un sondage sur les services liés aux études, les étudiants ne peuvent commenter les services dont ils ne font pas encore usage. C'est la raison pour laquelle nous créons des « **questions filtres** », qui orientent les répondants tout au long du sondage et ne les mènent qu'aux questions pertinentes pour eux. Parfois, les questions peuvent également revêtir une fonction de vérification. Ces **questions de vérification** servent à vérifier si les réponses fournies aux questions sont cohérentes. Nous posons ainsi deux questions sur un même point (si possible, avec un intervalle de temps). Les réponses sont cohérentes uniquement si elles sont les mêmes pour les deux questions. Ces questions de vérification doivent être utilisées avec précaution car elles peuvent avoir des effets négatifs. Par exemple, un répondant peut avoir l'impression de ne pas être pris au sérieux.

Fonctions
des questions

Dans les paragraphes suivants, nous découvrirons les différents types de question généralement utilisés dans les sondages (Neuman 2000, 261 ; Salant & Dillman 1994, 77 et seqq.).

Les questions ouvertes ne proposent pas de réponse fixe mais les répondants peuvent y répondre avec leurs propres mots. Les répondants ne sont pas contraints de répondre de la même manière, comme dans le cas des questions à choix multiple. Ces questions ouvertes laissent la possibilité au sondeur de découvrir des réponses qu'il n'avait pas encore envisagées. Elles permettent de découvrir de nouveaux domaines ou des domaines dans lesquels le sondeur a des connaissances limitées. Elles peuvent être utilisées afin de générer des réponses au format fixe. Les sondeurs ont besoin de beaucoup de temps pour administrer les questions ouvertes, et celles-ci doivent être codées. Elles nécessitent un effort supplémentaire de la part des répondants.

Types des
questions

Exemple (extrait d'un questionnaire sur l'évaluation d'un cours)

Quel est l'aspect du cours que vous avez préféré jusqu'ici ? (veuillez préciser) _____

Les questions fermées proposent au répondant différentes réponses formulées préalablement. Dans le cas des questions fermées, les réponses proposées doivent prendre en considération toutes les réponses possibles. On distingue les questions à choix unique, qui proposent une seule réponse, et les questions à choix multiple, qui proposent une variété de réponses possibles. Si plusieurs réponses sont possibles, la question doit le mentionner. On parle également de questions fermées lorsqu'on souhaite obtenir un avis sur un point spécifique en posant une série de questions (voir [sous-chapitre 5.3](#), Tableau 19).

Exemples (extraits d'un sondage sur l'évaluation d'un cours)

Choix unique

Dans le cadre de quel programme d'étude suivez-vous ce cours ?

- Bachelier en sciences Chimie
- Master en sciences Chimie
- Bachelier en sciences de l'eau
- Master en sciences de l'eau

Choix multiple

Pour quelles raisons suivez-vous ce cours ? (Plusieurs réponses possibles)

- horaire du cours
- intérêt spécifique pour la matière
- professeur
- pertinence du cours pour l'emploi futur
- cours axé sur la recherche
- cours obligatoire
- cours facultatif / optionnel
- préparation à l'examen
- pertinence du cours pour le travail de fin d'études
- autres raisons

Nous pouvons également poser des questions partiellement fermées. Ces questions sont un mélange de questions ouvertes et fermées. Outre les réponses formulées préalablement, les répondants ont la possibilité de répondre avec leurs propres mots. De manière générale, il est conseillé d'éviter les questions ouvertes lorsqu'il est possible de poser des questions fermées.

Exemples (extraits d'un questionnaire sur l'évaluation d'un cours)

Pour quelles raisons suivez-vous ce cours ? (Plusieurs réponses possibles)

- horaire du cours
- intérêt spécifique pour la matière
- professeur
- pertinence du cours pour l'emploi futur
- cours axé sur la recherche
- cours obligatoire
- cours facultatif / optionnel
- préparation à l'examen
- pertinence du cours pour le travail de fin d'études
- autres raisons (veuillez préciser) _____

Les réponses étant préalablement formulées, les questions (partiellement) fermées aboutissent essentiellement à des données quantitatives et sont utilisées dans les sondages structurés (voir [sous-chapitre 4.3](#)). En comparaison avec les questions ouvertes, les questions fermées offrent une meilleure comparabilité des réponses et une plus grande objectivité dans la mise en œuvre et l'analyse (voir [sous-chapitre 5.4](#)). Elles sont également plus faciles à compléter pour les répondants et plus faciles à analyser (voir [sous-chapitre 3.3](#)). Les questions fermées présentent néanmoins aussi des inconvénients : elles peuvent proposer des idées auxquelles les répondants n'auraient pas pensé. S'ils n'ont aucun avis ou connaissance sur un point déterminé, les répondants peuvent y répondre malgré tout, ou ils peuvent également être frustrés de ne pas retrouver leur choix dans les réponses proposées. L'ordre des réponses proposées peut par ailleurs influencer la réponse du répondant, en particulier dans le cas de questions à choix multiple.

Types de questions : avantages et inconvénients

Questions et tâches à accomplir

1. À quels obstacles doit-on s'attendre dans les sondages à questions ouvertes ?
2. Quelles sont les limites des questions fermées ?
3. Quels types de questions trouve-t-on généralement dans les entretiens structurés, et pourquoi ?

6.2 La formulation de bonnes questions

Si les questions d'un questionnaire sont mal formulées, cela peut nuire à la qualité des réponses fournies et, par conséquent, à l'utilité du sondage. Nous savons d'expérience que les répondants répondent généralement aux questions, mêmes si elles sont mal formulées. En tant que chercheurs, nous prenons parfois seulement conscience de cette mauvaise formulation à la lecture des réponses. Nous prenons conscience de ce problème, par exemple, lorsque de nombreux répondants n'ont pas répondu à une question en particulier, ou si les réponses aux questions ouvertes sont sans lien avec la question posée.

Il est facile de formuler de mauvaises questions.

Il est assez facile de formuler une mauvaise question, comme en témoigne l'exemple suivant, issu du contexte de l'enseignement supérieur. Imaginons que nous voulions obtenir des renseignements sur la mobilité des étudiants à l'étranger pendant leurs études. Nous formulons dès lors la question suivante : « Avez-vous passé une période de vos études à l'étranger ? » A priori, cette question semble acceptable. Mais est-ce vraiment la bonne question à poser pour obtenir des renseignements pour notre domaine de recherche, à savoir la mobilité des étudiants durant leurs études ? — La réponse est non, pour deux raisons au moins :

Tout d'abord, les étudiants peuvent ne pas savoir clairement à quoi il est fait référence : parle-t-on de vacances, de cours de langue, de séjour obligatoire dans le cadre du cursus, ou d'une visite chez un membre de la famille à l'étranger ? Chaque possibilité est envisageable, et le répondant peut interpréter la question comme il l'entend. Ensuite, les répondants ignorent la durée sous-entendue par « une période ». Un jour ? Plus de trois jours ? Plus d'une semaine ? Une fois encore, il revient à l'étudiant de décider de cette durée. Certains répondants pourraient dès lors ne pas répondre de la manière attendue. Par exemple, ils ont peut-être suivi un cours de langue de trois jours à l'étranger, mais ne le mentionnent pas dans le questionnaire parce qu'ils estiment la durée trop brève. Puisque la qualité des réponses dépend de la qualité des questions, il convient de ne pas perdre de vue une règle fondamentale :

Note : Les questions doivent toujours être comprises et abordées de la manière dont nous voulons qu'elles soient comprises et abordées.

Outre cette règle, les principales recommandations peuvent être résumées comme suit (Bryman 2004, 152 et seq.) :

Dix règles d'or pour la formulation de bonnes questions

« Dix règles d'or » pour la formulation de questions	Exemple et problème de formulation	Autre formulation possible
1. Utilisez une formulation simple, univoque, qui puisse être comprise de la même manière par tous les participants au sondage	« Si vous n'étiez pas intéressé par ET, esquiverez-vous ce cours ? » <i>Utilisation d'abréviations (ET) et d'expressions peu courantes (esquiver)</i>	« Si vous n'étiez pas intéressé par Ed Tech, éviteriez-vous ce cours ? »

« Dix règles d'or » pour la formulation de questions	Exemple et problème de formulation	Autre formulation possible
2. Évitez les questions hypothétiques	<p>« Imaginez que vous avez un fils de 16 ans et qu'il désire arrêter ses études pour devenir footballeur professionnel. Le soutiendriez-vous ? »</p> <p><i>Situation probablement éloignée de la réalité du répondant et difficile à imaginer.</i></p>	Posez des questions en lien direct avec ce que vous souhaitez savoir.
3. Évitez les questions longues et compliquées	<p>« De nombreuses personnes pensent qu'actuellement, les étudiants ont trop d'obligations en dehors de leurs études. Par exemple, ils doivent travailler pour avoir de quoi vivre, ou participent activement à plusieurs organes ou commissions d'étudiants, etc. au sein de leur université ou consacrent beaucoup de temps à leur passe-temps. Dans quelle mesure êtes-vous d'accord ? »</p> <p><i>Question trop longue et truffée d'informations superflues.</i></p>	« Aujourd'hui, les étudiants ont trop d'obligations (emploi, activités universitaires, passe-temps) en dehors de leurs études. Dans quelle mesure êtes-vous d'accord ? »
4. Évitez les questions doubles et les formulations négatives	<p>« Veuillez indiquer dans quelle mesure vous êtes d'accord avec la phrase suivante : Je ne me sens pas accepté(e) par mon patron et mes collègues. »</p> <p><i>formulation négative « je ne me sens pas accepté(e) » et double (patron et collègues).</i></p>	<p>« Veuillez indiquer dans quelle mesure vous êtes d'accord avec les phrases suivantes » :</p> <p>« Je me sens accepté(e) par mon patron. »</p> <p>« Je me sens accepté(e) par mes collègues. »</p>

« Dix règles d'or » pour la formulation de questions	Exemple et problème de formulation	Autre formulation possible
<p>5. Évitez les suppositions et questions suggestives</p>	<p>« Maintenant que vous connaissez les avantages des cours complémentaires pour améliorer des compétences clés (« compétences non techniques »), choisirez-vous encore un de ces cours au prochain semestre ? »</p> <p><i>Supposition / suggestif (vous connaissez les avantages)</i></p>	<p>« Si vous avez suivi un cours complémentaire au semestre passé pour améliorer des compétences clés (« compétences non techniques »), cela a-t-il porté ses fruits ? »</p> <p>N'oubliez pas que certains répondants n'ont peut-être suivi aucun cours. Prévoyez une réponse de type « pas d'application ».</p>
<p>6. Évitez les questions superflues</p>	<p>« Veuillez indiquer dans quelle mesure vous êtes d'accord avec la phrase suivante : Les indépendants travaillent dur. »</p> <p><i>Question intéressante, mais quelles informations en tirez-vous ?</i></p>	<p>Vérifiez toujours si la question est bien en lien avec votre projet de sondage ?</p>
<p>7. Utilisez des références (temporelles) claires</p>	<p>« Comment évaluez-vous les conseils fournis au sein de votre université par le passé ? »</p> <p><i>Enseignants, les services d'orientation, etc. ?) ni de pas de référence claire (conseils fournis par les référence temporelle claire (ces 3 dernières semaines, ces 3 dernières années ?)</i></p>	<p>Comment évaluez-vous les conseils fournis par le centre d'orientation des étudiants concernant vos options de cours au semestre dernier (année académique 14/15) ?</p>

« Dix règles d'or » pour la formulation de questions	Exemple et problème de formulation	Autre formulation possible
8. Utilisez des catégories de réponses exhaustives et disjonctives	<p><i>Quel âge aviez-vous à l'obtention de votre diplôme ?</i></p> <p><input type="radio"/> 22 – 24 ans</p> <p><input type="radio"/> 24 – 26 ans</p> <p><input type="radio"/> 26 – 28 ans</p> <p><i>Non disjonctif (où situer les personnes de 24 ou 26 ans ?) et incomplet (où situer les personnes de plus de 28 ans ?)</i></p>	<p>Quel âge aviez-vous à l'obtention de votre diplôme ?</p> <p><input type="radio"/> 22 – 24 ans</p> <p><input type="radio"/> 25 – 27 ans</p> <p><input type="radio"/> 27 ans ou plus</p>
9. Assurez-vous que le contexte d'une question n'influence pas le type de réponse	<p>Un sondage mené auprès de scientifiques de renom démontre que 90% d'entre eux n'aprouvent pas le « processus de Bologne ». Quel est votre avis ?</p> <p><i>Influence du contexte (un sondage mené auprès de scientifiques de renom démontre que...)</i></p>	<p>Comment évaluez-vous en règle générale le « processus de Bologne » ?</p>
10. Expliquez les expressions peu claires	<p>Quel est le revenu mensuel moyen net de votre ménage ?</p> <p><i>Tout le monde ne sait pas ce qu'est le revenu net moyen du ménage.</i></p>	<p>Quel est le revenu mensuel net moyen de votre ménage (à savoir la somme des revenus de tous les membres du ménage après imposition) ?</p>

Tableau 21 « Dix règles d'or » pour la formulation de questions (Bryman 2004, 152 et seqq.)

Ces règles ne sont toutefois pas absolues et se contredisent parfois. Voyons-y dès lors des recommandations, qu'il n'est parfois pas possible d'appliquer simultanément.

Pas de règle sans exception

Questions et tâches à accomplir

1. Citez un mauvais exemple pour chacune des dix règles d'or. Reformulez ensuite votre question en suivant la règle.

6.3 Conception du questionnaire

Le questionnaire est un outil essentiel de la collecte de données dans la recherche empirique en sciences sociales. Sa conception va bien au-delà de la simple formulation de bonnes questions (voir [sous-chapitre 6.2](#)). Selon qu'il s'agit d'un sondage par courrier électronique, en ligne ou téléphonique, nous devons également prendre en considération certains éléments, comme la conception des pages, le format et l'impression, la conception graphique et les représentations (par ex. pour les appareils mobiles) (Salant & Dillman 1994, 101 et seqq.). La conception d'un questionnaire doit, avant tout, respecter la règle d'or suivante :

Règle d'or
pour la
conception d'un
questionnaire

« Ne faites pas à vos répondants ce que vous ne voudriez pas qu'ils vous fassent! » (Trochim 2005, 86)

Enfin, le sondage est une charge supplémentaire dans la vie de nos répondants. Nous leur demandons de nous consacrer leur temps, leur l'attention, leur confiance et de nous divulguer des informations personnelles. N'oublions donc pas ceci :

La **règle d'or de la conception d'un questionnaire** dans la pratique (Trochim 2005, 86) :

- Remerciez d'emblée le répondant de vous permettre de mener votre étude et, à la fin, remerciez-le pour sa participation.
- Votre sondage doit être aussi bref que possible! N'y incluez que le strict nécessaire.
- Soyez attentif aux besoins de votre répondant et prêtez attention à tout signe de gêne de sa part.

Séquençage
des questions

Après avoir formulé les questions et réfléchi à la structure du questionnaire, nous savons que : les sondages suivent un schéma basé sur le séquençage des questions individuelles au sein du questionnaire (voir [sous-chapitre 6.1](#)). Si nous optons pour un schéma intelligent, cela peut avoir un effet positif sur le sondage. En revanche, un sondage perd de son efficacité si l'ordre des questions n'est pas adapté. Par exemple, les réponses aux questions peuvent être influencées par les questions précédentes. Certaines questions peuvent être posées trop tard ou trop tôt pour susciter l'intérêt. Une question peut également ne pas susciter l'attention requise en raison des questions qui l'entourent. Au vu de ces différents points, nous devons tenir compte des recommandations suivantes, indépendamment de la méthodologie choisie (Trochim 2005, 86) :

- Commencez par des questions faciles, non intimidantes, et placez les questions plus délicates, intimidantes et sociographiques en fin de sondage.
- N'entamez pas un sondage en ligne ou par courrier électronique par une question ouverte.
- Respectez l'ordre chronologique.
- Intéressez-vous à un thème à la fois et insérez une transition entre chaque thème.
- Si vous souhaitez interroger les répondants à propos de sujets délicats ou gênants, faites précéder ces questions de questions de mise en train, plus simples.

Effectuez un test
préalable de votre
questionnaire

Il est impératif de tester la fonctionnalité de votre questionnaire du point de vue de son contenu et de la technique utilisée et d'optimiser le questionnaire. À cet effet, vous devriez toujours réaliser un test préalable après avoir achevé le prototype du questionnaire, et ce même si vous copiez la plupart des questions d'un autre questionnaire. Ces tests préalables permettent de déterminer le temps de réponse nécessaire à un son-

dage et de déceler les problèmes techniques (comme les erreurs de filtrage) ou les questions trop vagues. Le plus simple est de demander à votre entourage de compléter le sondage et de les interroger ensuite sur cette expérience, sur les difficultés ou problèmes de compréhension éventuels rencontrés au cours du sondage (Babbie 2004, 256).

Les sondeurs (voir [sous-chapitre 4.3](#)) qui font usage d'un questionnaire lors d'un entretien jouent un rôle bien spécifique dans le processus de sondage : ils doivent motiver les répondants à prendre part au sondage (jusqu'à la fin) tout en restant neutres quant au contenu du sondage. Cette neutralité est essentielle pour garantir le niveau de normalisation requis. D'un point de vue administratif, les sondeurs doivent organiser les entretiens et tenir compte des exigences propres à l'échantillon. Il est dès lors judicieux de former les sondeurs avant le lancement de l'étude. Seuls les sondeurs bien formés sont à même d'effectuer un sondage correctement d'un point de vue technique et méthodologique (Babbie 2004, 263ff). L'administration du sondage doit par ailleurs être très précise et définir des exigences réalistes, telles que le nombre de sondeurs et le temps nécessaire à chaque entretien.

Formez vos sondeurs

Questions et tâches à accomplir

1. Expliquez la règle d'or pour la conception d'un questionnaire.
2. Décrivez un séquençage réussi des questions.
3. Pourquoi un questionnaire doit-il être testé avant son utilisation ?

6.4 Mesures pour l'obtention d'un taux élevé de réponse

Un faible taux de réponse est un autre grand problème des sondages. Le taux de réponse (également appelé taux de remplissage ou taux de retour) est le pourcentage de personnes qui prennent part à un sondage, divisé par le nombre de personnes choisies dans l'échantillon (Babbie 2004, 261). Il n'existe aucune valeur minimale raisonnable pour le taux de réponse car cette valeur dépend largement de la méthode utilisée. On peut toutefois affirmer ceci : plus le taux de réponse est faible, plus la probabilité est grande que les résultats soient basés sur l'auto-sélection des répondants ou sur l'absence de réponse à une question (voir [sous-chapitre 4.5](#)).

Par exemple, un établissement d'enseignement supérieur mène un sondage auprès de son personnel à propos du « leadership ». Seul(e)s quelques secrétaires prennent part au sondage par crainte de sanctions. Ces résultats peuvent fausser tout le sondage. Si le taux de réponse n'atteint pas 100%, l'échantillon risque toujours d'être biaisé (voir [sous-chapitre 4.4](#)), même si le taux de réponse est supérieur à 50%. Des taux de réponse plus élevés réduisent la probabilité de sources d'erreurs potentielles et renforcent les possibilités d'analyse statistique des données. Aussi importe-t-il de toujours viser le taux de réponse le plus élevé possible.

Les taux de réponse élevés sont privilégiés

Pour ce faire, il convient avant tout de tenir compte des recommandations relatives à la formulation des questions et à la conception du questionnaire (voir [sous-chapitres 6.2](#) et [6.3](#)). Selon la méthode conceptuelle

Méthode
conceptuelle
totale

totale (Dillman 1983), qui obtient des taux de réponse largement supérieurs à 50%, et la méthodologie y afférente, les éléments suivants doivent être pris en considération :

■ Conception du questionnaire

- Sondages papier : papier blanc et format brochure
- Sondages en ligne : un questionnaire doit pouvoir être illustré et complété à partir de n'importe quel navigateur, y compris à partir d'un navigateur pour appareil mobile, comme un smartphone ou une tablette.

■ Lettre d'invitation

- En règle générale, veillez à préciser les avantages du sondage et à souligner l'importance des répondants et de leur participation. Expliquez pourquoi les répondants ont été choisis et veillez à protéger les données fournies. Fixez un délai concret pour la participation et encouragez les questions éventuelles.
- Sondages papier : Limitez-vous au matériel officiel (par ex. du papier avec l'en-tête de l'établissement) ; le cas échéant : enveloppe correctement timbrée et mentionnant l'adresse du destinataire; signature personnelle (numérique).
- Sondages en ligne par courrier électronique : L'invitation doit être lisible à partir de n'importe quel gestionnaire de courrier électronique. Utilisez une adresse d'envoi fiable. Le cas échéant, ajoutez en pièce jointe l'invitation version papier.

■ Diffusion

- En règle générale : En plus de l'invitation, vous pouvez envoyer trois autres courriers de rappel. Le calendrier d'envoi de ces courriers dépend du calendrier du sondage.
- Sondages papier : Veillez à envoyer l'invitation afin qu'elle arrive à destination en milieu de semaine. Après une semaine, vous pouvez envoyer une carte afin de remercier les répondants pour leur participation / leur rappeler aimablement d'y prendre part. Après trois semaines, vous pouvez renvoyer un bref courrier contenant à nouveau le sondage.
- Sondages en ligne : des courriers électroniques de rappel peuvent être envoyés de la même manière. Si vous disposez également des adresses postales, vous pouvez envoyer une lettre d'information à propos du sondage en ligne. Si vous disposez des numéros de téléphone, vous pouvez relancer les répondants par téléphone plutôt que par courrier papier ou électronique.

■ Mesures incitatives

- En règle générale : Vous pouvez proposer une récompense financière ou autre pour inciter à participer à un sondage. Cette démarche doit toutefois faire l'objet d'une réflexion critique car elle peut entraîner des distorsions. C'est la raison pour laquelle elle ne doit pas être la motivation première pour la participation à un sondage. Elle doit également être la même pour tous les participants.

Questions et tâches à accomplir

1. Pourquoi un faible taux de participation peut-il être problématique et pourquoi des taux plus élevés le sont-ils moins ?
2. Expliquez comment vous concevriez un sondage électronique pour garantir un taux de participation aussi élevé que possible.



Lectures complémentaires

- Babbie, E. R. (2004). *The practice of social research* (10ème éd). Southbank : Wadsworth, 242-279.
- Bryman, A. (2004). *Social research methods* (2ème éd). Oxford : University Press, 144-162.
- Couper, M. (2008). *Designing effective web surveys*. New York : Cambridge University Press.
- Dillman, D. A. (1983). Mail and other self-administered questionnaires. In P. H. Rossi (Ed.), *Quantitative studies in social relations. Handbook of survey research* (pp. 359-378). Orlando : Academic Press.
- Groves, R. M., Fowler, F. J., Couper, M., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2009). *Survey methodology* (2ème édition). Hoboken, NJ : Wiley, 217-258.
- Salant, P & Dillman, D. A. (1994). *How to conduct your own survey*. New York : Wiley, 77-136.

Chapitre 7

Traitement et analyse des données : Des questionnaires aux résultats d'un sondage

7	Traitement et analyse des données : Des questionnaires aux résultats d'un sondage	103
7.1	Nettoyage et codification des données	103
7.2	Le livre de codes et le rapport d'enquête	104
7.3	Compte rendu des résultats d'un sondage	104



Au terme de ce chapitre, vous devriez pouvoir...

- disposer de connaissances fondamentales pour pouvoir procéder au nettoyage et au codage des données, dans le respect des règles générales en la matière ;
- expliquer les fonctions et citer les contenus des rapports d'enquête ;
- expliquer et interpréter les mesures de tendance centrale présentées ;
- disposer de connaissances fondamentales sur la préparation d'un premier rapport sur les résultats d'un sondage.

7 Traitement et analyse des données : Des questionnaires aux résultats d'un sondage

7.1 Nettoyage et codification des données

Une fois notre sondage terminé (ainsi que le processus de collecte de données), les données n'existent pas encore sous une forme que l'on puisse analyser sur-le-champ. Dans le cas d'un sondage en ligne ou téléphonique, l'enregistrement des données permet généralement leur extraction instantanée. Même ces ensembles de données devront toutefois encore être analysés et vérifiés afin de déceler d'éventuelles erreurs.

Si un ensemble de données n'existe pas encore au terme d'un sondage par courrier électronique, il convient d'en créer un. On utilise généralement un logiciel spécifique, qui permet la gestion et l'analyse des données quantitatives et, par exemple, au terme d'un sondage par courrier électronique, les données recueillies doivent être transférées dans un ensemble de données. Si cela ne peut se faire au moyen d'un logiciel utilisé pour la collecte de données (par ex. dans le cadre d'un sondage en ligne), il faut procéder manuellement. Toutes les questions d'un questionnaire doivent alors être codées, de même que toutes les réponses possibles aux questions fermées (en particulier) afin qu'une analyse des données puissent être réalisée ensuite.

Création d'un
ensemble
de données

Le nettoyage ultérieur des données est une étape fondamentale de l'analyse des données. Toute collecte de données, qu'elle soit automatique ou manuelle, peut comprendre des erreurs. Les erreurs les plus courantes sont **les codes sauvages** et **les valeurs aberrantes**. Les codes sauvages sont des valeurs qui sortent du cadre défini. Imaginons le cas d'une variable avec des valeurs situées entre 1 et 5. 1, 2, 3, 4, et 5 sont donc **des valeurs valides**. Si la matrice de données prévoit également la valeur 7, il s'agit là d'un code sauvage (peut-être le fruit d'une simple erreur de frappe). Les valeurs aberrantes sont des valeurs non plausibles. Par exemple, une variable sur les semestres inclut la valeur 85. Cette valeur n'est pas plausible et doit être éliminée des données recueillies. Les codes sauvages, valeurs aberrantes ou non-réponses (voir [sous-chapitre 4.5](#)) sont des valeurs non valides associées à un code spécifique (par ex. 8 = pas d'application et 9 = pas de réponse). On les appelle aussi « **valeurs manquantes** ». ⁴

Valeurs valides
et manquantes

Il se peut parfois que les données soient recueillies d'une manière plus différenciée que nécessaire. Comme nous l'avons déjà vu, les différentes variables peuvent être résumées en une nouvelle variable. Ce faisant, nous agrégeons les données (voir [sous-chapitre 4.3](#)). On peut ainsi résumer les valeurs d'une variable métrique (voir [sous-chapitre 5.3](#)) : Pour pouvoir illustrer graphiquement les réponses relatives aux revenus, il est judicieux de catégoriser ces réponses dans une nouvelle variable et de les recoder. Dans ce cas, la variable « revenus » est une série de valeurs numériques individuelles et peut être transférée dans une nouvelle variable recodée qui résume les valeurs numériques de portées différentes. La nouvelle variable peut avoir les valeurs

Contrôle qualité —
vérifiez vos
données

⁴ Pour de plus amples détails sur le traitement des données de sondage au terme du processus de collecte, veuillez consulter Groves et al. 2009, 329 et seqq.

suivantes : 1 (= pas de revenu), 2 (= revenus entre 0\$-250\$), 3 (= 751\$-1000\$) ; 6 (= 1001\$ et plus) ; 9 (= pas de réponse). Les variables métriques doivent toujours être catégorisées car ces catégories permettent d'autres possibilités d'analyse que les variables qu'elles contiennent (voir [sous- chapitre 7.3](#)).

Nous avons déjà vu un exemple de codification de données qualitatives au [sous-chapitre 3.3](#). Le traitement de données dans les études non structurées qui ne recueillent que des données qualitatives constitue un cas particulier. Nous n'aborderons pas ce cas spécifique dans le présent module en raison de sa complexité.⁵

7.2 Le livre de codes et le rapport d'enquête

Les listes de codes doivent être recensées dans un livre de codes, ou elles doivent être référencées. Cette étape est importante au cas où la saisie des données serait faite par une personne extérieure au projet. Il n'existe aucune exigence particulière, si ce n'est que toute personne qui lit ce livre de codes doit pouvoir comprendre comment traiter en tant que données les informations fournies dans le questionnaire.

Permettez
aux autres
de comprendre
votre démarche

Un **rapport d'enquête** comprend les motivations d'un sondage, l'élaboration (et peut-être le contexte théorique) des questions et du questionnaire, la conception méthodologique, la population ou l'échantillon, ainsi que le processus de collecte de données opérationnelles et administratives (par ex. les dates d'envoi des invitations et rappels). Grâce à ces renseignements, les personnes qui commandent un sondage ou prennent part à un processus d'évaluation peuvent avoir un aperçu du déroulement et de la qualité de l'étude. Un rapport d'enquête méthodologique doit dès lors inclure les informations suivantes :

- les parties prenantes qui commandent une étude, leurs motivations et les finalités de l'étude ;
- le plan et le test préalable des questions et du questionnaire ;
- le groupe cible et la conception de l'ensemble de données (le cas échéant) ;
- le cas échéant : les mesures de formation des sondeurs ;
- le taux de réponse, y compris les dates précises pour les invitations, rappels, etc. ;
- le calendrier approximatif des entretiens ;
- le cas échéant : les spécificités du sondage, par ex. les problèmes rencontrés et les limites.

Le rapport d'enquête fait partie intégrante du processus de compte rendu des résultats du sondage (voir [sous-chapitre 7.3](#)).

7.3 Compte rendu des résultats d'un sondage

Après la collecte et le traitement des données vient une étape plus intéressante de l'analyse des données, à savoir le compte rendu des résultats. À cet effet, les données sont généralement agrégées dans des tableaux, graphiques ou ce qu'il est convenu d'appeler des « **valeurs mesurées** ». Grâce à cela, les résultats peuvent

⁵ Pour de plus amples détails sur l'analyse qualitative de données (AQD) et le matériel d'apprentissage sur l'utilisation de logiciels d'analyse qualitative de données assistée par ordinateur ou CAQDAS (Computer Assisted Qualitative Data Analysis), veuillez consulter par ex. le site web <http://onlineqda.hud.ac.uk/index.php>.

être illustrés de manière claire et facile à lire pour les destinataires. C'est la raison pour laquelle nous privilégions les méthodes de « statistique descriptive ». Dans les paragraphes suivants, nous vous présenterons un large aperçu de ces statistiques et valeurs mesurées, courantes dans les rapports d'évaluation.

Note : Kumar (2005), 248 et seqq. et Salant & Dillman (1994), 206 et seqq. proposent des exemples utiles pour l'illustration des données en tableaux et graphiques.

Il est possible de créer **une distribution statistique** à partir des variables codées ou quantitatives. Nous obtenons alors des « distributions absolues » et les taux respectifs. Il est recommandé de créer des distributions statistiques pour tout type de variable car cela permet également de déceler les erreurs comme les codes sauvages ou les valeurs aberrantes lors du traitement des données (voir [sous-chapitre 7.1](#)).

Les indicateurs de tendance centrale servent à agréger les informations recueillies en un seul chiffre. La plupart des logiciels d'analytique peuvent aisément calculer ces indicateurs, qui fournissent des renseignements sur la valeur la plus fréquente (mode), sur la valeur située au milieu d'une distribution (médiane) et sur la valeur moyenne (moyenne).

La corrélation signale-t-elle un lien de cause à effet ?

Le **mode** peut déjà servir dans le cas des variables d'échelle nominale. Il décrit la valeur mesurée la plus courante. Pour calculer une **médiane**, nous avons au minimum besoin de variables ordinales. La médiane décrit la valeur située au milieu, si toutes les valeurs sont classées de la valeur la plus faible à la valeur la plus élevée. 50% des valeurs sont donc situées en dessous de cette médiane, et 50% au-dessus. La moyenne n'est calculée que dans le cas des variables d'échelle d'intervalle minimal. Elle est également souvent utilisée pour les échelles ordinales (voir [sous-chapitre 5.3](#)). La moyenne décrit la valeur moyenne et elle se calcule en divisant la somme de toutes les valeurs mesurées par le nombre de réponses. Cela présente un inconvénient pour les valeurs extrêmes, qui divergent fort des autres valeurs mesurées. L'**écart type** est une mesure de la dispersion des valeurs autour de leur moyenne. En d'autres termes, il mesure leur niveau de cohésion ou de dispersion. L'écart-type est la racine carrée de la variance. Celle-ci s'obtient en divisant la somme des écarts au carré par rapport à la moyenne par $n-1$ (n'étant le nombre d'unités constituant l'échantillon).

Bien souvent, on veut décrire plus que de simples valeurs ou distributions moyennes. Nous nous intéressons également aux corrélations statistiques entre les variables. Si tel est le cas, on dit que les variables **sont en corrélation** les unes avec les autres. Par exemple, une **corrélation positive** signifie que « plus la variable A est présente, plus la variable B l'est aussi. » S'agissant des variables « diplôme obtenu » et « revenus », une corrélation positive serait que plus le diplôme obtenu est élevé, plus les revenus sont élevés. Les corrélations sont des indicateurs, mais non une preuve de **causalité**, à savoir des corrélations avérées de lien de cause à effet. Pour illustrer la distribution statistique courante de deux variables dans le but de déceler des corrélations, nous pouvons créer des **tableaux croisés**. Par exemple, si nous analysons l'hypothèse « plus élevé est le diplôme obtenu, plus élevés sont les revenus », nous vérifions s'il existe une corrélation entre les variables « diplôme obtenu » et « revenus ». Dans ce cas, le diplôme obtenu est la « **variable indépendante** », parce que nous estimons qu'elle influence les revenus. Pour l'élaboration de tableaux croisés, nous devons tenir compte de la convention formelle en vertu de laquelle les variables indépendantes sont inscrites dans les colonnes,

et les variables dépendantes, horizontalement. Les résultats d'une analyse par tableaux croisés peuvent se résumer des trois manières suivantes :

- 1) nombres absolus
- 2) pourcentages des colonnes (la somme des pourcentages de chaque colonne est égale à 100)
- 3) pourcentages des lignes (la somme des pourcentages de chaque ligne est égale à 100)

Les programmes informatiques conçus pour l'analyse statistique calculent également une statistique appelée χ^2 . Cette valeur précise si deux variables sont en corrélation. Nous vérifions par la même occasion si la corrélation est **significative**. Notons toutefois que même si elle est significative, cette corrélation n'indique pas de lien de cause à effet. Pour ce faire, une analyse plus détaillée encore s'impose.

Après avoir clôturé le traitement des données et rédigé un rapport, nous pouvons utiliser une structure du type suivant (Salant & Dillman 1994, 203) :

- Résumé ou sommaire (y compris les principales conclusions)
- Énoncé du Problème (explique les motivations du sondage)
- Rapport d'enquête (voir [sous-chapitre 7.2](#))
- Conclusions (présentent les résultats pertinents par ordre logique)
- Implications (rassemblent les conclusions afin de répondre aux questions de départ et explorent les implications pour le processus décisionnel)
- Annexes (fournissent le matériel complémentaire, comme les questionnaires, etc.)

Pour décrire nos conclusions, nous pouvons nous reporter aux mesures calculées. Nous pouvons par ailleurs mettre au point des graphiques basés sur les distributions statistiques, plus faciles à lire que les tableaux.

Fonctions
du rapport

Notre rapport revêt dès lors trois fonctions : il décrit et documente notre traitement en matière de collecte et d'analyse des données ainsi que les résultats. Ces données peuvent étayer différents processus d'évaluation car la méthode scientifique appliquée garantit leur haut degré de qualité et les rend claires et compréhensibles pour tous.

Questions et tâches à accomplir

1. Comment procéderiez-vous pour analyser la qualité de votre matrice de données ?
2. Quelle est la différence entre des valeurs valides et non valides ? Veuillez citer un exemple.
3. Que sont les variables catégorisées ? Veuillez citer un exemple.
4. Citez les fonctions et le contenu des rapports d'enquête.
5. Quels indicateurs de tendance centrale sont souvent mentionnés dans les rapports de résultats ? Expliquez leur valeur explicative.
6. Décrivez les avantages des tableaux croisés.
7. Citez et expliquez les fonctions et la structure des rapports de résultats.



Lectures complémentaires sur l'analyse des données qualitatives

- Babbie, E. R. (2004). *The practice of social research* (10ème éd). Southbank : Wadsworth, 369-393.
- Bryman, A. (2004). *Social research methods* (2ème éd). Oxford : University Press, 398-416.
- Neuman, W. L. (2000). *Social research methods: Qualitative and quantitative approaches* (4ème éd). Boston : Allyn and Bacon, 417-442.



Lectures complémentaires sur l'analyse des données quantitatives

- Babbie, E. R. (2004). *The practice of social research* (10ème éd). Southbank : Wadsworth, 395-420.
- Bryman, A. (2004). *Social research methods* (2ème éd). Oxford : University Press, 218-242.
- Neuman, W. L. (2000). *Social research methods: Qualitative and quantitative approaches* (4ème édition). Boston : Allyn and Bacon, 313-343.
- Salant, P & Dillman, D. A. (1994). *How to conduct your own survey*. New York : Wiley, 175-200.
- Trochim, W. M. K. (2005). *Research methods: The concise knowledge base*. Mason, Ohio : Cengage Learning, 204-223.

Bibliographie

- Atteslander, P. (2010). *Methoden der empirischen Sozialforschung* (13ème édition). Berlin : Erich Schmidt Verlag.
- Babbie, E. R. (2004). *The practice of social research* (10ème édition). Southbank : Wadsworth.
- Bryman, A. (2004). *Social research methods* (2ème édition). Oxford : University Press.
- Couper, M. (2008). *Designing effective web surveys*. New York : Cambridge University Press.
- DeGEval – Gesellschaft für Evaluation (2001). *Evaluation standards (DeGEval-Standards)*. Récupérée de http://www.degeval.de/fileadmin/user_upload/Sonstiges/STANDARDS_2008-12_kurz_engl.pdf
- Dillman, D. A. (1983). Mail and other self-administered questionnaires, P. H. Rossi (Éd.), *Quantitative studies in social relations. Handbook of survey research* (359-378). Orlando : Academic Press.
- ENQA – European Association for Quality Assurance in Higher Education (2009). *Standards and guidelines for quality assurance in the European higher education area*. Récupérée de <http://www.enqa.eu/index.php/home/esg/>
- Groves, R. M., Fowler, F. J., Couper, M., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2009). *Survey methodology* (2ème édition). Hoboken, NJ : Wiley.
- Hussein, A. (2009). The use of triangulation in social sciences research : Can qualitative and quantitative methods be combined ? *Journal of Comparative Social Work* 2009/1.
- Kumar, R. (2005). *Research methodology : A step-by-step guide for beginners* (2ème édition). London : SAGE.
- McMillan, J. H. & Schumacher, S. (2010). *Research in education : Evidence-based inquiry* (7ème édition). Boston : Pearson.
- Neuman, W. L. (2000). *Social research methods : Qualitative and quantitative approaches* (4ème édition). Boston : Allyn and Bacon.
- Nickel, S. (2010). *Qualität in Hochschulen – was ist das ?* Präsentation im Rahmen des Moduls Qualitäts- und Personalmanagement im Masterstudiengang Hochschul- und Wissenschaftsmanagement. 4 June 2010.
- Pan, D. (5ème éd). (2008). *Learning to teach, teaching to learn : A handbook for NUS teachers*. Centre for Development of Teaching and Learning, NUS : Singapore.
- Rindermann, H. (1998). The Munich multifactorial model of course quality : Design, justification and verification. *Beiträge zur Hochschulforschung* 3, 189-224.
- Röbbken, H. (2012) : *Qualitätsmanagement in Bildungsorganisationen*. Oldenburg : Carl von Ossietzky Universität Oldenburg.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation : A systematic approach* (7ème édition). Thousand Oaks : SAGE.
- Salant, P & Dillman, D. A. (1994). *How to conduct your own survey*. New York : Wiley.
- Secolsky, C. & Denison, B. (Éd, 2012). *Handbook on measurement, assessment, and evaluation in higher education*. New York : Routledge.
- Stockmann, R. (2002). *Qualitätsmanagement und Evaluation – Konkurrierende oder sich ergänzende Konzepte?* CEval-Arbeitspapiere 3. Saarbrücken : Centrum für Evaluation.

- Stockmann, R. (2004). *Was ist eine gute Evaluation ?* CEval-Arbeitspapiere 9. Saarbrücken : Centrum für Evaluation.
- Stockmann, R. (2006). *Evaluation und Qualitätsentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement.* München : Waxmann.
- Stockmann, R. & Meyer, W. (2010). *Evaluation : Eine Einführung.* Opladen : Budrich.
- Trochim, W. M. K. (2005). *Research methods : The concise knowledge base.* Mason, Ohio : Cengage Learning.
- Yarborough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3ème édition). Thousand Oaks : SAGE.

Liste des tableaux

Tableau 1	Concept et perspective d'analyse des évaluations (adapté de Stockmann 2002, 6)	19
Tableau 2	Les normes d'utilité du JCSEE (adapté de JCSEE)	20
Tableau 3	Les normes de faisabilité du JCSEE (adapté de JCSEE)	20
Tableau 4	Les normes de déontologie du JCSEE (adapté de JCSEE)	21
Tableau 5	Les normes de précision du JCSEE (adapté de JCSEE)	22
Tableau 6	Les normes de responsabilité de l'évaluation du JCSEE (adapté de JCSEE) ..	22
Tableau 7	Atouts et inconvénients potentiels des évaluations internes et externes (tableau des auteurs)	28
Tableau 8	Trois étapes pour dégager un consensus et veiller à la transparence dans les évaluations de l'UDE (tableau des auteurs)	36
Tableau 9	Indice de qualité de l'enseignement de KIT (http://www.pst.kit.edu/230.php)	42
Tableau 10	Différences entre les données quantitatives et qualitatives (tableau des auteurs)	56
Tableau 11	Différences entre les stratégies de recherche quantitatives et qualitatives (tableau des auteurs, basé sur Kumar 2005, 17 et seqq.)	58
Tableau 12	Méthodes de recherche basées sur les résultats (tableau des auteurs)	60
Tableau 13	Différence entre les études des tendances, des cohortes et par panel (tableau des auteurs)	68
Tableau 14	Caractéristiques des sondages (tableau des auteurs)	73
Tableau 15	Sources d'erreurs dans les sondages (tableau des auteurs)	76
Tableau 16	Questions pour le choix de la méthode de sondage appropriée (tableau des auteurs)	79
Tableau 17	Opérationnalisation de la qualité du cours, (adapté de Rindermann 1998) .	84
Tableau 18	Catégorisation des variables (tableau des auteurs, basé sur Kumar 2005, 6)	85
Tableau 19	Recours à une série de questions pour l'analyse d'un seul thème (tableau des auteurs)	87
Tableau 20	Types d'erreur en recherche évaluative (tableau des auteurs, basé sur Neumann 2000, 105)	89
Tableau 21	« Dix règles d'or » pour la formulation de questions (Bryman 2004, 152 et seqq.)	97

Liste des figures

Figure 1	L'évaluation institutionnelle au sein de l'Université de Duisburg-Essen (illustration des auteurs)	35
Figure 2	Adaptation du modèle multifactoriel de Munich pour la qualité des cours (adapté de Rindermann 2001)	45
Figure 3	Questions à poser pour le choix d'une stratégie de recherche approprié (illustration des auteurs)	59
Figure 4	Les cinq phases idéales d'une étude empirique (tableau des auteurs)	65
Figure 5	Méthodologie « avant-après » (Kumar 2005, 95)	68
Figure 6	Population (étudiée) et échantillon (Kumar 2005, 164)	74
Figure 7	Étapes de l'opérationnalisation à la mesure (illustration des auteurs)	82



Avec le soutien financier du



Ministère fédéral de la
Coopération économique
et du Développement

Soutenu par

