



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

UNIVERSITE NANCY 2
Ecole doctorale "Langage, Temps, Société"
U.F.R. Connaissance de l'homme
Laboratoire de Psychologie

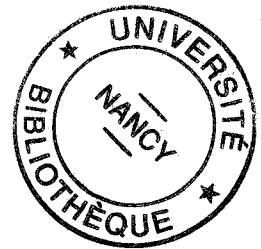
**LE BIEN-ETRE SUBJECTIF
VERS UNE MESURE DU BONHEUR**

Volume 1

Thèse présentée pour l'obtention du doctorat de psychologie et soutenue
publiquement le 22 Novembre 1994 à l'Université Nancy 2 par

Jean-Luc KOP

sous la direction du Professeur Paul DICKES



Membres du jury

Françoise BACHER

Professeur de psychologie, directrice honoraire de l'Ecole Pratique des Hautes Etudes

Paul DICKES

Professeur de psychologie à l'Université Nancy 2, directeur de la recherche

Winfried HUBER

Professeur émérite de psychologie à l'Université Catholique de Louvain, président du jury

Michel HUTEAU

Professeur de psychologie à l'Université Paris VIII, directeur de l'INETOP, rapporteur

Jacques JUHEL

Maître de Conférences de psychologie à l'Université Rennes 2, habilité à diriger des recherches, rapporteur

Remerciements...

au Professeur Paul Dickes, pour la confiance qu'il m'a toujours accordée, en espérant que cette thèse reflète un peu de l'esprit qui l'anime

à mes parents et ma famille, pour leur soutien sans faille

à mes amis, parce qu'ils étaient là

aux étudiants de maîtrise de psychologie différentielle de l'université Nancy 2 qui ont participé à certains recueils de données, toujours avec enthousiasme

au Ministère Luxembourgeois des Affaires Culturelles et de la Recherche Scientifique qui m'a soutenu financièrement pendant trois ans et au CEPS/INSTEAD (Wallerdange, Grand-Duché de Luxembourg) qui m'a accueilli durant cette période

à l'ADEPS (URA CNRS 1167) et à l'ensemble de ses membres qui m'ont offert des conditions et une ambiance de travail qu'envieraient beaucoup de chercheurs

quant à Jocelyne Journois, elle seule sait ce que lui doit ce travail

Nous devons mesurer ce qui est mesurable et rendre mesurable ce qui ne l'est pas.

(Galilée, 1564-1642)

If one is going to assess something, it might seem reasonable to expect that one would know what it is that one is assessing. Not so with personality, or with traits.

(Rorer & Widiger, 1983, p. 444)

RESUME

Cette thèse porte sur la mesure du bien-être subjectif et articule une réflexion théorique sur la mesure, une analyse critique et synthétique de la littérature et une approche empirique centrée sur les aspects définitoires et psychométriques. La réflexion théorique sur la mesure est l'occasion d'exposer une conception dans laquelle le processus de mesure s'intègre au sein du processus plus général de la démarche scientifique. La revue critique de la littérature permet de clarifier certaines confusions conceptuelles et présente un panorama organisé de l'ensemble des recherches sur le bien-être. Les différentes études empiriques questionnent la manière dont le bien-être est habituellement mesuré et utilisent les représentations afin d'enrichir la signification de ce concept. La conclusion offre un bilan de ces travaux et propose différentes orientations pour les recherches futures.

Mots-clés : bien-être, bonheur, mesure, psychométrie, questionnaire, représentations, satisfaction, validation.

SOMMAIRE

INTRODUCTION	11
---------------------------	----

PREMIERE PARTIE MESURE ET VALIDATION

Introduction	23
---------------------------	----

Chapitre I : La mesure comme représentation, comme processus et comme test de théories	26
---	----

1.1 Conceptions de la mesure	26
1.2 Le processus de mesure	31
1.3 Conclusion	39

Chapitre II : Validités et validation	41
--	----

2.1 Introduction	41
2.2 La notion de validité à travers son évolution historique	42
2.2.1 La validité comme valeur pronostique	42
2.2.2 Les trois grands types de validité	43
A. La validité critérielle	43
B. La validité de contenu	44
C. La validité de construit	47
2.2.3 Des validités à la validation	50
2.3 Validation intraconcept et validation interconcepts	53
2.3.1 Théorie nomologique et théorie définitoire	53
2.3.2 Logique et stratégies de la validation intraconcept et de la validation interconcepts	55
2.4 En guise de conclusion : validation et théorie de la mesure	60

Chapitre III : Une approche structurale de la mesure : la théorie en facettes ..	64
---	----

3.1 La mesure comme théorie structurale	64
3.2 Principes de l'approche en facettes	65
3.3 Un exemple d'approche en facettes : la satisfaction vis-à-vis de différents domaines de vie	70
3.4 Une évaluation critique de l'approche en facettes	73

**DEUXIEME PARTIE
LE BIEN-ETRE SUBJECTIF
UNE REVUE CRITIQUE DE LA LITTERATURE**

Introduction	78
Chapitre I : Définitions et conceptualisations générales du bien-être subjectif ..	81
1.1 Introduction	81
1.2 Une analyse philosophique et sémantique du bien-être	82
1.3 Le bien-être comme émotion, comme attitude, comme trait de personnalité	84
1.4 Synthèses théoriques et/ou empiriques	88
1.5 Les représentations du bien-être	91
1.6 Définition(s), mesure(s) et théorie(s)	95
1.7 Conclusion	98
Chapitre II : Les mesures de bien-être subjectif	100
2.1 Les instruments de mesure du bien-être	101
2.1.1 Une recension des questionnaires utilisés dans la littérature	101
2.1.2 Description de quelques échelles prototypiques	106
A. L'ABS de Bradburn (1969)	106
B. Le LSI de Neugarten et al. (1961)	109
C. La PGC-MS de Lawton (1972, 1975)	111
D. La MUNSH de Kozma & Stones (1980)	112
E. La SWLS de Diener et al. (1985)	113
2.1.3 D'autres mesures du bien-être subjectif	114
2.2 Evaluation psychométrique des mesures de bien-être	116
2.2.1 Fidélité des mesures de bien-être	117
2.2.2 Validité convergente et validité divergente	119
2.2.3 Validité de construit estimée par les modèles d'équations structurales	121
2.3 Les critiques adressées aux mesures de bien-être	123
2.3.1 Influence de l'humeur du moment	124
2.3.2 Influence de la désirabilité sociale	127
2.3.3 Effets de contexte	131
2.3.4 La distribution des réponses	132
2.4 Conclusion	134
Chapitre III : La structure du bien-être subjectif	138
3.1 Introduction	138
3.2 Les études structurales sur le bien-être	140
3.2.1 Une analyse structurale exploratoire sur des indicateurs généraux de bien-être subjectif : l'exemple d'Andrews & Withey (1976)	141
3.2.2 Analyses factorielles contraintes sur les échelles de bien-être pour personnes âgées : les travaux de Jersey Liang	143
A. Analyses structurales de la PGC-MS	144

B. Analyses structurales du LSI-A	146
3.2.3 Une analyse structurale des représentations du bien-être : l'étude de Thomas & Stock (1988)	149
3.2.4 Une approche en facettes du bien-être subjectif : l'étude de Levy & Guttman (1975)	152
3.3 Le couple positif-négatif dans l'étude du bien-être subjectif	158
3.3.1 Bien-être et mal-être : des dimensions différentes ?	158
3.3.2 Critiques et explications ponctuelles de l'indépendance entre affectivité positive et affectivité négative	160
A. Explications faisant intervenir des caractéristiques spécifiques de l'ABS ...	161
B. Explications faisant intervenir des artefacts méthodologiques	162
C. Explications faisant intervenir l'intensité	163
3.3.3 La structure de l'affectivité	168
3.3.4 L'indépendance entre affectivité positive et affectivité négative sur le plan théorique	173
3.3.5 Conclusion	175
3.4 Le couple affectif-cognitif dans l'étude du bien-être subjectif	178
3.5 Conclusion	183
Chapitre IV : Le réseau nomologique ou les corrélats du bien-être	186
4.1 Introduction	186
4.2 Problèmes théorico-méthodologiques dans l'étude des corrélats du bien-être ..	187
4.3 Les corrélats du bien-être	191
4.3.1 Bien-être et conditions de vie	193
A. Bien-être et société	195
B. Bien-être et place de l'individu dans la société	195
C. Bien-être et travail	199
D. Bien-être et relations inter-personnelles	200
4.3.2 Bien-être et caractéristiques individuelles	201
A. Ressources personnelles	205
B. Traits de personnalité	206
C. Style de vie	207
D. Aspirations, buts, valeurs	207
E. Evaluations de différents aspects de la vie	208
4.3.3 Les "antécédents" du bien-être	209
4.4 Conclusion	210
Chapitre V : Théories nomologiques et modèles explicatifs du bien-être	212
5.1 Introduction	212
5.2 Une proposition de classification des modèles explicatifs du bien-être	215
5.3 Les modèles du bien-être	223
5.3.1 Bien-être général et satisfactions spécifiques	223
5.3.2 Les modèles du bien-être basés sur les conditions de vie	226
A. Le modèle subjectif de la qualité de la vie de Campbell et al. (1976)	227
B. Le modèle des discordances multiples de Michalos (1985)	230

C. Critiques adressées aux modèles d'influence indirecte des conditions de vie	233
5.3.3 Les modèles du bien-être basés sur les changements des conditions de vie ..	240
A. Les modèles du bien-être basés sur l'influence directe des changements vécus	240
B. Les modèles du bien-être basés sur l'influence indirecte des événements vécus	241
* La théorie étendue-fréquence	242
* La théorie du niveau d'adaptation	247
C. Les modèles du bien-être basés sur l'influence directe des changements désirés	252
D. Les modèles du bien-être basés sur l'influence indirecte des changements désirés	256
5.3.4 Les modèles du bien-être basés sur la personnalité	257
5.3.5 Les modèles constructivistes du bien-être	260
5.4 Approches <i>top-down</i> et approches <i>bottom-up</i> du bien-être	264
5.4.1 Introduction	264
5.4.2 Le sens de la causalité : premiers éléments	265
5.4.3 Les approches <i>top-down</i> et <i>bottom-up</i> sur le plan des théories définitoires	268
5.4.4 Les approches <i>top-down</i> et <i>bottom-up</i> sur le plan des théories nomologiques	273
5.5 Le modèle d'équilibre dynamique de Headey & Wearing	283
5.6 Conclusion	290

TROISIEME PARTIE APPROCHE EMPIRIQUE

Chapitre I : Problématique, options théoriques et méthodologiques	294
1.1 Problématique	295
1.1.1 Introduction	295
1.1.2 Une problématique de la définition et de la mesure du bien-être	297
1.1.3 Présentation générale des études empiriques	304
1.2 Structure, jugements de similarité et hypothèse de distorsion systématique	308
1.3 Méthodes d'analyses structurales	313
1.3.1 Analyses spatiales et analyses en clusters	314
1.3.2 Convergences entre analyses factorielles et échelonnements multidimensionnels	315
1.3.3 Différences entre analyses factorielles et échelonnements multidimensionnels	316
1.3.4 Avantages respectifs des analyses factorielles et des échelonnements multidimensionnels	318
1.3.5 Analyses factorielles et échelonnements multidimensionnels de matrices de corrélations	319
1.3.6 La première composante principale dans les solutions d'échelonnement multidimensionnel : absence réelle ou présence virtuelle ?	321
* Etude Monte Carlo 1 : un facteur général et trois facteurs de groupe	325

* Etude Monte Carlo 2 : quatre facteurs de groupe bipolaires	326
* Etude Monte Carlo 3 : quatre facteurs de groupe unipolaires	329
* Conclusion	331
1.3.7 Options méthodologiques générales retenues	332

**Chapitre II : Les instruments de mesure classiques du bien-être :
une approche basée sur les items**

2.1 Introduction	336
2.2 Les échelles classiques de bien-être	338
2.2.1 Les échelles de bien-être pour adultes tout-venant	338
Les échelles	338
Le questionnaire	338
Echantillon	339
Analyses structurales	340
Etude de l'échelle de balance affective	346
Etude de l'échelle de satisfaction avec la vie	349
Relations inter-échelles	350
Etude de quelques corrélats du bien-être	350
Résumé - Conclusion	352
2.2.2 Les échelles de bien-être pour personnes âgées	353
Les échelles	353
Le questionnaire	354
Echantillon	355
Analyses structurales du LSI-A	356
Analyses structurales de la PGC-MS	361
Analyses structurales de la MUNSH	365
Analyses structurales sur l'ensemble des items	368
Analyses structurales de l'ensemble des items positifs des échelles pour personnes âgées	371
Analyses structurales de l'ensemble des items négatifs des échelles pour personnes âgées	375
Résumé des analyses structurales	379
Consistance interne des échelles	380
Relations inter-échelles	382
a. Relations entre le LSI-A, la PGC-MS et la MUNSH	382
b. Relations entre affectivité positive et affectivité négative	383
c. Relations entre les composantes de l'affectivité positive	384
d. Relations entre les composantes de l'affectivité négative	384
Relations entre les échelles de bien-être et les variables externes	385
2.2.3 Conclusion	392
2.3 Une analyse de contenu des items des échelles de bien-être	394
2.3.1 Introduction	394
2.3.2 Sélection du corpus à analyser	396
2.3.3 Grille d'analyse de contenu des items	397
2.3.4 Fidélité de la grille d'analyse	403
2.3.5 Description des items du corpus à l'aide de la grille d'analyse	405

2.3.6 Description des échelles de bien-être en fonction des caractéristiques de leurs items	412
2.3.7 Description des items des échelles de bien-être pour adultes et pour personnes âgées	413
2.3.8 Description des items des échelles de bien-être "affectif" et de bien-être "cognitif"	414
2.4 Effets des caractéristiques des items de bien-être sur les réponses des sujets	417
2.4.1 Introduction	417
2.4.2 Mise au point du questionnaire	418
2.4.3 Echantillons	421
2.4.4 Traitement des données par analyses de variance	421
Effets sur les moyennes	426
Effets sur les écarts-types	431
Effets sur les coefficients de discrimination	436
Résumé - Conclusion	440
2.4.5 Analyses structurales des réponses aux questionnaires	443
Analyses structurales des items sans domaine de l'étude 1	444
Analyses structurales des items sans domaine de l'étude 2	449
Analyses structurales des items avec domaine de l'étude 1	453
Analyses structurales des items avec domaine de l'étude 2	457
2.4.6 Conclusion	461
Chapitre 3 : Le domaine du bien-être : structures et représentations	465
3.1 Introduction	465
3.2 Une investigation du domaine du bien-être basée sur la technique des évocations libres	472
3.3 Etudes structurales des évocations libres	479
3.3.1 Matériel et méthode	479
3.3.2 Analyses des données du tri subjectif	481
3.3.3 Analyse des données du questionnaire	489
3.3.4 Comparaison entre les structures obtenues avec la méthode directe et la méthode indirecte	501
3.3.5 Niveau de bien-être et représentation du bien-être	507
3.3.6 Résumé - Conclusion	512
3.4 Niveau de bien-être et représentation : approfondissements	514
3.4.1 Présentation	514
3.4.2 Les résultats de la tâche d'évocations libres	516
3.4.3 Les évaluations en termes d'importance	520
Hiérarchies d'importance	520
Degré de consensus dans les évaluations d'importance	523
Analyses structurales des évaluations d'importance	524
3.4.4 Liens entre niveau de bien-être et représentations : considérations préalables	529
3.4.5 Différences interindividuelles en matière d'évocations libres	531
Thèmes évoqués (en termes de nombre) en fonction des groupes typologiques	532
Thèmes évoqués (en termes de pourcentage) en fonction des groupes typologiques	533

Nombre d'évocations et scores de bien-être	533
Evocations et variables socio-démographiques	533
3.4.6 Différences interindividuelles en matière d'évaluation d'importance	536
Scores factoriels d'importance en fonction des groupes typologiques	538
Scores factoriels d'importance et scores aux échelles de bien-être	539
Evaluations d'importance en fonction des groupes typologiques	540
Evaluations d'importance et scores aux échelles de bien-être	540
Scores factoriels d'importance et variables socio-démographiques	541
Evaluations d'importance et variables socio-démographiques	543
3.4.7 Résumé - Conclusion	544
Chapitre IV : Une approche du bien-être en termes de manifestations comportementales	549
4.1 Introduction	549
4.1.1 Présentation	549
4.1.2 L'approche par la fréquence des actes : aspects conceptuels et paradigmatiques	551
4.1.3 Critiques de l'approche par la fréquence des actes et implications	554
4.2 Nominations d'actes	559
4.2.1 Consigne	559
4.2.2 Echantillon	560
4.2.3 Le corpus recueilli	561
4.2.4 Actes et non-actes	561
4.2.5 Regroupement thématique des réponses	562
4.3 Evaluations de la prototypicalité des actes	569
4.3.1 Méthode	569
4.3.2 Accord inter-juges des évaluations de prototypicalité	570
4.3.3 Résultats	572
4.3.4 Evaluations de prototypicalité et nombre de citations	576
4.3.5 Différences interindividuelles en termes d'évaluation de prototypicalité	578
4.4 Fréquence de réalisation des actes	581
4.4.1 Méthode	581
4.4.2 Consistance interne du questionnaire	582
4.4.3 Relations entre prototypicalité et fréquence des actes	585
4.5 Analyses structurales des réponses au questionnaire de fréquence	590
4.5.1 Analyse en composantes principales du questionnaire de fréquence	591
4.5.2 Echelonnement multidimensionnel du questionnaire de fréquence	592
4.5.3 Analyse en clusters du questionnaire de fréquence	594
4.6 Résumé - Conclusion	595
CONCLUSION	601
BIBLIOGRAPHIE	613
ANNEXES	647
TABLE DES MATIERES	699

INTRODUCTION

INTRODUCTION

Le bonheur a été, pendant des siècles, l'un des sujets les plus abondamment traités par la philosophie. On a cherché à le définir, à en préciser les causes et les obstacles, à édicter des règles permettant de l'atteindre ; on a eu tendance à s'en méfier ou à le glorifier. Il a régulièrement été assimilé, depuis Aristote, au bien le plus suprême de l'homme et, aujourd'hui encore, dans un langage plus contemporain, le fait qu'il constitue le but principal des actions humaines est difficilement contestable.

L'impression générale qui se dégage de ces analyses préscientifiques est que tout a déjà été dit sur le bonheur. Certains psychologues actuels (e.g. Stones & Kozma, 1986 b) doutent même qu'il soit possible de faire autre chose que de reformuler d'anciennes conceptions. Il est vrai qu'à lire la très belle synthèse qu'a faite Tatarkiewicz (1976) des réflexions philosophiques, éthiques ou théologiques sur le bonheur, on ne peut qu'être étonné de la perspicacité de certaines analyses spéculatives qui trouvent, dans des travaux récents, des confirmations empiriques parfois inattendues. Mais l'arbre ne doit pas cacher la forêt car au-delà de rares convergences, c'est aussi une impression de confusion qui émane de cette littérature.

Si tout a été dit à propos du bonheur, son contraire l'a été aussi, avec moult arguments tout aussi convaincants les uns que les autres. Cette confusion s'exprime surtout à travers les conceptions contradictoires avancées à propos des causes du bonheur, que Tatarkiewicz a illustrées à partir de plusieurs dizaines d'exemples. Mais la signification même du bonheur n'est pas plus consensuelle : sous un même terme, on ne parle pas nécessairement de la même chose. Enfin, le sentiment de confusion est encore accru par le fait que certaines définitions du bonheur sont déjà des prises de position quant à la manière de l'atteindre.

Si le bonheur a été longtemps l'apanage des seuls philosophes, il n'en est plus de même aujourd'hui. Dès le XIX^{ème} siècle, le bonheur rejoint les préoccupations des économistes et des politiques. Mais l'avènement des sciences sociales, qui s'accompagne de l'adoption d'une épistémologie plus positiviste, suscite des réticences quant à l'étude d'un concept éminemment subjectif et propre aux spéculations les plus hasardeuses. Le bonheur est donc très peu présent dans la littérature psychologique de la première moitié de ce siècle.

En reprenant une expression de Kammann et al. (1984 a), on peut dire que le bonheur a été longtemps un non-concept en psychologie, situation que les auteurs attribuent au paradigme behavioriste qui a dominé une partie de l'histoire de la discipline. Il serait faux pourtant de croire que le bonheur a été totalement absent des préoccupations des psychologues. Il est présent dans certains travaux de la psychologie clinique et de la psychologie de la personnalité, mais souvent sous d'autres appellations.

C'est probablement le concept de santé mentale qui s'est substitué, pendant toute une période, au concept de bonheur. Ce concept présentait l'avantage de s'inscrire dans une tradition médico-psychologique qui entendait se démarquer des considérations philosophiques peu compatibles avec la nouvelle volonté d'aborder scientifiquement les phénomènes psychologiques et sociaux.

Mais le concept de santé mentale est trop pauvre pour rendre compte du bonheur. En assimilant le bonheur à l'absence de symptômes pathologiques, on se contente d'en donner une définition par la négative. Il n'est point besoin de longs discours pour reconnaître que des définitions de ce type s'avèrent très vite insuffisantes : être heureux, ce n'est pas seulement ne pas être malade.

Une autre raison incite à abandonner le concept de santé mentale pour étudier le bonheur. Elle tient à la distinction entre une approche objective, normative du bonheur et une approche subjective. Depuis les Grecs — à l'exception de Démocrite — jusqu'au XVII^{ème} siècle, c'est une conception objective du bonheur qui a prévalu. Beaucoup pensaient, à la suite d'Aristote, que le bonheur ne pouvait être atteint qu'en respectant certains préceptes moraux et en adoptant une manière de vivre guidée par la vertu. Cette conception du bonheur est normative car ces règles de "bonne vie" sont dictées par la société ; elle est objective parce qu'un observateur extérieur peut juger du bonheur d'une personne. Au XVIII^{ème} apparaît une conception du bonheur beaucoup plus subjective qui prévaut encore aujourd'hui. Bonheur et plaisir deviennent des quasi synonymes et les plaisirs n'ont pas besoin d'être vertueux. Ce faisant, seule la personne qui en fait l'expérience est à même de juger si elle est heureuse ou non. Dans l'idée de santé mentale, on trouve une conception qui se rapproche des définitions objectives. Certes, la philosophie morale a été remplacée par la psychologie. Il n'en reste pas moins que c'est maintenant le psychologue qui fixe les critères du bonheur. Or une personne souffrant de symptômes pathologiques aux yeux d'un psychologue peut très bien estimer qu'elle est heureuse. Des cas cliniques de ce type ne sont pas rares.

La réapparition explicite du concept de bonheur en psychologie est donc relativement récente. Bien qu'il ne soit pas possible de dater précisément cet événement, on peut considérer que le milieu des années 60 marque un réel renouveau. Depuis, le nombre de travaux ne cesse de croître.

Ce nouvel essor se manifeste tout d'abord par un changement terminologique. Trop chargée de connotations philosophiques, Bradbrun (1969) proposera de remplacer la notion de bonheur par celle de bien-être auquel viendra s'accoler l'adjectif subjectif, probablement pour entériner le changement de conception que nous venons d'évoquer. Bien que cette terminologie ne soit pas définitivement fixée, elle a tendance à s'imposer dans la majeure partie de la littérature contemporaine. Il semble, du moins sur le plan étymologique, que ce choix se justifie. En effet, alors que l'étymologie de bien-être est totalement transparente ("être bien"), celle de bonheur est plus ambiguë. En français, bonheur vient de la contraction de "bon" et de "heur" qui dérive du latin "*augurium*" signifiant "augure", "destin". Ces racines étymologiques se retrouvent dans beaucoup de langues et correspondent, selon Tatariewicz (1976), à une tradition pré-philosophique. Par exemple, en polonais, le terme bonheur s'écrit "*szczęście*" qui se décompose en deux parties, le préfixe positif "*su*" et le substantif "*część*", signifiant destin. En bref, c'est la notion de "bonne fortune" ("*εὐτυχία*") qui transparait dans le langage ; c'est d'ailleurs le premier sens que donnent les dictionnaires au bonheur.

L'intérêt pour ce tout nouveau "bien-être subjectif" n'est sans doute pas étranger à la volonté de développer ce que Warr (1978) a appelé une "psychologie normale", dont l'objectif serait d'étudier "la vie de tous les jours de ceux qui ne sont pas malades". Dans l'esprit des chercheurs de l'époque (e.g. Wilson, 1967), il y a le sentiment que le bonheur constitue le pôle positif d'un continuum de l'adaptation qui n'est connu que dans ses aspects négatifs. Aujourd'hui, on ne parle plus guère d'adaptation et l'existence d'un continuum qui irait de la détresse psychologique la plus profonde jusqu'à la félicité la plus absolue est largement contredite par les faits. Paradoxalement, c'est donc en montrant que les intuitions de l'époque s'avèrent inexactes que l'étude du bonheur assoit définitivement sa légitimité et sa spécificité.

Parallèlement, c'est aussi par le détour de l'adaptation que la gérontologie (re)découvre la notion de bonheur à travers la question suivante : comment définir ce qu'est une vieillesse réussie sans imposer, de l'extérieur, les critères d'évaluation ? N'ayant d'autre solution que de demander directement aux personnes si elles se sentent heureuses, la gérontologie a rejoint les préoccupations des psychologues intéressés par le pôle positif de l'adaptation. Mais cette communauté d'objectifs n'a pas débouché sur une collaboration très étroite. Aujourd'hui encore, la gérontologie semble marquée par l'origine de son intérêt pour le bien-être et se caractérise par une terminologie et des instruments de mesure qui lui sont propres. Les différences entre les deux traditions ont toutefois tendance à s'estomper dans les travaux les plus récents.

Apparue un peu plus tardivement, une troisième voie de recherche caractérise le renouveau de l'étude du bonheur dans les sciences sociales. A la différence des deux précédentes, qui s'inscrivent dans une perspective psychologique, celle-ci fait plutôt

référence à une tradition socio-économique, même si de nombreux psychologues y ont participé et y participent encore. Ce n'est plus en termes d'adaptation qu'est abordé le bien-être mais en termes de qualité de la vie. La volonté politique des gouvernements des pays occidentaux d'améliorer la qualité de la vie s'est accompagnée de la nécessité d'évaluer les résultats des programmes mis en oeuvre. C'est pour répondre à ces demandes pressantes qu'est né le mouvement des indicateurs sociaux dont le principal objectif est de développer des indicateurs permettant de décrire l'état du système social. Il n'a fallu que quelques années pour se rendre compte que les indicateurs objectifs et notamment les indicateurs monétaires étaient insuffisants pour pouvoir apprécier la qualité de la vie telle qu'elle est réellement perçue par les individus. C'est donc tout naturellement que le mouvement des indicateurs sociaux, sous l'impulsion de psychologues, a commencé à mettre au point des indicateurs de la qualité de la vie plus subjectifs qui s'apparentent très largement à la notion de bien-être.

Le bref survol historique que nous venons d'effectuer n'a d'autre intention que de situer approximativement les contours d'un domaine de recherche qui paraissent toujours un peu fuyants. Une véritable histoire du bonheur prenant en compte les développements contemporains reste à écrire, mais là n'est pas l'objectif de ce travail.

Si le bonheur n'a fait qu'une réapparition récente en psychologie, celle-ci a été fulgurante. Le nombre de publications, déjà fort important, continue à croître de manière exponentielle. Ainsi, pour ne citer qu'un exemple, Michalos (1986) rapporte qu'il a recensé 1238 publications en rapport avec le bonheur entre 1969 et 1972 : dix ans plus tard, entre 1979 et 1982, ce sont 2545 publications qui ont pu être repérées. Pourtant, malgré des avancées spectaculaires, la littérature scientifique sur le bonheur conduit à la même impression de confusion que le faisait la littérature philosophique.

Il est difficile de préciser d'emblée à quoi tient cette impression de confusion. Toutefois, ce qui transparait surtout à travers les très nombreuses études empiriques, c'est un manque de clarté conceptuelle. Autrement dit, on ne sait toujours pas aujourd'hui comment définir le bien-être ni même quelles sont exactement les limites de son champ d'études.

Cette thèse entend donc traiter des problèmes conceptuels qui se posent dans l'étude du bien-être. Mais comment aborder ces problèmes conceptuels ? Au risque de surprendre, nous dirons que c'est la mesure qui constitue le point d'achoppement idéal à cet égard. Bien évidemment ce n'est pas seulement d'une mesure instrumentale dont nous parlons ici. Il s'agit au contraire d'envisager la mesure comme un processus d'investigation dans lequel les aspects définitoires et théoriques occupent une place centrale. Comme cette conception peut soulever des interrogations légitimes, il importe que nous commençons à en préciser la logique avant de revenir plus en détail sur nos objectifs.

A l'instar de la plupart des processus de recherche, les études sur le bien-être distinguent habituellement les problèmes de définition, de mesure et de théorie. Ces trois aspects prennent la forme de questions pouvant être formulées schématiquement de la manière suivante :

- qu'est-ce que le bien-être ?
- comment peut-on le mesurer ?
- qu'est-ce qui fait qu'un individu est heureux ? (ou, plus formellement, quels sont les modèles explicatifs du bien-être ?)

Ces trois questions sont souvent envisagées comme des moments différents d'un processus de recherche conçu de manière séquentielle : le concept est d'abord défini ; cette définition permet d'élaborer un ou plusieurs instruments de mesure (dont on vérifie éventuellement les qualités psychométriques) ; les mesures mises au point servent à tester les théories (les modèles explicatifs). Ce fonctionnement, présenté ici de manière très schématique, méconnaît les liens qui unissent ces trois problématiques.

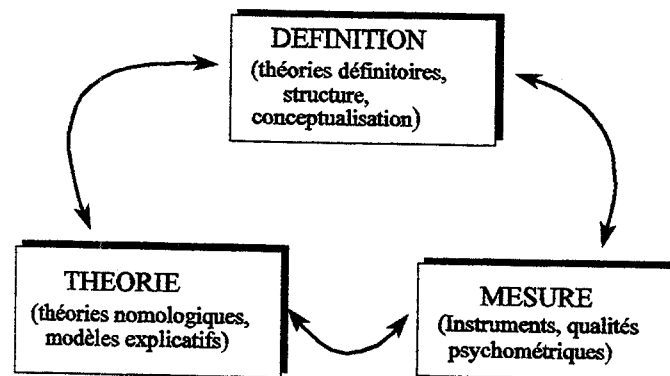


Figure 1 : Les interrelations entre définition, mesure et théorie

La figure 1 illustre ce qui nous semble être une des caractéristiques fondamentales du processus de recherche : l'interdépendance entre les étapes de définition, de mesure et de modélisation théorique. Pour comprendre ce schéma, il faut abandonner l'idée d'un enchaînement strictement séquentiel de ces étapes : la définition ne précède pas la mesure qui précéderait elle-même la mise au point de modèles théoriques.

A nos yeux, les problèmes conceptuels qui caractérisent nombre de domaines de la psychologie naissent principalement du fait que l'on ne tient pas compte de cette interdépendance. Ce faisant, les définitions, lorsqu'elles existent, sont vagues et

imprécises, les instruments de mesure sont utilisés sans que l'on sache très bien ce qu'ils mesurent et les modèles théoriques sont empreints d'ambiguïté.

Certes, ce tableau est un peu caricatural. Mais combien d'exemples peut-on citer dans lesquels cette interdépendance est effective ? Combien d'exemples illustreraient des situations dans lesquelles la mise au point de mesures plus précises a permis de tester de nouveaux modèles, l'évolution des théories a conduit à affiner ou à modifier les définitions, l'introduction de nouveaux paradigmes a changé les opérations de mesure et remis en cause l'intérêt heuristique de définitions antérieures ? Il y en a bien peu.

En revanche, si les étapes de mesure et de définition sont conçues comme faisant partie intégrante du processus de recherche, au même titre que la construction de modèles théoriques, les problèmes conceptuels peuvent être envisagés sous un aspect différent. Dans cette perspective, la mesure procède d'une démarche hypothético-déductive en se prêtant au test de propositions théoriques et les définitions perdent leur caractère statique. Il est alors possible de parler en termes de théories définitoires plutôt qu'en termes de définitions. L'intérêt de ce changement n'est pas que terminologique ; il souligne le fait que la définition d'un concept n'est plus seulement affaire de convention mais aussi de vérification empirique.

La perspective qui est exposée ici n'est pas entièrement originale. Elle rejoint assez largement les conceptions actuelles sur la validité des mesures. Dans celles-ci, le processus de validation présente un caractère intégratif qui doit concourir à l'interprétation et à la signification des mesures. La signification d'une mesure se construit notamment à partir du réseau nomologique entre des concepts, élaboré au fil des recherches. L'établissement de ce réseau nomologique dépend des modèles explicatifs qui sont proposés (et testés) : ainsi, théories nomologiques^[1] et mesures s'influencent réciproquement. Comme cette interdépendance concerne la signification des mesures, elle concerne aussi les problèmes de définition.

Il est relativement facile de distinguer, dans la littérature sur le bien-être subjectif, les études selon qu'elles relèvent de l'une ou de l'autre des trois étapes distinguées ci-dessus :

- Les tests empiriques de modèles explicatifs et, plus simplement, la recherche de corrélats du bien-être représentent la grande majorité des publications.
- Les travaux sur les mesures du bien-être occupent aussi une place importante. Toutefois, ces études adoptent une conception de la mesure plus restreinte que celle qui vient d'être évoquée, en se limitant le plus souvent à vérifier les qualités

[1] : Nous désignerons par théories nomologiques celles qui portent sur les relations entre les concepts, afin de les distinguer des théories définitoires. Les différences entre ces deux types de théories apparaîtront plus clairement dans la première partie.

psychométriques d'instruments existants (fidélité, validité convergente,...) ou à évaluer l'importance d'éventuels biais de réponse.

- Les travaux strictement définitoires sont beaucoup plus rares mais, comme il est d'usage, la plupart des publications réservent à ces questions quelques lignes ou, au mieux, quelques paragraphes en introduction. Il faut noter le cas particulier des études qui visent à dégager les composantes du bien-être à partir d'études structurales. Cet objectif justifie qu'on les considère comme des travaux définitoires, mais force est de constater qu'en fait de composantes du bien-être, ces études dégagent principalement les facteurs qui structurent les questionnaires existants.

Beaucoup d'auteurs regrettent que les problèmes de définition et de conceptualisation du bien-être soient si peu maîtrisés. Il est vrai qu'ils sont la source de nombreuses confusions :

- Confusions terminologiques : le bien-être est étudié, dans la littérature, sous des appellations différentes (bonheur, satisfaction, moral, santé mentale, qualité de la vie, "*welfare*" pour ne citer que les plus courants)^[2]. En l'absence de définitions claires de ces concepts, il est très difficile de savoir s'ils désignent ou non le même phénomène.

- Confusions conceptuelles : le bien-être est tantôt conçu comme une émotion, tantôt comme une attitude... ou comme un trait de personnalité, un état psychologique, un sentiment, une sensation, une humeur, un affect.

- Confusions théoriques : selon les études, un même concept sera soit considéré comme une composante du bien-être soit intégré dans son réseau nomologique (et donc considéré comme un concept distinct).

Loin d'être anecdotiques, ces confusions ont des conséquences importantes sur les recherches. Elles expliquent en particulier les difficultés que rencontre tout chercheur à synthétiser les résultats, à en présenter une analyse raisonnée et à adopter une démarche réellement cumulative. Elles expliquent aussi le cloisonnement relatif entre des disciplines ou des sous-disciplines qui adoptent des perspectives différentes : comment, en effet, intégrer les résultats de travaux qui considèrent le bien-être comme un état et ceux qui le considèrent comme une disposition ? Comment intégrer ce que l'on sait du bien-être en tant qu'attitude et en tant qu'émotion ?

[2] : Andrews & Robinson (1991, p. 61) proposent une liste — non exhaustive — encore plus longue : "*satisfaction, happiness, morale, positive affect, negative affect, affect balance, cognitive evaluation, elation, subjective well-being, sense of well-being, psychological well-being, perceived well-being, subjective welfare, (subjective / sense of / psychological / perceived) ill-being, anxiety, depression, distress, tension and perceived life quality*".

Il existe pourtant de rares tentatives dont l'objectif affiché a été de contribuer à cette clarification conceptuelle tellement nécessaire. Malheureusement, celles-ci restent peu convaincantes. C'est le cas par exemple du concept de "bonne vie" de Lawton (1983a, 1983b), de celui de qualité de la vie de Hörnquist (1982) ou du modèle systémique de la qualité de la vie de Shye (1989). Toutes ces tentatives proposent des conceptualisations plus ou moins formalisées visant à intégrer, dans un même schéma définitoire, différentes conceptions du bien-être. Malheureusement, ces conceptualisations deviennent tellement générales et abstraites que leur intérêt heuristique est limité. Aucune n'a d'ailleurs connu le succès auquel elle aurait pu prétendre.

Notre objectif n'est pas de proposer, nous aussi, un nouveau méta-concept qui ne ferait qu'ajouter à la confusion générale. Il nous semble plus important d'essayer de reprendre à la base les problèmes conceptuels qui se posent dans le cas du bien-être subjectif. Présenté ainsi, cet objectif n'est guère explicite. Il nous faut donc le préciser.

Ainsi que nous l'avons indiqué plus haut, la littérature scientifique sur le bien-être atteint un volume considérable qu'aucun chercheur ne peut embrasser totalement dût-il y consacrer toute sa carrière. Malheureusement, cette situation est préjudiciable à l'intégration des connaissances et elle contribue à la pérennité des confusions. Alors que telle question aura reçu une réponse satisfaisante dans une sous-discipline, elle continuera à être ignorée de chercheurs qui travaillent dans une autre sous-discipline. La conséquence évidente d'une telle situation est que l'on arrive à un stade où les chercheurs ne parlent plus du tout de la même chose. Caricature de la réalité ? Certainement pas. Nous aurons l'occasion de constater que non seulement de telles situations ne sont pas rares mais qu'elles dépassent parfois l'imagination.

Il nous semble donc que pour reprendre à la base les problèmes conceptuels posés par le bien-être subjectif, une revue critique de la littérature s'avère être un préalable indispensable. Certes, il ne s'agit pas, *a priori*, d'une proposition très originale. Pourtant, la dernière revue de la littérature sur le bien-être date de dix ans (Diener, 1984) et elle souffre, malgré ses qualités, de nombreuses imperfections sur lesquelles nous aurons l'occasion de revenir. La revue de la littérature que nous souhaitons faire n'existe pas. Elle sera abordée de manière à faire apparaître les liens entre définition, mesure et théorie. Bien plus qu'une simple présentation linéaire, il s'agit de faire un bilan raisonné de la littérature qui allie exhaustivité et critique, qui intègre les résultats de sous-disciplines qui s'ignorent, qui confronte conceptions et paradigmes, qui met en lumière les postulats implicites, les confusions et les erreurs méthodologiques. Conçue ainsi, cette revue de la littérature n'est pas seulement un exercice de style, elle est le résultat d'un travail de recherche qui comprend aussi des études empiriques ponctuelles pour questionner la validité de certaines démarches.

Ce sont certainement dans les considérations liées à la mesure que les problèmes conceptuels sont les plus marqués. On ne s'en étonnera pas : aucune opération de mesure ne peut échapper à la question de la signification de ce qui est mesuré. La partie empirique de cette thèse se proposera donc de préciser ce qui est mesuré par les instruments utilisés dans les recherches sur le bien-être. Surtout, elle visera à essayer de compléter notre connaissance du domaine du bien-être en termes de contenus. Mais si la signification conceptuelle est centrale dans la problématique de la mesure, elle renvoie nécessairement à des interrogations sur les représentations associées à ce concept, que ces représentations soient celles du psychologue ou celles de l'homme de la rue. Nos études empiriques tenteront donc de conjuguer les aspects concernant le domaine du bien-être et sa représentation.

Cette thèse comporte donc tout naturellement trois parties principales. La **première partie** se propose de présenter plus en détail notre conception de la mesure à la lumière du schéma plus général de la recherche illustré précédemment. Un *premier chapitre* nous permettra d'expliquer pourquoi la mesure peut être conçue à la fois comme activité de représentation, comme processus de recherche et comme test de théories. Ce faisant, nous insisterons surtout sur la démarche générale de l'opération de mesure au détriment de considérations plus spécifiques. Le *second chapitre* viendra combler cette lacune en proposant une réflexion autour de la notion de validation. Parce que c'est la notion qui renvoie la plus explicitement à la question de la signification des mesures, celle-ci présente un intérêt majeur pour notre propos. Le *troisième chapitre* illustrera comment il est possible d'aborder l'opération de mesure à travers une démarche hypothético-déductive qui se donne les moyens de poser des hypothèses réfutables par des données empiriques.

La **seconde partie** est entièrement consacrée à la revue critique de la littérature sur le bien-être. Nous avons déjà indiqué dans quel état d'esprit est effectuée cette synthèse. Rappelons simplement qu'elle vise à mettre en évidence les liens entre les aspects définitoires, psychométriques et théoriques de l'étude du bien-être et à clarifier certaines confusions fréquentes dans la littérature. Afin d'organiser la présentation, nous examinerons successivement : les tentatives de définitions et de conceptualisations (*chapitre 1*), les travaux sur les instruments de mesure (*chapitre 2*), ceux qui portent sur l'identification des composantes du bien-être et sur sa structure (*chapitre 3*), les très nombreux corrélats (*chapitre 4*) et les non moins nombreux modèles explicatifs (*chapitre 5*).

La **troisième partie** rapporte les études empiriques. Il est trop tôt pour en exposer les détails. Le *premier chapitre* sera entièrement consacré à l'exposé de la problématique et clarifiera les options théoriques et méthodologiques qui seront prises. Dans le *second chapitre*, nous aborderons la mesure du bien-être à partir des instruments utilisés dans la

littérature ; ce chapitre constitue en quelque sorte l'analyse du concept de bien-être sur la base des représentations des chercheurs. A l'inverse, les deux chapitres suivants aborderont la mesure du bien-être à partir des représentations qu'attribue l'homme de la rue à ce concept. Dans le *troisième chapitre* nous verrons comment ces représentations peuvent être utilisées afin d'enrichir la signification quelque peu étroite que les chercheurs confèrent au bien-être. Dans le *quatrième chapitre*, le domaine du bien-être sera enrichi sous un autre angle, celui des aspects comportementaux.

Nous n'avons ni l'ambition ni la prétention de résoudre définitivement les problèmes conceptuels relatifs au bien-être. Cela n'aurait d'ailleurs aucun sens étant donné notre conception des liens entre définition, mesure et théorie : la signification d'un concept dépend de la définition qu'on lui donne et celle-ci est toujours susceptible d'être remise en cause sur la base de nouveaux développements théoriques. Nous espérons toutefois que notre contribution s'inscrit à une juste place au sein d'un processus de recherche dans lequel rien n'est jamais figé et tout est continuellement remis en question.

PREMIERE PARTIE
MESURE ET VALIDATION

PREMIERE PARTIE MESURE ET VALIDATION

INTRODUCTION

Avant d'aborder l'étude du bien-être subjectif proprement dit, il importe de préciser notre conception de la mesure car celle-ci est au centre de ce travail. C'est à cela que nous allons nous attacher dans cette première partie.

Il n'y a guère de publications scientifiques qui ne fassent état de mesure et, en psychologie comme dans toute science, le progrès des connaissances passe par une dialectique complexe entre la qualité des mesures et l'élaboration théorique. En psychologie, la mesure se heurte à des problèmes qui paraissent spécifiques en raison de la subjectivité inhérente à son domaine d'études et de l'impossibilité fréquente d'obtenir des données ayant la précision de celles que l'on trouve, par exemple, en sciences physiques. Il n'est donc pas étonnant de constater que les réflexions générales les plus abouties sur la mesure soient le fait de psychologues ou, tout du moins, de chercheurs qui ont été confrontés, dans leurs travaux, à des questions de nature psychologique.

La mesure et la psychométrie, discipline qui en a fait son objet d'investigation, peuvent être envisagées de plusieurs manières. Au niveau le plus général, la mesure suscite des interrogations d'ordre épistémologique qui visent notamment à préciser les conditions devant être remplies pour qu'une mesure soit possible. De façon plus technique, la psychométrie a aussi comme objectif de développer des modèles mathématiques permettant d'améliorer la qualité des mesures. Mais la mesure est surtout présente dans un nombre considérable d'études empiriques, dès lors qu'il y a eu attribution de nombres à des objets.

Les liens entre ces trois approches, qui peuvent être désignées, pour simplifier, comme approches théorique, technique et pratique, ne sont pas toujours évidents. Les travaux théoriques explorent souvent des possibilités très éloignées des contraintes pratiques de la recherche ; les développements de modèles s'effectuent quelquefois en contradiction avec les principes énoncés au niveau théorique ; surtout, nombre d'utilisateurs ne portent qu'une attention limitée à une étape qui conditionne pourtant la qualité ultime des résultats produits.

Dans un ouvrage récent (Dickes, Tournois, Flieller & Kop, 1994), nous avons essayé d'offrir aux chercheurs une vision intégrée des approches théorique, technique et pratique en situant la mesure dans le contexte plus général de l'ensemble de la démarche

scientifique. Nous avons défendu, dans cet ouvrage, l'idée que l'on ne pouvait aborder la mesure qu'en termes de processus dont les règles et les enjeux ne différaient guère de ceux de la recherche scientifique. La mesure a, bien évidemment, ses propres spécificités qui s'expriment notamment par l'intermédiaire de ses techniques. Mais ces spécificités ne constituent nullement une rupture par rapport à cette démarche scientifique générale.

Il n'est pas question ici de reprendre en détail cet ouvrage. Nous souhaitons plutôt insister sur les aspects de la mesure qui nous semblent les plus fondamentaux ainsi que sur ceux qui sont directement pertinents pour ce travail. Aussi, privilégierons-nous, dans cette présentation, les considérations générales au détriment de considérations plus techniques dont le lecteur pourra trouver un exposé détaillé chez Dickes et al. (1994)^[1]. Cet exposé nous permettra aussi de faire état de nouveaux éléments de réflexion qui nous sont apparus depuis la rédaction de l'ouvrage cité.

Dans la première partie de cet exposé, nous présenterons plus en détail la conception de la mesure qui vient d'être esquissée ci-dessus. En reprenant les définitions classiques de la mesure, nous verrons, dans un premier temps, que celle-ci peut être conçue comme une activité de représentation au sens que donne la psychologie cognitive à ce terme. Surtout, nous insisterons sur les conséquences importantes qui découlent du fait que mesurer revient à tester une théorie, même si celle-ci n'est pas toujours explicitée avec la rigueur souhaitée. Dans une seconde section, la démarche de mesure sera précisée à travers la description des différentes étapes qui la caractérisent. Les problèmes de définition, qui occupent une place centrale dans cette démarche, feront l'objet d'une discussion approfondie.

Le second chapitre sera entièrement consacré à la notion de validation. Notion presque aussi diffuse que celle de mesure, elle en est indissociable. En retraçant l'évolution historique des stratégies développées pour s'assurer que l'on mesure bien ce que l'on est censé mesurer, nous verrons que la notion de validité s'est tout d'abord morcelée avant d'être réunifiée. Mais cette réunification a nécessité des choix qui rendent les conceptualisations actuelles encore insatisfaisantes. Nous proposerons donc une conception alternative de la validation qui a le mérite d'intégrer des aspects jusque là négligés. Bien qu'elle ne soit pas entièrement originale puisque l'on peut en trouver les prémisses chez certains auteurs, cette conception de la validité contribue à clarifier notablement les pratiques en ce domaine. Ce chapitre sera clos par quelques éléments de réflexion sur le fossé qui sépare les liens entre les approches théoriques de la mesure d'une part et les approches techniques et pratiques d'autre part.

[1] : Les aspects techniques des modèles de mesure qui seront mises en oeuvre dans cette thèse seront considérés ultérieurement.

Enfin, le troisième chapitre se propose de présenter la théorie en facettes qui constitue tout à la fois une conception de la mesure et une technique pour aborder les différentes phases du processus de mesure. Ce sont ces deux aspects qui justifient la place prise par cette approche dans cet exposé : elle illustre mieux que toute autre en quoi la mesure est un test de théorie et elle met l'accent sur des étapes du processus qui ont trop peu retenu l'attention des chercheurs jusqu'à présent.

CHAPITRE I

LA MESURE COMME REPRESENTATION, COMME PROCESSUS ET COMME TEST DE THÉORIES

1.1 Conceptions de la mesure

Il n'est presque plus besoin de rappeler la célèbre définition de la mesure attribuée à Stevens (1951) tellement celle-ci a été citée, commentée et critiquée : mesurer, c'est assigner des nombres à des objets selon certaines règles. Si le fait que la mesure concerne l'assignation de nombre à des objets ne semble guère être un sujet de polémique – quoique tout autre système formel ferait tout aussi bien l'affaire – on se doute que ce sont les règles utilisées pour cette assignation qui constituent l'essentiel des débats qui agitent les théoriciens de la mesure et les psychométriciens. L'assignation de nombres à des objets selon une certaine règle n'est pas suffisante pour réaliser une opération de mesure ainsi que le notent, ironiquement, Dawes & Smith (1985). Que l'on considère, par exemple, la règle suivante : le nombre attribué à un individu est égal au produit de sa taille, de la superficie de son logement et du temps moyen qu'il passe dans les transports pour se rendre à son travail. On admettra, sans grande difficulté, que cette assignation de nombres est conforme à la définition ci-dessus, mais qu'elle ne constitue pas une mesure malgré la précision de la règle utilisée^[1].

Il faut donc préciser cette définition. Notons tout d'abord que l'on ne mesure jamais des objets, mais certaines de leurs caractéristiques, de leurs propriétés : on ne mesure pas une table, mais sa longueur, sa masse, sa couleur, sa beauté ... ou son prix. Cette remarque est importante. Elle a comme conséquence que le fait d'attribuer des nombres identiques à des objets différents ne signifie pas que ces objets sont identiques : ils sont équivalents au regard de la caractéristique qui est mesurée *et* dans le cadre de référence spécifié par la règle d'assignation des nombres. Qu'en est-il alors de cette fameuse règle si ce n'est pas sa précision qui détermine une opération de mesure ? En reprenant la formulation de Jacoby (1991), qui reprend les idées de Stevens, on dira que pour qu'il y ait mesure, il faut que l'assignation de nombres à des objets possède une propriété importante : les différences entre les nombres assignés aux objets doivent refléter, selon une procédure qu'il faut spécifier, les différences entre les objets en ce qui concerne la caractéristique mesurée. Cette propriété qui vient d'être énoncée permet de mettre en évidence deux systèmes différents dont la mise en correspondance est à la base de la

[1] : Il faut d'ailleurs reconnaître que Stevens lui-même reviendra plus tard sur cette définition en considérant qu'elle est incomplète.

mesure : un système formel et un système empirique. Le système empirique est celui de la réalité empirique dans laquelle se situent les objets que l'on veut mesurer^[2]. Le système formel est celui des nombres auxquels sont attachées des propriétés mathématiques plus ou moins nombreuses selon le niveau de la mesure.

Cette conception de la mesure est assez largement acceptée à l'heure actuelle. On la retrouve sous une forme à peine différente dans les écrits des épistémologues qui ne sont pas nécessairement des spécialistes de la mesure. Que l'on en juge plutôt par cette citation de Granger dans un de ses ouvrages récents sur la science (1993, p. 93) : "La mise en correspondance de faits humains et de nombres n'a en effet de sens que si un système d'opérations portant sur ces faits peut être établi, qui ait la même organisation formelle que le système de nombres considéré."

Si mesurer consiste à attribuer des nombres à des objets de manière à ce que les différences entre les nombres reflètent les différences entre les objets selon la caractéristique qui est mesurée, mesurer revient à adopter une certaine représentation de la réalité. Autrement dit, mesurer c'est appliquer un modèle abstrait à des observations concrètes, ou encore, la mesure est une forme de représentation. Il est d'ailleurs frappant de noter à quel point certaines définitions des représentations se superposent aux définitions de la mesure^[3]. Cet aspect est particulièrement bien souligné chez Palmer (1978). Il a été repris par des auteurs comme Denis qui écrit (1989, p. 21) : "(...) nous dirons qu'il y a eu activité de représentation lorsqu'un objet ou lorsque les éléments d'un ensemble d'objets se trouvent exprimés, traduits, figurés, sous la forme d'un nouvel ensemble d'éléments, et qu'une correspondance systématique se trouve réalisée entre l'ensemble de départ et l'ensemble d'arrivée."

Mesurer est donc une activité de représentation : la plupart des théoriciens actuels de la mesure s'inscrivent d'ailleurs dans ce que beaucoup appellent la conception représentationnelle de la mesure ou, plus simplement, la théorie représentationnelle. Toute théorie de la mesure vise à établir les conditions dans lesquelles il est possible d'établir une correspondance entre un système formel et un système empirique (Aftanas, 1988). La théorie représentationnelle a été précédée, historiquement, par une théorie dite classique dans laquelle seules les grandeurs extensives^[4] sont mesurables et par une théorie opérationnaliste pour laquelle toute opération qui conduit à un nombre est une

[2] : Les débats philosophiques sur le statut de la réalité n'ont pas d'importance pour cette conception de la mesure : que celle-ci soit construite ou donnée, elle participe de la même manière à l'appréhension des objets.

[3] : Nous remercions Jocelyne Tournois d'avoir attiré notre attention sur ce point.

[4] : Une grandeur extensive suppose la possibilité de réaliser une addition empirique, appelée concaténation, des caractéristiques des objets (mettre deux objets sur le même plateau d'une balance pour la masse, aligner bout à bout deux objets pour la longueur, etc.).

mesure, à condition que cette opération soit suffisamment précise. La théorie représentationnelle se distingue de la théorie classique en diminuant le nombre de conditions requises pour réaliser une mesure (pour simplifier, disons que les axiomes de l'addition sont une condition nécessaire dans la théorie classique, mais pas dans la théorie représentationnelle) ; elle se distingue de la théorie opérationnaliste en refusant d'accepter que toute opération conduisant à un nombre soit une mesure. Il est donc devenu courant de désigner par le terme de théorie représentationnelle les conceptions défendues par les psychologues mathématiciens qui cherchent à établir les conditions nécessaires et suffisantes à l'établissement d'un homomorphisme entre un système empirique et un système formel. Mais le choix de ce terme n'est peut-être pas très heureux. La conception classique défend tout autant le caractère représentationnel de la mesure ; elle diffère sur le nombre de conditions requises pour établir la correspondance entre les deux systèmes.

L'activité de représentation que nous venons de mettre en évidence pour la mesure ne lui est pas spécifique. Elle constitue le fondement même de toute démarche scientifique. Ainsi, tout comme la démarche scientifique teste ses modèles théoriques à partir de prédictions qui doivent se vérifier dans le domaine empirique, la mesure se doit de vérifier que les relations entre les objets du système empirique sont conformes avec les hypothèses découlant de l'adoption d'un modèle (d'un système formel) censé représenter la propriété de ces objets. Ainsi, si, par exemple, on assigne à un objet A le nombre 4 pour représenter son poids (la propriété mesurée) et les nombres 6 et 10 respectivement à deux objets B et C pour représenter la même propriété, on peut, à partir du système formel, faire l'hypothèse que A et B contrebalancent C. Cette hypothèse concerne le système de relations empiriques entre ces éléments : il faut donc vérifier qu'une balance est en équilibre lorsque A et B sont sur l'un des plateaux et C sur l'autre. Si cette hypothèse n'est pas vérifiée, le système formel choisi sera jugé inadéquat. Notons au passage que la vérification de l'hypothèse ne prouve pas que le système formel choisi soit adéquat. Comme dans toute théorie, la non réfutation d'une (ou de plusieurs) hypothèse(s) ne signifie pas que toutes les hypothèses que l'on peut dériver du système formel sont vérifiées. Cet exemple montre seulement que l'opération de mesure consiste à faire des prédictions ; en cela, la mesure n'est rien d'autre qu'une théorie portant sur les relations entre les éléments d'un système empirique. La seule différence entre les théories liées à l'élaboration de mesures et celles de la science en général est que ces dernières portent sur les relations entre des concepts différents, alors que les premières portent sur les relations entre des éléments de même nature. Une première conséquence qui découle de l'assimilation de la mesure à la démarche hypothético-déductive caractérisant toute recherche scientifique est que, comme toute théorie, la mesure n'est jamais avérée : elle est sans cesse remise en question.

Mais la pratique de la mesure ne correspond pas toujours à cette conception idéale : le plus souvent, les démarches adoptées ne permettent pas de vérifier l'adéquation

entre le système formel choisi et le système empirique étudié. Les hypothèses sont irréfutables. C'est cette pratique que Dawes & Smith (1985) appellent la mesure d'indice ou la mesure non représentationnelle : dans ce cas, des nombres sont attribués à des observations empiriques sans que ces observations puissent remettre en cause cette assignation. Cette pratique est largement répandue : les cas les plus fréquents se rencontrent chez certains expérimentalistes qui "mesurent" leur variable dépendante avec un seul item. Mais cette situation n'est pas limitée au cas où un seul item est utilisé : si le chercheur utilise une échelle composée de plusieurs items qui sont additionnés sans autre précaution, le résultat est le même : l'hypothèse selon laquelle les items mesurent le même concept est acceptée comme postulat sans aucune possibilité de réfutation. Cette situation n'est pas rare : Dawes & Smith (1985) estiment que 90 % des mesures d'attitude sont des mesures non-représentationnelles. Ils notent cependant, à juste titre, que beaucoup de ces mesures sont utiles dès lors qu'elles permettent de faire des prédictions ; mais ces prédictions sont d'une autre nature : dans le cas des mesures non-représentationnelles, les prédictions sont essentiellement externes, dans le cas de mesures représentationnelles, les prédictions sont essentiellement internes. Nous nous expliquerons plus avant sur ce point dans le chapitre concernant la validation.

Cette pratique non-représentationnelle de la mesure est à rapprocher de la conception opérationnaliste défendue par le physicien Bridgman (1927), évoquée plus haut. Dans cette conception, la mesure n'est rien d'autre qu'une suite d'opérations clairement définies. Cette conception présuppose que le progrès scientifique dépend de la précision, de la clarté avec laquelle sont formulées les définitions et donc la suite des opérations qui conduisent à une mesure. Or, l'histoire de la physique en particulier et l'histoire des sciences en général ont montré que la précision des définitions opérationnelles de concepts pouvait être une conséquence du progrès scientifique (i.e. des modèles nomologiques) plutôt qu'une cause. On rejoint ici les relations entre définitions, mesures et théories esquissées en introduction.

Nous avons déjà souligné que si les opérations de mesure consistent à tester des théories, la mesure n'est jamais avérée. Mais cette conséquence n'est pas la seule. En effet, si un modèle nomologique n'est qu'un système formel représentant les relations entre des concepts distincts du monde empirique et si *chacun* de ces concepts est mesuré par un autre système formel visant à représenter les relations entre les éléments du concept, que se passe-t-il lorsqu'une hypothèse dérivée du système formel nomologique est réfutée ? Dans la grande majorité des cas, c'est le modèle théorique nomologique qui est remis en cause. Il est relativement rare que la mise en cause porte sur les systèmes formels adoptés pour représenter (pour mesurer) chacun des concepts, car ceux-ci sont acceptés comme définitifs. Or, ils sont tout autant susceptibles de révision, de réfutation que les systèmes formels nomologiques. Un exemple élémentaire viendra illustrer ce raisonnement. Supposons qu'un chercheur adopte un modèle théorique nomologique,

duquel il dérive l'hypothèse qu'un concept A est lié avec un concept B. La vérification empirique de cette hypothèse passe par le calcul d'un coefficient d'association (disons d'une corrélation linéaire, pour simplifier) entre *une mesure* de A et *une mesure* de B. Calculant une corrélation, le chercheur postule notamment que les concepts peuvent être représentés par un système formel unidimensionnel et que la fonction qui relie les objets et les nombres qui leur sont assignés est une fonction linéaire (ce qui signifie que le niveau de mesure est d'intervalles, cf. Jacoby, 1991). Si le postulat d'unidimensionnalité s'avère inadéquat pour l'une ou pour l'autre mesure (ou pour les deux), il est fort probable que la corrélation calculée sera faible^[5]. Ainsi, si le chercheur oublie que les données utilisées pour tester un modèle nomologique sont elles-mêmes un modèle de représentation des observations empiriques qui doit être vérifié, il tirera des conclusions abusives de son étude, à savoir que sa théorie nomologique est fautive (elle l'est peut-être, mais pas nécessairement).

Ces liens entre système formel nomologique et système formel de mesure trouvent un écho tout à fait remarquable dans l'évolution constatée depuis quelques années dans les modèles d'analyse de données. L'essor d'un modèle comme le LISREL, par exemple, est certainement lié en partie à la possibilité de tester *simultanément*^[6] un modèle de mesure et un modèle structural. Autrement dit, un tel modèle revient à tester simultanément l'adéquation du système formel de représentation des relations entre les éléments d'un même concept et l'adéquation du système formel de représentation des relations entre des concepts différents. C'est le cas également des modèles ALS/OS, moins bien connus, mais dont le développement nous semble assuré dans les années qui viennent. Ces modèles ("Alternative Least Squares/Optimal Scaling") ajustent le système formel utilisé pour la mesure d'un concept (notamment en changeant les règles initiales d'attribution des nombres aux objets) afin de maximiser l'adéquation du système formel nomologique aux observations (cf. Young, 1981 ; Jacoby, 1991). En tirant partie notamment de la plus grande flexibilité des variables mesurées au niveau ordinal ou nominal, ils illustrent encore plus nettement que le LISREL les influences réciproques entre système formel de mesure et système formel nomologique.

Nous avons dit plus haut que lorsqu'une hypothèse dérivée du système formel nomologique était réfutée, cela s'accompagnait, dans la plupart des cas, d'une remise en cause du modèle théorique nomologique. En fait, il faut nuancer cette affirmation.

[5] : C'est assez facile à comprendre : des mesures non unidimensionnelles ont tendance à avoir une fidélité basse et le coefficient de corrélation est sensible à la fidélité des mesures (même dans le cas d'une relation parfaite entre deux concepts, la corrélation entre les mesures de ces concepts ne peut être supérieure à la racine carrée du produit des coefficients de fidélité des deux mesures).

[6] : Bien que certains auteurs (e.g. Anderson & Gerbing, 1988) conseillent une approche en deux étapes : le test du modèle de mesure dans un premier temps et le test du modèle structural dans un second. Mais ces recommandations techniques n'enlèvent rien à la logique qui sous-tend ce modèle.

Nombre de controverses scientifiques utilisent, en effet, comme arguments une mauvaise opérationnalisation des concepts utilisés, des définitions divergentes ou encore la non prise en compte du caractère multidimensionnel des variables mesurées. Ces arguments concernent tous la qualité de l'adéquation entre le système formel utilisé pour la représentation du concept et les observations empiriques, mais ce sont, presque toujours, des arguments *ad hoc*. Autrement dit, ils sont évoqués *a posteriori* pour rendre compte, de manière ponctuelle, des inconsistances entre certains résultats. Ces arguments sont d'ordre théorique ou logique ; ils s'appuient rarement sur des résultats empiriques. Seule une conception de la mesure comme test de théories peut permettre de sortir de ces impasses.

1.2 Le processus de mesure

Après avoir souligné les liens entre la démarche de mesure et la démarche scientifique en général, précisons, de manière un peu plus concrète, ce qui fait la spécificité de la démarche de mesure. La démarche de mesure est un processus qui comporte plusieurs étapes présentées dans la figure 1.1. Pour des raisons de simplicité, ce processus est représenté isolément, mais il ne faut pas oublier qu'il s'intègre dans le schéma plus général des relations entre définition, théorie et mesure.

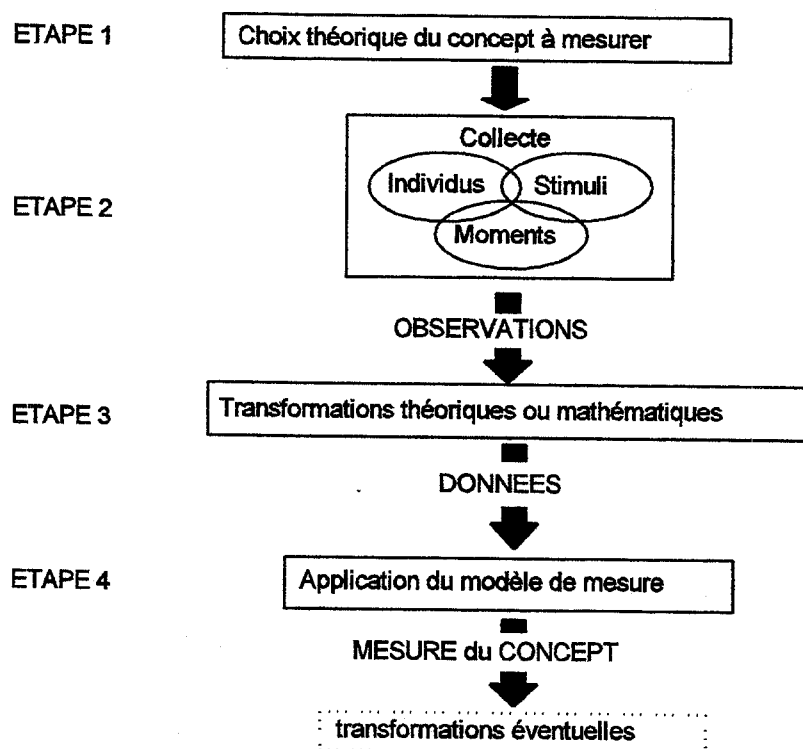


Figure 1.1 : Les étapes du processus de mesure (adapté d'après Dickes, Tournois, Flieller & Kop, 1994, p. 39)

On peut dégager quatre grandes étapes dans le processus de mesure, plus une étape facultative qui concerne la transformation des nombres attribués aux objets en d'autres nombres, le plus souvent afin d'en faciliter la présentation (e.g. suppression des valeurs négatives, standardisation...). Cette étape étant accessoire et n'ayant pas un intérêt direct pour la démarche de mesure en tant que telle, elle ne sera pas détaillée plus avant.

Ce schéma entend montrer clairement que la mesure est un résultat qui dépend de plusieurs étapes et, comme tout résultat, que sa qualité dépend du soin apporté à chacune d'entre elles. Il en découle que le processus de mesure est une démarche extrêmement flexible : à chaque étape, le chercheur est confronté à des choix qui doivent être justifiés. Selon les sujets choisis, les items retenus, la technique de collecte utilisée, les transformations pratiquées et le modèle appliqué aux données, différents résultats (différentes mesures) peuvent être obtenus. Comme dans tout autre processus de recherche, la convergence et la répliquabilité des résultats sont le garant de la qualité de la mesure qui aura été établie. Notons aussi que ce schéma, en présentant la mesure comme une suite d'opérations, réhabilite une version faible de la conception opérationnaliste de la mesure. Si la mesure ne peut se concevoir *seulement* comme une suite d'opérations définies de manière suffisamment précise, aucune mesure ne peut être établie sans ces opérations.

La première étape concerne le *choix théorique du concept à mesurer*. C'est sans doute à ce niveau qu'apparaissent le plus nettement les liens entre théorie et mesure. On ne mesure pas un concept simplement pour le plaisir de le mesurer. L'objectif est toujours, à plus ou moins long terme, de mettre en relation ce concept avec d'autres concepts. Ainsi, le choix du concept est fortement dépendant d'une théorie nomologique plus générale et dépend de l'état de nos connaissances à un moment donné. Le progrès scientifique suggère la création de nouveaux concepts, leur modification ou leur abandon. Ces nouveaux concepts, lorsqu'ils sont mesurés avec succès, contribuent eux-mêmes à enrichir nos théories scientifiques.

Mais choisir un concept n'est pas suffisant, il faut le définir afin de *spécifier les observations* qui relèvent de ce concept (et donc celles qui n'en relèvent pas). Comme la mesure entend rendre compte des relations observées dans le système empirique, cette étape est primordiale. Selon que certaines observations sont considérées comme relevant ou non du concept, le résultat de la mesure pourra être radicalement différent. Que l'on songe, à titre d'illustration, à la mesure de l'intelligence des enfants. Selon que les observations tiennent compte ou non des compétences sociales, le classement des enfants les uns par rapport aux autres (i.e. les relations entre eux) en fonction de la caractéristique "intelligence" ne sera pas forcément le même.

Une observation n'est rien d'autre que la réponse d'un individu (au sens statistique du terme) à un stimulus, à un moment donné. Il est bien évidemment impossible de disposer de la totalité des observations potentielles du concept qui seraient, en l'occurrence, les réponses de tous les sujets à tous les stimuli possibles à une infinité de moments. Le chercheur dispose donc d'un échantillon de ce qu'il est convenu d'appeler l'univers des observations. Comme dans tout passage d'un univers (ou d'une population) à un échantillon, on rencontre des problèmes de représentativité et d'inférence. En effet, comme le note très bien Kane (1982), aucun chercheur ne se satisfait de la mise en relation de deux observations : ce qui est important c'est de pouvoir généraliser la relation observée à des classes d'observation, généralisation que l'auteur considère comme l'objet même de la psychométrie.

Ce problème de représentativité se complique du fait que l'inférence doit porter à la fois sur les individus et sur les stimuli. Si les techniques d'échantillonnage des individus sont relativement bien maîtrisées, il n'en est pas de même de l'échantillonnage des stimuli^[7]. De plus, les stimuli – même s'ils constituent un tout homogène – peuvent être décomposés en deux parties qu'il est utile de distinguer : leur contenu et leur forme. Le contenu d'un stimulus fait référence à son sens, au-delà d'une mise en forme particulière. La forme d'un stimulus fait référence à la manière dont il est présenté au sujet ; plus généralement, il dépend du dispositif de collecte des informations qui a été mis en place. Lorsque l'on parle d'échantillonnage des stimuli, c'est surtout de l'échantillonnage des items en termes de contenu dont on parle. Mais le choix du dispositif d'observation ne doit pas être négligé pour autant : des dispositifs différents ne fourniront pas forcément la même mesure^[8].

Le problème de l'échantillonnage des items n'a jamais reçu de solution satisfaisante. On regrettera même, avec Reuchlin (1962, p. 138) qu'il est "remarquable que si peu d'importance ait été accordée jusqu'ici à ce problème d'échantillonnage des variables". Qu'en est-il aujourd'hui ? La situation n'a guère évolué même si on peut noter l'existence de quelques tentatives visant à aborder de front l'échantillonnage des items : c'est le cas des mesures à référence critérielle et, surtout, de l'approche en facettes que nous détaillerons plus loin. Mais, pour comprendre les difficultés posées par ce problème, il faut l'explicitier davantage.

On peut partir de l'idée simple selon laquelle, pour un concept donné, certains items sont pertinents et d'autres ne le sont pas. L'ensemble de tous les items pertinents par

[7] : Ni d'ailleurs celui de l'échantillonnage des occasions temporelles ; ce problème est généralement contourné, sur le plan empirique, en étudiant la stabilité des réponses à des moments différents.

[8] : Pour un exemple dans le domaine de la mesure des compétences lexicales, cf. Flieller, Delesmont & Thiébaud (1992).

rapport au concept mesuré constitue l'univers d'items. Cet univers est, dans la grande majorité des cas et à la différence de l'univers des individus, inconnu ; le plus souvent, il est impossible d'en lister les éléments qui sont trop nombreux. Pour se fixer les idées, il n'est pas inutile de penser à un concept psychologique qui a marqué l'histoire de la mesure, celui d'intelligence. Pour mesurer l'intelligence, les psychologues observent les réponses des sujets à des stimuli variés qui sont tous censés appartenir à l'univers des items d'intelligence. Mais considérons les items suivants : "Reproduire une figure avec 4 cubes formés de faces unicolores et de faces bicolores" ; "Connaître la date de la bataille de Marignan" ; "Réparer le pneu crevé d'une bicyclette" ; "Repérer, dans un groupe, des signes de tension" ; "Récupérer une clé tombée dans une bouche d'égoût". Font-ils partie ou non de cet univers ? Cela dépend, répondra-t-on, de la définition que l'on donne de l'intelligence. Sans doute, mais dans ce cas, une définition de l'intelligence devrait être suffisamment précise pour permettre de décider, sans ambiguïté possible, que tel item appartient à l'univers des items d'intelligence et que tel autre n'en fait pas partie. Autrement dit, existe-t-il, pour les concepts psychologiques, des définitions qui permettent de spécifier, de manière univoque, le domaine (i.e. l'univers des items) couvert par un concept ? Cette question fait l'objet d'un débat dans les travaux sur la catégorisation. Même si ces travaux n'ont pas de lien direct avec la problématique de la mesure, ils sont utiles pour notre propos. Ce débat oppose les partisans d'une approche classique des concepts psychologiques à ceux qui défendent l'idée d'une approche prototypique de ces concepts (cf. Mervis & Rosch, 1981 et, pour un exemple appliqué aux émotions, Clore & Ortony, 1991 ; Fehr & Russell, 1991 ; Russell, 1991). Dans l'approche classique, les membres d'une catégorie possèdent des caractéristiques nécessaires et suffisantes pour décider de leur appartenance à la catégorie. L'exemple traditionnel est la catégorie des nombres impairs qui peut facilement être définie en ces termes. Dans l'approche prototypique, on remet en cause la possibilité de telles caractérisations et on considère qu'aucun concept psychologique ne peut être défini de manière classique : l'appartenance d'un élément à une catégorie est affaire de degré (c'est une probabilité d'appartenance) et les frontières entre les concepts deviennent des frontières floues.

Quel est le lien entre ces travaux sur la catégorisation et la mesure ? Il nous semble que l'approche prototypique doit nous conduire à adopter un regard critique vis-à-vis de la notion d'univers d'items. S'il est impossible de définir un concept psychologique en termes de caractéristiques nécessaires et suffisantes, il faut abandonner l'idée de pouvoir déterminer, de manière définitive et absolue, l'univers de contenu en relation avec un concept. Mais cela ne signifie pas qu'on ne puisse essayer de le spécifier avec une démarche progressive dont les étapes se prêtent à la vérification empirique. De telles pratiques ne sont pas rares. C'est le cas par exemple de celle qui consiste à éliminer d'un instrument de mesure les items qui sont faiblement corrélés avec les autres. Mais ces pratiques sont généralement peu satisfaisantes. Elles résultent de décisions *a posteriori*

qui ne sont pas justifiées sur le plan théorique et qui dépendent du choix des autres items de l'instrument. Aurait-on éliminé le même item s'il avait été inclus dans un autre ensemble d'items administré à un autre échantillon de sujets ? Rien n'est moins sûr.

Plus intéressante est, selon nous, une démarche qui s'attache à déterminer, avec toutes les méthodes dont dispose le psychologue, quelles sont, parmi un ensemble d'items, les sources de variation qui entraînent des différences dans les réponses des individus. C'est grâce à l'identification de ces sources de variation et à leur combinaison que l'univers de contenu peut être spécifié, défini. Si on accepte cette conception, la définition d'un concept n'est plus un simple énoncé verbal vague et arbitraire. Une définition devient une proposition théorique qui spécifie les sources de variation des réponses et cette définition est susceptible d'être critiquée, réfutée, modifiée... ou abandonnée. Ce qui est important c'est que, par le biais d'une telle définition, les relations entre les éléments du concept sont explicitées. C'est en ce sens que la mesure n'est "qu'une forme d'expérimentation visant à mettre à l'épreuve des faits des hypothèses structurales sur les relations entre les éléments du concept" (Dickes et al., 1994, p. 60). Bien évidemment, toute proposition théorique définitoire d'un concept s'appuie sur des résultats issus de l'application de modèles de mesure dont la fonction principale est de tester la validité des hypothèses concernant les relations entre éléments. Il est tout aussi essentiel pour une définition de tenir compte des modèles théoriques qui intègrent ce concept dans des réseaux nomologiques, car, ce faisant, ils en précisent la signification.

Formulée autrement, cette conception revient à dire que l'on contribue à définir un concept en même temps qu'on le mesure. Autrement dit, s'il serait exagéré de prétendre que définition et mesure sont complètement confondues, il n'en est pas moins vrai qu'elles partagent des liens extrêmement étroits. Parmi les auteurs qui se sont le plus intéressés à cette question, Roskam (cf. en particulier Roskam, 1989) défend un point de vue quelque peu divergent. Selon lui, il faut distinguer définition et mesure, la première précédant la seconde car elle en spécifie le contenu. Il n'y aurait donc ni bonne ni mauvaise définition, mais seulement des définitions heuristiques. Mais, pour nous, toute définition, même si elle ne précise que le contenu, est déjà une hypothèse car elle doit spécifier l'univers des items. En effet, et pour reprendre la terminologie mathématique, mis à part les cas — très rares — où il est possible de définir un univers en extension^[9], il n'y a pas d'autre solution que de le définir en compréhension. Pour cela, il faut spécifier des facettes qui organisent cet univers et dont les combinaisons permettent de générer les items possibles. Or, s'il est vrai que l'hypothèse de la couverture exhaustive du domaine ne peut donner lieu qu'à des vérifications empiriques indirectes, l'hypothèse selon laquelle les facettes qui sont retenues dans la définition organisent l'univers est une hypothèse que l'on peut vérifier directement : il s'agit de montrer que ces facettes sont des sources de

[9] : e.g. les 1296 configurations à quatre faces de type cubes de Kohs ; les quatre opérations élémentaires appliquées aux cent premiers nombres entiers, etc.

variation des réponses. Dans ce cas de figure, on peut trouver des critères qui permettent de confronter des définitions différentes, notamment lorsqu'elles spécifient le même univers d'items. La meilleure définition sera, selon les cas, celle qui prend en compte les facettes qui entraînent les variations les plus importantes dans les réponses, celle qui spécifie le plus petit nombre de facettes (critère d'économie) ou encore celle dont les facettes ont déjà été mises en évidence pour d'autres concepts.

Ces longs commentaires sur la seconde étape du processus de mesure étaient indispensables. Les problèmes de définition, d'univers d'items sont au centre de la mesure et de la démarche scientifique et nous ne manquerons pas de revenir sur cette question. L'absence de base définitionnelle stable est, en effet et comme l'écrit Shye (1978 a), l'un des obstacles principaux à la réplication des résultats dans les sciences du comportement. Ce problème est particulièrement marqué dans le domaine de la personnalité ainsi que le rappelle Fiske (1986) pour qui tant qu'un concept est mal défini, on ne peut jamais montrer qu'il est inadéquat : c'est sans doute, selon cet auteur, la raison qui explique que tant de traits de personnalité ont réussi à survivre malgré des dizaines d'années de recherche qui n'ont apporté, le plus souvent, que des résultats équivoques.

Les autres aspects de cette seconde étape sont des aspects plus techniques : ils concernent le *choix d'un échantillon de sujets*, le *choix d'un dispositif de collecte* des observations et les problèmes posés par *l'interaction entre les sujets et le dispositif*. Mieux connus, ces aspects ne seront pas développés ici^[10]. Nous rappellerons simplement qu'il s'agit de ne pas en négliger l'importance car la possibilité de généraliser les résultats dépend des choix faits à ce stade : si cela est évident en ce qui concerne les sujets, on a peut-être encore trop tendance à oublier que c'est tout aussi vrai en ce qui concerne le dispositif de collecte des observations. Il nous faut encore souligner que le choix d'un dispositif d'observation participe aussi, le plus souvent implicitement, à la définition d'un concept. C'est le cas lorsque le dispositif introduit des contraintes rendant irréfutables certaines hypothèses quant à la structure du concept. L'exemple le plus classique est celui de l'unidimensionnalité qui est imposée *a priori* dans le dispositif.

La troisième étape du processus de mesure concerne la *transformation des observations en données*. Alors que les observations sont, rappelons-le, les réponses données par des individus à des stimuli à un certain moment, les données constituent les informations qui vont être analysées par le modèle de mesure. Il est courant que les modèles de mesure requièrent des calculs sur les observations : le calcul de coefficients de corrélation ou, plus généralement, de coefficients d'association en est un exemple familier. Mais c'est à Coombs (1964) que revient le mérite d'avoir souligné que, même en l'absence de calculs, les observations subissent une transformation qui en change le

[10] : Nous renvoyons le lecteur à la présentation synthétique que nous en avons faite (Dickes et al., 1994, pp. 78-122).

statut. C'est une chose que d'observer qu'un individu fournit telle réponse à un énoncé en faveur de la politique économique du gouvernement. C'est est une autre que de considérer que cette réponse doit être interprétée comme un indicateur de l'attitude favorable ou défavorable du sujet en faveur de cette politique. Cette distinction entre observation et donnée pourra paraître subtile. Elle est pourtant bien réelle au sens où le passage de l'une à l'autre implique un choix théorique dont l'évaluation de la pertinence doit s'inscrire dans le cadre plus général de l'ensemble du processus de mesure, aussi bien en aval qu'en amont (cf. Coombs, 1964).

C'est dans la quatrième étape du processus que se situe l'*application du modèle de mesure*. Les modèles de mesure sont trop nombreux pour qu'on puisse les décrire ici. Il est plus intéressant d'essayer de repérer, sous leur diversité, leurs caractéristiques communes. Certes, nous avons conscience de tout ce que cet exercice comporte comme pièges, étant donné le manque d'unité et l'absence de conception générale suffisamment aboutie sur cette question. Paradoxalement, alors qu'un consensus relatif existe sur le fait que tel modèle est un modèle de mesure, il n'en existe pas de définition précise. On peut s'accorder sur le fait qu'il s'agit de la fonction qui met en correspondance les relations entre les données avec celles du système formel qui les représente, mais les caractéristiques de cette fonction, qui pourraient servir à élaborer une classification des modèles, restent vagues. Plusieurs travaux montrent les liens qui unissent des modèles considérés jusqu'alors comme différents^[11] : c'est grâce à de tels travaux que l'on pourra comprendre la logique fondamentale qui sous-tend l'utilisation de tout modèle de mesure en en précisant les spécificités à travers leurs postulats, les hypothèses testées, la manière dont les hypothèses sont mises à l'épreuve des faits ou encore le type de relations entre variables latentes et variables manifestes. Cet effort d'abstraction et de formalisation est nécessaire pour permettre de clarifier des pratiques dont la logique échappe souvent même aux utilisateurs les plus avertis.

Si la mesure consiste à fournir une représentation des relations observées dans le système empirique, l'une des fonctions d'un modèle de mesure est d'explicitier les hypothèses sous-jacentes à la mise en correspondance du système formel et du système empirique. Toute transformation de données en mesure repose sur de telles hypothèses, qui sont, dans les modèles les plus frustes, de simples postulats acceptés comme vrais sans possibilité de réfutation. Bien évidemment, la mesure sera d'autant plus satisfaisante que le modèle s'appuiera sur peu de postulats et sur plusieurs hypothèses clairement formulées pouvant être réfutées ; il faut pourtant admettre que la distinction entre les postulats et les hypothèses d'un modèle n'est pas toujours très nette.

[11] : C'est le cas notamment des liens entre analyse factorielle des correspondances, analyse en classes latentes, modèle de Rasch, modèle du jugement comparatif, modèle hiérarchique de Guttman, modèle du dépliage, analyse log-linéaire (mais est-ce un modèle de mesure ?) qui ont été soulignés, sous des formes diverses, par des auteurs tels que Jansen (1984), McCutcheon (1987), Andrich (1988) ou Langeheine & Rost (1988).

La possibilité de vérifier que les données sont compatibles avec les hypothèses à l'aide de tests d'adéquation est l'une des caractéristiques les plus importantes des modèles de mesure car cette possibilité assure à la démarche de mesure le statut de démarche scientifique. La logique de ces tests d'adéquation repose sur la confrontation entre le système empirique et sa représentation formelle. Ce n'est que lorsque les écarts entre les données observées et les données reconstituées à partir des valeurs qui leur sont attribuées dans le système formel sont suffisamment faibles que la mesure peut être acceptée. Autrement dit, lorsque les écarts sont peu importants entre données observées et données estimées, on a trouvé un système formel dans lequel les relations entre les nombres peuvent être considérées comme une représentation correcte des relations entre les objets empiriques.

Le modèle de mesure est appliqué à des données en provenance d'observations. Celles-ci ne se justifient que par le fait que le concept que l'on veut mesurer n'est pas accessible directement : on n'observe pas l'intelligence ou l'extraversion mais leurs "manifestations". Les observations prennent donc le statut d'indicateurs manifestes d'un concept qui, lui, est latent, inobservable. Le modèle de mesure doit donc aussi spécifier les relations entre ces variables manifestes et cette variable latente. Ces relations diffèrent notamment selon qu'elles sont déterministes ou probabilistes. Dans le premier cas, la variable latente détermine totalement la variable manifeste ; dans le second, la variable latente ne permet de prédire la variable manifeste qu'avec une certaine probabilité.

Mais ces relations ne diffèrent pas seulement sur ce point. Dans des constructions théoriques comme celle de Lazarsfeld (e.g. Lazarsfeld, 1959), c'est la spécification de ces relations qui occupe une position centrale. Préciser sous forme mathématique des hypothèses concernant les relations entre variables manifestes et variables latentes est, en effet, l'objectif principal de l'analyse en structure latente présentée par cet auteur. Il faut souligner ici que l'analyse de structure latente ne doit pas être assimilée à un modèle de mesure. C'est une conceptualisation très générale qui, à partir du seul axiome d'indépendance locale, permet de dériver des modèles spécifiques qui eux, nécessitent des postulats supplémentaires. Si nous évoquons l'analyse de structure latente ici, c'est que nous pensons qu'elle peut être ce schéma général dont nous regrettons l'absence plus haut. Elle dispose pour cela de plusieurs atouts. Tout d'abord, elle repose sur un schéma de départ simple et explicite qui permet des développements totalement formalisés, sans nuire pour autant à la richesse des modèles qui en découlent. De plus, certains modèles classiques semblent pouvoir être dérivés directement de l'analyse de structure latente : dès 1952, Green montrait que l'analyse factorielle pouvait être assimilée à un cas particulier de l'analyse de structure latente ; depuis, on a mis en évidence les similitudes entre l'analyse en classes latentes (qui dérive directement de l'analyse de structure latente), le modèle de Rasch et l'analyse factorielle des correspondances (cf. Langeheine & Rost, 1988). Si de telles filiations se confirment avec d'autres modèles de mesure, on pourra

clarifier la distinction entre les postulats et les hypothèses d'un modèle, les postulats seront les axiomes nécessaires aux dérivations algébriques et mathématiques sous-tendant la transformation des données, les hypothèses porteront sur les relations entre variables manifestes et variables latentes. Une telle clarification serait, à nos yeux, une contribution majeure à la problématique de la mesure.

1.3 Conclusion

Dans ce chapitre, dont l'objectif était d'exposer notre conception générale de la mesure, nous avons insisté, dans un premier temps, sur le fait que mesurer était une activité de représentation, car le but de toute mesure est d'établir une correspondance, un homomorphisme entre des objets du système empirique et les symboles — des nombres en général — d'un système formel. Cet homomorphisme ne prend sens que par rapport à une caractéristique spécifique des objets, par rapport à un concept. Les théories de la mesure cherchent à établir les conditions — les axiomes — grâce auxquelles il est possible d'établir cette correspondance. Le système d'axiomes développé dans le cadre de la théorie représentationnelle de la mesure permet de donner une justification théorique et logique à la mesure dès lors que celle-ci s'inscrit dans cette optique.

Dans un deuxième temps, nous avons souligné que la mesure ne se distinguait pas de la démarche scientifique en général, qui consiste à vérifier empiriquement des hypothèses qui découlent de modèles théoriques. Lorsque des nombres sont attribués à des objets, on peut, en appliquant les règles du système formel choisi, faire des hypothèses quant aux relations entre les objets (e.g. si le nombre attribué à l'objet A est plus petit que le nombre assigné à l'objet B et si le nombre attribué à l'objet B est plus petit que le nombre attribué à l'objet C, on peut faire l'hypothèse que, dans le système empirique et en fonction de la caractéristique mesurée, A est plus petit que C). Ces hypothèses n'exploitent pas forcément toute la puissance des propriétés du système numérique et en particulier les règles de l'addition. Du nombre de règles qu'il est possible d'appliquer aux éléments du système formel dépend le niveau de la mesure et non pas sa qualité. Si la mesure est une démarche scientifique comme une autre, elle en a les avantages, mais aussi les contraintes : les hypothèses qui la sous-tendent doivent être réfutables et donner lieu à des vérifications empiriques.

Dans un troisième temps, enfin, nous avons décrit la mesure comme un processus composé de quatre étapes : la définition du concept à mesurer, la collecte des observations, la transformation des observations en données et l'application d'un modèle de mesure. Nous avons longuement insisté sur le passage entre la définition du concept et la collecte des observations qui est l'une des étapes les plus délicates de ce processus. La définition doit permettre de circonscrire l'univers des observations dont une petite

partie seulement sera collectée. Comme la mesure n'a pas pour objectif d'explicitier les relations entre des observations factuelles, mais de rendre compte des relations entre l'ensemble des éléments de cet univers, il est essentiel de pouvoir généraliser les résultats et de savoir jusqu'où cette généralisation est légitime. Pour résoudre le problème posé par le caractère souvent arbitraire et spécifique des définitions de concepts psychologique, nous avons proposé que l'étape définitoire fasse l'objet de vérifications empiriques. En spécifiant les sources de variation des réponses, une définition devient une proposition théorique susceptible d'être réfutée. L'univers des observations est alors complètement déterminé à partir de la combinaison de toutes les sources de variation qui ont été mises en évidence. Seule une telle démarche permet, à notre avis, de répondre de manière empirique aux controverses définitoires dont le caractère stérile a été maintes fois souligné à propos des concepts psychologiques. La question des modèles de mesure n'est pas non plus un problème simple. Si nous avons pu mettre en relief quelques-unes de leurs caractéristiques, il est vrai qu'il manque encore à l'heure actuelle une conceptualisation générale satisfaisante de ces modèles, qui ferait apparaître leurs similitudes et leurs différences. On remarquera d'ailleurs à ce sujet le hiatus qui existe entre les modèles de mesure classiques (e.g. analyses factorielles, modèle du score vrai, etc.) et les modèles développés par les théoriciens de la mesure (e.g. modèle de la mesure conjointe). Si, dans le dernier cas, l'axiomatisation de ces modèles est bien connue, il n'en est pas de même pour les premiers. Nous avons suggéré que l'analyse de structure latente de Lazarsfeld (1959) pouvait être le point de départ d'une telle conceptualisation générale, mais celle-ci reste à faire.

Les conceptions exposées dans ce chapitre sont principalement celles que nous devons aux travaux théoriques sur la mesure. Mais la mesure a également été envisagée de manière plus pratique, notamment à travers les travaux sur les tests psychologiques. C'est à ce courant, que nous qualifions de psychométrie pour le distinguer du courant théorique, que nous allons à présent nous intéresser. La présentation s'appuiera sur la notion de validité qui en constitue le concept central. Et nous tenterons de faire apparaître les liens potentiels entre ces deux courants qui, bien que leur objet soit commun, s'ignorent encore largement (cf. Cliff, 1992).

CHAPITRE II

VALIDITES ET VALIDATION

2.1 Introduction

La définition de la validité est presque aussi célèbre que la définition classique de la mesure... mais tout aussi insatisfaisante. Dire qu'un instrument est valide dès lors qu'il mesure bien ce qu'il est censé mesurer met en lumière l'enjeu de ce concept, mais n'en précise guère la signification. Pour Kane (1982), cette définition implique que l'on suppose l'existence d'une "valeur vraie" pour toute mesure, mais la difficulté est de spécifier ce que cette valeur vraie représente. On a plutôt tendance à dire aujourd'hui (e.g. Messick, 1989) que la validité consiste à rechercher quelle est la signification d'une mesure et à en donner une interprétation.

Le concept de validité est apparu dans la théorie des tests en psychologie et, aujourd'hui encore, c'est dans cette littérature que se trouvent les discussions les plus abouties qui s'y rapportent (cf. Messick, 1989 ; Anastasi, 1990 ; Cronbach, 1990 ; Moss, 1992 ; Shepard, 1993). On retrouve pourtant la notion de validité dans la démarche expérimentale (cf. Cook & Campbell, 1979) et, plus généralement, dans toutes les sciences du comportement (cf. Kirchgässler, 1991). Certains ont même tenté, dans un effort louable mais très abstrait, de le généraliser à l'ensemble du processus de recherche (cf. Brinberg & McGrath, 1985) : cette tentative ne fait qu'illustrer un point très important sur lequel nous reviendrons, à savoir que la recherche de validité ne diffère pas sensiblement de la démarche scientifique en général.

L'existence de la notion de validité dans les sciences physiques et biologiques est plus controversée. Pour Kane (1982), par exemple, la validité n'y a pas d'équivalent, car elle est directement établie à travers l'opération de mesure. Pourtant, on retrouve dans certains écrits concernant ces sciences l'idée d'une confrontation entre les résultats d'une mesure et une valeur vraie. Dans son texte introductif à une célèbre conférence sur *l'histoire de la signification de la mesure dans les sciences naturelles et sociales*, qui s'est tenue en 1959 à l'université du Wisconsin, Wilks (1961) insiste notamment sur le caractère indirect du processus de validation. Comme il n'existe pas, le plus souvent, de procédure indépendante permettant de déterminer la valeur vraie, la validité reste un problème qui ne peut être résolu de manière directe. Il illustre ses propos à travers l'exemple de la mesure de la vitesse de la lumière réalisée par Michelson en 1935. Selon lui, la validité de la procédure mise en oeuvre à cette occasion repose sur le fait que les

estimations auxquelles Michelson aboutit ne conduisent à aucun résultat contradictoire ou inconsistant avec d'autres phénomènes physiques connus. Ce serait donc bien grâce à une procédure indirecte que la validité de cette mesure serait établie. On verra que cette démarche ne diffère pas sensiblement de celles qui sont utilisées en psychologie.

Ce chapitre présentera tout d'abord l'évolution historique des conceptions de la validité : la validité, d'abord conçue comme une notion unitaire, s'est subdivisée en différents types de validité, la tendance actuelle étant de les intégrer dans un processus unique appelé validation. Il exposera ensuite une conception de la validation qui distingue une validation intraconcept et une validation interconcepts. C'est dans le cadre de cette discussion qu'apparaîtra le plus nettement le hiatus entre la tradition psychométrique de la validité et la tradition théorique de la mesure. Cette discussion aura aussi comme objectif de commencer à combler le fossé qui existe entre ces deux traditions.

2.2 La notion de validité à travers son évolution historique

On peut distinguer trois étapes dans l'évolution du concept de validité. Dans la première, la validité est synonyme de valeur pronostique d'un instrument de mesure. Dans la seconde, la validité se scinde en validités spécifiques ayant chacune un sens bien particulier ; la valeur pronostique abandonne sa prééminence et devient un type de validité parmi d'autres. Enfin, la validation devient un concept unitaire qui intègre ces différents types de validité. Parler de validation, c'est mettre l'accent sur un processus de recherche continu dont l'objectif est d'établir la signification des scores obtenus grâce à un instrument de mesure : "La validation d'un test n'est rien d'autre, par essence, qu'une recherche scientifique de la signification des scores, rien de plus, mais également rien de moins." (Messick, 1989, p. 56). Parallèlement à cette évolution, on insiste sur le fait que l'on ne valide plus un instrument de mesure, mais les réponses à cet instrument (Cronbach, 1971 ; Messick, 1975) : en effet, les réponses à des stimuli sont non seulement fonction des caractéristiques de ces stimuli, mais aussi des caractéristiques des sujets et du contexte dans lequel l'instrument est utilisé.

2.2.1 La validité comme valeur pronostique

Une phrase de Guilford (1954, p. 398) illustre parfaitement la conception de la validité comme valeur pronostique : "un test est valide dès lors qu'il corrèle avec quelque chose". C'est cette conception que l'on retrouve encore dans la sixième édition du vocabulaire de la psychologie de Piéron (1979) : le coefficient de validité d'un test est le coefficient de corrélation entre la performance des sujets au test et leur performance dans

une activité que le test est censé prévoir. Seule compte dans cette conception la possibilité de prévoir la performance d'un sujet, quel que soit l'instrument utilisé. Ainsi, rien n'empêche de considérer, si on accepte cette conception, que la longueur des cheveux est une mesure valide si elle permet de prévoir la réussite scolaire.

Cette conception conduit effectivement à des difficultés logiques. Lorsque l'on veut mettre en relation les réponses à un instrument de mesure avec un critère à prédire, il faut également mesurer ce critère et s'assurer de sa validité. La seule possibilité de sortir de ce cercle vicieux est d'admettre que certaines mesures sont valides par essence, ce qui n'est guère satisfaisant, en dehors de situations pour lesquelles il est possible de déterminer un critère bien défini, dont la mesure immédiate s'apparente à une mesure opérationnelle, sa signification faisant l'objet d'un consensus. De telles situations sont rares, notamment en dehors du domaine restreint des tests de performance.

2.2.2. Les trois grands types de validité

Cette conception unitaire s'est avérée peu à peu insuffisante pour apprécier la qualité d'une mesure. D'autres stratégies ont été proposées à tel point que l'on trouve, au début des années 80, plus de 35 types de validité (selon la recension de Brown, 1980, citée par Lynn, 1986). Cette seconde période est marquée par la prise de conscience que des instruments de mesure différents n'ont pas forcément les mêmes fonctions. Ces diverses fonctions appellent des techniques différentes de validation ; autrement dit, différents types de validité sont nécessaires.

Mais étant trop spécifiques, la plupart des types de validité ont eu une existence éphémère. Seules la validité critérielle, la validité de contenu et la validité de construit se sont imposées.

A. La validité critérielle

La validité critérielle reprend et complète la conception de la validité comme valeur pronostique. Elle la reprend parce qu'elle met en avant la possibilité de prédire un critère externe à partir d'un instrument de mesure. Elle la complète parce qu'elle propose une distinction entre validité prédictive et validité concourante.

La validité prédictive et la validité concourante ne diffèrent que sur un point. Dans le cas de la validité prédictive, on cherche à prédire un critère ultérieur qui n'est pas disponible au moment où s'effectue la mesure ; elle est utile lorsque l'on mesure des caractéristiques d'individus en vue d'une sélection. Dans le cas de la validité concourante,

on cherche à prédire un critère dont on connaît l'existence au moment où la caractéristique est mesurée : ce type de validité est particulièrement utile lorsque les instruments de mesure ont une fonction diagnostique et lorsqu'ils sont censés se substituer à d'autres techniques diagnostiques plus coûteuses.

Il n'y a pas de rupture entre la validité critérielle et la première conception de la validité que nous avons vue. La démarche reste exclusivement empirique. Le changement réside dans le fait que ce n'est plus la seule stratégie pour établir la validité d'une mesure.

B. La validité de contenu

Avec la notion de validité de contenu se trouve affirmée l'importance de la signification des mesures. Un instrument de mesure ne doit plus seulement permettre de prédire un critère, il doit aussi être composé d'items qui sont à la fois pertinents et représentatifs par rapport à l'ensemble des items du domaine défini par le concept que l'on veut mesurer. Alors que la validité critérielle est essentiellement empirique, la validité de contenu repose sur une analyse théorique du domaine.

La validité de contenu s'est toujours heurtée à un problème majeur. Pour vérifier la pertinence et la représentativité des items qui composent un instrument de mesure, il faut connaître l'univers des items. Même dans les cas les plus simples, la validité de contenu est difficile à apprécier. Que l'on songe, par exemple, à un enseignant qui élabore un test destiné à évaluer les connaissances qu'ont ses élèves d'un cours qu'il a donné. L'enseignant a une connaissance parfaite du domaine sur lequel porte son évaluation. Pourtant, si sa conscience professionnelle le pousse à s'interroger sur la validité de contenu de son test, il est peu probable qu'il soit à même de répondre à des questions du type : les items retenus pour l'examen échantillonnent-ils correctement tous les aspects qui ont été abordés dans le cours ? N'y a-t-il pas certaines questions qui font appel à des connaissances qui n'ont pas été abordées directement dans le cours ?, etc.

C'est dans le domaine éducationnel pourtant que l'on trouve une approche originale pour aborder la question de l'univers des items. Le courant des tests à référence critérielle (cf. Hambleton, 1990) s'est développé en partie avec l'intention de "construire des instruments dont les items dérivent de manière explicite du domaine défini par rapport aux objectifs pédagogiques" (Dickes et al., 1994, p. 68). C'est en fonction de cet objectif qu'ont été mises au point des techniques permettant de générer, de manière systématique, tous les items d'un univers. On s'aperçoit malheureusement que ces techniques, qui sont très puissantes lorsque le domaine est bien délimité, n'ont qu'une utilité restreinte dans le cas de la plupart des concepts psychologiques. Dans ce cas, la théorie en facettes (cf. Canter, 1985) est une autre approche dans laquelle l'univers des items est défini de

manière explicite. Plus générale que les tests à référence critérielle, elle connaît des applications dans le domaine des attitudes, de l'intelligence ou de la personnalité. Elle sera présentée en détail dans le prochain chapitre.

Le problème central de la validité de contenu réside donc dans la définition du concept et dans la spécification de l'univers des items. Les deux techniques que nous venons d'évoquer ont le mérite de le faire apparaître clairement. Elles suggèrent toutes deux qu'un domaine n'est pas défini par un listage des éléments de l'univers, mais par la spécification des caractéristiques de ses éléments. Ces caractéristiques sont essentiellement formelles dans le cas des tests à référence critérielle, elles ont une portée théorique beaucoup plus importante dans la théorie en facettes qui les considère comme des sources de variation des réponses et qui essaie de préciser comment elles sont reliées entre elles. Ces conceptions rejoignent celles que nous avons présentées à propos de la définition du concept dans le processus de mesure.

Lorsqu'elle n'est pas envisagée en termes de démarche hypothético-déductive visant à établir la validité d'une proposition théorique définitoire, la vérification de la validité de contenu d'une mesure est quasi impossible. On trouve, dans la littérature, deux grandes stratégies pour tenter de la vérifier, aucune d'elles n'étant pourtant satisfaisante.

La première stratégie consiste à faire évaluer, par des personnes choisies pour leur connaissance du domaine (i.e. des "experts"), les items de l'instrument en fonction de leur pertinence et de leur représentativité (cf. Lawshe, 1975 ; Lynn, 1986). En général, aucune définition du domaine n'est fournie aux experts. En fait, cette procédure permet surtout de repérer des items perçus comme ne faisant pas partie du domaine investigué. C'est à nouveau dans le domaine éducationnel que cette technique a connu le plus d'applications, pour des raisons qui s'expliquent assez facilement (les domaines peuvent être définis assez clairement, les experts potentiels sont nombreux, etc.). Mais il est bien évident que cette démarche ne donne pas de garantie sur la couverture exhaustive du domaine.

La seconde stratégie utilise les résultats de l'application d'un modèle de mesure. Différentes méthodes sont utilisées (analyses factorielles, modèles de réponse à l'item, etc.), mais elles reposent toutes sur la même logique : la recherche de l'homogénéité des contenus. La méthode la plus simple est l'analyse d'items, qui découle de l'application du modèle du score vrai. Cette stratégie consiste à examiner les corrélations entre les réponses à un stimulus (à un item) et les mesures obtenues à l'ensemble de l'instrument. Lorsque ces corrélations sont faibles ou négatives, on décide que les items en question ne font pas partie du domaine^[1]. L'objectif principal de cette démarche, tout comme celui

[1] : Cette technique est très connue ; elle est présentée ici de manière très schématique, donc caricaturale, afin de mettre en relief ses liens avec la problématique de la validité de contenu. Cette prise de position n'enlève rien à l'utilité générale (à suivre...)

de la précédente, est de détecter les items inadéquats par rapport au domaine, et non pas d'assurer la validité de contenu. Cette démarche soulève de plus une difficulté clairement soulignée par Reuchlin (1962). En utilisant les résultats de l'application d'un modèle de mesure, on raisonne – implicitement – de la manière suivante : les items non conformes au modèle de mesure ne font pas partie de l'univers des items. Cette pratique est tellement répandue qu'elle ne choque plus personne. Pourtant, il paraîtrait tout à fait inconvenant qu'un chercheur décide, *a posteriori*, que des individus ne fassent pas partie de la population car ils ne se conforment pas au modèle théorique testé. Une telle stratégie est tout aussi contestable dans le cas des items que dans celui des individus.

Compte tenu de ces difficultés, il n'est pas étonnant que la validité de contenu ait été discutée, certains allant même jusqu'à remettre en cause l'utilité de cette notion. En fait, trois grandes thèses sont présentes dans la littérature psychométrique (cf. Fitzpatrick, 1983). Pour certains, la validité de contenu est suffisante pour assurer la validité d'une mesure, pour d'autres, c'est une condition nécessaire mais non suffisante, d'autres enfin estiment que la validité de contenu ne concerne pas la validation.

Les deux premières thèses ont tendance à être minoritaires aujourd'hui, même si plusieurs auteurs continuent à en souligner l'importance. C'est le cas, par exemple de Robinson, Shaver & Wrightsman (1991) qui, dans leur ouvrage de synthèse sur les mesures d'attitudes et de personnalité, indiquent (p. 6) que le premier critère qu'ils ont utilisé pour choisir les meilleures échelles est un échantillonnage approprié du contenu (même s'ils notent qu'il s'agit d'une tâche très difficile)^[2].

Nombreux sont les auteurs qui considèrent que la validité de contenu n'est pas un concept pertinent pour évaluer la validité des mesures. La plupart d'entre eux (e.g. Tenopyr, 1977 ; Yalow & Popham, 1983 ; Shepard, 1993) s'appuient sur l'idée selon laquelle la validité de contenu n'est qu'une analyse conceptuelle des items d'un instrument de mesure et qu'elle n'a donc rien à voir avec la recherche de la signification des mesures (sic !) qui constitue le principal objectif de la validation, selon les conceptions actuelles (nous reviendrons plus bas sur ce point). Ainsi, on aboutit à une situation paradoxale dans laquelle la signification d'une mesure dépend des corrélations entre celle-ci et d'autres mesures, mais pas du contenu de l'instrument qui a permis de l'obtenir ! Même des auteurs comme Anastasi (1990 ; cf. également Kirchgässler, 1991, qui défend une position similaire pour l'ensemble des sciences sociales) sont réticents à l'égard de la validité de

[1] : (...suite)

de cette pratique pour la construction d'échelles unidimensionnelles, ce qui constitue, souvent, l'objectif immédiat des chercheurs.

[2] : Voir également à ce sujet le débat entre Nicholls, Pearl & Licht (1982, 1983), Spence & Helmreich (1983) et Friedman (1983) qui reconnaissent, au-delà de leurs divergences, que les problèmes de contenu occupent une place primordiale dans la recherche psychologique.

contenu, en particulier dans des domaines flous comme ceux de la personnalité ou des attitudes. De telles positions ne se comprennent que si la validité de contenu est limitée à une évaluation *a posteriori* de la représentativité et de la pertinence d'un ensemble d'items par rapport à un univers d'items restreint dont la définition fait l'objet d'un relatif consensus. Dans le cas de domaines plus flous, il est vrai que de telles procédures sont vouées à l'échec et que le chercheur qui éprouve un certain malaise devant la manière d'aborder le problème de la validité de contenu puisse être tenté de s'en remettre à d'autres stratégies qu'il connaît mieux (validité critérielle, validité de construit). Nous croyons qu'il est nécessaire de changer de perspective par rapport à la validité de contenu : en considérant que les problèmes de définition et de contenu doivent être résolus de la même manière que tout problème scientifique, c'est-à-dire en avançant des propositions théoriques définitoires réfutables qui s'appuient sur les sources de variation des réponses, des débats comme ceux que nous venons d'évoquer apparaîtront totalement vains.

C. La validité de construit

Le terme de validité de construit a été introduit pour la première fois dans un comité de réflexion sur les tests mis en place par l'Association Américaine de Psychologie au début des années cinquante. Lee Cronbach et Paul Meehl l'ont rendu célèbre dans un article paru en 1955^[3]. Les travaux psychométriques lui confèrent aujourd'hui une place centrale.

La validité de construit entendait, à l'origine, s'appliquer aux instruments de mesure qui se développaient en l'absence de critère (au sens de la validité critérielle) et en l'absence d'univers d'items (au sens de la validité de contenu). La majeure partie des mesures des traits de personnalité se trouvait dans ce cas. L'idée principale sous-jacente à la validité de construit est l'hypothèse de l'existence d'un construit psychologique inobservable qui rend compte des différences interindividuelles observées dans les réponses des individus à un instrument de mesure. Autrement dit, et comme l'écrivent Cronbach & Meehl (1955, p. 28), "la validité de construit est invoquée chaque fois qu'un test doit être interprété comme un attribut ou comme une qualité qui n'est pas définie de manière opérationnelle". Il est clair que l'originalité de cette validité est que la mesure du concept ou du construit psychologique est posée comme l'intérêt principal du processus de mesure. La mesure n'est plus un moyen indirect de connaître ou de prédire la performance dans une autre tâche ; ce qui importe ici, c'est la relation entre l'objet mesuré et le construit et l'interprétation de ce construit.

[3] : Cronbach explique en 1992 que l'idée de la validité de construit doit être attribuée à Paul Meehl. C'est sur l'insistance de ce dernier que l'ordre entre les deux auteurs a été déterminé par un tirage au sort... qui a été gagné par le premier.

Pour aborder cette validité, on a recours à des informations externes. Le construit va acquérir une signification en fonction du réseau d'association au sein duquel il se trouve. Cette façon de procéder justifie l'appellation "validité de construit" (ou *construct* en anglais). L'interprétation du concept hypothétique censé rendre compte des différences interindividuelles dans les réponses va se *construire* au fur et à mesure des relations qui sont mises en évidence entre ce concept et d'autres variables (i.e. son réseau nomologique).

On distingue, en général, les procédures de validation de ce type selon qu'elles conduisent à une "validation faible" ou à une "validation forte". On parle de "validation faible" lorsque l'on se contente de "collectionner"^[4] les variables corrélées avec le construit en question. Il s'agit ensuite de donner une signification à toutes ces relations, ce qui n'est pas toujours très simple. Dans la "validation forte", la démarche est hypothético-déductive. Les hypothèses portant sur le réseau nomologique découlent d'une théorie sur le construit et sont ensuite testées. Bien que cette distinction soit utile, il faut reconnaître que, dans la pratique, les procédures de validation sont souvent un mixte de ces deux approches.

Si la validité de construit indique dans quelle mesure un test mesure bien un construit ou un trait psychologique comme le souligne Anastasi (1990), il faut pouvoir montrer que ce construit est indépendant de la procédure de mesure utilisée et qu'il est différent d'autres construits déjà existants. Autrement dit, il faut pouvoir montrer que des mesures d'un même construit obtenues par des méthodes différentes conduisent à des résultats équivalents (validité convergente) et qu'un construit n'est pas trop fortement corrélé avec un autre construit (validité discriminante). C'est à Donald Campbell et Donald Fiske (1959) que l'on doit les notions de validité convergente^[5] et de validité discriminante. Parallèlement, ces auteurs ont proposé une méthode afin d'évaluer ces deux types de validité. Cette méthode fait appel à une matrice multitraits-multiméthodes (matrice MTMM). Comme nous y ferons référence à plusieurs reprises dans la suite du texte, nous allons la décrire brièvement.

[4] : Pour reprendre une expression de Landy (1986) qui assimile cette démarche à celle du philatéliste.

[5] : La validité convergente doit être comprise comme la vérification de la convergence des mesures avec des instruments différents. Elle a tendance à être réduite aujourd'hui au simple calcul d'une corrélation entre un questionnaire et un autre questionnaire. S'il s'agit bien de mesures différentes d'un même construit, il faut reconnaître que les différences ne concernent que le choix des items. Dans la conception de Campbell & Fiske, il y a tout intérêt à choisir des mesures qui diffèrent non seulement sur les items mais, plus généralement, sur l'ensemble du dispositif de collecte des observations.

Une matrice MTMM présente toutes les intercorrélations possibles obtenues lorsque plusieurs concepts (plusieurs traits) sont mesurés avec des méthodes différentes^[6]. Les applications les plus courantes impliquent trois concepts, mesurés chacun avec trois méthodes différentes. La figure suivante illustre une telle matrice qui présente toutes les corrélations possibles entre les neuf mesures.

	A ₁	B ₁	C ₁	A ₂	B ₂	C ₂	A ₃	B ₃	C ₃
A ₁	<u>.80</u>								
B ₁	.15	<u>.82</u>							
C ₁	.20	.80	<u>.85</u>						
A ₂	<u>.70</u>	.06	.04	<u>.90</u>					
B ₂	.03	<u>.72</u>	.64	.12	<u>.92</u>				
C ₂	.05	.67	<u>.75</u>	.10	.82	<u>.95</u>			
A ₃	<u>.75</u>	-.02	.02	<u>.72</u>	.06	.03	<u>.91</u>		
B ₃	-.01	<u>.68</u>	.61	.01	<u>.70</u>	.62	.11	<u>.92</u>	
C ₃	.00	.65	<u>.69</u>	-.07	.63	<u>.71</u>	.09	.81	<u>.89</u>

Figure 1.2 : Un exemple fictif de matrice multitraits-multiméthodes (X_i désigne la mesure du concept X avec la méthode i ; sur la diagonale principale se trouvent les corrélations des mesures avec elles-mêmes)

Pour vérifier la validité convergente et discriminante de ces mesures, il faut que les hypothèses suivantes soient confirmées :

- les corrélations entre les mesures d'un même trait, mesuré par des méthodes différentes, sont élevées (il s'agit des corrélations monotrait-hétérométhodes qui apparaissent soulignées avec un trait dans la figure) ;
- les corrélations entre les mesures d'un même trait par différentes méthodes sont supérieures aux corrélations entre les mesures de traits différents avec des méthodes différentes (les corrélations monotrait-hétérométhode, soulignées, sont supérieures aux corrélations hétéroatrait-hétérométhode, en italique) ;

[6] : La note précédente s'applique également ici. On constate trop fréquemment aujourd'hui des matrices MTMM construites à partir de mesures dont les méthodes diffèrent très peu. C'est le cas par exemple, lorsque les méthodes diffèrent uniquement quant au nombre d'échelons de l'échelle de réponse proposée aux sujets. De telles applications constituent, à notre avis, une dérive par rapport à l'esprit de cette technique qui requiert des méthodes les plus dissimilaires possibles. Lorsque les méthodes sont très semblables, la vérification de la validité convergente est triviale.

- les corrélations monotrait-hétérométhode (soulignées) sont supérieures aux corrélations hétérotrait-monométhode (e.g. A_1 avec B_1 ; B_2 avec C_2 , etc.) ;
- les corrélations entre les mesures de concepts différents doivent être du même ordre de grandeur, quelle que soit la méthode utilisée (cf. les corrélations en italique).

L'exemple fictif présenté dans la figure 1.2 permet d'émettre des doutes quant à la validité discriminante des concepts B et C ; pour le reste, les hypothèses semblent confirmées.

L'analyse des matrices MTMM à longtemps reposé sur des critères subjectifs. Cette situation a changé depuis quelques années grâce au développement des modèles généraux d'analyse des matrices de covariance tel le LISREL. Ces méthodes permettent de tester l'adéquation empirique d'un modèle théorique qui serait, en l'occurrence, un modèle factoriel avec trois facteurs-traites et trois facteurs de méthode (cf. Widaman, 1985 ; Schmitt & Stults, 1986 ; Marsh, 1989 ; Marsh & Bailey, 1991). Ces modèles ont deux avantages importants pour l'analyse des matrices MTMM : les critères d'acceptation ou de rejet des hypothèses deviennent explicites et les hypothèses elles-mêmes peuvent être beaucoup plus précises (corrélations entre facteurs, etc.).

Les conceptions de Campbell & Fiske et celles de Cronbach & Meehl sont complémentaires. Alors que les premières concernent surtout l'adéquation des instruments comme mesure d'un construit, les secondes concernent plutôt l'adéquation de la mesure d'un construit par rapport à la construction scientifique en général. D'autres conceptions théoriques générales se rattachant à la validité de construit ont été proposées. Les plus intéressantes sont celle de Loevinger (1957) et celle d'Embretson (1983). Comme celles-ci font apparaître une distinction fondamentale dans notre conception de la validité (cf. Dickes et al., 1994), elles seront présentées plus loin.

2.2.3. Des validités à la validation

L'abandon de la conception tripartite de la validité est relativement récent ; on peut le dater du début des années 80. Dans la nouvelle conception, deux points importants peuvent être dégagés. D'une part, le concept de validité prend un sens plus précis : il s'agit de rechercher quelle est la signification des scores (des mesures) obtenus à l'aide d'un instrument. D'autre part, la notion de validation se substitue à celle de validité. Ce dernier changement met en relief une conception dynamique de la validation qui devient un processus continu obéissant aux règles générales de la démarche scientifique. Comme

l'écrit Messick (1989, p. 31), il s'agit "d'un processus de recherche à propos du caractère approprié des interprétations et des actions basées sur les scores aux tests".

Malgré quelques petites divergences, cette conception est largement partagée dans la littérature psychométrique, qu'il s'agisse d'articles de synthèse (cf. Hogan & Nicholson, 1988 ; Moss, 1992 ; Shepard, 1993) ou des dernières éditions des manuels de référence sur la question (cf. Linn, 1989 ; Anastasi, 1990 ; Cronbach, 1990). Pour Cronbach, il s'agit notamment d'expliquer les scores et les relations entre les scores ; pour Anastasi, la validité nous renseigne sur ce que l'on mesure ; pour Messick enfin (in Linn, 1989), la validité concerne la signification et l'interprétation de la mesure.

Tout élément permettant de préciser la signification d'une mesure constitue un aspect du processus de validation. La diversité des démarches et des arguments pouvant être utilisés dans le processus de validation est pratiquement illimitée. Les différents types de validité décrits au paragraphe précédent en constituent bien évidemment des exemples. Savoir que des items retenus dans un instrument sont représentatifs de l'univers des items, qu'une mesure permet de prédire un certain critère, que les mesures obtenues sont indépendantes des méthodes utilisées ou encore que les mesures d'un concept sont liées aux mesures d'autres concepts sont autant d'éléments qui concourent à donner une signification à une mesure. Mais l'abandon du cadre contraignant et restreint de la validité tripartite permet d'imaginer bien d'autres approches qui prennent sens par rapport à une théorie du construit.

Les aspects théoriques occupent une place centrale dans la conception unitaire de la validation. A partir d'une théorie sur un construit, on peut élaborer des hypothèses et les soumettre à des vérifications empiriques ; de plus, lorsqu'une hypothèse est invalidée, la théorie peut être modifiée. C'est pourquoi la validation n'est rien d'autre qu'un processus de recherche au sens le plus général du terme. On comprend également, de ce point de vue, que le processus de validation ne se limite pas à des approches déterminées *a priori*.

L'importance accordée à la théorie sur le construit permet de souligner l'intérêt de certaines stratégies qui n'ont pas toujours été intégrées dans le processus de validation. C'est ainsi, par exemple, que pour Messick (1989, p. 37), le domaine ne peut être défini sans avoir une théorie sur le construit. La théorie doit spécifier quelles sont les différentes composantes du construit et quel est le degré d'homogénéité ou d'hétérogénéité de ces composantes. Des analyses structurales permettent de vérifier si ces composantes théoriques sont retrouvées empiriquement. Des indicateurs de consistance interne permettent de juger du degré d'homogénéité d'une mesure. Les valeurs trouvées ne doivent pas être interprétées dans l'absolu, mais en fonction de la théorie du construit. Si cette dernière fait référence à un construit hétérogène, un indicateur de consistance

interne relativement faible pourra être considéré comme un indice en faveur de la validité. Il en est de même en ce qui concerne la stabilité temporelle des mesures. Si le construit est défini comme un état psychologique, on doit s'attendre à trouver des corrélations faibles entre deux mesures successives du même construit.

La théorie sur le construit, et donc le processus de validation, peuvent aussi bénéficier des travaux menés dans le cadre des recherches relatives au traitement de l'information. L'identification des processus mis en oeuvre pour répondre aux stimuli contribue à enrichir la signification du construit et à en affiner la théorie. Si cette approche concerne surtout, pour l'instant, la mesure des capacités cognitives (cf. Snow & Lohman, 1989 ; Sternberg, 1990), on notera avec intérêt l'existence de travaux du même type dans le domaine de la personnalité ou des attitudes^[7].

La conception unitaire de la validation offre de nombreux avantages. En mettant l'accent sur l'idée d'un processus global, continu et en réhabilitant l'importance des aspects théoriques, elle permet d'intégrer la démarche de validation dans l'ensemble du processus de recherche. Néanmoins, cette conception soulève des questions qui ne peuvent trouver de réponses à l'intérieur de ce schéma.

Tout d'abord, le statut de la validité de contenu dans le processus de validation est toujours discuté. Alors que certains considèrent que cette validité fait partie intégrante du processus de validation, d'autres, nous l'avons vu, sont beaucoup plus réticents. Les arguments utilisés dans ce débat tiennent au fait que la validité de contenu — dans son acception classique — fait uniquement référence à l'instrument de mesure alors que l'on considère aujourd'hui que ce n'est pas l'instrument de mesure qui doit être validé, mais les réponses aux stimuli.

De plus, le statut même du processus de validation est questionné. En tant que processus de recherche identique à tous les autres, la validation n'aurait plus de caractère spécifique. En raisonnant ainsi, on oublie que les hypothèses qui découlent d'une démarche scientifique classique visent à préciser les relations entre des concepts différents, alors que dans la validation, on trouve deux classes d'hypothèses : celles qui concernent les relations entre des concepts différents (comme dans la validité critérielle) et celles qui concernent les relations entre les éléments d'un même concept (comme dans les analyses structurales dont l'objectif est d'identifier les composantes d'un concept). En poussant ce raisonnement jusqu'à l'absurde, la démarche scientifique deviendrait donc un

[7] : Comme ces travaux sont moins connus, il n'est pas inutile d'en donner quelques références : Cliff (1968, 1977) ; Nowakowska (1970) ; Cliff, Bradley & Girard (1973) ; Kuncel (1973) ; Rogers (1974a, 1974b) ; Voyce & Jackson (1977) ; De Boeck (1978, 1981) ; Jackson (1986) ; Krosnick & Alwin (1987) ; Hastie (1987) ; Tourangeau & Rasinski (1988) ; Helmes & Jackson (1989).

cas particulier du processus de validation des mesures ! Cette conclusion provocatrice a pour seul objet d'illustrer les problèmes posés par la conception actuelle de la validation.

Enfin, et toujours dans le même ordre d'idée, si la validation est un processus unitaire et global, l'infirmité des hypothèses théoriques pose un problème d'interprétation. Lorsque les résultats empiriques ne confirment pas une hypothèse portant sur des relations entre des construits différents, deux interprétations — au moins — sont possibles. Dans la première, c'est l'hypothèse théorique qui est considérée comme fautive. Dans la seconde, c'est la mauvaise qualité des instruments utilisés pour mesurer les construits qui est incriminée. Le plus souvent, ces deux interprétations sont aussi plausibles l'une que l'autre. Rares sont les auteurs qui reconnaissent l'existence de cette difficulté ; ceux qui le font (e.g. Messick, 1989) ne proposent pas de solutions, si ce n'est de chercher à mettre au point des stratégies ou des heuristiques permettant de trancher entre les deux interprétations.

A notre avis, les difficultés que nous venons de souligner par rapport à la conception unitaire de la validation tiennent à l'absence de distinction explicite entre les aspects de la validation qui se rapportent aux relations entre des concepts différents (validation interconcepts) et ceux qui concernent les relations entre les éléments du concept que l'on veut mesurer (validation intraconcept). L'objectif du prochain paragraphe est de présenter une conception de la validation qui s'appuie sur cette distinction.

2.3 Validation intraconcept et validation interconcepts

2.3.1 Théorie nomologique et théorie définitoire

La distinction proposée entre validation intraconcept et validation interconcepts (cf. Dickes et al., 1994) a comme principal objectif de résoudre les problèmes posés par la conception unitaire de la validation. Elle ne constitue pas une rupture par rapport à cette conception puisqu'elle en reprend les idées de base : la validation est un processus continu qui adopte les mêmes règles que la démarche scientifique ; les arguments utilisés dans le processus de validation peuvent être de nature variée et ne se limitent pas aux grands types de validité ; la validation concerne la vérification empirique de propositions théoriques à propos du construit. La spécificité de notre conception repose sur une différenciation introduite au niveau de ces propositions théoriques. Les propositions qui concernent les relations entre des construits différents relèvent de la théorie nomologique et s'inscrivent dans une validation interconcepts ; elles ne diffèrent pas des propositions théoriques qui caractérisent la démarche scientifique en général. Mais le processus de validation inclut également des aspects plus spécifiques qui portent sur le seul concept

à mesurer. En précisant l'univers des items associés à un concept et les relations entre ces items, la validation intraconcept s'appuie sur une théorie définitoire. Ce sont donc ces aspects qui font la spécificité de la démarche de validation par rapport à la démarche scientifique ; on ne sera donc pas surpris d'y reconnaître des considérations qui ont été développées à propos de la problématique de la mesure au début de cette première partie.

La distinction entre validation interconcepts et validation intraconcept offre plusieurs intérêts. Elle permet, en particulier, de résoudre les contradictions qui apparaissaient avec la conception unitaire de la validation. Le statut de la validité de contenu ne peut plus être discuté : celle-ci fait partie intégrante de la validation intraconcept au même titre que d'autres stratégies visant à définir ce qui est mesuré. Les problèmes d'interprétation des résultats sont également résolus. Si des résultats infirment des hypothèses concernant des relations entre des construits, ils remettent en cause la théorie nomologique ; s'ils infirment des hypothèses concernant les relations entre les éléments du concept, ils remettent en cause la théorie définitoire et donc la mesure. En insistant sur l'importance de la théorie définitoire, notre conception de la validation rejoint les préoccupations des auteurs (cf., par exemple, Schriesheim et al., 1993) qui soulignent l'ambiguïté des résultats issus de travaux dans lesquels la qualité des mesures utilisées ne donne lieu à aucune vérification. Le problème le plus délicat reste celui de la relation entre validation interconcepts et démarche scientifique. Si la validation interconcepts s'apparente au test de théories nomologiques, n'importe quelle recherche scientifique est une entreprise de validation de mesures ; on parle d'ailleurs, au niveau le plus général, de validation de théories scientifiques. En fait, il faut resituer ce problème dans le cadre des relations entre définition, mesure et théorie, dont il a déjà été question. La validation interconcepts participe à l'élaboration des mesures, au sens où elle permet d'en préciser la signification à travers la modification des théories définitoires. Lorsque l'on montre, par exemple, que certains aspects de l'intelligence des enfants sont fortement liés au niveau socio-culturel des parents alors que d'autres le sont moins, on affine la définition de l'intelligence en y distinguant une composante fluide et une composante cristallisée. Inversement, la validité d'une théorie définitoire de l'intelligence distinguant ces deux composantes se trouve renforcée par des résultats qui montrent que le réseau nomologique associé à chacune de ces composantes n'est pas le même. C'est donc en termes d'influences réciproques qu'il faut comprendre les relations entre théorie nomologique et théorie définitoire. C'est en ce sens que la validation interconcepts s'intègre dans le processus de validation des mesures.

Pour utile qu'elle soit, la distinction entre validation intraconcept et validation interconcepts, entre théorie définitoire et théorie nomologique n'a aucun caractère absolu. C'est le contexte de la recherche et la problématique du chercheur qui lui confèrent un statut. Les théories nomologiques et les théories définitoires ne sont que des hypothèses à propos de relations entre des éléments. Si un chercheur s'intéresse aux relations entre

des éléments qui représentent, pour lui, des construits différents, sa démarche est nomologique. Si un autre chercheur s'intéresse aux mêmes relations, mais considère que les éléments sont des composantes différentes d'un même construit, sa démarche est définitoire. Il n'y a aucune contradiction entre ces deux démarches dès lors que l'on accepte d'en préciser les limites et les enjeux et que l'on reconnaît que dans une démarche nomologique, les construits mis en relation doivent être mesurés.

2.3.2 Logique et stratégies de la validation intraconcept et de la validation interconcepts

Nous ne reviendrons pas en détail, dans ce paragraphe, sur les techniques utilisées de manière habituelle dans le processus de validation, évoquées à propos des types de validité. Nous nous contenterons de les resituer dans le cadre général de notre conception. En revanche, nous nous attarderons plus longuement sur certaines contributions originales dans la littérature sur la validité. Les travaux de Loevinger (1957), de Embretson (1983) et, plus récemment, ceux de Shepard (1993) font apparaître des différenciations qui s'apparentent à la distinction entre validations intraconcept et interconcepts. Aucune de ces conceptions ne s'oppose à celle que nous défendons ici ; nous pensons au contraire qu'elles en éclairent des aspects spécifiques. A un niveau plus général, la distinction entre validation intraconcept et validation interconcepts est celle qui permet le mieux de comprendre le processus de validation.

Chez Loevinger (1957), on trouve trois composantes de la validité : une composante substantive, une composante structurale et une composante externe. La composante externe est identique à ce que nous appelons la validation interconcepts. Ce sont donc les deux autres composantes qui nous intéressent ici.

La composante substantive de la validité inclut à la fois la logique de la validité de contenu – au sens classique – et la vérification empirique d'hypothèses portant, d'une part, sur les interrelations entre les éléments du concept et, d'autre part, sur les relations entre ces éléments et des variables externes. La composante substantive relève donc plutôt de la validation intraconcepts, même s'il est fait référence à des variables externes. En fait, ces variables externes ont un rôle bien spécifique dans ce cas de figure. Les hypothèses ne portent pas tant sur les relations entre le construit et des variables externes que sur le fait que des items censés mesurer la même facette du concept doivent avoir des corrélations comparables avec des variables externes (cf. à ce sujet Piazza, 1980 qui a présenté une conception de la validité de construit insistant largement sur cet aspect). Autrement dit, la composante substantive concerne les relations intraconcept. Le recours à des variables externes n'est qu'une manière détournée pour vérifier que des items jugés équivalents par la théorie définitoire le sont réellement : si les réseaux nomologiques

attachés aux items ne sont pas les mêmes, il est fort probable que l'hypothèse d'équivalence soit fausse. Cette remarque illustre une des caractéristiques les plus fondamentales d'une théorie définitoire, celle de permettre d'engendrer des items conceptuellement équivalents.

La composante structurale relève également de la validation intraconcept. Il y a validité structurale, pourrait-on dire, lorsque la mesure correspond à sa définition. Au-delà du caractère trivial de cet énoncé, on y reconnaîtra l'idée, déjà souvent évoquée, de la nécessité d'une correspondance entre la définition théorique d'un concept, les hypothèses qui en découlent et les résultats empiriques qui viennent mettre à l'épreuve des faits ces hypothèses. Cette mise en correspondance peut prendre des formes diverses en fonction des caractéristiques retenues dans la définition. Moskowitz (1986), par exemple, considère que montrer une faible stabilité intersituationnelle d'une mesure est un argument en faveur de la validité structurale de ce concept à condition que celui-ci soit défini comme un état psychologique^[8]. De manière plus générale, la validité structurale implique une correspondance parfaite entre la structure inter-items spécifiée par la théorie et la structure empirique des observations. On ne sera pas surpris de retrouver ici ce qui constitue l'un des fondements de la mesure au sens de la théorie représentationnelle : la correspondance entre un système formel et un système empirique. On retrouve d'ailleurs chez certains théoriciens de la mesure des positions qui se rapprochent très clairement de la notion de composante structurale de la validité. Ainsi, pour Narens & Luce (1993), la mesure est un moyen de décrire la structure des observations.

A la différence de Loevinger, Embretson (1983) ne distingue que deux composantes de la validité. Une autre différence réside dans le fait que la validité est uniquement envisagée au sens de la validité de construit. Embretson (1983) distingue donc entre la "représentation du construit" et le "réseau nomologique" (traduction bien imparfaite de "*nomothetic span*"). Le réseau nomologique est très semblable à notre validation interconcepts : il concerne le réseau de relations entre des mesures différentes. Une différence doit tout de même être soulignée. Chez Embretson, le réseau nomologique inclut toutes les relations entre des mesures différentes, y compris les mesures différentes d'un même concept. Lorsqu'un même concept est mesuré avec des méthodes différentes, on s'intéresse à la convergence des mesures. Pour nous, montrer que la mesure d'un concept est indépendante des méthodes utilisées (choix des items, types de collecte des observations, etc.) relève de la validation intraconcept. Si on excepte cette différence, la notion de réseau nomologique est identique à la notion de validation interconcepts.

[8] : En fait, cet argument pourrait tout aussi bien relever de la composante substantive si on considère que l'on étudie les relations entre les items et des variables externes qui sont, en l'occurrence, les caractéristiques des situations. Ce qui est important ici, c'est que l'on peut tester empiriquement des hypothèses dérivées de la théorie définitoire et que cela relève d'une validation intraconcept.

La principale originalité de la contribution de Embretson réside dans la notion de représentation du construit. Tout en partageant des caractéristiques communes avec les composantes structurale et substantive de la validité de Loevinger, elle les dépasse en les précisant. La représentation du construit vise à identifier les processus théoriques qui sous-tendent les réponses à un stimulus. Alors que dans le réseau nomologique, ce sont les différences interindividuelles (et donc la variabilité entre les sujets) qui constituent le principal intérêt, la représentation du construit se focalise sur la variabilité des stimuli. Les études qui portent sur la représentation du construit s'attachent à décomposer les stimuli proposés aux sujets de manière à faire apparaître les facteurs de variation des réponses, ceux-ci étant assimilés aux facettes théoriques du concept. La pertinence de cette décomposition est évaluée à partir de modélisations qui s'appuient sur des généralisations des modèles de réponse à l'item. On voit tout l'intérêt d'une telle décomposition dans le processus de validation intraconcept. En identifiant les sources de variation des réponses, on contribue, comme nous l'avons déjà souligné, à définir le concept mesuré. De plus, cette conception est compatible avec le souhait, exprimé par plusieurs auteurs travaillant dans des perspectives différentes, de s'intéresser aux propriétés des stimuli, souvent négligées dans la littérature psychométrique.

Les travaux de Angleitner et al. (1986) s'inscrivent dans une telle perspective. A partir d'une analyse de contenu d'items de questionnaires de personnalité, ils étudient l'influence des caractéristiques des items sur les réponses. Il existe aussi des études expérimentales qui font varier de manière systématique la formulation des items. De telles études, courantes dans la mesure des attitudes et des opinions (cf., entre autres, Schuman & Presser, 1981 ; Grémy, 1987 ; Bradburn & Sudman, 1991 ; Brignier, 1991 ; Molenaar, 1991), ont souvent pour objectif de repérer les caractéristiques des items qui introduisent des biais dans les réponses. De telles études permettent aussi, à notre avis, de mieux comprendre les processus qui sont à l'oeuvre dans les réponses (cf. Hippler, Schwarz & Sudman, 1987 ; Jobe & Mingay, 1991) et participent ainsi à la représentation du construit. Une logique identique se retrouve dans la théorie en facettes, sur laquelle nous reviendrons dans le prochain chapitre, puisque les facettes retenues pour définir un concept doivent être des sources de variation des réponses. Dans le domaine cognitif, les travaux de ce type sont plus rares, même si l'on peut noter quelques tentatives intéressantes. Solano-Flores (1993), par exemple, propose une technique qui tient compte des propriétés structurales des items pour développer un test de lecture des chiffres romains. La difficulté des items et leurs interrelations doivent pouvoir être prédites à partir de ces propriétés structurales. La formulation la plus claire de cette problématique se trouve chez Stenner et al. (1983), qui parlent d'ailleurs de "théorie de définition du construit". Pour ces auteurs, l'objectif primordial est d'arriver à expliquer les valeurs d'échelle attribuées aux items (c'est-à-dire leur mesure) ; tant que cet objectif n'est pas atteint, il n'y a pas de compréhension de ce qui est mesuré.

Au total, la représentation du construit réhabilite l'importance des items dans le processus de mesure. De plus, l'identification des propriétés structurales des items justifie que l'on complète les études corrélationnelles par des études plus expérimentales, ce que plusieurs auteurs réclament depuis longtemps (cf. Cronbach, 1957 ; Landy, 1986). L'idée de la représentation du construit est donc fondamentale dans le processus de validation. Elle n'est d'ailleurs pas très éloignée de l'objectif qu'assigne Roskam (1981) aux études qui portent sur la mesure : pour lui, il s'agit en effet d'établir des modèles algébriques et des règles de composition qui relient les réponses des sujets aux attributs des stimuli.

La contribution de Shepard (1993) n'apporte rien d'original aux idées de Loevinger et de Embretson. Mais, plus récente, elle offre une synthèse qui, tout en reprenant l'essentiel du travail de Messick (1989), insiste davantage sur la distinction entre les composantes internes et externes de la validité. C'est cette synthèse que nous reprenons ici. Pour Shepard, les études de validation se caractérisent de la manière suivante (p. 417) :

- elles doivent aborder à la fois la structure interne des mesures^[9] et les relations externes avec d'autres variables ;
- les études corrélationnelles classiques doivent être complétées par des études expérimentales ;
- la démarche de validation ne se contente pas d'une confirmation empirique des hypothèses théoriques ; il faut aussi que les théories puissent engendrer des hypothèses opposées qu'il est possible de soumettre à des investigations empiriques ;
- les études de validités convergente et divergente sont des outils utiles dans le processus de validation, mais leurs résultats ne doivent pas être interprétés de manière mécanique, en dehors de toute démarche théorique.

L'ensemble de ces points, avec lesquels nous sommes d'accord, ne sont pas incompatibles avec notre propre conception. Il n'est peut-être pas inutile de la résumer à partir d'un schéma général qui situe les différentes stratégies de validation dans le processus de mesure.

La figure 1.3 présente, sous une forme schématique, l'essentiel de notre conception de la validation. Celle-ci est conçue comme une démarche continue, totalement intégrée dans le processus de mesure. Il est utile de distinguer entre une validation intraconcept et une validation interconcepts afin de préciser la démarche.

[9] : "The internal model should reflect all aspects of the theory that defines a construct, including its subdomains or subconstructs, the expected interrelationships among dimensions of the construct, and the processes believed to underlie test performance." (Shepard, 1993, pp. 417-418).

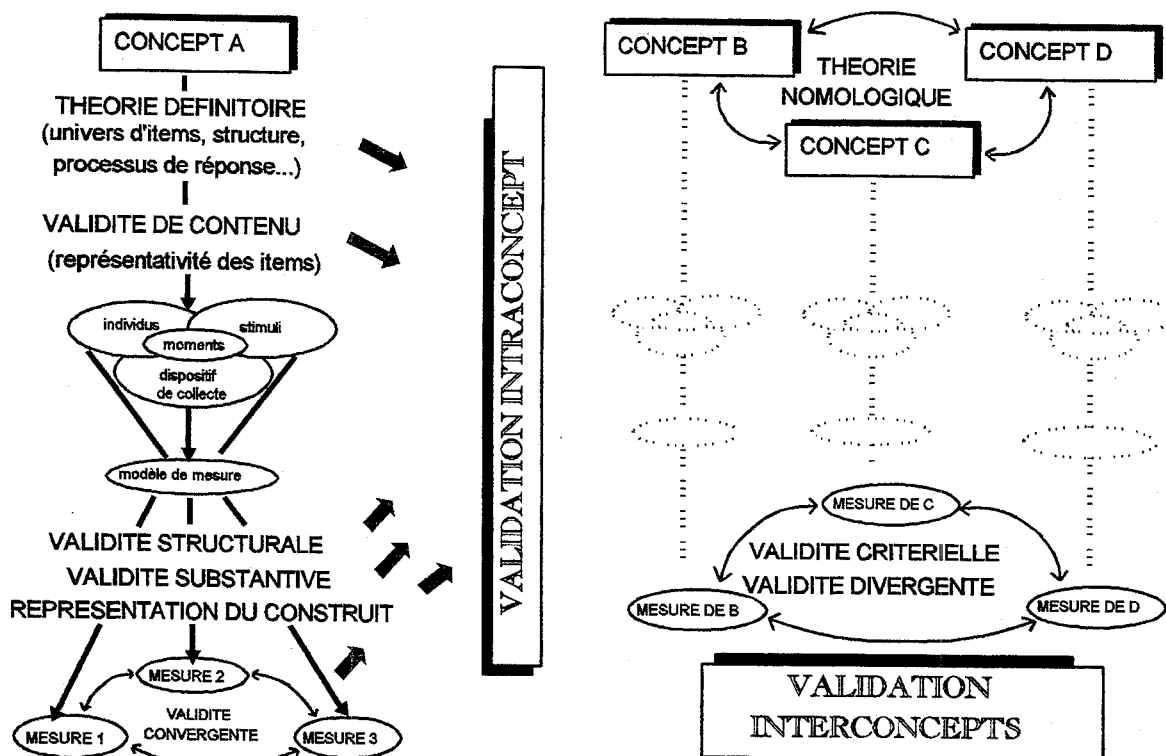


Figure 1.3 : La démarche de validation dans le processus de mesure

La validation interconcepts est très utilisée dans la littérature. Associée à une théorie nomologique sur le réseau de relations entre des construits différents, elle ne se distingue pas, dans la pratique, de la démarche scientifique en général. Mais elle s'en distingue au niveau conceptuel. La logique de la validation interconcepts est d'améliorer la signification de la caractéristique mesurée par l'intermédiaire du réseau de relations avec d'autres caractéristiques ; la validation interconcepts ne se substitue donc pas à l'ensemble de la démarche scientifique.

La validation intraconcept, plus embarrassante pour les chercheurs, a été longtemps négligée dans les travaux psychométriques. Trop souvent réduite à une conception très instrumentale de la validité de contenu, ce n'est que récemment qu'elle a trouvé ses lettres de noblesse grâce à des contributions comme celles de Jane Loevinger (1957)^[10] ou de Susan Embretson (1983). La validation intraconcept a en commun avec la validation interconcepts une logique hypothético-déductive. Mais ce ne sont plus des théories nomologiques qui sont testées, ce sont des théories définitives qui précisent non

[10] : Malgré l'ancienneté de cet article, ce n'est que depuis quelques années que les idées qui y sont développées ont un réel impact sur la communauté scientifique.

seulement quelles sont les différentes facettes du concept, mais aussi leurs interrelations, les items qui peuvent être engendrés par la combinaison de ces facettes et les processus de réponse à ces items. Toute hypothèse qui découle d'une théorie définitoire doit, comme dans toute recherche, être réfutable. C'est aux modèles de mesure, lorsqu'ils sont suffisamment élaborés, que revient la fonction de tester la plausibilité de ces hypothèses.

2.4 En guise de conclusion : validation et théories de la mesure

Nous venons de présenter deux chapitres portant sur la question fondamentale de la mesure en psychologie. Dans le premier, ce sont surtout les aspects théoriques de l'opération de mesure qui ont été présentés ; dans le second, l'accent a été mis plutôt sur les stratégies permettant de valider des mesures. Pourtant, malgré une communauté d'intérêt évidente, les liens entre ces deux approches n'apparaissent pas au fil du texte. Cela s'explique par le simple fait que ces liens n'existent pas, ou pratiquement pas. Nous allons essayer de fournir quelques éléments de réflexion sur cette situation paradoxale. Ces éléments resteront spéculatifs. Il n'est pas exagéré, en effet, d'affirmer que personne n'est à même, à l'heure actuelle, d'apporter une réponse claire à cette question dont la difficulté n'est pas tant technique qu'épistémologique. Pour simplifier la présentation, nous parlerons, dans le premier cas, de l'approche théorique de la mesure et, dans le second, de l'approche psychométrique.

Les auteurs ayant réfléchi à la question du fossé qui sépare l'approche théorique et l'approche psychométrique sont peu nombreux ; ceux qui ont essayé de les intégrer dans leur pratique le sont encore moins (cf. Schwager, 1991 ; Cliff, 1992 ; Narens & Luce, 1993 ; Borg & Shye, in press)^[11]. On tentera de dégager de ces travaux les points qui nous paraissent les plus importants pour comprendre la raison de la si faible influence mutuelle des deux approches.

Rappelons tout d'abord ce qui caractérise ces deux approches. Nous avons vu que pour l'approche théorique, l'objectif de la mesure est d'établir un homomorphisme entre un système empirique et un système formel. Une caractéristique est mesurable dès lors que l'on peut trouver un système formel (numérique le plus souvent) structurellement semblable au système empirique (cf. Schwager, 1991). Cela signifie qu'une caractéristique est mesurée si on peut montrer que les propriétés associées à un système

[11] : Cette liste n'est pas exhaustive. Des auteurs comme Coombs ou Roskam abordent aussi cette question dans leurs publications, mais de manière plus incidente.

formel sont vérifiées dans le système empirique^[12]. Dans l'approche psychométrique, une mesure est jugée "acceptable" (i.e. est validée) si les hypothèses définitives et nomologiques associées à la caractéristique ne sont pas infirmées par des résultats empiriques. Ces hypothèses ne portent pas tant sur l'homomorphisme entre le système formel et le système empirique, que sur la cohérence du système de relations inter- et intraconcept qui est construit par le chercheur.

Le premier problème qui apparaît dans la confrontation entre ces deux approches est celui de la signification des mesures. Dans l'approche psychométrique, celle-ci est fondamentale ; la signification se construit au fur et à mesure que le réseau de relations inter- et intraconcept s'enrichit. Cette notion est totalement absente de l'approche théorique^[13], tout comme, d'ailleurs, la notion de validation. En fait, dans la théorie axiomatique, la vérification empirique d'axiomes associés à un système formel semble être une condition nécessaire et suffisante pour que le système empirique soit représenté (mesuré). Mais, dans ce cas, quelle est la caractéristique qui a été mesurée par cette opération ? La réponse pourrait très bien être : aucune ! En effet, on peut très bien imaginer un système empirique qui respecte parfaitement des axiomes comme celui de la transitivité sans que la mesure qui en découle ait une signification. Il suffit de songer, par exemple, à un questionnaire absurde composé des questions dichotomiques suivantes : (a) avez-vous déjà été sur la lune ? (b) êtes-vous déjà monté dans un avion ? (c) avez-vous déjà mangé du pain ? Il ne fait guère de doute que ces trois questions puissent être représentées par un système numérique caractérisé par des axiomes correspondant à un niveau de mesure ordinal ; elles se conformeraient parfaitement au modèle hiérarchique de Guttman. On est pourtant en droit de s'interroger sur la caractéristique qui serait mesurée par un tel questionnaire.

Le problème de la signification des mesures est lié à leur intérêt pratique et théorique. L'approche psychométrique l'a parfaitement compris en insistant, dans la démarche de validation, sur la validation interconcepts. Il n'en est pas de même pour l'approche axiomatique qui a tendance à s'abstraire du contexte scientifique dans lequel les mesures sont développées. Ceci est particulièrement manifeste dans l'analyse que font

[12] : Ces propriétés sont en fait des axiomes et l'approche théorique est quelquefois qualifiée de "théorie axiomatique de la mesure". Le terme "axiomes" renvoie au système formel et se justifie par le fait que, à l'intérieur de ce système, ils ne sont pas démontrés. En revanche, ces axiomes sont testés dans le système empirique et deviennent, par là même, des propriétés vérifiables. La transitivité est un exemple classique d'axiome habituellement utilisé (e.g. Schwager, 1991). La transitivité empirique des relations entre des éléments en fonction d'une certaine caractéristique est une condition nécessaire pour que cette caractéristique puisse être représentée, dans le système formel des nombres, par la relation "est supérieur à". Si la transitivité empirique n'est pas vérifiée, la caractéristique ne peut être représentée (i.e. mesurée) par un système formel auquel est associée la relation "est supérieur à". En d'autres termes, la caractéristique ne peut être mesurée au niveau ordinal, mais elle pourrait très bien être mesurée d'une autre manière, avec un système formel moins exigeant.

[13] : Bien que le terme "signification" soit présent dans l'approche théorique, il ne renvoie pas du tout à la même notion. Il représente l'invariance dans toutes les transformations admissibles des valeurs d'échelle (cf. Suppes & Zinnes, 1963).

les théoriciens de la mesure de concepts comme la longueur, la masse ou la température. Cette analyse se limite à dégager les axiomes du système formel choisi pour les représenter ; elle ne permet pas d'expliquer pourquoi certaines procédures de mesure ont été utilisées plutôt que d'autres. Schwager (1991) note à ce sujet que la mesure de ces concepts ne peut être envisagée sans référence aux développements et à l'évolution des théories scientifiques. Ce sont ces dernières qui confèrent une signification aux mesures : la distinction entre le poids et la masse n'a été possible que grâce à une évolution des théories physiques ; le baromètre n'a pu être compris qu'à partir des théories et des expériences de Pascal ; la mesure de la température se confond avec l'évolution de la thermodynamique (cf. à ce sujet, le très bel article posthume de Choppin, 1985).

Il ne faudrait pas croire pour autant, au vu des remarques précédentes, que l'approche psychométrique est à l'abri de toute critique. En privilégiant la signification des mesures par l'intermédiaire d'une validation interconcepts, elle ne fait qu'assurer une cohérence ponctuelle à la construction scientifique. Quelle que soit l'utilité des mesures développées de cette manière, elles manquent d'assises formelles et risquent d'être très dépendantes des évolutions théoriques. Ce faisant, aucune mesure psychologique ne fait l'objet d'un consensus, à l'inverse de ce que l'on peut observer, par exemple, en physique. Cela explique peut-être le caractère non-cumulatif des connaissances que certains croient discerner en psychologie.

La manière d'aborder la mise en relation d'un système formel et d'un système empirique distingue également ces deux approches. Cette question renvoie d'ailleurs au problème complexe du statut des modèles de mesure et des erreurs de mesure. Pour Borg & Shye (*in press*), l'approche théorique adopte une stratégie descendante ("*top-down*") en essayant de définir a priori les propriétés algébriques que doivent posséder les données pour pouvoir être représentées. Cette idée est exprimée de la manière suivante par Kyburg (1984, cité par Schwager, 1991) : "*How must the world be — what must be its structure — in order for measurement of a certain sort to be possible ?*". Dans l'approche psychométrique, du moins lorsqu'elle prend la peine de s'intéresser aux relations intraconcept, la stratégie est plutôt ascendante ("*bottom-up*") : on part des données et on essaie de maximiser l'adéquation de ces données à un certain modèle mathématique. La seconde stratégie a l'avantage de pouvoir tenir compte des erreurs qui sont inévitables dans toute observation. C'est beaucoup moins évident pour la première stratégie (cf. Cliff, 1992) dans laquelle les axiomes du système formel doivent être totalement vérifiés pour que la caractéristique soit mesurée. Comme cette situation idéale n'est jamais atteinte, sauf dans des cas triviaux, il faudrait pouvoir accepter un certain niveau d'erreurs ; il reste à savoir quelles erreurs tolérer et comment les prendre en compte.

Plus généralement, se pose le problème des modèles de mesure mis en oeuvre. Si on excepte le cas des modèles d'échelonnement multidimensionnel ou de la loi du

jugement comparatif dont l'axiomatisation est relativement bien comprise (cf. Michell, 1990), on n'a guère d'indications quant à la plausibilité des fondements théoriques des autres modèles utilisés dans l'approche psychométrique. Cliff (1989, 1992) cite le cas des modèles de réponse à l'item à trois paramètres qui ne peuvent permettre, selon lui, d'établir des échelles d'intervalles, car les postulats de ces modèles sont en contradiction avec les axiomes associés à ce niveau de mesure. Cet exemple montre tout l'intérêt qu'il peut y avoir à confronter ces deux approches ; malheureusement, de tels exemples sont rares.

Le courant psychométrique pouvait attendre beaucoup des progrès accomplis depuis une trentaine d'années dans la compréhension des aspects théoriques de la mesure. Pourtant, cette "révolution n'a jamais eu lieu", comme l'exprime très bien Cliff (1992) à travers le titre de son article. On peut espérer que la confrontation tant attendue entre les deux courants aura lieu dans les prochaines années. A l'heure où l'approche théorique commence à être critiquée (cf. Schwager, 1991 ; Michell, 1990), cette confrontation ne pourra qu'être profitable, même s'il faut, pour cela, remettre en cause les fondements de l'une et de l'autre approche.

CHAPITRE III

UNE APPROCHE STRUCTURALE DE LA MESURE : LA THEORIE EN FACETTES

3.1 La mesure comme théorie structurale

C'est en tant que président de la Société de Psychométrie que Louis Guttman délivra, en 1971, une conférence intitulée "*La mesure en tant que théorie structurale*". Le texte de cette conférence, publié en 1979 dans un ouvrage collectif, est sans doute la contribution la plus synthétique de Guttman au problème de la mesure. La spécificité de la mesure (dans ses théories, dans ses démarches) en occupe la majeure partie. C'est en faisant apparaître les différences avec la statistique que les spécificités de la mesure apparaissent le plus nettement. Alors que la première concerne l'inférence des observations obtenues sur un échantillon à la population, la seconde se caractérise par la construction d'hypothèses structurales et par l'examen de ces hypothèses en fonction des relations entre les variables.

Pour Guttman, la mesure est donc essentiellement une démarche théorique et la mesure psychologique ne diffère pas, sur ce point, de la mesure physique. Dans les deux cas, on utilise une théorie structurale qui spécifie comment les observations sont obtenues et quelles hypothèses sont faites sur les relations entre les variables (1979, p. 57). Mais une théorie, ce n'est pas seulement un système de propositions ; la construction théorique et l'investigation empirique constituent deux pans indissociables de la recherche scientifique. Cette conception se retrouve dans une phrase de Guttman qui a été abondamment citée : "*A theory is a hypothesis of correspondance between a definitional system for a universe of observations and an aspect of the empirical structure of those observations, together with a rationale for such an hypothesis.*" (Guttman, 1981, p. 50).

L'approche structurale de la mesure se distingue à la fois de l'approche axiomatique de la mesure et de l'approche psychométrique. Malgré leurs différences, on peut relever des points communs qui se retrouvent surtout dans le processus de validation intraconcept de l'approche psychométrique. Plus généralement, l'idée selon laquelle la mesure est une démarche théorique qui inclut l'élaboration d'hypothèses et leur test empirique est une idée partagée par les trois approches. La spécificité de l'approche structurale tient à la place accordée aux problèmes de définition des concepts et d'univers des items.

C'est dans ce contexte général de la mesure que doit être considérée l'approche en facettes^[1] : celle-ci propose un cadre général à l'intérieur duquel on peut proposer des définitions, générer des univers d'items et tester des hypothèses théoriques concernant la structure des items.

3.2 Principes de l'approche en facettes

Historiquement, on peut concevoir avec Canter (1983, cité par Brown, 1985) que l'approche en facettes est née du fait que Guttman était confronté à des problèmes de sélection d'items, à la faiblesse des techniques d'analyse factorielle et au manque de clarté caractérisant les aspects définitoires des problématiques scientifiques.

L'approche en facettes peut être abordée à partir de perspectives différentes. On peut choisir, par exemple, dans une perspective très large, d'insister sur le fait qu'elle n'est rien d'autre qu'une approche générale de la recherche scientifique (cf. Brown, 1985), qui généralise la notion de plan d'expérience de Fisher à n'importe quel type d'observations (cf. Levy, 1993). Mais nous préférons mettre en relief ce qui constitue l'enjeu principal de l'approche en facettes. Celui-ci concerne les problèmes de définition des concepts psychologiques en l'absence d'un univers d'items connu : "l'objectif principal de l'approche en facettes est d'explicitier l'implicite" écrivent très joliment Borg & Shye (*in press*) ; pour Shye (1978 a, p. 5), l'approche en facettes se caractérise par le "développement de techniques pour construire et présenter des systèmes définitoires pour des univers d'observations".

Le système définitoire s'élabore autour de la notion de facettes. Une facette peut être conçue comme une collection d'attributs (de modalités) qui caractérisent les propriétés des variables. Mais ces facettes, et leurs attributs, ne sont pas arbitraires : elles doivent être, théoriquement et empiriquement, pertinentes par rapport à l'objet d'investigation. Les facettes sont souvent identifiées lors d'études préalables ayant montré que la manipulation de certaines caractéristiques des stimuli entraînait des différences dans les réponses. La combinaison formelle de toutes les facettes retenues s'effectue dans une phrase en facettes : celle-ci permet de définir de manière très précise l'univers des observations associées à un concept. Mais les facettes ne sont jamais fixées définitivement : leur nombre et leurs modalités évoluent en fonction de l'avancement des théories et des recherches empiriques.

[1] : On dit aussi théorie en facettes, mais dans ce cas, il faut comprendre le terme théorie au sens de la définition de Guttman citée plus haut.

On distingue en général (cf. Levy, 1981a) trois grands types de facettes. Le premier type, habituellement désigné par le symbole X, spécifie la population des individus étudiés. Le second concerne le contenu des variables. Ensemble, ces deux premiers types de facettes constituent le domaine. Le troisième spécifie l'étendue, c'est-à-dire l'ensemble des modalités de réponse qui sont considérées comme pertinentes par rapport au concept mesuré. Une phrase en facettes peut donc être représentée schématiquement comme suit (figure 1.4).

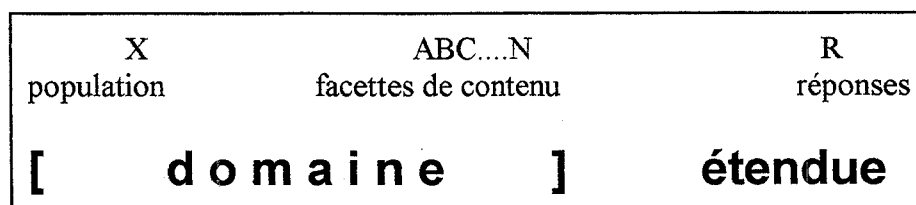


Figure 1.4 : Les trois types de facettes dans une phrase en facettes (d'après Levy, 1981a, p. 79)

Parmi ces trois types de facettes, ce sont bien évidemment les facettes de contenu qui sont les plus intéressantes ; ce sont celles qui permettent de définir l'univers des items. Il n'est sans doute pas utile de revenir sur les problèmes posés par la définition d'un univers d'items : nous avons eu l'occasion, à plusieurs reprises, d'en souligner les difficultés. En revanche, il faut insister sur l'originalité de l'approche en facettes sur ce point : l'univers d'items est totalement défini par la spécification des facettes, de leurs modalités et des effets des relations entre les facettes sur la structure des réponses. En offrant la possibilité de réfuter empiriquement une telle définition, l'approche en facettes autorise une démarche définitoire hypothético-déductive qui supprime certains aspects arbitraires des définitions classiques^[2].

Pour que l'approche en facettes permette effectivement de définir un univers d'items, le choix des facettes et de leurs modalités doit répondre à un certain nombre de règles qui ont été spécifiées, par exemple, par Runkel & McGrath (1972, cités par Brown, 1985) :

- les items doivent pouvoir être classés en fonction de toutes les facettes retenues ;
- chaque facette est divisée en un ensemble exhaustif de modalités ou d'éléments ;
- les éléments, à l'intérieur de chaque facette, sont mutuellement exclusifs ;
- les relations logiques entre les facettes doivent être spécifiées ;
- les facettes doivent couvrir exhaustivement le domaine d'étude.

[2] : Qui en supprime certains aspects seulement. En effet, des phrases en facettes différentes peuvent très bien définir les mêmes ensembles d'items (pour un exemple dans le domaine des valeurs au travail, cf. Borg, 1990). Si les différentes phrases en facettes sont toutes validées empiriquement, aucune ne pourra être considérée comme meilleure qu'une autre (du moins sur des critères internes).

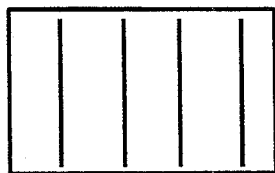
Il est presque superflu de noter que ces règles sont extrêmement contraignantes et correspondent à un idéal rarement atteint dans les applications pratiques.

Les facettes sont des caractéristiques qui permettent de définir de manière univoque le contenu des items appartenant à un même domaine, à un même univers. Chaque facette contient un nombre variable d'éléments mutuellement exclusifs et exhaustifs qui servent à structurer, au niveau théorique, l'univers de contenu.

A ce stade, une phrase en facettes n'est rien d'autre qu'un moyen commode d'émettre une proposition structurale à propos du domaine étudié. L'intérêt de l'approche en facettes réside dans le test de cette proposition. C'est la mise en correspondance entre la structuration théorique et des données empiriques qui assure la validité de la démarche. La logique de cette mise en correspondance repose sur le principe de contiguïté (cf. Foa, 1958) : plus les variables sont similaires au niveau de la structure en facettes, c'est-à-dire plus nombreuses sont les modalités de facettes qu'elles partagent, plus ces variables doivent être liées empiriquement. Encore faut-il que cela se vérifie.

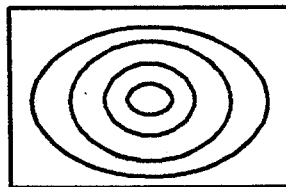
Cette vérification empirique s'effectue généralement par l'intermédiaire de modèles d'échelonnement multidimensionnel (EMD), mais l'utilisation de modèles linéaires est également possible (analyse de variance, analyse discriminante ; cf. Borg & Shye, *in press* ; analyse factorielle tri-modale, modèles généraux d'analyse des matrices de covariance de type LISREL ; cf. Mellenbergh et al., 1979). La préférence des utilisateurs de l'approche en facettes pour les modèles d'EMD s'explique par leur caractère non linéaire et les possibilités de représentation géométrique qu'offrent ces modèles. En effet, les facettes peuvent jouer des rôles qui contribuent à structurer la position des variables dans un espace de manière bien spécifique. On distingue habituellement trois grands rôles qui sont représentés dans la figure 1.5.

Les éléments de la facette sont ordonnés

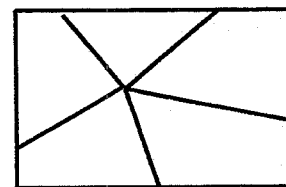


Rôle axial

Les éléments ne sont pas ordonnés



Rôle radial
(ou modulaire)



Rôle angulaire
(ou polaire)

Figure 1.5 : Le rôle des facettes (d'après Levy, 1981a).

Lorsque l'espace des variables peut être découpé en bandes relativement parallèles (quelle que soit leur orientation) de manière à ce que l'on ne trouve, dans une même bande, que des variables correspondant à un même élément d'une facette, la facette a un rôle axial. L'ordre de ces bandes correspond à un ordre dans la signification des modalités de la facette. Mais, lorsque la facette est ordonnée, on peut aussi trouver une figure constituée de cercles concentriques, chaque cercle étant délimité à partir des variables partageant le même élément de la facette en question (rôle radial ou modulaire). Enfin, si les modalités de la facette ne portent aucune notion d'ordre, on parle de rôle angulaire : les variables occupent dans l'espace des régions contiguës délimitées à partir d'une même origine. Des noms ont été donnés à ces figures typiques : on appelle respectivement ces trois figures un simplex, un radex et un circumplex^[3].

En général, les items ne sont pas définis par une seule facette mais par la combinaison de plusieurs. A ces combinaisons de facettes correspondent des figures plus complexes que celles qui ont été envisagées jusqu'à présent. Les plus courantes sont présentées dans la figure 1.6.

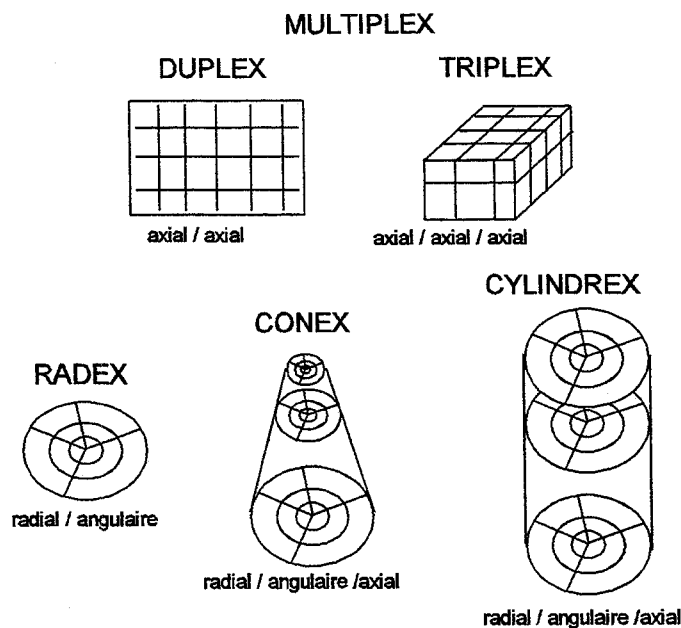


Figure 1.6 : Exemples de structures classiques obtenues par la combinaison du rôle de plusieurs facettes.

La logique de la construction de ces figures est la même que précédemment. Dans chaque région de l'espace délimitée par la combinaison de plusieurs facettes ne doivent

[3] : Le circumplex correspond aussi à une organisation circulaire spécifique : celle d'une seule dimension totalement incurvée (e.g. le cercle des couleurs).

se trouver que des variables partageant les mêmes modalités sur les différentes facettes retenues.

Les études empiriques qui se réclament de l'approche en facettes utilisent régulièrement certaines facettes de contenu. Afin de concrétiser quelque peu la présentation de ces principes théoriques, nous proposons, dans le tableau suivant, une sélection de facettes fréquemment utilisées dans les applications empiriques, le rôle qui leur est habituellement associé et quelques références bibliographiques les présentant. Ce tableau ne prétend nullement à l'exhaustivité. Son seul objectif est de proposer au lecteur intéressé une vue très schématique des applications potentielles de la théorie en facettes.

Facettes	Rôle	Références
Modalité comportementale (cognitive, affective, instrumentale)	polaire modulaire	Elizur & Sagie (1993) Tziner & Plourde (1993)
Domaines de vie	polaire	Edmundson et al. (1993) Levy & Guttman (1975) Dickes & Tournois (1993) Levy (1993)
polarité	axial	Edmundson et al. (1993) Dickes & Tournois (1993) Dancer (1985)
environnement primaire vs secondaire	axial ou modulaire	Bilsky & Wetzels (1993)
systèmes fonctionnels (personnalité, physique, social, culturel)	polaire	Denesh & Shye (1993)
éléments internes vs externes	modulaire	Meyer-Schweizer (1993)
éléments généraux vs spécifiques	modulaire	Meyer-Schweizer (1993)
dépend de soi vs dépend de l'extérieur	modulaire	Borg (1990)
état vs ressources	modulaire	Levy (1981b) Roskam (1981)
récompenses vs ressources	modulaire	Meyer-Schweizer (1993)
domaine matériel vs physique vs psychologique	axial	Bilsky & Wetzels (1993)
temporalité (passé, présent, futur)	axial	Levy (1981 a)
fréquence d'un comportement	axial	Edmundson et al. (1993)

Tableau 1.1 : Exemples de facettes utilisées dans les applications empiriques, rôle et références bibliographiques.

3.3 Un exemple d'approche en facettes : la satisfaction vis-à-vis de différents domaines de vie

L'exemple présenté ci-dessous est emprunté à Levy (1981b). Il a été choisi d'une part pour sa simplicité, afin d'illustrer les différents concepts de la théorie en facettes, d'autre part pour son domaine d'application: comme il porte sur la satisfaction, il se rapproche de notre problématique générale sur le bien-être subjectif. Néanmoins, il faut reconnaître que, de par sa simplicité, les résultats qu'il rapporte peuvent être jugés triviaux sur certains aspects. Nous le concédons bien volontiers et nous y reviendrons dans la prochaine section.

Contrairement à la démarche habituelle qui consiste à construire une phrase en facettes puis à en dériver les items et à recueillir les observations, l'étude de Levy est basée sur des données pré-existantes : l'objectif est de comparer la structure des réponses à des items de satisfaction dans deux pays, les Etats-Unis et Israël. Nous ne retiendrons pour illustration que les résultats obtenus aux Etats-Unis.

Dans une vaste enquête nationale, les sujets devaient répondre, sur une échelle graduée, à des questions portant sur leur niveau de satisfaction relativement à 15 domaines de vie spécifiques. Levy (1981b, p. 118) propose la phrase en facettes suivante pour "définir" les observations recueillies dans cette étude (figure 1.7).

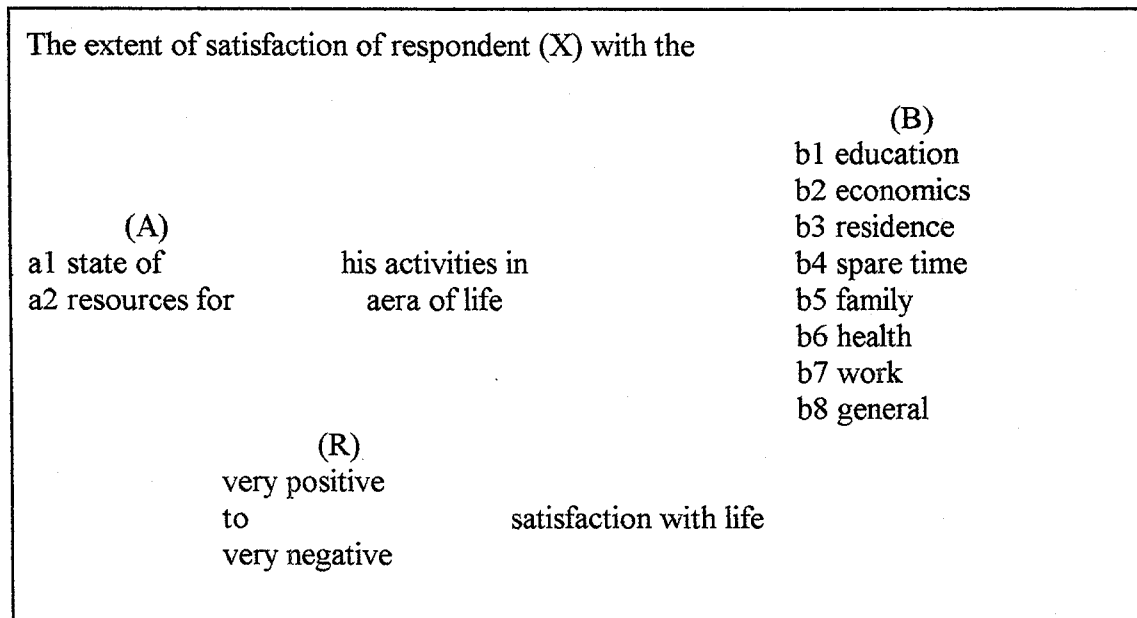


Figure 1.7 : Phrase en facettes définissant la satisfaction vis-à-vis de différents domaines de vie (d'après Levy, 1981b, p.118).

Cette phrase comprend quatre facettes reliées entre elles par des éléments de discours qui en facilitent la lecture. La première facette (X) désigne la population des sujets ; les deux suivantes, (A) et (B), sont les facettes de contenu ; la dernière (R) spécifie l'étendue des réponses possibles. C'est la combinaison des deux facettes de contenu qui permet de générer les items. On les désigne habituellement par leur *structuple* (a1b1, a2b2, ..., a2b8). Normalement, il devrait y avoir autant d'items que le produit cartésien du nombre de modalités de chaque facette de contenu. Ce n'est pas le cas ici, puisqu'il n'y a que 15 items pour 16 combinaisons possibles^[4].

Pour valider cette phrase en facettes, il faut montrer que les réponses aux items se structurent empiriquement en fonction des facettes. Autrement dit, il faut montrer que les relations entre les items obéissent à des règles liées au rôle joué par les facettes et leur combinaison. Nous avons vu que l'approche en facettes privilégie l'utilisation des modèles d'échelonnement multidimensionnel pour tester les relations entre les variables. Ces modèles permettent de représenter les items dans un espace euclidien de telle sorte que les items sont d'autant plus proches que leur corrélation est élevée (cf. Tournois & Dickes, 1993). Pour que la définition ne soit pas infirmée, il faut que chaque facette partitionne l'espace en autant de régions que de modalités dans la facette. En liaison avec cette analogie géométrique, ces hypothèses sont assimilées à des hypothèses régionales^[5]. Le résultat bidimensionnel d'un échelonnement multidimensionnel réalisé sur les 15 items de satisfaction est présenté dans la figure 1.8.

Cette représentation géométrique des relations entre les items permet de reconnaître un radex : la facette (A) joue un rôle radial (modulaire) ; elle est ordonnée, en termes de distance psychologique, du centre vers la périphérie. Près du centre de la configuration se trouvent les items qui évaluent la satisfaction avec les ressources des activités. Les cercles concentriques apparaissant sur la configuration n'ont d'autre but que d'en faciliter la lecture ; ils n'apparaissent pas, bien évidemment, sur les résultats ; c'est au chercheur de faire apparaître cette structure. La facette (B), quant à elle, joue un rôle angulaire (polaire). Les domaines de vie (i.e. les modalités de cette facette) divisent l'espace en régions émanant d'une même origine. Cette facette n'est pas ordonnée, mais l'arrangement circulaire des domaines n'est pas fortuit, même s'il n'est pas toujours très facile d'en comprendre la logique. Cet arrangement rapproche les domaines conceptuellement proches (c'est ainsi, par exemple, que le domaine de l'habitation est

[4] : En fait, dans cet exemple, certains structuples sont représentés par plusieurs items alors que d'autres ne sont pas du tout représentés. Cela apparaîtra clairement dans les résultats présentés.

[5] : Rappelons encore une fois que les modèles d'échelonnement multidimensionnel ne sont pas les seuls modèles qui peuvent être utilisés pour valider un phrase en facettes, même s'ils sont les plus fréquemment utilisés. Pour un exemple d'utilisation de méthodes linéaires appliquées à la validation d'une définition de l'anxiété dentaire, on pourra consulter Stouthard et al. (1993).

voisin du domaine économique ou que le domaine du travail est voisin de celui de l'éducation).

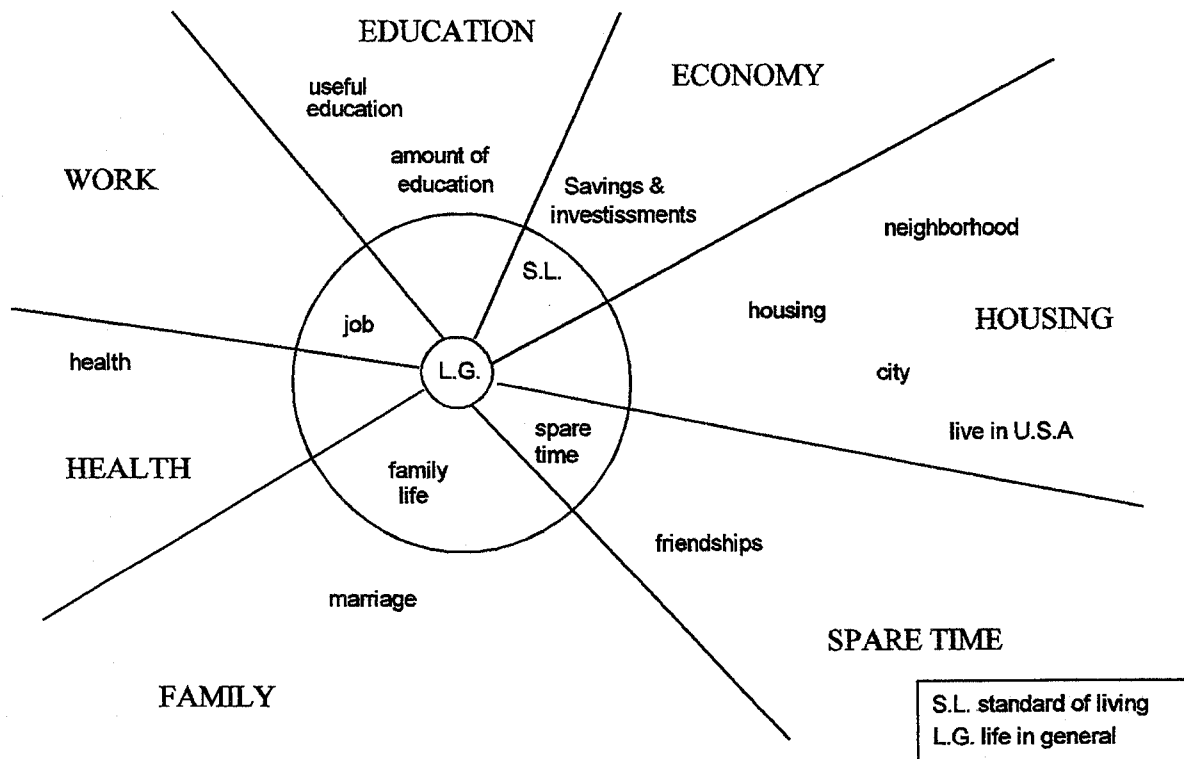


Figure 1.8 : Un radex de la satisfaction vis-à-vis de différents domaines de vie (d'après Levy, 1981b, p. 122) (les items sont reportés en minuscules, les domaines de vie en majuscules).

Ces résultats montrent, selon Levy, que la phrase en facettes qui a été proposée est validée (en fait, on devrait dire qu'elle n'est pas infirmée). En effet, les facettes permettent de partitionner l'espace en régions bien délimitées à l'intérieur desquelles on ne trouve que les items attendus. Ceci montre que l'approche en facettes procède d'une démarche confirmatoire. Comme dans toute démarche confirmatoire, tout résultat non conforme aux hypothèses remet en cause la théorie qui a présidé à leur élaboration.

Cet exemple simple ne doit pas cacher certaines difficultés inhérentes à l'application de l'approche en facettes. Tout en passant en revue ces difficultés, nous essaierons de faire de cette approche une évaluation raisonnée.

3.4 Une évaluation critique de l'approche en facettes

L'approche en facettes est séduisante, tant par ses objectifs que par sa méthode. C'est l'une des seules approches qui abordent de front le problème de l'univers des items (en l'absence d'un listage exhaustif des éléments), en proposant une solution originale qui consiste à s'appuyer sur les sources de variation des items. La méthode associée à cet objectif est en parfaite cohérence avec la démarche scientifique traditionnelle : en soi, elle n'est pas criticable. Plus généralement, Bohrnstedt (1990, cité par Borg & Shye, in press) note que l'approche en facettes peut être assimilée à une méthodologie des plans d'expérience ("*design methodology*") et que personne ne peut être contre les plans d'expérience !

Pourtant, l'approche en facettes ne connaît pas, dans la communauté scientifique, le succès qu'elle paraît mériter. Surtout utilisée en Israël, sous l'impulsion de Louis Guttman, elle reste peu développée en Europe et aux Etats-Unis^[6]. Les raisons de cette situation sont plus à rechercher, à notre avis, dans les problèmes posés par les applications empiriques que dans l'approche théorique générale qui n'est guère sujette à controverses.

Ce qui frappe tout d'abord lorsque l'on examine en détail les applications empiriques qui s'appuient sur cette théorie, c'est le peu d'attention qui est portée aux contenus des items (cf., à ce sujet, Swanborn, 1993). Cette remarque peut paraître paradoxale pour une approche qui se focalise sur les aspects définitoires et sur l'univers des items. Pourtant, on ne peut que constater que de trop nombreuses publications omettent d'annexer la liste des items utilisés. Le plus souvent, seuls quelques items sont fournis à titre d'exemple ou sont simplement repérés par des mots-clés. De plus, dans les rares cas où on dispose des items, leur mise en correspondance avec les structuples théoriques est difficile, même pour un lecteur averti, et quelquefois contestable^[7]. On citera, à titre d'exemple, la facette "modalité comportementale" qui figure dans de nombreuses applications. Cette facette distingue entre les aspects cognitif, affectif et instrumental des conduites. Si la spécificité de la composante instrumentale semble poser le moins de problèmes, il est souvent plus difficile de distinguer *a priori* entre les composantes affective et cognitive^[8]. Dans les études empiriques, les résultats peuvent

[6] : Les seuls pays européens qui dérogent à cette règle sont l'Allemagne, les Pays-Bas et, dans une moindre mesure, l'Angleterre.

[7] : La vérification de l'accord interjuges de la mise en correspondance entre les items et les structuples devrait être un préalable indispensable à toute étude utilisant l'approche en facettes. Malheureusement, cet accord interjuges n'est jamais vérifié.

[8] : Dans certains cas (e.g. Levy & Guttman, 1985), les auteurs eux-mêmes ne parviennent pas à discriminer parfaitement ces trois modalités et formulent des appartenances mixtes pour certains items.

servir d'indices permettant d'affiner cette distinction mais cette démarche ne peut plus alors être qualifiée de confirmatoire.

Une autre difficulté est liée à la démarche analytique qu'implique l'approche en facettes. Dès lors que le nombre de facettes et/ou de modalités augmente, le nombre de structuples, correspondant au produit cartésien des modalités des facettes, croît de manière considérable^[9]. Cette remarque a plusieurs conséquences.

Tout d'abord, il est fréquent que le nombre d'items utilisés soit inférieur au nombre de structuples théoriques. Il est clair que dans ce cas, ce n'est qu'une partie de la phrase en facettes qui pourra être validée par une étude empirique. On peut regretter aussi que les études n'utilisent que trop rarement plusieurs items correspondant à un même structuple. En effet, c'est une chose que de montrer que des items correspondant à des structuples différents occupent des régions distinctes de l'espace. C'en est une autre que de montrer que des items correspondant à un même structuple occupent des régions identiques. Ainsi conduite, la validation d'une phrase en facettes aurait plus de poids.

Globalement, la grande majorité des applications utilise un nombre d'items très réduit. Mais ceci n'est pas propre à la théorie en facettes ; on peut le constater pour toute démarche confirmatoire^[10]. Comme ce type de démarche conduit à invalider la théorie dès qu'un seul item n'est pas conforme aux hypothèses, il est d'autant plus difficile de confirmer un modèle que celui-ci concerne beaucoup d'éléments. Il est vrai aussi que toute théorie qui porte sur les relations entre des items est difficile à valider. Les réponses à un item peuvent être affectées par beaucoup de facteurs qui ne sont pas prévus par le chercheur. En psychométrie, la théorie classique du score vrai a depuis longtemps mis en évidence la part importante d'erreur de mesure associée à un item unique. De plus, chaque item étant la combinaison d'un nombre plus ou moins important de modalités de facettes, des effets interactifs très complexes peuvent se produire et obscurcir de manière notable les solutions empiriques.

Dans le cas où le nombre d'items est important, les solutions deviennent complexes à interpréter. Borg & Shye (*in press*) conseillent, dans ce cas, d'effectuer plusieurs analyses portant sur des sous-ensembles d'items. Il ne s'agit pas seulement d'alléger la représentation géométrique, mais de se donner plus de chances de parvenir à interpréter directement la structure résultante. En effet, il est possible qu'aucune structure évidente n'apparaisse lorsque les items sont projetés dans un plan défini par un certain système

[9] : Ainsi, par exemple, la définition du bien-être proposé par Levy & Guttman (1975) permet d'engendrer 5824 structuples. Dans l'étude empirique qui suit, et sur laquelle nous reviendrons, seuls 24 items sont créés à partir de cette définition.

[10] : L'utilisation de l'analyse factorielle confirmatoire conduit aux mêmes restrictions.

d'axes de coordonnées, mais qu'une structure puisse paraître évidente dans un autre système d'axes. Que l'on pense, par exemple, à un cylindre (cf. figure 1.6 *supra*) occupant un espace à trois dimensions ; si l'axe du cylindre ne correspond à aucun des axes de coordonnées du système, il sera très difficile de le percevoir dans des projections sur des plans orthogonaux^[11]. Ce problème soulève d'ailleurs une question technique fort embarrassante, qui n'a reçu à ce jour aucune réponse satisfaisante. Dans le résultat d'un échelonnement multidimensionnel, seules les distances entre les items sont déterminées par le modèle. A partir de ces distances, et pour faciliter la lecture des résultats, on peut choisir parmi un nombre infini de systèmes d'axes de coordonnées orthogonaux. Théoriquement, ce choix peut être aléatoire puisque n'importe quel système permet, avec la même précision, de localiser les points dans l'espace sans que la distance qui les sépare n'en soit affectée. Mais, si le choix d'un système d'axes ne change rien à la localisation des points dans l'espace, il n'offre pas la même perspective sur les résultats. Dans un espace à deux dimensions, il est assez facile de changer de perspective, simplement en tournant sa feuille de papier. Mais dans des solutions à trois dimensions et plus, la projection des points dans des espaces bidimensionnels orthogonaux impose l'angle de vue. En fait, le choix des axes de coordonnées ne se fait pas totalement au hasard. Les programmes d'échelonnement multidimensionnel utilisent, en général, comme axes de coordonnées, les composantes principales du nuage de points. Pourquoi ? Parce que l'on a remarqué, de manière totalement empirique que, le plus souvent, ce choix était efficace. Autrement dit, ce système d'axes de coordonnées est celui qui permet au mieux de retrouver des structures significatives. Malheureusement, ce n'est pas toujours le cas et rien n'impose que cela le soit.

En plus de ces remarques techniques qui ne remettent pas fondamentalement en cause les applications empiriques de la théorie en facettes, on notera une difficulté théorique concernant les facettes. L'approche en facettes repose largement sur l'identification de sources de variation des réponses. Ces sources de variation ont une importance théorique considérable. La mise en évidence de nouvelles facettes permet d'enrichir ou d'amender d'anciennes théories ou d'en proposer des nouvelles. Or, force est de constater que la productivité de la théorie en facettes en ce domaine est pauvre. Les applications se cantonnent souvent dans un système de facettes largement éprouvé lors de travaux antérieurs. Cette situation peut tenir à la difficulté de mettre en évidence de nouvelles sources de variation à l'aide d'une procédure aussi exigeante que l'approche en facettes ou à un certain manque d'originalité. Mais il paraît certain que la diffusion de cette approche dans la communauté scientifique serait largement facilitée si un intérêt plus grand était porté à cette question.

[11] : Pour des structures supérieures à trois dimensions, on procède *nécessairement*, pour l'interprétation des résultats, à une projection des items sur des plans orthogonaux définis pas les axes de la solution.

Les remarques critiques que nous venons de formuler à propos de l'approche en facettes peuvent paraître bien sévères. Elle reflètent pourtant, à notre avis, l'état des applications empiriques. Ces remarques n'enlèvent rien aux fondements théoriques et épistémologiques qui sous-tendent cette approche. Il n'est pas inutile de réaffirmer que nous les partageons pleinement et que cette thèse s'inscrit tout à fait dans l'esprit de cette approche.

SECONDE PARTIE

LE BIEN-ETRE SUBJECTIF

UNE REVUE CRITIQUE DE LA LITTERATURE

DEUXIEME PARTIE

LE BIEN-ETRE SUBJECTIF : UNE REVUE CRITIQUE DE LA LITTERATURE

INTRODUCTION

Présenter à l'heure actuelle une synthèse de la littérature sur le bien-être subjectif ressemble à une gageure. La dernière revue en date a été publiée il y a exactement dix ans par Ed Diener dans le *Psychological Bulletin*, avec comme objectif de faire le point sur l'évolution des travaux depuis la précédente synthèse due à Wilson en 1967. Diener (1984) rapporte que plus de 700 publications ont été recensées entre ces deux dates (cf. Diener & Griffin, 1984), mais ce chiffre est certainement très en deça de la réalité. On sait en effet que Michalos (1986) a, quant à lui, recensé 2545 publications entre 1979 et 1982 avec les seuls mots-clés bonheur et satisfaction^[1] !

La différence entre ces deux recensions tient sans doute au choix des mots-clés, mais Diener n'est pas très précis sur ce point. Il est vrai que Michalos (1986) indique que parmi ces publications, 50% concernent la satisfaction avec le travail. Mais même si on ne tient compte que des 16% qui concernent la satisfaction vis-à-vis de la vie dans son ensemble, une rapide estimation permet de penser qu'il y a au bas mot plus de 1500 travaux publiés entre 1967 et 1984 sur cette question. Depuis, même si on ne dispose pas d'informations chiffrées, il n'y a aucune raison de penser que le nombre de publications diminue, bien au contraire !

Golberg indiquait, en 1974, que le simple fait de publier sa liste de références sur le MMPI suffirait à dépasser le nombre limite de pages de l'article de synthèse qu'on lui avait demandé sur l'ensemble des instruments de diagnostic de la personnalité. Même si nos contraintes en nombre de pages ne sont pas aussi strictes que celles de Goldberg, nul doute que la liste des travaux sur le bien-être subjectif occuperait un volume d'une épaisseur conséquente.

Notre objectif est pourtant de proposer une vue d'ensemble des travaux sur le bien-être. Cette synthèse a l'ambition d'être exhaustive, mais bien évidemment pas dans le sens

[1] : Le terme bien-être lui-même n'apparaît que depuis une période très récente dans les index des revues bibliographiques spécialisées alors que le terme bonheur existe, par exemple, depuis 1973 dans l'index de la revue "*Psychological Abstracts*". Notons aussi, à titre de comparaison, que dix ans plus tôt, entre 1969 et 1972, on trouve 1238 publications à partir des mêmes mots-clés.

où elle ferait référence à tous les travaux publiés. Il nous semble plus intéressant d'essayer d'aborder l'ensemble de la problématique du bien-être à travers des études qui peuvent être considérées comme prototypiques plutôt que d'accumuler une liste de références dont les résultats spécifiques n'apportent souvent guère d'informations supplémentaires. L'exhaustivité à laquelle nous prétendons se situe donc au niveau des grands thèmes qui se dégagent des recherches sur le bien-être, pas au niveau des publications.

Cette manière de procéder présente de nombreux avantages. Elle offre en particulier une plus grande liberté dans l'exposé et, surtout, elle permet d'ouvrir le champ parfois un peu restreint des études sur le bien-être à des domaines voisins. Nous utiliserons largement ces possibilités car nous verrons qu'il existe des pans entiers de la psychologie scientifique qui étudient de manière presque totalement indépendante des phénomènes très proches de celui du bien-être mais dans des traditions différentes. Cette mise en perspective apporte un éclairage nouveau sur certains aspects débattus dans le cadre du bien-être, et l'inverse est vrai aussi. Il est assez étonnant de constater que ces traditions ont tendance à s'ignorer largement. Nous espérons donc que cette revue de synthèse sera l'occasion d'en faire apparaître les convergences de problématique.

Le plan adopté pour cette synthèse est classique : définition, mesure, structure, corrélat et théorie. Pourtant, ces divisions sont largement arbitraires étant donné les liens qu'elles entretiennent. La mesure du bien-être ne peut être envisagée en dehors de considérations définitoires et théoriques ; les corrélats participent à la construction du réseau nomologique et donc à ce que nous avons appelé la validation interconcepts ; les études structurales relèvent, quant à elles, de la validation intraconcept. C'est pourquoi, bien que cette thèse porte sur la mesure du bien-être, il nous semble indispensable de ne pas nous limiter à un exposé sur les seuls aspects qui relèvent directement de considérations psychométriques au sens restreint. Nous verrons d'ailleurs que les modèles explicatifs mis au point dans le cadre du bien-être soulèvent des questions fondamentales quant à sa conceptualisation générale et donc quant à l'élaboration d'une théorie définitoire.

Cette synthèse ne se veut pas simplement narrative. Nous essaierons d'adopter une démarche critique en soulignant les difficultés soulevées par tel ou tel résultat ou par telle ou telle conception. La confrontation d'approches différentes sera l'occasion de faire apparaître les convergences mais aussi les contradictions et les postulats implicites adoptés dans les recherches sur le bien-être. Au-delà des critiques spécifiques, nous nous interrogerons aussi sur la pertinence de certains choix méthodologiques et conceptuels.

Le premier chapitre est consacré aux aspects définitoires du bien-être et, plus généralement, à tout ce qui a trait à sa conceptualisation. Nous débiterons par un rapide rappel de considérations philosophiques et sémantiques qui entourent ce concept avant

d'examiner plus en détail les conceptions psychologiques auxquelles il peut être rattaché. Les rares tentatives de synthèses théoriques et/ou empiriques seront ensuite discutées ainsi que les études ayant proposé d'aborder le bien-être à partir de ses représentations. Cette revue nous conduira à la question fondamentale des liens entre définition, mesure et théorie.

Le second chapitre est principalement instrumental. Il se propose de présenter les instruments mis au point pour mesurer le bien-être, de mettre en lumière leur logique et d'en évaluer les qualités psychométriques. Une partie importante de ce chapitre sera consacrée aux critiques émises à l'encontre des mesures de bien-être.

Le troisième chapitre se propose d'étudier le bien-être en termes structuraux. Nous dégagerons tout d'abord les principaux résultats rapportés dans la littérature à travers la description d'études prototypiques. Cette analyse permettra de s'apercevoir que la structure du bien-être a principalement été envisagée à partir de la distinction entre affectivité positive et affectivité négative d'une part et de la distinction entre composantes affectives et cognitives d'autre part. Les travaux concernant ces deux couples de dimensions feront donc l'objet d'une analyse détaillée.

Le quatrième chapitre présentera, sous une forme très synthétique, les corrélats du bien-être. A elle seule, cette question pourrait faire l'objet de plusieurs volumes. Nous avons plutôt choisi de résumer les résultats à partir de revues de synthèse existantes et de porter plus d'attention aux problèmes plus généraux que pose ce type d'études.

Enfin, le dernier chapitre est entièrement consacré aux modèles explicatifs développés pour le bien-être. Ces modèles sont extrêmement nombreux et il s'avèrera indispensable d'essayer d'en organiser la présentation. Nous proposerons donc un système classificatoire original qui permet de situer clairement les modèles les uns par rapport aux autres. C'est en suivant ce schéma que seront décrits et critiqués les différents modèles. Cette revue critique nous permettra notamment de montrer que l'une des familles de modèles les plus répandues repose sur des hypothèses qu'il est impossible de réfuter. Au-delà de la spécificité de chacun des modèles, il apparaîtra que le sens de la causalité entre le bien-être et les variables incluses dans le réseau nomologique constitue une question fondamentale sur le plan théorique qui méritera une attention toute particulière. Enfin, nous terminerons ce chapitre par une présentation d'un modèle général qui est sans doute le plus plausible en l'état actuel des connaissances et qui permet de répondre à plusieurs critiques soulevées antérieurement.

CHAPITRE I

DEFINITIONS ET CONCEPTUALISATIONS GENERALES DU BIEN-ETRE SUBJECTIF

1.1 Introduction

Alors que les questions concernant la mesure ou les modèles théoriques du bien-être subjectif commencent à recevoir des réponses empiriques relativement satisfaisantes, c'est la définition du bien-être qui paraît poser le problème le plus sérieux. Ce premier chapitre entend traiter de ces questions conceptuelles à un niveau très général, étant entendu que les discussions concernant certains points spécifiques trouveront leur place dans les prochains chapitres.

Nombreux sont les auteurs qui réclament une clarification conceptuelle de la notion. Les citations suivantes en témoignent de manière très claire^[1] :

- "(...) nous avons désespérément besoin d'une plus grande clarification conceptuelle de tout ce que l'on regroupe sous le terme bien-être." (Lawton et al., 1984, p. 68).
- "Le problème le plus critique que rencontrent les utilisateurs des différentes mesures et construits du bien-être est le besoin de clarification et de consistance des concepts." (George, 1981 cité par Horley, 1984, p. 124).
- "Le manque de modèles conceptuels est un problème continu dans la recherche sur le bien-être subjectif (...)." (Stock et al., 1986, p. 91).

Malgré ces appels répétés, la situation ne s'est guère améliorée. Alors que l'on aurait pu attendre que l'accumulation des résultats empiriques permette d'en clarifier la signification, on a plutôt l'impression que ceux-ci n'ont fait qu'amplifier le désarroi : "l'histoire des recherches sur le bien-être est l'histoire de la confusion" écrit Veenhoven (1991 a, p. 8).

Il faut toutefois reconnaître que les chercheurs n'ont pas toujours su tirer profit de ces nombreuses études empiriques. Certains résultats sont suffisamment établis, à l'heure actuelle, pour qu'ils aient une influence décisive sur l'élaboration d'une théorie définitoire du bien-être. Malheureusement, les étapes de définition, de mesure et de théorisation

[1] : C'est nous qui traduisons.

(nomologique) sont encore trop souvent conçues comme des étapes séparées. De plus, la majorité des travaux continue à utiliser des instruments de mesure élaborés, pour beaucoup, il y a plus de vingt ans, à partir d'une conceptualisation implicite du bien-être, différente de celle que l'on pourrait avoir aujourd'hui. A cela s'ajoute le fait que les recherches sur le bien-être s'inscrivent dans des traditions caractérisées par des terminologies, des techniques et des modèles différents et que la communication entre ces courants n'est pas toujours très aisée.

Ces difficultés rendent très hasardeuse toute tentative de clarification conceptuelle. Il nous semble pourtant que ce problème mérite que l'on s'y attarde. A travers la littérature, on peut mettre en lumière plusieurs lignes d'investigation ayant comme dénominateur commun une réflexion générale sur le concept de bien-être. Certains de ces travaux sont relativement spécifiques, d'autres sont beaucoup plus diffus. On peut les regrouper selon qu'ils :

- effectuent une analyse philosophique et/ou sémantique du concept de bien-être ;
- essaient de rattacher le bien-être à des concepts psychologiques de base comme les émotions, les attitudes ou les traits de personnalité ;
- proposent une synthèse théorique ou empirique ;
- abordent le bien-être à partir des représentations.

C'est à l'examen de ces travaux que nous allons nous attacher avant d'essayer d'en tirer les principaux enseignements.

1.2 Une analyse philosophique et sémantique du bien-être

Parmi les tentatives de clarification conceptuelle du bien-être, deux contributions originales sont à noter : celle de Tatarkiewicz (1976) et celle de Veenhoven (1984).

L'ouvrage de Tatarkiewicz intitulé "L'analyse du bonheur" présente sans doute la synthèse la plus complète de tous les travaux philosophiques, littéraires et "pré-scientifiques" sur le bien-être^[2]. L'analyse linguistique et philosophique menée par l'auteur lui permet de distinguer quatre sens différents du mot bonheur. Le premier fait référence à un état temporaire de joie intense. Le second insiste sur la notion de satisfaction : il est fait référence à la vie en général et non plus à un épisode particulier ;

[2] : Ecrit pendant la seconde guerre mondiale en Pologne, cet ouvrage n'a été publié en anglais qu'en 1976. Il ne couvre donc pas les travaux psychologiques et psychosociologiques menés après la guerre.

c'est dans ce cas un jugement, plus qu'une émotion, écrit Tatarckiewicz (1976, p. 2)^[3]. Le troisième sens apparaît surtout dans les usages linguistiques de différents pays. Dans les langages slaves et allemands, le mot bonheur désigne également la chance alors que le latin, le français ou l'anglais ont pour ces deux sens des mots différents. Enfin, le quatrième sens s'origine dans la philosophie grecque. Le bonheur, chez Aristote ou chez Platon, c'est l'"*eudaemonia*" (*εὐδαιμονία*) qui signifie la possession des plus grands biens disponibles pour l'homme. Le bonheur doit donc s'apprécier par la quantité des biens possédés. La nature de ces biens diffère selon les auteurs ; chez certains, il s'agit de biens moraux, chez d'autres, de biens hédonistes, chez d'autres encore, d'un mixte des deux. Ce qui est intéressant, dans cette dernière conception, c'est qu'elle propose une approche objective du bien-être (ou normative, comme le dit Diener, 1984). On peut lui opposer une approche subjective dont la première formulation est sans doute due à Démocrite, pour qui une vie heureuse est une vie qui procure de la joie et de la satisfaction. Démocrite préfère utiliser des termes comme bien-être, satisfaction ou bonne humeur, qui impliquent moins les circonstances : l'important n'est pas ce qu'un individu possède, comme dans la conception d'Aristote, mais la manière dont il réagit à ce qu'il possède. Pour Tatarckiewicz, l'idée d'un bonheur objectif a persisté, quasi identiquement, jusqu'au XVII^{ème} siècle^[4]. Ce n'est que depuis, et cela rejoint les conceptions actuelles, que le bonheur acquiert une connotation plus subjective.

A ces considérations philosophiques, Veenhoven (1984) ajoute des réflexions concernant le concept de bonheur dans le langage courant et dans les sciences sociales. A partir des définitions issues de dictionnaires de différentes langues européennes, différents sens peuvent être mis en évidence. Ces définitions insistent tantôt sur la notion de bonne fortune^[5], de prospérité, tantôt sur l'idée d'une expérience agréable ou d'une émotion positive intense. Certaines accentuent l'aspect temporaire du bonheur alors que d'autres considèrent un état plaisant plus permanent (cf. Veenhoven, 1984, p. 14). L'analyse des conceptions scientifiques permet de dégager plusieurs distinctions significatives dont certaines apparaissent déjà dans les analyses précédentes. Les définitions objectives qui insistent sur les conditions nécessaires à une "bonne vie" s'opposent aux définitions subjectives qui font référence à l'expérience, à l'appréciation des individus. Parmi les définitions subjectives, on peut distinguer celles qui visent à décrire les

[3] : Au XVIII^{ème} siècle, en France, on distinguait ces deux sens. On réservait au premier le terme félicité et au second le terme bonheur. Il en était de même en Allemagne (*Glückseligkeit* vs *Glück*).

[4] : Il est d'ailleurs intéressant de noter que dans les idées utopistes, comme la république de Platon, le bonheur individuel n'est jamais le but. Il s'agit plutôt de construire une société idéale basée sur un critère de "ce qui est bien" (cf. Tatarckiewicz, 1976, p. 285). Il en est de même dans l'utopie de Thomas More au début du XVI^{ème} siècle : ses idées s'exprimaient en termes de morale, pas en termes de satisfaction individuelle. La création d'une société idéale ne laisse que peu de place au bonheur individuel. Ce n'est qu'au XIX^{ème} siècle qu'apparaissent des utopies dont l'objectif est de rendre les personnes heureuses.

[5] : Comme l'étymologie du terme français le suggère.

caractéristiques de l'expérience de bonheur et celles qui se proposent d'expliquer les conditions de sa réalisation. Ces dernières partagent certaines caractéristiques avec les définitions objectives, ce qui amène l'auteur à regretter une relative confusion entre définition et prédiction du bonheur. Nous reviendrons plus loin sur cet aspect. Les définitions diffèrent aussi selon qu'elles font référence à une expérience affective, en grande partie inconsciente et automatique, ou à une expérience cognitive qui nécessite une activité consciente de jugement, d'évaluation. Une autre différence concerne le niveau de spécificité véhiculée par la définition : il y a un bonheur général et des bonheurs plus spécifiques. Enfin, la dernière distinction reprend une conception d'un bien-être momentané ou durable.

Nous verrons, au fil du texte, que les sens différents du bonheur dégagés par Veenhoven et Tatarikiewicz n'ont pas qu'un intérêt historique. La plupart de ces distinctions sont toujours d'actualité dans la littérature scientifique actuelle. Le prochain paragraphe en donnera déjà quelques exemples.

1.3 Le bien-être comme émotion, comme attitude, comme trait de personnalité

De nombreux auteurs ont assimilé le bien-être à l'un des trois concepts psychologiques de base que sont les émotions, les attitudes ou les traits de personnalité. Chez certains, cette assimilation n'est qu'implicite ; chez d'autres, elle est beaucoup plus explicite et peut être considérée comme une tentative de clarification conceptuelle.

Le bien-être a été rattaché aux théories sur l'émotion et, plus particulièrement, aux théories cognitives de l'émotion (e.g. Cherlin & Reeder, 1975 ; Stock et al., 1986). Nous verrons notamment, dans le chapitre sur la structure du bien-être, que de nombreuses études sur l'émotion ou sur l'affectivité sont directement pertinentes pour l'étude du bien-être.

Le bien-être est souvent aussi conçu comme une attitude (cf. Levy & Guttman, 1975 ; Andrews & McKennell, 1980 ; Andrews & Robinson, 1991). Cela n'est pas sans fondement à la lumière de certaines définitions des attitudes^[6]. Malgré leurs différences, les définitions et les conceptions des attitudes partagent pratiquement toutes l'idée que la dimension évaluative est la principale caractéristique d'une

[6] : Cf., par exemple, Ajzen (1989, p. 241) : "Une attitude est une disposition individuelle à répondre de manière favorable ou défavorable envers un objet, une personne, une institution, un événement ou envers n'importe quel aspect discriminable du monde d'un individu" ou Petty & Cacioppo (1981, cités par Breckler & Wiggins, 1989b) : une attitude est "une sensation positive ou négative générale et durable à propos d'une personne, d'un objet ou d'un thème."

attitude. Il deviendra clair au fil du texte que cette dimension est également fondamentale pour le bien-être.

Pour d'autres, enfin, le bien-être a lieu d'être considéré comme un trait de personnalité. C'est ainsi, par exemple, que dans leur classification conceptuelle des traits de personnalité, Buss & Finn (1987) définissent le bien-être comme un trait affectif, non-social et dirigé vers soi. Toujours dans le domaine de la personnalité, certains considèrent même que le bien-être doit être assimilé à des dimensions beaucoup plus générales que les traits. Stones & Kozma (1986a, 1986b) parlent, à ce sujet, d'une "propension" qui influencerait les variables attitudinales et comportementales de niveau plus spécifique (satisfaction, croyances à propos de son efficacité personnelle, styles de comportements susceptibles d'induire des renforcements positifs de la part de l'environnement, etc.). Dans le modèle de la personnalité de Tellegen (e.g. Tellegen et al., 1988), le bien-être – appelé dans ce cas émotionnalité positive – constitue l'un des trois grands facteurs de second ordre et influence de multiples autres traits.

Malgré des aspects séduisants, aucune de ces trois grandes lignes de pensée n'est entièrement satisfaisante. En effet, on peut déjà s'interroger sur la nature même des concepts d'émotion, d'attitude ou de trait, chaque assimilation du bien-être à l'un ou l'autre posant ses problèmes propres. On peut aussi s'interroger sur les similitudes et les différences entre ces concepts.

Il ne suffit pas de considérer que le bien-être est une émotion pour que sa signification soit précisée, d'autant que le concept d'émotion lui-même suscite de nombreuses questions. En fait, il n'y a pas non plus d'accord sur la définition de termes comme émotion, affect, sentiment ou humeur. Si tous ces termes renvoient au très général "champ affectif" (cf. Cosnier, 1994), la frontière qui les sépare est floue. Pour Forgas (1994, p. 3), l'affect fait référence à la fois aux humeurs et aux émotions. L'humeur est un état affectif d'intensité faible, diffus et relativement durable n'ayant pas de cause saillante et donc pas de contenu cognitif. A l'inverse, une émotion est plus intense, plus courte, a une cause bien définie et un contenu cognitif clair. On trouve une conception qui s'en rapproche chez Cosnier (1994, p. 14), pour qui les émotions désignent l'ensemble des "événements ou états du champ affectif qui se caractérisent par un ensemble d'«éprouvés» psychiques spécifiques accompagnés, de façon variable en intensité et en qualité, de manifestations physiologiques et comportementales". Les éprouvés psychiques constituent, pour Cosnier, les affects qui sont définis comme "des éprouvés subjectifs qualifiant une représentation, une situation, une relation ou un état mental et/ou corporel. Pris dans un sens général, les affects comprennent les émotions, les sentiments et les humeurs." (ibidem, p. 160). Il semble donc y avoir un accord parmi les spécialistes des émotions au moins sur le fait que l'humeur est moins intense, plus diffuse que l'émotion

(cf. Taylor, 1991 ; Ekman, 1992 ; Cosnier, 1994 ; Forgas, 1994). A partir de ces définitions, on peut donc se demander si le bien-être n'est pas plus proche du concept d'humeur que de celui d'émotion. De plus, Clore et al. (1987) notent que, si toutes les émotions sont des sentiments et des conditions affectives, l'inverse n'est pas vrai et que l'affectivité renvoie à une évaluation de la signification d'un objet en termes de valence. On retrouve ici la dimension évaluative dont nous avons déjà souligné l'importance à propos de la conception du bien-être en tant qu'attitude. Mais la question soulevée par cette remarque est plus fondamentale : en réduisant le bien-être à une émotion qui n'est qu'une condition affective parmi d'autres, on aboutit à une conceptualisation extrêmement étroite qui devient incapable de rendre compte d'une bonne partie des régularités empiriques observées.

Dans le cas où le bien-être est conçu comme une attitude, deux questions, au moins, se posent. Tout d'abord, une attitude s'accompagne toujours de référents comportementaux^[7]. Or, dans le cas du bien-être, ceux-ci n'ont pratiquement jamais été étudiés (cf. Andrews & McKennell, 1980). On se doute que certaines manifestations telles que le sourire, un timbre de voix agréable y sont liées (cf. Veenhoven, 1984, p. 39). Néanmoins, il n'est pas sûr que ce soient des comportements dirigés vers un objet comme le sont ceux qui sont traditionnellement étudiés dans les travaux sur l'attitude. La seconde question concerne l'objet de l'attitude lui-même. Levy & Guttman (1975) ont discuté en détail du statut de cet objet dans le cas du bien-être, ce qui les conduit à s'interroger sur l'existence d'un objet d'attitude unique. A priori, l'objet le plus évident est le "soi". Pour Greenwald (1989 a), le soi constitue d'ailleurs l'objet d'attitude le plus important pour la plupart des gens. Mais l'attitude à l'égard de soi et des caractéristiques du soi définit l'estime de soi (cf. Shavitt, 1989), pas le bien-être. Si le concept de bien-être ne peut être réduit à l'estime de soi^[8], tout comme il ne pouvait être réduit à une émotion, il faut soit trouver un objet d'attitude qui, tout en étant plus large que le soi réussisse à l'intégrer, soit considérer que le bien-être est associé à plusieurs objets d'attitude différents. Dans un cas comme dans l'autre, il n'existe pas, à notre connaissance, de proposition satisfaisante.

Enfin, le bien-être n'échappe pas, lorsqu'il est assimilé à un trait de personnalité, au débat trait-état. Autrement dit, le bien-être caractérise-t-il un attribut d'une personne ou un attribut d'une personne dans une situation donnée (cf. Kane, 1982) ? En reprenant les distinctions faites habituellement entre traits et états (cf. Fridhandler, 1986), cette

[7] : La question des comportements se pose aussi dans le cas où le bien-être est conceptualisé comme un trait de personnalité. Les traits de personnalité sont habituellement définis comme des construits psychologiques hypothétiques, censés rendre compte des covariations observées entre les comportements (e.g. Rorer & Widiger, 1983). Dans le cas du bien-être, la question est de savoir de quelles régularités comportementales il s'agit.

[8] : Certaines études se sont attachées à étudier les relations empiriques entre bien-être et estime de soi (cf. Dobson et al., 1979 ; Headey et al., 1985 ; McNeil et al., 1986 a), mais aucune d'entre elles n'a abordé cette question en termes de similarité conceptuelle.

question peut être précisée : le bien-être a-t-il une durée longue ou courte ? S'agit-il de manifestations continues ou réactives ? S'agit-il d'une entité abstraite ou concrète, accessible à l'observation ? Est-on en présence d'une causalité personnelle ou situationnelle ? Bien évidemment nous ne répondrons pas à ces questions dans ce paragraphe ; elles nécessitent un examen approfondi qui passe par l'analyse critique de nombreux résultats empiriques. Nous y reviendrons donc. Nous noterons simplement pour l'instant que le débat n'est pas clos. Si on excepte les auteurs qui prennent clairement position pour l'une ou l'autre conception, on peut noter que pour certains (e.g. Headey et al., 1989), le bien-être est trop stable pour être un état, mais trop peu stable pour être un trait, alors que pour d'autres, le bien-être comprend à la fois une composante à long terme et une composante à court terme (e.g. Kammann & Flett, 1983 ; Kozma et al., 1990 ; Watson, 1988a).

En ce qui concerne les différences entre les concepts d'émotion, d'attitude et de trait, il semble qu'elles relèvent plus de circonstances historiques que d'arguments conceptuels. La littérature présente de plus en plus de commentaires allant dans ce sens.

C'est le cas pour les attitudes et les traits. Ajzen (1987), par exemple, note que la distinction entre traits et attitudes repose sur une histoire et une tradition différentes^[9], mais que, sur le plan conceptuel, il est très difficile de les distinguer. A titre illustratif, considérons les quatre arguments utilisés par Greenwald (1989a) pour justifier l'importance du concept d'attitude. Selon lui, les attitudes sont omniprésentes, elles permettent de prédire le comportement, elles constituent une force sélective pour la perception et la mémoire et elles ont des fonctions psychiques variées (adaptation sociale, évaluation des objets, défense de l'ego, etc.). Si on remplaçait, dans les phrases qui précèdent, le terme "attitudes" par "traits de personnalité", il est peu probable qu'un spécialiste de la personnalité s'y oppose.

Les ressemblances entre attitude et émotion sont également soulignées par Greenwald (1989b). Ces concepts, mais aussi les affects, les pulsions, les besoins, les valeurs, etc., sont tous, selon cet auteur, des concepts motivationnels qui ne diffèrent que quant au niveau de généralité des éléments auxquels ils s'appliquent. L'attitude est le concept le plus large : il s'applique au niveau des objets (e.g. un ami, une automobile), des catégories (e.g. les automobiles, les serpents), des propositions (e.g. utiliser des moyens contraceptifs) et des schémas (e.g. la théorie psychanalytique). Les émotions ne concerneraient que les objets et les schémas alors que les affects ne s'appliqueraient qu'au niveau le plus élémentaire que constituent les caractéristiques des objets (leur couleur, leur brillance, etc.).

[9] : Respectivement la psychologie différentielle de la personnalité et la psychologie sociale, dont d'autres ont aussi noté, à un niveau plus général, les convergences (cf. Blass, 1984).

Les émotions et les traits de personnalité connaissent aussi des rapprochements conceptuels et empiriques. Plusieurs auteurs ont noté l'identité structurale entre ce qui est habituellement étudié sous le nom d'états émotionnels ou d'états affectifs et certains traits de personnalité (cf. Larsen & Ketelaar, 1991). C'est ainsi, par exemple, que Watson & Tellegen (1985) montrent qu'aux dimensions de base des états affectifs correspondent deux facteurs primaires de la personnalité ; Meyer & Shack (1989) réussissent à superposer la structure en circumplex de l'affectivité avec deux structures bidimensionnelles de la personnalité, celle de Eysenck (introversion-extraversion et névrosisme) et celle de Gray (1981) qui distingue l'impulsivité et l'anxiété. Cette identité s'exprime également à travers les opérations de mesure : Watson (1988a, cf. aussi Meyer & Shack, 1989) montre que les deux facteurs de base de l'expérience émotionnelle peuvent être mesurés comme des traits de personnalité. Certains travaux assimilent d'ailleurs de manière plus ou moins explicite le bien-être au niveau moyen de l'humeur sur une période temporelle suffisamment longue (cf., entre autres, Underwood & Froming, 1980 ; Diener & Emmons, 1985).

On peut conclure, au moins provisoirement, de ces remarques, que le bien-être ne se réduit ni à une émotion, ni à une attitude, ni à un trait de personnalité. Il partage toutefois avec ces concepts des caractéristiques et, selon le cas, il peut être heuristique d'envisager le bien-être selon l'une ou l'autre de ces perspectives. En tout état de cause, si cela ne résoud pas tous les problèmes conceptuels, cela fait apparaître la richesse du concept de bien-être.

1.4 Synthèses théoriques et/ou empiriques

Les tentatives de synthèses théoriques et/ou empiriques sont rares et, mises à part peut-être les contributions de Ryff (1989a, 1989c) dont il sera question plus bas, elles n'apportent guère d'éléments nouveaux par rapport à ce qui a été dit.

Parmi ces tentatives, celle de Stones & Kozma (1980 a) s'inscrit délibérément dans une volonté de clarification conceptuelle. Ces auteurs ont confronté des informations provenant de sources diverses : définitions du bien-être trouvées dans des dictionnaires, analyse de contenu (sommaire) des questionnaires utilisés pour mesurer ce concept, corrélations empiriques entre échelles de bien-être et analyse de la littérature philosophique. Ils concluent à la difficulté qu'il y a de distinguer théoriquement et empiriquement les différentes composantes du bien-être. Ils proposent cependant de différencier ce qu'ils assimilent à une affectivité momentanée et une composante plus dispositionnelle qui serait la satisfaction des désirs importants dans la vie.

La synthèse de Lawton (Lawton, 1983a, 1983b ; Lawton et al., 1984) articule revue de la littérature et études empiriques. L'analyse théorique de ces auteurs leur permet de proposer le méta-concept de "bonne vie" (*good life*) pour rendre compte des liens qui peuvent être dégagés entre les concepts habituellement considérés comme relevant du bien-être :

- l'environnement objectif dans lequel vit le sujet, qui peut être décrit à partir d'indicateurs objectifs de qualité de la vie ;
- la compétence comportementale qui s'exprime à travers la qualité adaptative des comportements, telle qu'elle est évaluée par des standards sociaux dans les domaines de la santé, de la sensation et de la perception, du comportement moteur et de la cognition ;
- la qualité de vie perçue qui transite par une auto-évaluation de sa satisfaction dans différents domaines de vie ;
- le bien-être psychologique qui est une évaluation subjective de la qualité générale de son expérience intérieure et qui peut être décomposé de la manière suivante :
 - * névrosisme ou affectivité négative (dépression, agitation, pessimisme, anxiété) ;
 - * affectivité positive (sentiments de plaisir, description d'un état émotionnel positif limité dans le temps) ;
 - * bonheur (jugement plus cognitif que le précédent, qui couvre une durée plus longue) ;
 - * congruence entre buts atteints et buts désirés.

Les études empiriques (cf. Lawton et al., 1984) qui accompagnent cette synthèse théorique ne confirment que très partiellement cette décomposition. Après avoir rassemblé 144 items issus de la littérature, les auteurs les soumettent à une analyse factorielle dont les résultats, combinés avec une analyse de la similitude conceptuelle entre les items, permettent de dégager 12 facteurs relativement homogènes. Une analyse factorielle de second ordre est alors menée, qui met en évidence deux facteurs interprétés, le premier comme un facteur interne de bien-être psychologique et le second comme un facteur externe de qualité de la vie. En fait, ces deux facteurs ne font que distinguer, ainsi que le notent McNeil et al. (1986b), des aspects objectifs et des aspects subjectifs du bien-être. Il existe d'autres études ayant procédé par analyses structurales de plusieurs échelles de bien-être. Toutes concluent à l'existence d'un facteur général (e.g. Kammann et al., 1984 a). Pour Stones & Kozma (1985), cela indique que lorsque l'on cherche un facteur général de bien-être, on le trouve !

Les études de Carol Ryff (1989a, 1989c) partent du constat que les travaux sur le bien-être n'ont que peu de bases théoriques. Pourtant, il existe toute une littérature qui

s'est longuement intéressée au fonctionnement mental positif, que l'auteur assimile au bien-être, même si elle a toujours été négligée. L'objectif de Ryff est donc de s'appuyer sur cette littérature afin de proposer un modèle synthétique du bien-être, en en décrivant les caractéristiques essentielles. Les référents théoriques utilisés sont divers. On peut les regrouper dans trois grandes catégories. La première porte sur des théories développementales du cycle de vie. On y trouve la notion de stades psycho-sociaux de Erikson, les tendances de base de la vie de Bühler et la description du processus d'intériorisation et du processus exécutif de la personnalité de Neugarten. Dans la seconde, qui regroupe différentes théories cliniques relatives à la croissance personnelle, on rencontre l'actualisation de soi de Maslow, la "personne fonctionnant pleinement" (*fully functioning person*) de Rogers, le processus d'individuation de Jung et de von Franz, et la conception de la maturité chez Allport. Enfin, la littérature sur la santé mentale constitue la dernière catégorie : elle comprend surtout la contribution de Jahoda qui a exposé six critères de la santé mentale positive.

L'analyse de cette littérature théorique permet à Ryff d'identifier six thèmes principaux qui se retrouvent en général dans différentes conceptions, sans que l'ensemble des six ne soit présent dans une seule conception. Ces thèmes constituent les six dimensions que retient l'auteur dans son modèle du bien-être. C'est par la caractérisation d'individus hypothétiques se situant aux pôles opposés de chaque dimension que chacune est définie. Nous reprenons cette présentation ici, en la simplifiant (cf. tableau 2.1).

Ce modèle a permis de mettre au point un questionnaire de 120 items (Ryff, communication personnelle) qui semble, en première analyse, montrer des qualités psychométriques satisfaisantes (Ryff, 1989c). Malheureusement, ce modèle conceptuel du bien-être n'est pratiquement jamais cité dans la littérature ; il souffre donc d'un manque de travaux empiriques pour en assurer la validation. Il est clair pourtant, et nous y reviendrons, que cette étude théorique permet d'enrichir notablement certaines conceptions un peu étroites du bien-être. La pertinence de plusieurs de ces dimensions apparaîtra d'ailleurs dans certaines de nos études empiriques (cf. troisième partie).

DIMENSION	Pôle positif	Pôle négatif
Acceptation de soi	a une attitude positive envers soi ; accepte et apprécie les multiples aspects du soi, qu'ils soient positifs ou négatifs ; attitude positive envers le passé	se sent insatisfait avec soi ; désappointé par rapport au passé ; troublé par rapport à certaines de ses qualités personnelles ; souhaite être différent de ce qu'il est
Relations positives avec autrui	a des relations satisfaisantes, chaleureuses et de confiance avec autrui ; est concerné par le bien-être des autres ; est capable d'empathie, d'affection et d'intimité ; manifeste de la compréhension dans les relations humaines	a peu de relations intimes et de confiance en autrui ; trouve qu'il est difficile d'être chaleureux, ouvert et intéressé par autrui ; isolé et frustré dans les relations interpersonnelles ; ne veut faire aucun compromis pour entretenir des liens avec autrui
Autonomie	auto-déterminé et indépendant ; capable de résister aux pressions sociales afin de penser et d'agir d'une certaine manière ; auto-régule son propre comportement ; s'évalue par rapport à ses standards personnels	attache de l'importance aux attentes et aux évaluations des autres ; se base sur le jugement d'autrui pour prendre des décisions importantes ; se conforme aux pressions sociales pour penser et pour agir
Maîtrise de l'environnement	a un sentiment de maîtrise et de compétence dans la gestion de son environnement ; contrôle un réseau complexe d'activités externes ; utilise efficacement les opportunités du milieu ; capable de choisir ou de créer des contextes adaptés à ses besoins personnels et à ses valeurs	a des difficultés à gérer ses affaires quotidiennes ; se sent incapable de changer ou d'améliorer le milieu dans lequel il est ; n'a pas conscience des opportunités du milieu ; manque de contrôle sur le monde externe
Sens de la vie	a des buts dans la vie ; trouve une signification dans la vie présente et passée ; maintient des croyances qui donnent du sens à la vie ; a des buts et des objectifs	n'accorde aucune signification à la vie ; a peu de buts et d'objectifs ; n'accorde pas de sens à sa vie passée ; n'a pas de croyance qui donne une signification à la vie
Croissance personnelle	a le sentiment de se développer ; ouvert à de nouvelles expériences ; a le sentiment de réaliser son propre potentiel ; perçoit des améliorations chez soi et dans son comportement au cours du temps ; change dans le sens d'une meilleure connaissance de soi et d'une meilleure efficacité	a le sentiment d'une stagnation personnelle ; n'a pas le sentiment d'amélioration ou d'évolution au fil du temps ; se sent incapable d'adopter de nouvelles attitudes ou de nouveaux comportements ; se sent triste et désintéressé par rapport à la vie

Tableau 2.1 : Description des six dimensions du bien-être d'après Ryff (1989a)

1.5 Les représentations du bien-être

Différents arguments peuvent être évoqués pour justifier le recours aux représentations que se font les individus d'un concept. Il ne s'agit pas, bien évidemment, de remplacer une psychologie scientifique par une psychologie naïve ou populaire. Toutefois, outre l'intérêt qu'il y a d'étudier les représentations du bien-être pour elles-

mêmes, il nous semble que celles-ci peuvent faire apparaître des éléments qui, conjugués avec les connaissances scientifiques, participent à une entreprise de définition et de clarification conceptuelle.

Certains auteurs (e.g. Thomas & Chambers, 1989) argumentent en faveur d'une approche phénoménologique du bien-être. L'utilisation de démarches plus qualitatives que celles qui sont habituellement utilisées doit permettre d'appréhender de manière plus fine un concept aussi subjectif que le bien-être. C'est une idée semblable que l'on retrouve chez Ryff (1989b) lorsqu'elle se demande si les indicateurs qui guident les recherches sur le bien-être correspondent effectivement à ce que les individus considèrent comme pertinents pour leur propre bonheur. A ce sujet, on trouve chez Felton (1987) un début de confirmation empirique de l'hypothèse selon laquelle certaines variations dans les niveaux de bien-être en fonction de l'âge pourraient être dues à des changements dans la signification que les individus attachent au bien-être. Ryff (1989b) considère aussi que le recours aux représentations peut permettre d'identifier de nouvelles composantes du bien-être, pertinentes au niveau théorique, mais qui auraient été négligées dans les travaux empiriques. De manière plus générale, on peut noter aussi tout l'intérêt qu'il y a à connaître les informations sur lesquelles les individus se basent pour décider s'ils sont plutôt heureux ou non. C'est d'ailleurs dans cette large perspective que s'inscrivent les travaux qui visent à comprendre comment se construit la connaissance de soi (cf. Klein & Loftus, 1993). Enfin, il n'est guère besoin de rappeler l'importance qu'a eu, dans l'émergence du modèle de la personnalité à cinq facteurs (le "*big-five*") l'approche lexicale, qui se fonde partiellement sur les représentations. C'est grâce à la convergence entre cette approche lexicale et des approches qui ne se basent pas sur les représentations que le consensus actuel sur ce modèle de la personnalité a pu apparaître (cf., au sujet de l'approche lexicale, John et al., 1988 ; John, 1990 ; Goldberg, 1993).

Malgré l'intérêt potentiel que présentent les études sur les représentations, elles sont peu nombreuses dans la littérature. De plus, elles sont, pour la plupart, tellement spécifiques que les informations dont on peut disposer ne sont pas généralisables. C'est ainsi que l'étude de Thomas & Chambers (1989) porte uniquement sur des personnes âgées et que l'entretien est dirigé de telle manière que les thèmes dégagés doivent plutôt être assimilés au mal-être qu'au bien-être. Les questions ayant servi au recueil des observations sont rarement les mêmes et sont parfois très indirectement liées au bien-être. Ryff (1989b), par exemple, pose plusieurs questions ouvertes dont aucune ne contient le mot bien-être^[10]. Le plus souvent, on ne dispose pas de la totalité des réponses, mais seulement des thèmes les plus fréquents. C'est le cas de l'étude de Thomas & Stock (1988) qui a procédé par évocations libres en utilisant "bonheur" comme mot inducteur ; c'est le

[10] : Parmi ses questions, on trouve notamment : "Qu'est-ce qui est le plus important pour vous dans votre vie en ce moment ?" ; "Qu'est-ce que cela signifie d'être bien adapté ?" ; "Qu'est-ce que l'accomplissement personnel signifie pour vous ?" ; "Comment définiriez-vous la personne idéale ?".

cas de celle de Gallup (1976) qui a recensé, dans une enquête mondiale, les thèmes liés aux espoirs et aux aspirations des individus ; c'est aussi celui du travail de Ostroot et al. (1982), les auteurs ayant demandé aux habitants de Springfield (Etats-Unis) et d'Aix-en-Provence (France), ce que la qualité de la vie signifiait pour eux ; c'est le cas enfin de trois études recensées par Veenhoven (1984), dans lesquelles les auteurs ont interrogé des sujets sur la signification du mot bonheur. Les résultats les plus importants de ces études sont synthétisés dans le tableau 2.2.

références et descriptions	thèmes les plus souvent cités
Gallup (1976) : espoirs, aspirations des individus et perceptions de la vie idéale	sécurité économique, santé pour soi et sa famille, vie de famille heureuse, développement de soi, désir de paix et d'un monde meilleur, religion
Campbell (1981) : qu'est-ce que la qualité de la vie ? de quoi dépend-elle ?	sécurité économique, vie de famille, qualités personnelles (honnêteté, "force d'âme", intelligence), amitié, attractivité de l'environnement physique
Ostroot et al. (1982) : qu'est-ce que la qualité de vie ? qu'est-ce qui y contribue ?	relations familiales, relations amicales
Veenhoven (1984) : trois études sont citées qui portent sur la signification du mot bonheur	humeur plaisante, changement vers de meilleures conditions, contentement
Felton (1987) : " <i>why is this a happy time ?</i> ", pourquoi vous sentez-vous heureux ?	thèmes économiques et matériels, enfants, mariage, travail, indépendance, santé, qualités personnelles
Thomas & Stock (1988) : évocations libres à partir du mot bonheur	amour, famille, amis, santé, enfants, travail, liberté
Thomas & Chambers (1989) : entretiens semi-directifs (beaucoup de thèmes sont spécifiques à la vieillesse et font référence au mal-être)	satisfait avec la vie, désir d'être actif, famille, conjoint, travail
Ryff (1989b) : que signifie "être bien adapté" ? ^[11]	être orienté vers les autres, confiant-affirmatif, se développer, joie de vivre, acceptation-connaissance de soi, expériences familiales, sens de l'humour, accepter le changement, tolérant

Tableau 2.2: Recension des études ayant porté sur les représentations du bien-être et principaux résultats

Ces quelques résultats sont trop frustes pour apporter une réponse satisfaisante à la question des représentations du bien-être, même si on peut constater la récurrence de certains thèmes (c'est le cas de la famille, de la sécurité économique et matérielle, des relations avec autrui). Des informations plus précises peuvent toutefois être trouvées dans une publication — peu connue — de Flanagan (1978). Celui-ci présente les résultats

[11] : Cette question est retenue dans ce tableau plutôt que d'autres parce qu'elle reflète, selon l'auteur, les conceptions du bien-être.

d'une étude menée à l'initiative des directeurs des Instituts de Recherche Américains qui avaient décidé de développer un effort important en faveur de l'amélioration de la qualité de la vie. C'est la première étape de cette étude qui nous intéresse ici^[12]. On a demandé à un échantillon très hétérogène de 3000 personnes de penser à la dernière fois qu'ils avaient vécu une expérience particulièrement satisfaisante et/ou importante dans leur vie et de la décrire. Le corpus recueilli de cette manière comporte 6500 items qui ont été regroupés en 15 grandes catégories. Le contenu de ces catégories est résumé dans la figure 2.1.

1. Bien-être physique et matériel
 - A. Bien-être matériel et sécurité financière (avoir de la bonne nourriture, logement, confort et attentes de choses positives dans le futur. L'argent et la sécurité financière sont particulièrement importants et, pour la plupart des gens, satisfaire ces besoins constitue le but primaire de leurs efforts)
 - B. Santé et sécurité personnelle (ne pas être malade, éviter les accidents, problèmes liés à la drogue, à l'alcool, à la vieillesse, etc.)
2. Relations avec les autres personnes
 - C. Relations avec le conjoint (amour, satisfaction sexuelle, compréhension, communication, etc.)
 - D. Avoir et élever des enfants (les regarder se développer, passer du temps avec eux, les guider, les aider, les apprécier, apprendre d'eux et leur apprendre)
 - E. Relations avec les parents, les frères et sœurs et les autres proches (communiquer et faire des choses avec eux, rendre visite, partager, comprendre, aider et être aidé, avoir quelqu'un avec qui discuter)
 - F. Relations avec les amis (partager des activités, des intérêts et des points de vue, être accepté, rendre visite, être aidé et aider, amour, confiance, etc.)
3. Activités sociales, locales et civiques
 - G. Activités liées à l'aide et à l'encouragement des autres. (ces activités peuvent être individuelles ou associatives : église, club, association caritative, etc.)
 - H. Activités liées au gouvernement local et national (participer en votant ou par d'autres moyens, avoir et apprécier la liberté politique, sociale et religieuse)
4. Développement personnel et accomplissement de soi
 - I. Développement intellectuel (apprendre, s'éduquer, etc., améliorer ses connaissances et sa compréhension)
 - J. Compréhension personnelle et orientation (développer et acquérir les principes guidant sa vie, devenir plus mature, être conscient de ses qualités et de ses défauts, avoir l'expérience et prendre conscience de son développement et de sa croissance personnelle, avoir des expériences et activités religieuses ou spirituelles)
 - K. Rôle professionnel (utiliser ses compétences, faire bien les choses, apprendre et produire, être reconnu)
 - L. Créativité et expression personnelle (montrer de l'ingéniosité, de l'originalité, de l'imagination dans la musique, dans l'art, dans des activités scientifiques et techniques ou dans les activités quotidiennes, s'exprimer à travers un projet personnel)
5. Loisirs
 - M. Socialisation (rencontrer d'autres personnes, faire des choses avec elles, participer à des fêtes)
 - N. Activités de loisirs passives (regarder la télé, écouter de la musique, lire, aller au cinéma, apprécier l'art et la beauté dans plusieurs aspects de la vie)
 - O. Activités de loisirs actives (sport, voyages, jouer d'un instrument, chanter, etc.)

Figure 2.1 : Les catégories thématiques des représentations associées au bien-être selon Flanagan (1978)

[12] : D'ailleurs, l'ambitieux programme de recherche annoncé n'a, à notre connaissance, jamais dépassé ce stade.

On retrouve, dans ces grands thèmes, la plupart de ceux qui ont déjà été évoqués dans les études précédentes. Ce qui est intéressant, c'est que l'on y trouve aussi des évocations de dimensions du bien-être qui ont été dégagées par l'intermédiaire d'analyses plus théoriques (cf. le grand thème du développement personnel et de l'accomplissement de soi). Les résultats de cette étude illustrent parfaitement, à notre avis, le fait que les représentations peuvent enrichir la signification d'un concept et en éclaircir la conceptualisation. Il n'en est pas moins vrai que ces catégories thématiques sont moins à même de faire apparaître des distinctions qui ont été évoquées plus haut comme la différence entre les conceptions objectives et subjectives ou celles entre les aspects cognitifs et affectifs.

1.6 Définition(s), mesure(s) et théorie(s)

Avant d'essayer de tirer les principaux enseignements de ces travaux qui ne présentent pas, à première vue, d'homogénéité très nette, il faut relever deux aspects très importants qui apparaissent en filigrane dans cette rapide revue de la littérature. Ces deux aspects ne sont pas totalement indépendants mais leur présentation s'en trouvera simplifiée s'ils sont distingués.

Le premier aspect concerne la confusion entre composantes ou dimensions du bien-être et corrélats. Nous avons vu que plusieurs auteurs, afin de clarifier la signification conceptuelle du bien-être, essaient d'en spécifier les caractéristiques (les composantes, les dimensions). Mais la question se pose très vite de savoir à quoi correspondent ces caractéristiques. Sont-elles des composantes fondamentales du bien-être ou n'en sont-elles que des corrélats (auquel cas elles renverraient à d'autres concepts) ? Cette question a rarement été posée de manière explicite dans la littérature^[13]. Pourtant, elle s'est posée plusieurs fois au niveau empirique ainsi qu'on peut le voir à partir de trois exemples typiques. Les rapports entre santé (physique) et bien-être ont été beaucoup étudiés. Certains considèrent (e.g. Conte & Salomon, 1982) que la santé est une dimension fondamentale du bien-être général ; d'autres (e.g. Pierce & Clark, 1973) s'opposent à cette conception en arguant du fait qu'elle constitue un domaine conceptuellement différent de celui du bien-être. Il en est de même pour le sens de la vie qui apparaît dans la conceptualisation de Ryff (1989a, 1989c). Pour Harlow & Newcomb (1990), bien-être et sens de la vie ne sont qu'un seul et même concept ; d'autres (e.g. Chamberlain & Zika, 1988 ; Zika & Chamberlain, 1992 ; Korotkov, 1993) sont plus réservés à ce sujet, même s'ils notent un important recouvrement conceptuel. Le dernier

[13] : Mais cf. Kozma & Stones (1978, p. 242) : "*adequate formulation of well-being must clearly differentiate between components and correlates*".

exemple concerne l'estime de soi dont nous avons déjà parlé. La plupart du temps, l'estime de soi est étudiée comme un corrélat du bien-être (e.g. Headey et al., 1985 a). Pourtant, il est tout à fait frappant de constater les similitudes conceptuelles et théoriques entre l'estime de soi et le bien-être. Au niveau conceptuel, Blascovich & Tomaka (1991) définissent l'estime de soi comme une évaluation affective globale de sa propre valeur : l'estime de soi est la composante évaluative du concept de soi qui comprend aussi une composante cognitive (plus descriptive qu'évaluative) et une composante comportementale. Au niveau théorique (nomologique), les modèles de l'estime de soi font souvent référence à la différence entre le soi réel et le soi idéal ; plusieurs modèles du bien-être s'appuient, quant à eux, sur la différence entre une évaluation de sa situation actuelle et celle à laquelle on aspire. Tout comme dans le cas de la santé et du sens de la vie, on trouve des auteurs qui assimilent l'estime de soi à un corrélat du bien-être alors que d'autres (e.g. George & Bearon, 1980) en font une de ses composantes.

Le second aspect que nous voulons souligner généralise les commentaires précédents. On peut se demander, en effet, dans quelle mesure les théories définitoires du bien-être — aussi peu élaborées soient-elles — ne sont pas déjà des théories nomologiques. Autrement dit, quand on pose la question "qu'est-ce que le bien-être ?", on a l'impression qu'on répond souvent à une autre question "qu'est-ce qui fait que les individus sont heureux ?". Nous avons vu que Veenhoven (1984) s'est posé cette question en d'autres termes en différenciant les définitions descriptives et les définitions explicatives. En nous appuyant sur les conceptions développées dans la première partie, nous parlerons plutôt de confusion entre théorie définitoire et théorie nomologique. Quelques exemples typiques nous serviront à nouveau d'illustration. Parmi les nombreuses théories élaborées pour rendre compte du bien-être, certains auteurs (e.g. Brock, 1993 ; Scanlon, 1993) croient pouvoir distinguer trois grands types^[14] : les théories hédonistes pour lesquelles le bien-être est causé par des expériences de joie, de plaisir ; les théories des idéaux qui côtoient les philosophies morales, pour lesquelles certains idéaux doivent être atteints pour être heureux ; les théories des préférences (ou théories des désirs)^[15]. Ces trois théories nomologiques pourraient tout aussi bien être considérées comme des théories définitoires. En particulier, la première implique une conception subjective du bien-être alors que la troisième implique une définition objective. Lorsque

[14] : On verra plus loin que cette classification est beaucoup trop sommaire pour rendre compte de la diversité des approches développées ces dernières années. Cette classification décrit plutôt les travaux pré-scientifiques sur le bien-être, mais cela n'enlève rien à l'argumentation développée ici.

[15] : Ces théories ne sont pas très faciles à décrire. Elles rejettent la nécessité d'expériences, mise en avant dans les théories hédonistes. Le bien-être d'un individu est déterminé en fonction du nombre de préférences ou de désirs qui sont satisfaits, que ces préférences le touchent personnellement ou non, que la personne en soit consciente ou non. On considérera, avec Scanlon (1993), que ces théories n'ont pas de caractère spécifique par rapport aux théories hédonistes ou aux théories des idéaux. En effet, soit une personne désire quelque chose parce que cela va lui apporter du plaisir et on se retrouve dans le cadre des théories hédonistes ; soit une personne désire quelque chose parce que cela constitue un bien essentiel, une "vertu" idéale et on se retrouve dans le cadre des théories des idéaux.

Allardt (1993) défend une conception du bien-être basée sur une version simplifiée de la théorie des besoins de Maslow^[16], la même question se pose. Est-ce que le bien-être est une conséquence de la satisfaction des besoins, ou est-ce que le bien-être *est* la satisfaction des besoins ? Chez George & Bearon (1980), le bien-être se définit comme l'accomplissement des besoins, l'évaluation cognitive des progrès vers un but et l'évaluation de la différence entre sa situation actuelle et ses aspirations. Les trois termes de cette définition peuvent, à nouveau, être considérés comme des hypothèses visant à expliquer le bien-être plutôt qu'à le définir. On pourrait passer en revue bien d'autres définitions et aboutir à des conclusions identiques. Cela ne semble pas nécessaire. En revanche, on peut encore se demander si cette interférence entre théorie définitoire et théorie nomologique se limite à la conceptualisation du bien-être ou si elle se reflète également dans sa mesure. D'après ce que nous savons des liens entre définition, théorie et mesure, la réponse ne fait guère de doute.

Ces influences peuvent être mises en évidence à deux niveaux différents. Il y a tout d'abord des influences implicites dans la construction des instruments de mesure. Il y a aussi des stratégies plus explicites qui peuvent encore être distinguées selon qu'elles sont de nature empirique ou plus théorique. Parmi les influences implicites^[17], on prendra deux exemples d'items faisant partie d'échelles couramment utilisées pour la mesure du bien-être. Le premier item appartient à trois échelles de bien-être pour personnes âgées (le LSI-A de Neugarten et al., 1961 ; la PGC-MS de Lawton, 1972, 1975 ; la MUNSH de Kozma & Stones, 1980). Il est formulé ainsi : "Je suis aussi heureux que lorsque j'étais plus jeune". Seules les réponses positives à cet item sont considérées comme des indicateurs de bien-être. Ainsi, ce codage des réponses conduit à admettre, implicitement, l'hypothèse que le bien-être est corrélé négativement avec l'âge. En effet, ce codage n'admet pas qu'une personne réponde négativement à cet item parce qu'elle est plus heureuse que lorsqu'elle était plus jeune. Le second item "Dans l'ensemble, ma vie est proche de mon idéal" est emprunté à l'échelle de satisfaction avec la vie de Diener (SWLS, Diener et al., 1985 a). On reconnaîtra sans difficulté dans cette formulation la comparaison entre situation actuelle et situation idéale mentionnée plus haut.

Les stratégies explicites sont encore plus démonstratives. On peut en trouver de nombreux exemples. Au niveau théorique tout d'abord, plusieurs auteurs (e.g. Atkinson, 1982 ; Headey et al., 1984 a) affirment qu'une des qualités principales des mesures de

[16] : Allardt ne retient que trois grands types de besoin qui sont, tout comme chez Maslow, hiérarchisés. Il y a tout d'abord les besoins matériels et impersonnels ("*having*" : ressources économiques, conditions de logement, travail, conditions de travail, santé, éducation, etc.). On trouve ensuite les besoins sociaux ("*relating*" : famille, amis, collègues, participation associative, etc.), puis les besoins de croissance personnelle ("*being*" : participer aux décisions concernant sa vie, s'engager dans des activités politiques, pouvoir profiter de ses loisirs, de la nature, avoir un travail gratifiant, etc.).

[17] : Implicite, dans ce contexte, fait référence au fait que les auteurs n'en font pas état, soit qu'ils ne perçoivent pas qu'une s'influence s'exerce, soit qu'ils ne jugent pas utile de la mentionner.

bien-être doit être leur sensibilité aux changements de conditions de vie. Cette recommandation ne fait qu'affirmer l'hypothèse nomologique selon laquelle les conditions de vie influencent le bien-être. Plus explicite encore est la démarche de Meadow et al. (1992) qui proposent une mesure du bien-être dérivant directement d'une théorie du bien-être basée sur l'idée que ce dernier est le résultat d'une comparaison entre sa situation et différents standards de référence (sa situation passée, celle de ses proches, etc.). Il en est de même pour Smith (1992) dont l'échelle est composée d'une liste d'événements de vie négatifs dont la somme, pondérée par leur gravité, est censé fournir une mesure de bien-être. Il est fait référence, dans ce dernier cas, à une théorie situationniste du bien-être, qui le fait dépendre des événements de vie vécus par les individus. Au niveau empirique ensuite, certains auteurs construisent des échelles de bien-être qui incluent des items mesurant des variables dont on a montré qu'elles corrélaient avec le bien-être. C'est le cas notamment chez Conte & Salomon (1982) ou chez Fordyce (1985)^[18].

Cette confusion entre théorie nomologique, théorie définitoire et mesure n'est pas dommageable en soi. Nous avons assez souvent répété qu'elles étaient inextricablement liées. Il faut pourtant se préserver de l'erreur qui consisterait à valider ses mesures par des procédures de validation interconcepts alors même que les mesures sont construites de telle manière que les hypothèses testées ne peuvent être réfutées.

1.7 Conclusion

Que pouvons-nous tirer comme enseignements de cette discussion sur la conceptualisation du bien-être ? Malgré les nombreuses difficultés qui subsistent en ce domaine, quelques points méritent de retenir l'attention.

Cette revue de la littérature permet de mettre en relief une distinction assez nette entre une conception objective du bien-être et une conception subjective. Cette distinction apparaît surtout dans la littérature philosophique. La conception objective a longtemps prédominé. Aujourd'hui, il y a un relatif consensus sur l'idée que le bien-être ne peut être appréhendé sans référence à la subjectivité des individus qui ne partagent pas les mêmes valeurs et qui évaluent, selon des critères qui leur sont propres, les circonstances dans lesquelles ils vivent. Malgré la prévalence des conceptions subjectives, nous avons vu

[18] : Il n'est pas inutile de décrire la stratégie suivie par Fordyce (1985) pour construire son "PsychHap Inventory", notamment pour illustrer le fait que certaines procédures de validation risquent de devenir triviales. Dans une première étape, une revue de la littérature permet à l'auteur d'identifier les caractéristiques qui distinguent les personnes heureuses des personnes malheureuses. Ensuite, des items censés mesurer ces caractéristiques sont construits et administrés à un échantillon de sujets en même temps que deux items censés mesurer le bien-être. La corrélation entre les caractéristiques retenues et les items de bien-être sert de critère empirique pour construire l'échelle définitive ; cette dernière est ensuite validée en la corrélant avec une multitude d'autres échelles mesurant, en particulier, les caractéristiques retenues au préalable. Doit-on ajouter que l'auteur conclut à la validité de son échelle ?

(e.g. Lawton, 1983a, 1983b ; Lawton et al., 1984) que certaines conceptualisations intègrent à la fois des composantes objectives et des composantes subjectives. C'est aussi le point de vue de George & Bearon (1980) pour qui le bien-être se compose d'évaluations subjectives et de conditions objectives telles que la santé ou le statut socio-économique. Pourtant quelques auteurs refusent encore d'accorder un statut subjectif au bien-être. Erickson (1993) est de ceux-là. Il contraste une approche descriptive du bien-être et une approche évaluative. Cette dernière ne fait qu'appréhender l'adaptation des individus à leurs conditions de vie alors que l'approche descriptive, bien plus intéressante selon lui, permet de se centrer sur les ressources dont disposent les individus pour satisfaire leurs besoins. Même si ce débat n'est pas entièrement clos et s'il existe encore des tentatives relativement isolées pour construire des indicateurs objectifs du bien-être (e.g. Johnston, 1988), nous nous focaliserons essentiellement, dans la suite de ce travail, sur les conceptions subjectives qui constituent la position largement dominante actuellement.

La distinction entre composantes affectives et composantes cognitives du bien-être (cf. Diener, 1984 ; Veenhoven, 1984 ; Andrews & Robinson, 1991) paraît également majeure. Campbell (1976) va jusqu'à considérer que les conceptions du bien-être comme expérience affective et comme expérience cognitive constituent deux traditions de recherche relativement indépendantes. Aujourd'hui, elles ont tendance à s'intégrer, d'autant que cette distinction satisfait aussi bien une conception du bien-être en tant qu'attitude qu'en tant qu'émotion. Nous verrons pourtant que cette distinction n'est pas très facile à établir ni au niveau théorique ni au niveau empirique^[19].

Nous serons plus circonspects quant aux autres conclusions qui peuvent être tirées de cette analyse. S'il apparaît bien, dans certaines conceptions, une distinction entre un bien-être immédiat, instantané et un bien-être à long terme, on ne sait pas trop si ces deux approches sont compatibles ni de quelle manière. Quant aux travaux théoriques de Ryff (1989a, 1989c) et aux études sur les représentations du bien-être qui mettent en évidence certaines caractéristiques négligées du bien-être, ils méritent que l'on s'y intéresse, mais ils manquent d'assises empiriques.

Enfin, nous avons insisté, à partir de plusieurs exemples, sur les confusions entre théorie définitoire et théorie nomologique. Il ressort de cette réflexion qu'il n'est pas possible de décider, de manière univoque, ce qui relève uniquement de l'une et ce qui relève de l'autre. A un moment ou à un autre, le chercheur doit choisir à quel niveau il se situe. Mais ce choix doit être fait en toute connaissance de cause afin de ne pas construire des réseaux nomologiques triviaux, dans lesquels les composantes du bien-être sont aussi ses propres prédicteurs.

[19] : D'aucuns parlent même (e.g. Liang, 1984, p. 614) d'"évaluation cognitive de l'affectivité".

CHAPITRE II

LES MESURES DE BIEN-ETRE SUBJECTIF

Nous aborderons, dans ce chapitre, les mesures du bien-être dans une perspective instrumentale. Nous nous intéresserons donc surtout aux instruments de mesure proposés pour le mesurer. Une telle présentation peut être fait sous un angle analytique ou synthétique. La démarche analytique consisterait à décrire chaque instrument en détail et présenter toutes les études dans lesquelles il a été utilisé afin d'en dégager les faiblesses et les qualités. Etant donné son caractère fastidieux, nous avons retenu une démarche plus synthétique. Elle vise à mettre en relief les principaux résultats en s'appuyant, notamment, sur des revues préalables de la littérature et en les complétant.

Ce chapitre se divise en trois grandes parties. La première est consacrée à la présentation, sous la forme d'un tableau synthétique, de l'ensemble des instruments de mesure du bien-être recensés dans la littérature. Quelques échelles prototypiques seront ensuite choisies afin de les décrire de manière plus détaillée. Tous ces instruments sont en fait des questionnaires auto-administrés, cette technique ayant recueilli la préférence de l'immense majorité des chercheurs. Toutefois, d'autres types de mesures ont quelquefois été utilisés, souvent à des fins de validation. Ils feront l'objet d'une présentation spécifique. La seconde partie traitera des qualités psychométriques des instruments de mesure. Dans la troisième et dernière partie les critiques les plus fréquentes auxquelles ont dû répondre les constructeurs d'échelles de bien-être seront analysées..

Précisons que certaines questions importantes relatives à la mesure du bien-être ne seront pas traitées dans ce chapitre dont le caractère est plutôt instrumental. C'est le cas notamment de la structure du bien-être dont l'aspect fondamental requiert une présentation détaillée (cf. chapitre 3). C'est le cas aussi d'aspects plus spécifiques, comme la stabilité temporelle des mesures, qui renvoient à des considérations théoriques plus générales qui seront largement discutées dans le chapitre sur les théories nomologiques (cf. chapitre 5).

2.1 Les instruments de mesure du bien-être

2.1.1 Une recension des questionnaires utilisés dans la littérature

Il y a encore quelques années, on aurait pu présenter un panorama pratiquement exhaustif des mesures de bien-être subjectif. Aujourd'hui, cette entreprise est beaucoup plus délicate. Les instruments sont nombreux, quelquefois décrits dans des publications plus ou moins confidentielles et n'ont pas toujours donné lieu à des analyses empiriques suffisantes.

Les instruments présentés dans le tableau qui suit ont été recensés à partir de synthèses de la littérature (cf. Larson, 1978 ; George & Bearon, 1980 ; Diener, 1984 ; Andrews & Robinson, 1991) et complétés par notre propre revue de la question. Il n'est pas possible d'affirmer que cette liste est exhaustive ; il est certain toutefois qu'elle est la plus complète qui ait été établie jusqu'à ce jour.

La difficulté majeure de ce travail de recension réside dans la décision d'intégration des instruments : il n'est pas toujours très facile de décider si telle échelle mesure le bien-être ou non. Ainsi que l'on s'en rendra compte, certaines échelles portent des noms dont le rapport avec le bien-être n'est pas toujours immédiat. Le plus souvent pourtant, on trouve, dans le titre des échelles, des synonymes du bien-être comme la satisfaction, le bonheur, le moral, l'affectivité ou encore la qualité de la vie. Ont été exclues de cette recension toutes les échelles mesurant la santé mentale négative comme les échelles de dépression ou d'anxiété, même si certains considèrent qu'elles mesurent le pôle négatif du bien-être^[1]. Toutes les mesures ponctuelles du bien-être élaborées dans le cadre d'études spécifiques et n'ayant pas été utilisées par la suite en ont également été exclues.

Les échelles présentées ci-dessous sont très diverses. Certaines, même, ne méritent pas le nom d'échelles puisqu'elles ne comportent qu'un seul item. Malgré les faiblesses psychométriques de telles mesures, elles ont été retenues parce qu'elles font partie des premières tentatives pour mesurer le bien-être et qu'elles sont encore utilisées, de nos jours, dans de vastes enquêtes internationales. Certaines échelles sont spécifiques aux personnes âgées : elles apparaissent en italique dans le tableau. Les échelles prototypiques qui seront décrites de manière plus détaillée apparaissent en caractères gras. Signalons enfin que ces échelles sont classées par ordre chronologique d'apparition.

[1] : Cette question sera traitée en détail dans le chapitre 3.

Nom	Références	Nbre items	Description	Remarques
Chicago Attitude Inventory	Cavan et al. (1949)	70	satisfaction dans différents domaines de vie ; sentiment général de bonheur ; sentiment d'utilité	concerne surtout l'adaptation à la vieillesse des personnes âgées
Kutner Morale Scale	Kutner et al. (1956)	7	échelle unidimensionnelle de moral	se conforme à une échelle hiérarchique de Guttman
Echelle de Thompson	Thompson et al. (1960)	9	satisfaction dans la vie ; découragement ; désespoir (réponses ouvertes)	chacune des trois sous-échelles se conforme à une échelle hiérarchique de Guttman
Gurin Scale	Gurin et al. (1960)	1	Taking all things together, how would you say things are these days. Would you say you are : very happy, pretty happy or not so happy ?	
Life Satisfaction Index (LSI)	Neugarten et al. (1961)	20	enthousiasme vs apathie ; détermination et "force d'âme" ; congruence entre buts atteints et buts désirés ; concept de soi ; humeur	de nombreuses versions existent (cf. texte)
Self-Anchoring Ladder	Cantril (1965)	1 (3)	Here is a picture of a ladder. Suppose we say that the top of the ladder represents the best possible life for you and the bottom represents the worst possible life. Where on the ladder do you feel you personally stand at the present time ? Where on the ladder would you say you stood five years ago ? And where do you think you will be on the ladder five years from now ?	
Affect Balance Scale (ABS)	Bradburn (1969)	10	deux sous-échelles indépendantes mesurant les composantes affectives positive et négative du bien-être	
Philadelphia Geriatric Center - Morale Scale (PGC-MS)	Lawton (1972, 1975)	22 puis 17	enthousiasme ; attitude envers sa propre vieillesse ; satisfaction ; émotion , optimisme ; désappointement lié à la solitude puis : attitude envers la vieillesse ; émotion ; solitude	la première version de 1972 a très vite été abandonnée au profit d'une version révisée en 1975
Index of General Affect	Campbell et al. (1976)	8	items construits selon la technique du différenciateur sémantique	

Nom	Références	Nbre items	Description	Remarques
Delighted-Terrible Scale	Andrews & Whitey (1976)	1	How you feel about your life as a whole ? : terrible, unhappy, mostly dissatisfied, mixed, mostly satisfied, pleased, delighted	cet item doit sa célébrité aux étiquettes extrêmes de son échelle de réponse
Happiness Measures	Fordyce (1977, 1988)	2	In general, how happy or unhappy do you usually feel ? On the average, what percent of time do you feel happy ?	
General Well-Being Schedule (GWB)	Dupuy (1978), Fazio (1977)	18	santé ; énergie ; vie intéressante ; bonne humeur ; détendu ; contrôler ses émotions et son comportement	
Self-description Inventory	Fordyce (1978)	?	bonheur atteint ; personnalité heureuse ; attitudes et valeurs face au bonheur ; style de vie heureux	on dispose de peu d'informations sur cette échelle
Affectomètre	Kammann et al. (1979) Kammann & Flett (1983)	96 puis 40	mesure la fréquence des affects positifs et négatifs. Echelle unidimensionnelle.	deux version existent ; la seconde comporte deux formes
Differential Personality Questionnaire	Tellegen (1979)	21	affectivité positive ; attitudes positives ; optimisme	une seule sous-échelle de ce questionnaire mesure le bien-être
Mood Survey	Underwood & Froming (1980)	34	niveau général de l'humeur ; fréquence des changements d'humeur ; intensité de l'humeur	
<i>Memorial University of Newfoundland Scale of Happiness (MUNSH)</i>	<i>Kozma & Stones (1980)</i>	<i>24</i>	<i>affectivité positive et affectivité négative à partir d'expériences générales et spécifiques</i>	<i>comprend beaucoup d'items du LSI, de la PGC-MS et de l'ABS</i>
<i>Salamon-Conte Life Satisfaction in the Elderly Scale (SCLSES)</i>	<i>Salamon & Conte (1981) Conte & Salamon (1982)</i>	<i>40</i>	<i>plaisir dans les activités quotidiennes ; sens de la vie ; congruence entre buts désirés et buts atteints ; humeur ; concept de soi ; santé ; sécurité financière ; contacts sociaux</i>	
Mental Health Inventory	Veit & Ware (1983)	38	anxiété ; dépression ; perte de contrôle du comportement et de l'émotion ; affectivité positive générale ; attaches émotionnelles	construit en partie à partir du GWB

Nom	Références	Nbre items	Description	Remarques
Affect Intensity Measure	Larsen (1983) Larsen & Diener (1987)	40	mesure de l'intensité de l'affectivité positive et de l'affectivité négative	
<i>Perceived Well-Being Scale</i>	<i>Recker & Wong (1984)</i>	?	<i>bien-être psychologique et bien-être physique</i>	
PsycHap Inventory	Fordyce (1985)	80	cf. Self-description Inventory	4 formes différentes existent
Satisfaction With Life Scale (SWLS)	Diener et al. (1985 a) Pavot & Diener (1993 a)	5	mesure la composante cognitive générale du bien-être	
Positive and Negative Affect Schedule (PANAS)	Watson et al. (1988)	20	mesure de l'affectivité positive et de l'affectivité négative basée sur le circumplex de l'humeur (Watson & Tellegen, 1985)	
Oxford Happiness Inventory	Martin et al. (en préparation) cités par Argyle et Martin (1991)	30	items du questionnaire de dépression de Beck formulés positivement	
Scales of Psychological well-being	Ryff (1989a, 1989c, communication personnelle)	120	autonomie ; maîtrise de l'environnement ; croissance personnelle ; relations positives avec autrui ; sens de la vie ; acceptation de soi	cf. seconde partie, chapitre 1 il existe une forme abrégée de 84 items
Quality of Life Questionnaire	Evans & Cope (1989) Evans et al. (1993)	192	bien-être matériel ; bien-être physique ; croissance personnelle ; relations maritales ; relations parents-enfants ; relations familiales ; relations extrafamiliales ; comportement altruiste ; comportement politique ; caractéristiques du travail ; relations avec les collègues ; satisfecit du travail ; comportement créatif ; activités sportives ; activités de loisirs	
Need Satisfaction Inventory	Lester (1990)	50	mesure du degré de satisfaction des 5 besoins hiérarchiques de Maslow (besoins physiologiques, de sécurité, d'appartenance, d'estime et d'actualisation de soi)	
Index of societal well-being	Smith (1992)	60	Somme d'événements négatifs pondérés par leur gravité	

Nom	Références	Nbre items	Description	Remarques
Life Satisfaction Measure	Meadow et al. (1992)	10	Evaluation de la satisfaction en fonction de 10 standards de référence	s'inscrit dans les théories relativistes du bien-être (cf. chapitre 5)
Depression-Happiness Scale	McGreal & Joseph (1993)	25	fréquence des affects positifs et négatifs. Echelle unidimensionnelle	
World Health Organization Quality of Life Instrument (WHOQOL)	WHO (1993)	en cours	6 domaines évalués : physique, psychologique, niveaux d'indépendance, relations sociales, environnement, spiritualité	en cours de construction et de validation

Tableau 2.3 : Recension des échelles de bien-être utilisées dans la littérature

2.1.2 Description de quelques échelles prototypiques

Les questionnaires recensés dans le tableau précédent ont connu des fortunes diverses. Certains sont encore trop récents pour que l'on puisse les évaluer ; d'autres (c'est le cas des plus anciens) ont été abandonnés ; d'autres encore n'ont guère été utilisés que par leur concepteur (c'est le cas, notamment, des instruments mis au point par Fordyce). Mais, parmi ces instruments, quelques-uns ont connu un tel succès qu'ils sont utilisés dans l'immense majorité des publications. Ces instruments sont l'"*Affect Balance Scale*" (ABS) de Bradburn (1969) et le "*Life Satisfaction Index*" (LSI) pour les personnes âgées. Il faut ajouter la "*Satisfaction With Life Scale*" (SWLS) de Diener et al. (1985 a), plus récente, qui commence à être largement diffusée (cf. Pavot & Diener, 1993 a). Nous adjoindrons volontiers deux autres échelles pour personnes âgées, un peu moins connues, mais qui sont associées généralement au LSI : la "*Philadelphia Geriatric Center - Morale Scale*" (PGC-MS) de Lawton (1972, 1975) et la "*Memorial University of Newfoundland Scale of Happiness*" (MUNSH) de Kozma & Stones (1980). Comme nous serons amené à reparler de ces échelles par la suite, nous les présenterons en même temps que le LSI.

A. L'"*Affect Balance Scale*" (ABS) de Bradburn (1969)

Les travaux de Norman Bradburn sont à l'origine de ce que beaucoup considèrent comme le "premier vrai inventaire de bien-être" (Kammann et al., 1984 a, p. 94). Le succès de cette échelle ne s'est pas démenti depuis, même si de nombreuses critiques ont été formulées à son égard. Ce succès n'est sans doute pas étranger à un résultat qui allait occuper pendant de nombreuses années la recherche sur le bien-être, à savoir l'indépendance totale entre ce qu'il est convenu d'appeler l'affectivité positive et l'affectivité négative. Ce résultat obtenu par Bradburn (1969, mais cf. aussi Bradburn & Clapovitz, 1965) signifie qu'une personne peut être à la fois heureuse et malheureuse ou, en termes plus techniques, qu'un score élevé d'affectivité positive ne permet pas de prédire un faible score d'affectivité négative. On se doute de l'effet qu'a pu produire un résultat aussi contraire à l'intuition. Aujourd'hui encore, après 25 années de travaux, certains émettent des doutes quant à cette indépendance. Nous traiterons en détail de ce débat dans le chapitre 3.

Le cadre de référence de Bradburn est psycho-sociologique. Il abandonne délibérément le concept de santé mentale, trop vague et trop large, et ses racines médico-psychologiques. Ce qui intéresse Bradburn (1969, p. 1), c'est la "relation entre la situation de vie d'un individu et ses réactions psychologiques face à cette situation". Le concept qui lui paraît le mieux à même d'appréhender ces réactions psychologiques est celui de

bonheur, qui est donc conçu, soulignons-le, comme un état psychologique relativement instable.

Pour Bradburn, le bonheur est une qualité de l'expérience individuelle et non pas du comportement. Il faut donc interroger les individus de manière à savoir combien d'expériences plaisantes et déplaisantes ils ont connu au cours d'un laps de temps déterminé. Le contenu spécifique de chacune de ces expériences n'est pas important : seule importe la tonalité affective qui leur est associée. Cette conception a un avantage évident quant à la construction d'un instrument de mesure. En effet, comme il n'est pas possible de constituer un échantillon représentatif de l'ensemble de toutes les expériences positives et négatives possibles, il suffit de formuler les items en termes suffisamment généraux pour que les individus puissent répondre en termes de tonalité affective.

A partir de ces considérations générales, 10 items sont rédigés, la moitié couvrant des expériences positives, l'autre moitié des expériences négatives. A l'origine, l'échelle de réponse est dichotomique ; on interroge les individus sur l'occurrence d'une expérience. Enfin, comme Bradburn s'intéresse aux réactions psychologiques des individus, le questionnaire spécifie un cadre de référence temporel assez restreint (ces dernières semaines) afin de pouvoir appréhender d'éventuels changements en fonction des circonstances. L'ABS est présentée dans la figure 2.3.

We are interested in the way people are feeling these days. During the past weeks, did you ever feel :

1. Pleased about having accomplished something ?
2. That things were going your way ?
3. On the top of the world ?
4. Proud because someone had complimented you on something you had done ?
5. Particularly excited or interested in something ?
6. Very lonely or remote from other people ?
7. Upset because someone criticized you ?
8. So restless that you couldn't sit long in a chair ?
9. Bored ?
10. Depressed or very unhappy ?

Figure 2.3 : L'"Affect Balance Scale" (ABS) de Bradburn (1969)

Chaque réponse positive se voit attribuer un point. La somme des items positifs permet d'obtenir un score d'affectivité positive, la somme des items négatifs, un score d'affectivité négative. Jusque là, rien de bien étonnant. Mais Bradburn propose aussi de calculer un score de bien-être général en faisant la différence entre les scores d'affectivité positive et d'affectivité négative. Cette pratique, largement justifiée par le bon sens

comme le souligne Bradburn lui-même^[2], participe à l'une des confusions les plus importantes de toute l'histoire de la recherche sur le bien-être. Expliquons-nous.

Bradburn est l'un des premiers à montrer que le bien-être n'est pas unidimensionnel, mais qu'il faut distinguer deux dimensions indépendantes, l'affectivité positive et l'affectivité négative. Ceci étant établi, à la fois sur le plan conceptuel et sur le plan empirique, on a la surprise de lire, sous la plume de Bradburn (1969, p. 9), que "la position d'un individu sur **la dimension** du bien-être psychologique est conçue comme étant la résultante de la position de cet individu sur **deux dimensions indépendantes**, l'affectivité positive et l'affectivité négative."^[3] Il ne fait aucun doute que le bien-être d'un individu puisse être représenté par sa position dans un espace bidimensionnel orthogonal. Mais l'opération qui consiste à reconstruire une dimension unique à partir de deux dimensions indépendantes n'a aucun sens, ni au niveau théorique, ni au niveau psychométrique. Tout psychologue sait que le calcul d'un score à partir de plusieurs items est soumis à un certain nombre de précautions préalables visant à vérifier que tous les items mesurent le même concept. Ce qui est totalement fascinant dans l'histoire de l'ABS, c'est que personne — du moins à notre connaissance — n'a relevé cette erreur flagrante. Tout se passe comme si la conception populaire d'un bien-être unidimensionnel était tellement prégnante que même les scientifiques les plus éminents l'admettaient, malgré l'évidence^[4].

Malgré des qualités psychométriques satisfaisantes (consistance interne, stabilité temporelle, validité convergente, ... cf. Andrews & Robinson, 1991) plusieurs critiques ont été adressées à l'ABS, principalement dans un contexte visant à remettre en cause l'indépendance entre affectivité positive et affectivité négative. Contrairement aux objectifs de Bradburn, plusieurs auteurs (e.g. Brenner, 1975 ; Cherlin & Reeder, 1975 ; Diener, 1984) ont souligné des différences entre les niveaux de spécificité des expériences évoquées dans les items. Il suffit en effet de lire les items (cf. figure 2.3) pour constater que certains sont très généraux (e.g. items 2, 10) alors que d'autres font référence à des expériences beaucoup plus spécifiques (e.g. items 1, 4). Cherlin & Reeder (1975) notent aussi que les items positifs constituent plus une mesure du niveau d'activité ou d'engagement alors que les items négatifs expriment plutôt des sensations de (dé)plaisir. Sur un autre plan, les mêmes auteurs (Brenner ; Cherlin & Reeder ; Diener ; mais aussi Kamman et al., 1984 a) font remarquer qu'en interrogeant les sujets sur

[2] : Effectivement, le bon sens veut qu'une personne sera heureuse si elle éprouve à la fois beaucoup de sentiments positifs et peu de sentiments négatifs.

[3] : C'est nous qui soulignons.

[4] : On peut bien évidemment discuter de l'indépendance réelle entre ces deux dimensions. Mais, répétons-nous, l'opération qui consiste à calculer un score global n'a aucun sens à partir du moment où on accepte l'indépendance totale entre les dimensions comme le suggèrent les résultats de Bradburn.

l'occurrence d'un certain type d'expériences, on accordait, implicitement, autant de poids à un sentiment éprouvé une seule fois qu'à un sentiment éprouvé souvent. Cette critique conduira d'ailleurs à des utilisations de l'ABS avec une échelle de réponses formulée en termes de fréquence (e.g. Beiser, 1974 ; Andrews & Whitey, 1976 ; Dickes, 1989 a). Enfin, signalons que la période de référence temporelle utilisée par Bradburn a quelquefois été remise en cause (e.g. Diener, 1984). Tout comme pour l'échelle de réponses, rien n'empêche un chercheur de la changer, ce qui a d'ailleurs été fait à plusieurs reprises.

B. Le "*Life Satisfaction Index*" (LSI) de Neugarten et al. (1961)

Les travaux de Neugarten et al. (1961) s'inscrivent dans un cadre de référence gérontologique. Comme le note de manière un peu ironique Ryff (1989c), cette échelle n'a pas été construite pour mesurer le bien-être, mais pour aboutir à une définition opérationnelle de la "vieillesse réussie" qui ne se base pas sur des critères sociaux externes de succès et de compétence comme la participation sociale. Pourtant, cette échelle connaîtra, tout comme l'ABS, un grand succès et plusieurs échelles plus récentes s'en inspireront directement.

Les auteurs se sont basés sur une revue de la littérature concernant l'adaptation sociale et le moral pour distinguer cinq composantes relativement indépendantes :

- l'enthousiasme (vs l'apathie) ;
- la détermination et la "force d'âme" ;
- la congruence entre les buts atteints et les buts désirés ;
- le concept de soi positif ;
- l'humeur.

Ces cinq composantes permettent, dans un premier temps, d'établir un guide pour des entretiens qui sont menés par des psychologues cliniciens et évalués, de manière indépendante, par deux juges disposant de critères de notation explicite. Après une vérification de l'accord inter-juges, les auteurs "valident" cette première échelle (appelée LSR, *Life Satisfaction Rating*) en examinant la corrélation entre les évaluations moyennes des deux juges et celles du psychologue clinicien qui a mené l'entretien.

Mais cet instrument ne satisfait pas aux objectifs de facilité d'administration que s'étaient fixés les auteurs. Ceux-ci vont donc relever, dans les entretiens recueillis, les phrases qui différencient le plus les individus en fonction de leur score au LSR. A ces phrases, les auteurs ajoutent des items rédigés de manière à refléter les composantes théoriques et quelques items issus de l'échelle de moral de Kutner (cf. tableau 2.3). Deux

échelles sont alors mises au point : le LSI-A composé de 25 items à réponses dichotomiques et le LSI-B composé de 19 questions ouvertes. Après des analyses supplémentaires visant à éliminer les items les moins discriminatifs, le LSI-A — le seul qui nous intéresse ici — est finalement ramené à 20 items qui sont reproduits dans la figure 2.4. La validité du LSI-A est établie à partir des corrélations avec le LSR et avec les évaluations du psychologue clinicien.

1. As I grow older, things seem better than I thought they would be.
2. I have gotten more of the breaks in life than most of the people I know.
3. This is the dreariest time of my life.
4. I am just as happy as when I was younger.
5. My life could be happier than it is now.
6. These are the best years of my life.
7. Most of the things I do are boring or monotonous.
8. I expect some interesting and pleasant things to happen to me in the future.
9. The things I do are as interesting to me as they ever were.
10. I feel old and somewhat tired.
11. I feel my age, but it does not bother me.
12. As I look back on my life, I am fearly well-satisfied.
13. I would not change my past life even if I could.
14. Compared with other people my age, I've made a lot of foolish decisions in my life.
15. Compared with other people my age, I make a good appearance.
16. I have made plans for things I'll be doing a month or a year from now.
17. When I think back over my life, I didn't get most of the important things I wanted.
18. Compared to other people, I get down in the dumps too often.
19. I've gotten pretty much what I expected out of life.
20. In spite of what people say, the lot of the average man is getting worse, not better.

Figure 2.4 : Les items du "*Life Satisfaction Index - A*" (LSI-A) de Neugarten et al. (1961)

Cette échelle originale a connu de nombreuses versions. Adams (1969) propose, sur la base de nouvelles analyses, de supprimer deux items (les items n° 11 et 14). Wood et al. (1969) proposent le LSI-Z, qui ne comprend plus que 13 items, les autres étant éliminés à la suite d'analyses d'items^[5]. Bigot (1974) offre, quant à lui, une version à 8 items qu'il baptise le LSI-W^[6], en se basant notamment sur les résultats d'analyses factorielles. Il ne semble pas qu'il y ait dans ces réductions successives de logique théorique : les seuls critères adoptés sont purement empiriques.

[5] : Il s'agit des items n° 5, 8, 10, 11, 13, 14 et 15.

[6] : Ce sont les items n° 4, 5, 6, 9, 13, 14, 17, 18.

Le LSI-A a été soumis à de nombreuses analyses factorielles, exploratoires ou contraintes^[7] (e.g. Adams, 1969 ; Hoyt & Creech, 1983 ; Liang, 1984, 1985 ; James et al., 1986 ; Stock et al., 1986). Aucune d'elles ne permet de retrouver les cinq dimensions théoriques de cette échelle. Seules les dimensions de congruence entre les buts atteints et les buts désirés, d'humeur (appelée bonheur par Liang, 1985) et, à un moindre degré, celle d'enthousiasme semblent pouvoir être établies de manière consistante^[8]. Stock et al. (1986) proposent une solution différente à deux facteurs affectifs (affectivité positive et affectivité négative) et un facteur cognitif. Il semble clair dans tous les cas (cf. Liang, 1984, 1985) qu'un facteur de second ordre puisse rendre compte des corrélations entre les facteurs de premier ordre ; ce facteur peut être assimilé au bien-être général.

Les qualités psychométriques du LSI-A apparaissent satisfaisantes en termes de consistance interne (entre .62 et .84 : cf. Wood et al., 1969 ; Dobson et al., 1979 ; Kozma & Stones, 1980 ; Stock & Okun, 1982 ; Himmerfalb & Murrell, 1983) et en termes de validité convergente (entre .60 et .80 : cf. Lohmann, 1977, 1980).

C. La "*Philadelphia Geriatric Center - Morale Scale*" (PGC-MS) de Lawton (1972, 1975)

Mise au point par Lawton (1972), la PGC-MS est destinée à pouvoir être utilisée avec des personnes très âgées hébergées dans des institutions (cf. George & Bearon, 1980). La première version comprend 22 items, dichotomiques pour la plupart, mesurant six dimensions : enthousiasme, attitude envers sa propre vieillesse, satisfaction, émotion, optimisme et désappointement lié à la solitude. Très vite, Lawton (1975) sera amené à en proposer une version révisée ne comprenant que 17 items pour 3 dimensions, la structure factorielle mise en évidence lors de la première étude n'étant pas répliquée (ces résultats ont été confirmés dans une étude indépendante de Morris & Sherwood, 1975). Moins utilisée que le LSI-A, les indicateurs classiques de qualité psychométrique (consistance interne, stabilité temporelle, validité convergente) sont satisfaisants (cf. Lohman, 1977 ; Larson, 1978 ; George & Bearon, 1980). Plusieurs analyses factorielles ont été menées avec cette échelle ; elles confirment approximativement la structure dégagée par Lawton (cf. Morris & Sherwood, 1975). Les analyses factorielles contraintes effectuées par Liang et ses collaborateurs (Liang & Bollen, 1983, 1985 ; Liang et al.,

[7] : Nous utilisons l'expression "analyses factorielles contraintes" à la place de la traditionnelle "analyses factorielles confirmatoires" comme le suggèrent Hattie & Fraser (1988) ou Juhel (1990, qui parle d'analyses factorielles restrictives), pour souligner que, dans la pratique, ces analyses procèdent rarement d'une démarche réellement confirmatoire (cf. également à ce sujet Junker & Pilkonis, 1993).

[8] : Hoyt & Creech (1983) proposent d'interpréter ces trois facteurs comme, respectivement, la satisfaction avec le passé, avec le présent et avec le futur.

1987 a) permettent de mettre en évidence, tout comme pour le LSI-A, un facteur général de second ordre. Le libellé de ces items est présenté dans la figure 2.5.

1. Little things bother me more this year.
2. I sometimes worry so much that I can't sleep.
3. I am afraid of a lot of things.
4. I get mad more than I used to.
5. I take things hard.
6. I get upset easily.
7. Things keep getting worse as I get older.
8. I have as much pep as I had last year.
9. As you get older you are less useful.
10. As I get older, things are better/worse than I thought they would be.
11. I am as happy now as when I was younger.
12. How much do you feel lonely ?
13. I see enough of my friends and relatives.
14. I sometimes feel that life isn't worth living.
15. Life is hard for me much of the time.
16. How satisfied are you with your life today ?
17. I have a lot to be sad about.

Figure 2.5 : Les items de la "*Philadelphia Geriatric Center - Morale Scale*" (PGC-MS) de Lawton (1972, 1975)

D. La "*Memorial University of Newfoundland Scale of Happiness*" (MUNSH) de Kozma & Stones (1980)

La MUNSH, proposée par Kozma & Stones (1980), se veut être une tentative pour mesurer les composantes affectives du bien-être chez les personnes âgées. La stratégie adoptée pour construire cette échelle est totalement empirique. Les auteurs partent d'un ensemble de 72 items, composé à partir des items de la PGC-MS, du LSI-Z, de l'ABS et de nouveaux items rédigés selon le principe de l'ABS, en incluant notamment des items destinés à mesurer l'état affectif à long terme. La sélection des items s'effectue en fonction de leur corrélation avec deux questions portant sur le bonheur général des individus au moment de l'enquête et au cours du mois précédent. La version finale de la MUNSH comporte 24 items dichotomiques, la moitié d'entre eux étant positifs, l'autre moitié, négatifs. Selon les auteurs, les items se distinguent aussi selon qu'ils font référence à des expériences générales (14 items) ou à des expériences plus spécifiques (10 items). Les indicateurs de qualité psychométrique de cette échelle sont jugés satisfaisants (cf. Kozma & Stones, 1980 ; Kozma et al., 1985 ; Stones & Kozma, 1986 a). Une seule analyse factorielle a été menée, à notre connaissance, sur une sélection de 18 items de cette échelle (Kop, 1993a). Les résultats font apparaître un facteur d'affectivité positive et un facteur d'affectivité négative. Les items de la MUNSH sont présentés dans la figure 2.6.

On pourra constater que celle-ci comporte plusieurs items en commun avec le LSI-A, la PGC-MS et l'ABS.

- In the past months, have you been feeling :
1. On the top of the world ?
 2. In high spirits ?
 3. Particular content with your life ?
 4. Lucky ?
 5. Bored ?
 6. Very lonely or remote from other people ?
 7. Depressed or very unhappy ?
 8. Flustered because you didn't know what was expected of you ?
 9. Bitter about the way your life has turned out ?
 10. Generally satisfied with the way your life has turned out ?
-
11. This is the dreariest time of my life.
 12. I am just as happy as when I was younger.
 13. Most of the things I do are boring or monotonous.
 14. The things I do are as interesting to me as they ever were.
 15. As I look back on my life, I am fairly well satisfied.
 16. Things are getting worse as I get older.
 17. How much do you feel lonely ?
 18. Little things bother me more this year.
 19. If you could live where you wanted, where would you live ?
 20. I sometimes feel that life isn't worth living.
 21. I am as happy now as I was when I was younger.
 22. Life is hard for me most of the time.
 23. How satisfied are you with your life today ?
 24. My health is the same or better than most of people's my age.

Figure 2.6 : Les items de la "Memorial University of Newfoundland Scale of Happiness" (MUNSH) de Kozma & Stones (1980)

E. La "Satisfaction With Life Scale" (SWLS) de Diener et al. (1985 a)

L'objectif affiché par Diener et al. (1985 a) est de construire une échelle qui mesure la composante cognitive du bien-être. En effet, selon ces auteurs, le bien-être subjectif comporte deux composantes affectives (l'affectivité positive et l'affectivité négative) et une composante cognitive qui reflète la satisfaction avec la vie. Or, alors qu'il existe plusieurs instruments destinés à mesurer les composantes affectives, aucun n'est spécifiquement destiné à mesurer la composante cognitive. La démarche suivie pour mettre au point cette échelle est empirique. Les auteurs établissent, dans un premier temps, une liste de 48 items positifs et négatifs dont le contenu renvoie à la satisfaction avec la vie en général. Une analyse factorielle effectuée sur ces items permet de dégager trois facteurs que les auteurs interprètent respectivement comme des facteurs d'affectivité positive, d'affectivité négative et de satisfaction. Seuls sont retenus les 10 items ayant les

saturations les plus élevées sur ce dernier facteur. Finalement, l'élimination des items trop similaires sur le plan sémantique permet d'aboutir à une échelle de cinq items. Les sujets doivent indiquer, sur une échelle en 7 points, leur degré d'accord avec chaque item. On trouvera chez Pavot & Diener (1993 a) une synthèse de l'ensemble des travaux ayant porté sur cette échelle dont les qualités psychométriques apparaissent tout à fait acceptables. La figure 2.7 présente cette échelle.

1. In most ways, my life is close to my ideal.
2. The conditions of my life are excellent.
3. I am satisfied with my life.
4. So far I have gotten the important things I want in my life.
5. If I could live my life other, I would change almost nothing.

Figure 2.7 : Items de la "*Satisfaction With Life Scale*" (SWLS) de Diener et al. (1985)

2.1.3 D'autres mesures du bien-être subjectif

Même si les questionnaires auto-administrés constituent l'immense majorité des instruments mis au point pour mesurer le bien-être, il faut signaler l'existence de quelques techniques originales qui permettent d'obtenir des "mesures non traditionnelles du bien-être", comme les appelle Diener (1994). C'est d'ailleurs dans le laboratoire de ce dernier, à l'Université d'Illinois, qu'ont été élaborées la plupart de ces techniques qui ont servi, le plus souvent, à des études de validation des questionnaires.

La première de ces techniques consiste à mesurer quotidiennement l'humeur des individus pendant plusieurs jours (ou semaines) consécutifs à partir d'adjectifs soigneusement sélectionnés. Bien qu'étant toujours basée sur des auto-évaluations, cette technique permet d'éviter certains inconvénients liés à l'utilisation de questionnaires, comme les effets de mémoire ou l'influence de la situation. En effet, on utilise comme mesure du bien-être le score moyen des sujets sur l'ensemble de la période d'observation. On a eu recours à plusieurs variantes de cette technique, dont on trouvera une présentation générale chez Hormuth (1986). Dans certains cas (e.g. Diener & Emmons, 1985 ; Sandvik et al., 1993), les sujets remplissent les questionnaires en fin de journée, avant de se coucher ; dans d'autres, le moment de la journée est déterminé aléatoirement par la sonnerie d'un petit appareil qu'ils portent sur eux (e.g. Diener & Emmons, 1985 ; Csikszentmihalyi & Wong, 1991). Outre la lourdeur de cette technique, il faut signaler qu'elle repose sur le postulat implicite selon lequel le bien-être ne serait rien d'autre que la somme des humeurs quotidiennes.

Méthode traditionnelle d'évaluation de la validité des questionnaires de personnalité (Moskowitz, 1986), l'utilisation des réponses fournies par des personnes proches du sujet a également été pratiquée dans le cadre du bien-être (e.g. Kammann et al., 1984 b ; Sandvik et al., 1993). On demande à des personnes qui connaissent bien le sujet de remplir le même questionnaire ou de donner des évaluations plus globales.

Nous avons déjà évoqué, avec le LSR de Neugarten et al. (1961), la possibilité de recourir à des approches plus qualitatives se basant sur des entretiens menés par des psychologues cliniciens. Cette technique a été utilisée entre autres par Thomas & Chambers (1989), Sandvik et al. (1993) ou Shedler et al. (1993).

Seidlitz & Diener (1993) ont proposé d'utiliser comme mesure du bien-être la différence entre le nombre d'événements personnels positifs qu'un sujet est capable de se rappeler pendant une période de trois minutes et le nombre d'événements négatifs rappelés pendant une période de durée égale^[9]. L'idée sous-jacente à cette procédure est que les personnes ont tendance à oublier les événements non congruents avec leur niveau de bien-être. Cette technique n'est pas sans rappeler certaines tentatives pour obtenir des mesures indirectes des attitudes. C'est le cas, par exemple dans l'étude de Hammond (1948, cité par Pratkanis, 1989) qui utilise le fait que les attitudes introduisent des erreurs dans les jugements. Dans cette étude, on demande à des sujets de choisir l'une des deux propositions, toutes deux fausses, qu'on leur présente. Comme les sujets ont tendance à choisir la proposition en accord avec leur attitude, on en obtient, indirectement, une mesure.

De rares tentatives visant à obtenir des indicateurs comportementaux du bien-être ont été faites. Weinstein (1982) a mesuré, pendant l'entretien, le nombre de rires et de sourires. Kammann et al. (1984 a) ont utilisé un questionnaire basé sur des comportements. Malheureusement, dans ce dernier cas, les auteurs ne fournissent aucune indication quant à la manière dont les comportements ont été choisis ; de plus, certains items retenus peuvent difficilement passer pour des comportements (e.g. optimiste, irrité, attitude positive envers les gens...).

Pour répondre à la question de l'influence de la désirabilité sociale sur les mesures de bien-être (cf. section 2.3.2), Sandvik et al. (1993) ont élaboré un questionnaire dans lequel les individus sont forcés de choisir entre 54 couples d'adjectifs composés d'un adjectif désignant un trait de personnalité et d'un autre censé mesurer le bien-être ; les couples ont été constitués de telle sorte que les deux adjectifs ont le même niveau de désirabilité sociale.

[9] : Une variante consiste à demander au sujet de se rappeler les événements positifs et négatifs ayant eu lieu pendant l'année écoulée (cf. Sandvik et al., 1993).

Récemment, Burgess & Haaga (1994) ont proposé un instrument d'évaluation des "pensées positives automatiques", qui pourrait devenir une mesure indirecte du bien-être. Les sujets se voient proposer une liste de pensées positives et doivent indiquer, pour chacune d'entre elles, s'ils ont eu souvent ou non cette pensée. Les premiers résultats montrent que cette échelle est corrélée positivement avec une échelle d'affectivité positive et négativement avec une échelle de dépression. Il reste bien évidemment à confirmer ces résultats.

Enfin, Diener (1994) évoque la possibilité d'utiliser des indices physiologiques (mesures électromyographiques, taux hormonaux, etc.) et/ou comportementaux (consommation d'alcool, manque d'appétit, troubles du sommeil...).

L'intérêt de développer des techniques alternatives pour la mesure du bien-être ne semble guère faire de doute. Il faut toutefois se garder, comme on peut le constater dans certains articles, de faire de ces nouvelles mesures des indicateurs absolus, sans erreur. Les résultats obtenus par ces techniques sont également des mesures et celles-ci doivent, comme toutes les mesures, être validées. Oublier cette évidence reviendrait à considérer que, par essence, certaines mesures sont plus "vraies" que d'autres.

2.2 Evaluation psychométrique des mesures de bien-être

Tous les auteurs qui ont entrepris d'analyser les mesures de bien-être semblent conclure — quelquefois avec surprise — à la bonne qualité psychométrique de ces instruments (cf. Diener, 1984 ; Larsen et al., 1985 ; McNeil et al., 1986 a ; Okun & Stock, 1987 ; Andrews & Robinson, 1991 ; Sandvik et al., 1993 ; Diener, 1994), même si certains notent quelques problèmes non encore totalement résolus (cf. Diener & Fujita, 1994). Les évaluations psychométriques des instruments de mesure du bien-être se sont toujours heurtées à un problème majeur, l'absence d'un critère objectif univoque, comportemental ou non. Les conditions (objectives) de vie ne pouvaient pas, à l'évidence, jouer ce rôle étant donné que les indicateurs subjectifs du bien-être se sont développés en partie à cause de l'incapacité des indicateurs objectifs à rendre compte de l'expérience vécue des individus. Il n'est certes pas très facile d'imaginer des procédures de validation critérielle des mesures de bien-être. Mais il faut noter que pendant très longtemps le climat scientifique des études du bien-être n'était guère propice à de telles investigations. En effet, et nous le verrons en détail, pendant toute la période qui a connu le développement rapide des études sur le bien-être, les chercheurs n'étaient intéressés que par sa prédiction. Autrement dit, pendant très longtemps, le bien-être subjectif a été exclusivement appréhendé en tant que variable dépendante, endogène. Sans que cela ne soit écrit nulle part, il aurait semblé tout à fait inconvenant d'envisager que le bien-être

puisse être la cause d'une autre variable. Ce n'est que depuis peu de temps que cette situation commence à évoluer. Il est encore trop tôt, pourtant, pour qu'un tel changement dans les mentalités ait un impact sur les procédures de validation.

Les études psychométriques des mesures de bien-être sont restées assez classiques. Elles se déclinent autour d'analyses de fidélité, de validité convergente et divergente, d'analyses factorielles. Ces dernières seront traitées séparément dans le chapitre 4. Les travaux de Andrews et de ses collègues de l'Université du Michigan (e.g. Andrews & Crandall, 1976 ; Andrews & Whitey, 1976 ; Rodgers et al., 1988 ; Rodgers, 1989) constituent une application originale des modèles d'équations structurales à la validité de construit. La validité de contenu a été tout particulièrement négligée dans la littérature^[10]. Certains reconnaissent d'ailleurs que le choix des items se fait de manière arbitraire (e.g. Campbell et al., 1976, à propos du choix d'items de satisfaction dans différents domaines de vie). L'évolution des instruments de mesure constitue une autre illustration de ce désintérêt : nous avons vu que les questionnaires les plus récents s'appuient souvent sur les items des questionnaires antérieurs. L'exception que constitue l'étude de Roberts & Clifton (1992), dans laquelle il a été demandé à des juges d'évaluer la pertinence et la représentativité d'items par rapport à six dimensions destinées à mesurer le bien-être d'étudiants, ne change pas grand chose : dans ce cas c'est le choix des six dimensions théoriques qui n'est pas justifié.

Cette section comprendra donc trois parties : fidélité, validité convergente et divergente, validité de construit par les modèles d'équations structurales. Nous serons relativement brefs car les résultats ne soulèvent pas beaucoup de questions : tous les indices convergent pour montrer que les questionnaires de bien-être mesurent bien quelque chose. Il reste à savoir quoi !

2.2.1 Fidélité des mesures de bien-être

La question de la fidélité des mesures de bien-être est traditionnellement abordée à partir de leur consistance interne et de leur stabilité temporelle. La première fait principalement référence à l'homogénéité du contenu des items. C'est une condition nécessaire mais non suffisante pour assurer la validité des mesures. L'homogénéité du contenu est évidemment une question théorique, mais elle peut être obtenue de manière mécanique en augmentant la redondance des items. La seconde pose, quant à elle, des problèmes théoriques plus fondamentaux qui ne seront qu'évoqués ici. En effet, la

[10] : A titre purement anecdotique, citons l'étude de Larsen et al. (1985) dont l'objectif est tout simplement d'évaluer les mesures de bien-être subjectif. Dans le plan de l'article annoncé en introduction, la validité de contenu y figure en bonne place...avant de disparaître totalement sans aucune explication.

stabilité temporelle des mesures n'est désirable qu'à partir du moment où le bien-être est conceptualisé comme un trait relativement stable. En revanche, si le bien-être est plutôt conçu comme un état transitoire oscillant de manière plus ou moins importante, la stabilité des mesures est contraire aux attentes. Il faut bien reconnaître, pourtant, que cette question de la stabilité a rarement été envisagée en termes théoriques. Elle donne lieu à des vérifications routinières^[11], parce que l'évaluation de la fidélité test-retest fait partie de la démarche psychométrique classique, y compris chez des auteurs pour qui le bien-être est plus un état qu'un trait. La vérification de la stabilité temporelle fait effectivement partie de la démarche de validation lorsque celle-ci est envisagée comme le test d'hypothèses théoriques. Mais ces hypothèses doivent être rendues explicites pour éviter à la démarche toute contradiction interne.

La consistance interne a été vérifiée pour pratiquement toutes les échelles publiées, avec des procédures standards telles que le calcul d'un coefficient α de Cronbach (1951). Dans leur méta-analyse, Okun & Stock (1987)^[12] rapportent un coefficient moyen de .81 ($\sigma = .08$; $n = 31$). Il y a bien évidemment des variations selon les échelles. Ainsi, par exemple, les échelles d'affectivité positive et d'affectivité négative de l'ABS qui comprennent chacune cinq items semblent moins consistantes (cf. Larsen et al., 1985) que la SWLS qui comprend également cinq items (cf. Pavot & Diener, 1993 a). Malgré ces différences, la consistance interne des mesures de bien-être est jugée très satisfaisante.

En ce qui concerne la stabilité temporelle, Okun & Stock (1987) font état d'un impressionnant coefficient moyen de .74 ($\sigma = .11$; $n = 10$) pour des échelles comportant plusieurs items et d'un coefficient de .69 ($\sigma = .04$; $n = 5$) pour des mesures du bien-être basées sur un seul item. Les corrélations semblent d'autant plus élevées que les périodes temporelles sont courtes, ce qui n'est pas étonnant. Mais même sur de longues périodes, les corrélations sont encore élevées (e.g. entre .40 et .56 après 4 ans chez Palmore & Kivett, 1977 ; .61 après 3 ans chez Baur & Okun, 1983 ; .71 après 18 mois chez Kozma et al., 1990 ; etc.).

[11] : Un exemple particulièrement caricatural peut être cité à ce sujet. Dans un article publié en 1991, Yardley & Rice se proposent d'étudier les relations entre l'humeur et le bien-être à l'aide d'une seconde enquête répétée 10 semaines après la première. Les corrélations "test-retest" calculées pour la mesure de bien-être et celle de l'humeur sont respectivement de .50 et de .48, soit une stabilité temporelle identique pour un trait et pour un état psychologique, ce qui ne semble pas émouvoir les auteurs !

[12] : Cette méta-analyse porte sur un corpus de 556 sources antérieures à 1980 dont les titres comprennent l'un des termes suivants : satisfaction, bonheur, moral, qualité de la vie, bien-être.

2.2.2 Validité convergente et validité divergente

La validité convergente a été largement étudiée à partir des corrélations entre plusieurs questionnaires. Chez Okun & Stock (1987), on trouve un coefficient moyen de .52 ($\sigma = .20$; $n = 77$), ce qui reflète assez fidèlement ce que l'on trouve habituellement dans la littérature. Il y a deux façons d'interpréter ce coefficient : on peut le trouver acceptable ou trop faible. Il est clair en tout cas qu'il est inférieur aux estimations de consistance interne, ce qui semble indiquer que certaines échelles mesurent des aspects spécifiques qui ne sont pas pris en compte par d'autres.

Le procédé utilisé pour assurer la validité divergente des mesures de bien-être consiste le plus souvent à examiner les corrélations avec les mesures de concepts voisins tels que l'adaptation (corrélation moyenne = .38 ; écart-type des corrélations = .19 ; $n = 25$) et le névrosisme ($m = -.33$; $\sigma = .13$; $n = 31$) (cf. Stock & Okun, 1987), l'optimisme et l'estime de soi (cf. Diener, 1994) ou encore la sociabilité et l'activité (cf. Larsen et al., 1985). Il est important de noter que ce procédé repose sur une logique de validation interconcepts. Il s'agit à la fois de montrer que les mesures de bien-être corrélaient avec d'autres concepts proches et quelles possèdent une part de variance spécifique. Le plus souvent pourtant, ces analyses ne reposent pas sur une démarche théorique suffisamment explicitée^[13]. De plus, nous avons vu que dans certains cas, avec des concepts comme le sens de la vie, les corrélations sont tellement élevées que certains auteurs sont amenés à s'interroger sur l'identité des deux concepts.

Les réponses aux questionnaires de bien-être ont aussi été confrontées avec les résultats obtenus en utilisant les "mesures non traditionnelles" du bien-être brièvement décrites dans le paragraphe précédent. Sandvik et al. (1993) ont effectué plusieurs études allant dans ce sens. Les corrélations entre des questionnaires de bien-être et des évaluations issues d'entretiens cliniques varient entre .60 et .71 ; celles qui portent sur la différence entre le rappel d'événements positifs et négatifs, entre .20 et .41 ; celles qui utilisent le questionnaire élaboré pour contrôler la désirabilité sociale entre .27 et .50. Les corrélations entre les questionnaires d'auto-évaluation du bien-être et les réponses fournies par des personnes proches du sujet sont relativement élevées : elles oscillent entre .34 et .58 (Sandvik et al., 1993). Pourtant, Kamman et al. (1984 b) ont montré que l'accord soi-autrui dans l'évaluation du bien-être était en général assez faible, ce qui ne les étonne pas outre mesure. En effet, si le bien-être est une expérience subjective individuelle, il n'y a aucune raison pour qu'elle soit accessible à une autre personne, aussi proche soit-elle. Les divergences entre les résultats de ces deux études peuvent s'expliquer partiellement par le fait que dans celle de Sandvik et al. (1993), ce sont des évaluations moyennes d'un minimum de sept personnes proches du sujet qui ont été utilisées. Ceci

[13] : De manière plus générale, on peut tirer arguments de tous les travaux étudiant les corrélations entre le bien-être et d'autres variables. Ceux-ci sont très nombreux ainsi qu'on le verra dans le chapitre 4.

rejoindrait les explications de certains spécialistes de la personnalité (e.g. Moskowitz, 1986), qui considèrent que les corrélations augmentent en même temps que le nombre des personnes qui effectuent ces évaluations, car chacune d'elles perçoit le sujet dans des situations différentes.

La littérature présente aussi plusieurs résultats, souvent disséminés, sur la convergence entre les réponses à des questionnaires rétrospectifs classiques et les réponses moyennes à des interrogations répétées quotidiennement. En général, les auteurs (cf. Larsen et al., 1985 ; Sandvik et al., 1993 ; Diener, 1994) concluent à une convergence satisfaisante entre les résultats obtenus grâce à ces deux méthodes. Pourtant, quelques réserves peuvent être émises vis-à-vis de cette conclusion. A travers les publications de Diener et de son équipe, on trouve en effet, dans certains cas, des corrélations beaucoup plus faibles que celles qui sont rapportées, en général, dans les articles traitant de la convergence des mesures. Chez Emmons & Diener (1986), par exemple, la corrélation entre les niveaux moyens d'affectivité positive et les résultats à la sous-échelle positive de l'ABS de Bradburn n'est que de .32 ; pour l'affectivité négative, elle est encore plus faible ($r = .22$). Il n'est pas facile d'expliquer ces divergences de résultats et les auteurs eux-mêmes se gardent bien de les relever. Il n'est pas exclu que les procédures utilisées n'aient pas été les mêmes dans tous les cas. Chez Sandvik et al. (1993), par exemple, on ne dispose pas de la liste des items du questionnaire administré quotidiennement. De plus, toujours dans ce même article, l'affectivité positive et l'affectivité négative ne sont pas distinguées, ni pour le questionnaire quotidien, ni pour l'ABS : la légitimité de cette procédure a déjà été questionnée plus haut. Il est dès lors difficile d'évaluer réellement la convergence des données obtenues par questionnaire rétrospectif et par calcul du niveau moyen d'affectivité sur plusieurs jours consécutifs. Cette mise en correspondance empirique pose aussi, par ailleurs, des questions théoriques plus fondamentales qui ont déjà été évoquées.

Enfin, en ce qui concerne les indices comportementaux, les résultats sont rares mais intéressants. Stones & Kozma (1980 b, cités par McNeil et al., 1986 a) auraient trouvé une corrélation de .60 entre une échelle de bien-être et des indicateurs comportementaux comme le rire et le sourire. Weinstein (1982), quant à lui, a montré que le nombre de rires et de sourires relevé pendant un entretien de 45 minutes était lié aux réponses des individus à une question sur leur niveau de bien-être. Ces deux études sont plutôt encourageantes et incitent à considérer le lien entre évaluation du niveau de bien-être et mesures de comportement avec plus d'attention.

2.2.3 Validité de construit estimée par les modèles d'équations structurales

C'est à Andrews (cf. Andrews & Crandall, 1974 ; Andrews & Withey, 1976) et à certains de ses collègues (cf. Rogers et al., 1988 ; Rodgers, 1989) que l'on doit une application des modèles d'équations structurales à l'estimation de la validité de construit des mesures de bien-être. Selon cette perspective, la validité de construit peut être exprimée par un coefficient qui est la corrélation entre la variable manifeste et la variable latente du modèle de mesure élaboré.

L'idée de base de cette technique consiste à essayer de décomposer la variance des réponses à un questionnaire de bien-être^[14] en trois parties : une partie due au trait latent hypothétique qui est censé sous-tendre les réponses (c'est la variance "vraie") ; une partie due à la méthode spécifique utilisée ; une partie due à des facteurs aléatoires. Pour pouvoir distinguer entre la variance due à la méthode et la variance due aux facteurs aléatoires, il faut que des observations aient été recueillies à partir de méthodes différentes. C'est le cas, par exemple, lorsque l'on analyse une matrice multitraits-multiméthodes (cf. section 2.2.2, première partie). Dans ce cas, on peut obtenir des estimations de ces trois composantes de la variance totale. La figure 2.8 illustre un modèle de mesure dans lequel on trouve neuf variables observées mesurant trois traits obtenus avec trois méthodes différentes. Y_{ij} désigne la variable observée correspondant au trait i mesuré avec la méthode j ; ξ_i désigne le facteur latent correspondant au trait i ; ξ_j^* désigne le facteur latent correspondant à l'utilisation de la méthode j ; δ_{ij} désigne la part d'erreur aléatoire (i.e. qui n'est due ni au trait, ni à la méthode) associée à la variable observée Y_{ij} .

Ce type de modèles permet à Andrews & Crandall (1976) d'estimer que 60% de la variance des réponses est due au trait mesuré, ce qui est très satisfaisant^[15]. Ce chiffre est supérieur à celui que l'on obtient habituellement, avec cette méthode, pour des construits psychologiques. A titre de comparaison, Buckley et al. (1990, mais cf. aussi Cote & Buckley, 1987) estiment, après une analyse secondaire de 56 études ayant utilisé des matrices multitraits-multiméthodes, qu'en moyenne, 42% de la variance peut être attribuée au trait, ce pourcentage étant plus faible dans le cas de traits de personnalité (33% en moyenne).

Avec une technique identique, Rodgers (1989) confronte différents items de bien-être en fonction du nombre de propositions de réponse qui sont offertes aux sujets. Les résultats montrent que les items assortis d'échelles de réponses en 7 ou en 11 points ont

[14] : Ou, tout simplement, à un item censé mesurer le bien-être.

[15] : Buckley et al. (1992) ont effectué des analyses secondaires de matrices multitraits-multiméthodes en relation avec la satisfaction avec le travail. Celles-ci permettent d'attribuer 50% de la variance des réponses au trait.

une part de variance d'erreur aléatoire et systématique (i.e. due à la méthode) plus faible que les items pour lesquels les échelles de réponses sont moins fines (3 ou 5 points). De plus, les items dont tous les points de l'échelle de réponses sont labellisés semblent moins valides que les autres. Mais de manière générale, toutes ces études montrent que la part de variance due à la méthode est importante pour les items de bien-être (cf. Rodgers et al., 1988)^[16].

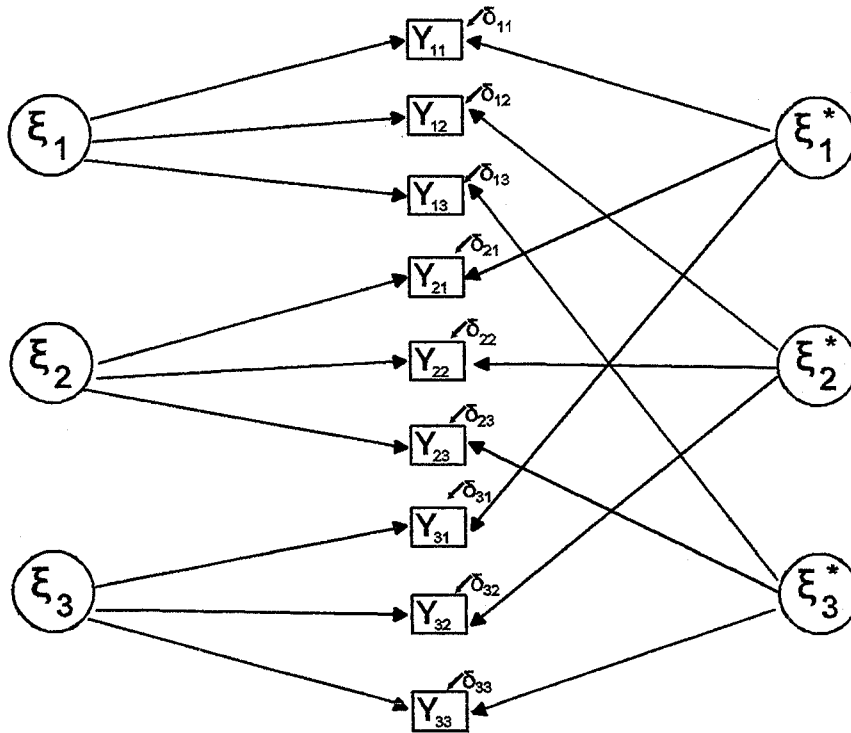


Figure 2.8 : Exemple de modèle de mesure pour neuf variables mesurant trois traits avec trois méthodes différentes

Notons enfin que les résultats obtenus avec cette technique doivent être relativisés. En effet, dans la plupart des cas, en fait de méthodes différentes, les auteurs se contentent de changer l'échelle de réponses des items. Les résultats auraient beaucoup plus d'impact si les matrices MTMM analysées étaient construites à partir de méthodes réellement

[16] : Ces mises en garde ont conduit certains auteurs à se demander si les corrélations entre différentes mesures de concepts liés au bien-être n'avaient pas été surestimées à cause de l'utilisation de mêmes méthodes. Lance & Sloan (1993) ont entrepris de comparer les corrélations "brutes" (entre les variables observées) et les corrélations vraies (entre les variables latentes) entre différents traits. A leur grande surprise, les corrélations brutes se sont avérées identiques aux corrélations vraies. Ce résultat peut être interprété en faisant référence à deux influences dont les effets semblent se compenser : l'absence de fidélité des mesures tend à faire baisser les corrélations brutes alors que le fait de partager une même méthode tend à les faire augmenter.

différentes comme, par exemple, les mesures non traditionnelles du bien-être dont nous avons parlé.

2.3 Les critiques adressées aux mesures de bien-être

Malgré les arguments accumulés en faveur de la qualité psychométrique des mesures de bien-être, celles-ci ont essuyé de nombreuses critiques qu'il importe de décrire et d'évaluer. Certaines de ces critiques sont très générales. Elles contestent la possibilité même de mesurer le bien-être du fait qu'il revêt un caractère totalement subjectif et idiosyncrasique. Ces caractéristiques rendraient impossible toute approche interindividuelle et quantitative du bien-être. Dans le même ordre d'idée, certaines critiques portent sur la conscience des états de bien-être. Les individus n'auraient pas les informations suffisantes ou le recul nécessaire pour évaluer leur propre bien-être. Plus prosaïquement, certains suggèrent que la plupart des individus n'ont jamais évalué s'ils étaient heureux ou non. En réponse à ces critiques, Veenhoven (1984) rapporte plusieurs études qui montrent que les items des questionnaires de bien-être sont bien acceptés par les sujets (les non-réponses sont très peu nombreuses) et que le bien-être constitue un thème central pour beaucoup d'individus. Il est plus difficile de répondre aux critiques concernant le caractère idiosyncrasique du bien-être. Cependant, ce débat concerne bien d'autres concepts de la psychologie.

D'autres critiques sont plus spécifiques et s'appuient sur des résultats empiriques. Les nombreux travaux de Norbert Schwarz et ses collaborateurs, qui adoptent le cadre de référence de la cognition sociale, s'inscrivent dans cette perspective. Pour Schwarz & Strack (1991), par exemple, les réponses aux questionnaires de bien-être ne peuvent être considérées comme des indicateurs valides de l'état interne des sujets. En effet, celles-ci sont des jugements, construits par les sujets, en fonction de nombreuses influences transitoires, telles que l'humeur ou le contexte. Ces travaux ont le mérite d'éclairer certains des processus utilisés par les sujets pour répondre aux items qui leur sont proposés. Menés dans une tradition expérimentale, ils en ont les qualités, mais aussi les défauts. En particulier, les variables dépendantes sont trop souvent réduites à un seul item de bien-être. Or, on connaît les faiblesses des mesures basées sur un seul item (cf. Epstein, 1979) : la composante d'erreur de mesure y est très élevée et on maximise ainsi la probabilité de mettre en évidence des effets expérimentaux. D'ailleurs, Pavot & Diener (1993 b) ont répliqué certaines études de Schwarz et de ses collaborateurs en utilisant une mesure de bien-être composée de plusieurs items. Leurs travaux montrent que si les résultats de ces effets transitoires ne disparaissent pas totalement, ils sont tout de même fortement réduits.

Ainsi, il faut envisager les résultats de ces études avec prudence. On ne peut guère les ignorer car elles posent des questions pertinentes par rapport à la mesure du bien-être. Mais ne pas les ignorer ne signifie pas adhérer complètement au modèle du bien-être présenté par Schwarz & Strack (1991)^[17]. Celui-ci présente un schéma cohérent permettant d'intégrer les régularités expérimentales ayant été mises en évidence en termes d'influences transitoires. Toutefois, en se limitant à des influences de ce type, ce modèle est incapable de rendre compte de résultats comme celui de la stabilité des mesures de bien-être.

Dans cette section, nous examinerons plus en détail les critiques principales qui ont été adressées au bien-être. Il s'agit de l'influence de l'humeur du moment sur les réponses aux questionnaires, de l'influence de la désirabilité sociale et de celle du contexte. Le dernier point n'est pas vraiment une critique, même si certains l'ont envisagé ainsi : il concerne la forme de la distribution des réponses aux items de bien-être.

2.3.1 Influence de l'humeur du moment

Plusieurs études expérimentales ont montré que les réponses à des items de bien-être étaient influencées par l'humeur du moment. Schwarz et ses collaborateurs ont notamment montré, avec des manipulations expérimentales de l'humeur, que les sujets sont plus heureux^[18] : lorsqu'il fait beau (Schwarz et Clore, 1983), lorsqu'ils trouvent de l'argent sur une photocopieuse (Schwarz, 1983), lorsqu'ils doivent se rappeler des événements positifs (Strack et al., 1985), lorsqu'ils reçoivent une barre de chocolat (Münkel et al., 1987), lorsqu'ils se trouvent dans une pièce agréable (Schwarz et al., 1987) ou lorsqu'ils regardent leur équipe nationale gagner un match de football (Schwarz et al., 1987). Des résultats similaires sont aussi obtenus dans des études non expérimentales (cf. Moum, 1988 ; Kozma et al., 1990 ; Yardley & Rice, 1991).

Malgré ces nombreuses études, il semble que l'effet de l'humeur soit moindre que ce que ces résultats laissent suggérer, notamment dans des situations naturelles et lorsque les variables dépendantes comportent plusieurs items (cf. Diener et al., 1991 a ; Pavot & Diener, 1993 b). De plus, il s'agit sans doute d'effets complexes dont on ne sait pas grand chose. Schwarz & Clore (1983) montrent, par exemple, que l'effet des conditions

[17] : Cf. chapitre 5.

[18] : C'est-à-dire, bien évidemment, que les moyennes des réponses aux items de bien-être sont plus élevées.

atmosphériques disparaît s'il est fait référence, lors de l'entretien, au temps qu'il fait^[19]. Sur un plan plus général, en ce qui concerne l'influence de l'humeur sur la valence des descriptions de soi, Sedikides (1994) a montré que la valence de ces auto-descriptions n'est congruente avec l'humeur induite expérimentalement que dans un premier temps. Après un certain délai, cette valence a tendance à s'opposer à l'humeur.

Même si les effets de l'humeur sur les réponses aux items de bien-être sont faibles, il est difficile de nier totalement leur existence. On peut, dans ce cas, se demander pourquoi l'humeur a cette influence. Pour Schwarz, l'humeur constitue un indice peu coûteux de son bien-être général (Schwarz & Clore, 1983 ; Schwarz & Strack, 1991). En effet, selon cet auteur, la tâche consistant à faire une évaluation globale de sa vie est extrêmement complexe et il est probable que les sujets fassent un tel effort lorsqu'ils répondent à un questionnaire^[20]. Selon les principes de la cognition sociale, c'est l'information la plus directement accessible qui est utilisée par les sujets. Dans le cas du bien-être, l'humeur constitue l'une des informations les plus directes et les plus pertinentes.

Une autre explication fait intervenir les liens entre humeur et mémoire. De nombreuses études expérimentales ont montré qu'il existait une certaine congruence entre l'humeur du moment et la valence du matériel retrouvé en mémoire^[21]. Deux explications (cf. Blaney, 1986 ; Bower, 1981) sont avancées pour expliquer un tel résultat. Dans la première, l'humeur a une influence sur l'apprentissage, sans en avoir sur le rappel : les sujets apprennent mieux les informations congruentes avec leur humeur au moment de l'apprentissage. Dans la seconde, l'humeur facilite le rappel d'informations congruentes avec l'humeur actuelle. De ces deux explications, seule la seconde peut permettre de rendre compte des effets de l'humeur sur les réponses aux items de bien-être. Dans ce cas, en effet, on peut penser que les sujets, avant de répondre, se remémorent des expériences congruentes avec leur humeur du moment et utilisent ces expériences pour évaluer leur bien-être.

Que pouvons-nous conclure de cette discussion ? Il semble établi que l'humeur peut influencer les auto-évaluations du bien-être. Néanmoins, ces influences semblent

[19] : Cette manipulation expérimentale doit amener le sujet, selon les auteurs, à se dire que son humeur provient du temps qu'il fait. Dans ce cas, l'humeur est attribuée à des causes externes transitoires et ne peut plus être utilisée comme indicateur de son bien-être général (cf. plus bas, les explications possibles de l'influence de l'humeur). Plus récemment, Schwarz et ses collègues (cf. Levine et al., 1994) ont admis que cette influence était plus probable dans le cas d'évaluations générales que dans celui d'évaluations plus spécifiques.

[20] : A moins de considérer que les individus effectuent cette évaluation en permanence, en la réajustant en fonction de leurs nouvelles expériences. Cette hypothèse n'est jamais évoquée dans la littérature.

[21] : Mais cf. Parrot & Sabini (1990, cités par Seidlitz & Diener, 1993) qui ne retrouvent pas ce résultat classique dans des conditions naturelles (c'est-à-dire lorsque l'humeur n'est pas induite expérimentalement).

limitées et n'ont pas encore reçu d'explication définitive. Par ailleurs, les résultats des études corrélationnelles ne confirment qu'en partie les résultats des études expérimentales dans lesquelles l'humeur n'est pas mesurée mais induite par des procédures dont on peut questionner l'efficacité (cf. Gerrardshesse et al., 1994). De plus, les études expérimentales ne tiennent jamais compte du fait que si l'humeur peut influencer les réponses à des items de bien-être, on peut tout aussi bien faire l'hypothèse que le niveau de bien-être d'un individu peut influencer son humeur et sa susceptibilité aux manipulations expérimentales.

La formalisation la plus explicite des liens réciproques entre humeur et bien-être (appelé, dans ce cas, "humeur-trait") est sans doute celle de Eid et al. (1994) qui s'appuie sur les modèles "traits-états latents" développés par Steyer & Schmitt (*in press a*, *in press b*). Dans ces modèles, la variance des réponses des sujets est décomposée en une partie due au trait, une partie due à la situation (i.e. à l'état du moment) et à l'interaction entre le trait et la situation, une partie due à la méthode et une partie due à des facteurs aléatoires. Dans l'application de cette méthode à la mesure de l'humeur que proposent Eid et al. (1994), deux échelles sont utilisées. La première (ML) est destinée à mesurer l'humeur-trait (le bien-être), la seconde (PU) est destinée à mesurer l'humeur-état. Ces deux échelles sont proposées à quatre reprises, à trois semaines d'intervalle, à un échantillon de 488 sujets. Le modèle testé dans cette étude est présenté dans la figure 2.9. Il est estimé au moyen d'un modèle d'équations structurales (LISREL).

Ce qui est intéressant pour notre propos, dans cet exemple, c'est que le modèle présenté fait intervenir des facteurs de situation tant pour les réponses au questionnaire destiné à mesurer l'humeur que pour celles au questionnaire destiné à mesurer le bien-être. De plus, ce modèle inclut explicitement un facteur de bien-être général qui rend compte des relations entre les facteurs de situation, y compris pour le questionnaire-état. Autrement dit, on trouve dans ce modèle l'idée que le bien-être peut influencer l'humeur, ce qui – intuitivement du moins – semble être une hypothèse plausible. Enfin, signalons que le facteur latent-trait du questionnaire-trait (ξ) est corrélé à .78 avec le facteur latent-trait du questionnaire-état (ξ^*), ce qui semble indiquer qu'au-delà de leurs différences, les deux questionnaires mesurent des construits relativement similaires.

Les relations entre humeur et bien-être ont donc tout lieu d'être reconsidérées avec plus d'attention dans la littérature. Celles-ci sont trop souvent limitées à la recherche d'hypothétiques biais dans les réponses. Sans négliger pour autant l'importance de ces derniers, il nous paraît plus important d'essayer de clarifier les relations théoriques qui existent entre humeur et bien-être.

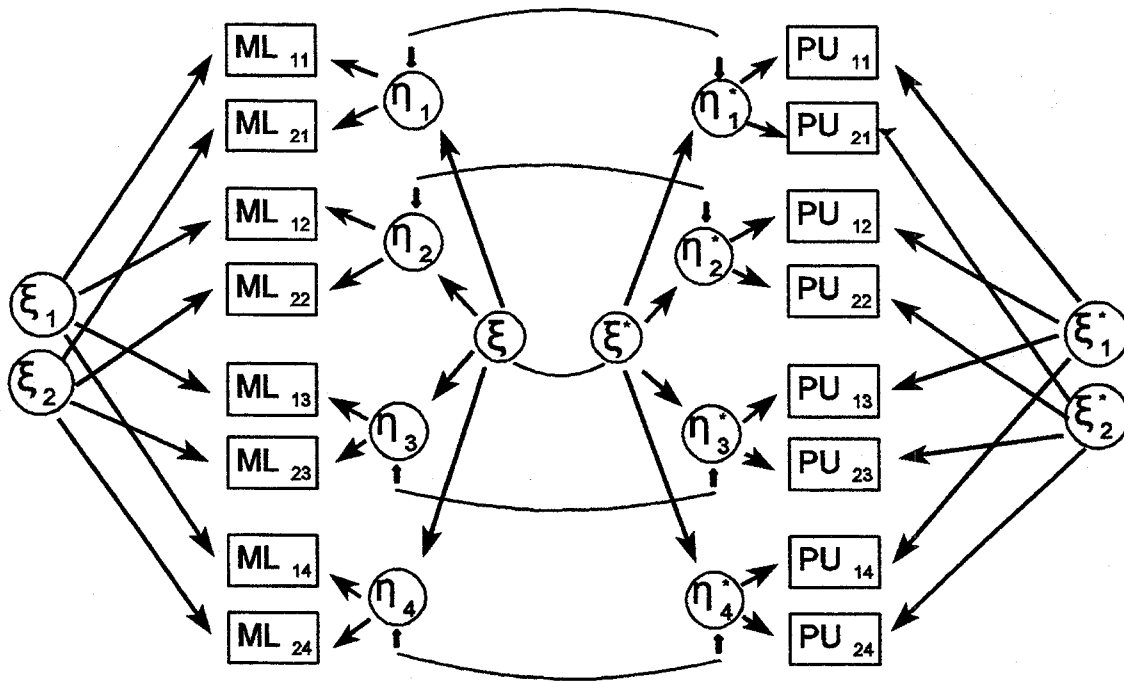


Figure 2.9 : Un modèle "traits-états latents" de l'humeur selon Eid et al. (1994)

Légende

ML_{ij} désigne les réponses observées à l'occasion j à la partie i du questionnaire trait

PU_{ij} désigne les réponses observées à l'occasion j à la partie i du questionnaire état

ξ et ξ^* désignent les facteurs latents-trait correspondant respectivement au questionnaire-trait et au questionnaire-état

ξ_i et ξ_i^* désignent les facteurs de méthodes correspondant respectivement au questionnaire-trait et au questionnaire-état

η_i et η_i^* désignent les facteurs latents-états correspondant respectivement au questionnaire-trait et au questionnaire-état

2.3.2 Influence de la désirabilité sociale

Lorsque Carstensen & Cone rapportent, en 1983, des corrélations supérieures à .60 entre le bien-être et la désirabilité sociale, leur article fait l'effet d'une petite bombe dans le milieu de la recherche sur le bien-être. Immédiatement, un débat virulent s'instaure par l'intermédiaire du *Journal of Gerontology* (cf. Kozma & Stones, 1985 ; McCrae & Costa, 1985 ; Cone & Carstensen, 1985). Depuis cette date, de nombreuses recherches ont permis de mieux comprendre les relations entre bien-être et désirabilité sociale et de les reconsidérer sous un aspect quelque peu différent.

La désirabilité sociale a longtemps été envisagée comme un style de réponse ayant pour effet de biaiser les résultats obtenus par l'intermédiaire de questionnaires. Dans ce cas, et si les résultats de Carstensen & Cone sont vérifiés, la validité des mesures de bien-être serait sérieusement mise en cause.

Les premiers arguments contestant ces conclusions pessimistes font référence à l'identité de contenu entre les items des échelles de bien-être et ceux des items de l'échelle de désirabilité sociale d'Edwards (1957), utilisée par Carstensen & Cone. Kozma & Stones (1985) ainsi que McCrae & Costa (1985) font remarquer que les items de l'échelle d'Edwards ont un contenu psychopathologique évident et qu'ils font référence de manière quasi exclusive à la santé physique et mentale. Il n'est pas étonnant dès lors de trouver des corrélations entre bien-être et désirabilité sociale. Edwards lui-même réaffirme en 1990 que son échelle est composée d'items du MMPI qui ont été jugés fortement désirables (ou fortement indésirables) par un échantillon de sujets. Si on admet que les différents aspects du bien-être sont plutôt jugés désirables, on doit trouver, dans l'échelle d'Edwards des items qui y font référence. Nous n'en donnerons qu'un seul exemple, particulièrement démonstratif : l'item "Je me sens heureux la plupart du temps". Il pourrait tout aussi bien faire partie d'une échelle de bien-être.

On peut trouver, dans la littérature, deux approches complémentaires qui étudient les relations entre bien-être et désirabilité sociale. Dans la première, il s'agit de montrer que la désirabilité sociale n'influence pas autant les réponses que le laissent suggérer les résultats de Carstensen & Cone (1983). Dans la seconde, on montre que la désirabilité sociale – du moins sous une certaine forme – est une composante fondamentale du bien-être ; ces travaux conduisent à s'interroger sur la nature même de la désirabilité sociale.

Il paraît relativement bien établi, aujourd'hui, que les corrélations entre bien-être et désirabilité sociale diminuent de manière très importante lorsque l'on utilise une autre échelle de désirabilité sociale que celle d'Edwards. La plus connue et la plus utilisée est celle de Marlowe-Crowne (Crowne & Marlowe, 1960). Elle présente l'avantage, par rapport au débat qui nous occupe ici, de comporter des items dont les contenus sont relativement indépendants de la santé physique et mentale. Les travaux qui se sont basés sur cette échelle, ou sur des versions abrégées, rapportent des corrélations de l'ordre de .30 (e.g. Larsen et al., 1985 ; Kozma & Stones, 1987 ; Kop, 1993 a). Un argument plus indirect est fourni par Veenhoven & Ouwenel (1989) qui utilisent des données de bien-être agrégées au niveau national. Leur raisonnement est le suivant : si les réponses aux items de bien-être sont biaisées par la désirabilité sociale, on devrait observer les niveaux de bien-être les plus élevés dans les pays qui valorisent le plus les valeurs liées au bien-être. Les résultats qu'ils obtiennent infirment cette hypothèse.

Dans la seconde approche, on essaie de contrôler les effets de la désirabilité sociale afin d'examiner ce que devient le niveau de corrélation entre le bien-être et d'autres variables. En effet, si la désirabilité sociale est considérée comme un biais de réponse, cela signifie que la variance des réponses aux questionnaires de bien-être peut être décomposée en deux parties (au moins) : une part de variance vraie (due au bien-être) et une part de variance artefactuelle (due à la désirabilité sociale). Il suffit donc de

contrôler la part de variance due à la désirabilité sociale pour améliorer la validité des réponses aux questionnaires de bien-être. Or, toutes les études qui ont utilisé ce paradigme aboutissent à des résultats contraires aux attentes : non seulement le contrôle de la désirabilité sociale n'améliore pas la validité des mesures de bien-être mais, le plus souvent, il la fait même diminuer.

C'est le cas de l'étude de McCrae (1986) dans laquelle, après contrôle de la désirabilité sociale, les corrélations entre le bien-être et l'évaluation du névrosisme par les proches^[22] diminuent. C'est le cas aussi chez Kozma & Stones (1987) qui utilisent, comme critère externe de validation, l'évaluation du bien-être par les proches. Diener et al. (1991 a), utilisent, quant à eux, trois types de mesure : une mesure de bien-être construite pour que les résultats soient indépendants de la désirabilité sociale^[23] ; des mesures non-traditionnelles du bien-être qui sont moins sensibles à l'influence de la désirabilité sociale (e.g. évaluation par des personnes proches, etc.) ; des mesures classiques obtenues par questionnaire. Il ressort de cette dernière étude deux résultats particulièrement importants. Tout d'abord, les mesures pour lesquelles la désirabilité sociale est contrôlée sont plus faiblement corrélées aux mesures de bien-être obtenues par des méthodes non-traditionnelles. Ensuite, la mesure de désirabilité sociale utilisée (en l'occurrence l'échelle de Marlowe-Crowne) corrèle à un degré non négligeable ($r = .36$) avec un score composite de bien-être élaboré à partir de diverses mesures non-traditionnelles.

Ces résultats sont particulièrement impressionnants, parce qu'ils suggèrent que non seulement la désirabilité sociale n'est pas un biais de réponse, mais qu'en plus, elle constitue une composante essentielle du bien-être, qui contribue à augmenter la part de variance vraie dans les mesures. En fait, le débat sur l'influence de la désirabilité sociale sur les mesures de bien-être n'est qu'une résurgence spécifique des débats qui ont agité, pendant les années 60, la communauté scientifique à propos de la désirabilité sociale comme style de réponse. Mais "les habitudes conceptuelles meurent difficilement", ainsi que l'écrivent McCrae & Costa (1983, p. 887), qui considèrent que la désirabilité sociale doit être étudiée pour elle-même, comme une source de différences interindividuelles véritable et non comme un style de réponse (cf. aussi Furnham, 1986).

Dans les travaux actuels sur la désirabilité sociale, on peut discerner certaines conceptualisations qui permettent de mieux comprendre les relations entre bien-être et désirabilité sociale. Certains insistent, en effet, sur la nécessité de distinguer deux composantes de la désirabilité sociale (cf. Paulhus, 1986, 1991). Il y aurait une

[22] : Cette mesure est considérée ici comme un critère externe de validation totalement indépendant de la désirabilité sociale, puisque ce ne sont pas des auto-évaluations.

[23] : Cf. section 2.1.3. Les sujets doivent choisir entre deux adjectifs ayant le même niveau de désirabilité sociale, l'un décrivant un trait de personnalité, l'autre décrivant le bien-être.

composante largement consciente et accessible aux manipulations expérimentales appelée "gestion de son impression" ("*management impression*") et une composante largement inconsciente, imperméable aux manipulations classiques, appelée "auto-duperie" ("*self-deception*"). L'auto-duperie est la composante la plus intéressante, car elle se laisse difficilement appréhender par des techniques classiques. Paulhus la rapproche à la fois du concept de mauvaise foi au sens de Sartre et du concept de résistance de l'ego de Block (1965). L'auto-duperie serait (Paulhus, 1986) l'absence motivée de la conscience de cognitions conflictuelles. Elle aurait une valeur adaptative importante et serait inextricablement liée au contenu de mesures telles que l'anxiété, la réussite, la motivation, l'estime de soi ou encore le bien-être (Paulhus, 1991, p. 23).

On ne peut manquer de faire le lien entre le concept d'auto-duperie et le concept d'illusion (Taylor & Brown, 1988 ; Shelder et al., 1993) qui renvoie à la notion de mécanismes de défense et de déni psychologique. La proximité des notions apparaît très clairement dans l'étude de Shelder et al. (1993). Ceux-ci considèrent en effet que des scores faibles à des questionnaires de santé mentale (dépression, anxiété, etc.)^[24] n'indiquent pas forcément l'absence de symptômes pathologiques : certains individus maintiendraient une illusion (largement inconsciente) de santé mentale. Pour identifier ces individus, les auteurs ont recours à une procédure en double aveugle dans laquelle les sujets répondent à des questionnaires de santé mentale et sont évalués, lors d'entretiens, par des psychologues cliniciens. Trois groupes de sujets sont ainsi déterminés. Le premier comprend des individus jugés en bonne santé mentale à la suite des entretiens cliniques et qui rapportent peu de symptômes dans leurs réponses aux questionnaires ; le second est composé d'individus jugés en mauvaise santé mentale par les cliniciens et qui obtiennent des scores élevés de pathologie aux questionnaires. Le troisième groupe est le groupe discordant : il comprend les individus qui ont des scores non pathologiques aux questionnaires alors qu'ils sont jugés en mauvaise santé mentale lors des entretiens. Ce qui est intéressant dans cette étude, c'est que les individus de ce troisième groupe se différencient de manière assez nette des deux autres sur des indicateurs physiologiques : on constate notamment des niveaux élevés de réactivité coronarienne dans des tâches induisant du stress (e.g. augmentation de la pression sanguine). Ces résultats restent bien évidemment à confirmer^[25]. Ils corroborent toutefois le fait que le bien-être pourrait être lié positivement à certains mécanismes d'illusion ou d'auto-duperie qui ne seraient plus, dans ce cas, des biais de réponse, mais des techniques de maintien de son propre bien-être, même dans des conditions défavorables. Ces résultats viennent, par ailleurs, mettre

[24] : Malheureusement, ces auteurs ont une conception limitée de la santé mentale qui est définie comme l'absence de symptômes pathologiques. Ce faisant, on ne trouve pas, dans leur étude, de questionnaires de bien-être, ce que l'on peut regretter, mais ce qui n'enlève rien à l'intérêt de leurs résultats.

[25] : D'autant que certaines options méthodologiques prises par les auteurs de l'étude, comme l'élimination de plusieurs sujets, sont discutables.

en cause certaines conceptions humanistes du bien-être (e.g. Rogers, 1968), qui l'associent à la justesse de la conscience de soi.

2.3.3 Effets de contexte

La notion d'effets de contexte renvoie à des mécanismes assez complexes par lesquels le contexte, envisagé dans un sens très large (le contexte social créé par la situation dans laquelle se trouve le sujet au moment où il répond au questionnaire, le contexte créé par les autres questions du questionnaire ou encore celui qui est créé par les modalités de réponse), peut avoir une influence sur les réponses des sujets (cf. Schwarz & Sudman, 1992). La plupart des travaux menés dans cette perspective sont dus à Norbert Schwarz et ses collaborateurs. Nous avons vu que pour ces auteurs, l'humeur pouvait être une source d'information utilisée par les sujets pour répondre à des items de bien-être ; le contexte constitue une autre source potentielle d'information. Dans les deux cas, cela revient à considérer que les réponses des sujets ne sont pas des indicateurs valides de leur état interne, car elles reposent sur des influences transitoires. Néanmoins, il faut signaler encore une fois que l'intensité de ces effets est fortement amoindrie lorsque les variables dépendantes sont mesurées par plusieurs items, ainsi que le montrent Pavot & Diener (1993 b).

En ce qui concerne le contexte social, Strack et al. (1990) montrent que le niveau de bien-être rapporté par les sujets est plus élevé lorsque ceux-ci se trouvent en présence d'une personne handicapée. Celle-ci joue alors le rôle d'un standard de référence auquel les sujets se comparent pour évaluer leur propre bien-être. Mais le contexte social ne se réduit pas à un processus de comparaison sociale comme le suggère ce résultat. Les réponses sont aussi influencées par un souci de présentation de soi qui s'exprime, par exemple, dans les résultats d'une étude de Smith (1979) ; elle montre que les sujets sont plus heureux lorsqu'ils répondent à un questionnaire administré par un enquêteur que lorsqu'ils répondent de manière anonyme. Ce souci de présentation de soi n'est pas étranger, bien évidemment, à la composante de gestion de son impression mentionnée à propos de la désirabilité sociale.

Le contexte peut aussi être fourni par la position des items dans le questionnaire. Si on interroge des individus sur leur satisfaction relative à leur vie en général puis sur leur satisfaction relative à leur mariage, les réponses aux deux items sont peu liées (cf. Bradburn & Sudman, 1991). En revanche, si l'ordre des questions est inversé, la corrélation entre les deux devient plus importante. Dans ce cas, la réponse à la question sur le mariage deviendrait un indice utilisé par les sujets pour répondre à la question générale. Un résultat similaire est fourni par Strack et al. (1988) qui demandent à des

sujets d'évaluer la fréquence avec laquelle ils sortent avec leurs amis et leur niveau de bien-être. La corrélation entre ces deux questions est égale à $-.12$ lorsque la question sur le bien-être est posée en premier, à $.66$ dans le cas inverse. Ce résultat est interprété de la même manière que précédemment : lorsque la question spécifique est posée en premier, elle sert d'information directement accessible aux individus pour évaluer leur bien-être.

Il existe d'autres types d'effets de contexte mis en évidence dans des domaines différents de celui du bien-être et qui concernent notamment les modalités de l'échelle de réponse^[26]. Ceux-ci ne seront pas discutés ici. De ces travaux, il ressort que des effets de contexte peuvent effectivement biaiser les réponses à certains items. Mais il y a deux façons de prendre en compte ces résultats. Soit on considèrera, comme le font Schwarz & Strack (1991), que ces effets invalident les réponses aux items de bien-être, soit on remarquera, avec Pavot & Diener (1993 b) que ces effets sont de faible importance lorsque les mesures de bien-être comportent plusieurs items. De plus, lorsque les questionnaires comprennent de nombreux items, les influences potentielles du contexte sont tellement nombreuses qu'il devient impossible d'en connaître les effets. Dans ce cas, il n'est pas impossible que ces influences aient plus ou moins tendance à s'annuler. C'est pourquoi nous ne croyons pas que ces expériences remettent fondamentalement en cause les réponses des individus aux items de bien-être. S'il est plus que vraisemblable que des effets de contexte existent effectivement, il nous semble plus intéressant de les utiliser pour mieux comprendre comment les sujets répondent aux questions qui leur sont posées que d'en faire un argument pour rejeter toute possibilité d'évaluation du niveau de bien-être des individus.

2.3.4 La distribution des réponses

Depuis les premières études quantitatives sur le bien-être, on a remarqué que les distributions des réponses étaient biaisées négativement. Autrement dit, les gens ont tendance à s'estimer plutôt heureux que malheureux. De plus, on constate en général que les individus considèrent, en moyenne, qu'ils sont plus heureux que les gens qui les entourent. C'est ainsi, par exemple, que Beckham, dès 1929, note que seulement 33% des individus estiment que les noirs sont heureux, alors que 75% d'entre eux se considèrent eux-mêmes comme heureux. Ce type de résultat a été constamment répliqué depuis. Campbell et al. (1976) trouvent, par exemple, que les individus s'attribuent, en moyenne, un niveau de satisfaction plus élevé que celui qu'ils attribuent à l'américain moyen.

[26] : Cf., par exemple, Schwarz & Hippler (1987) qui montrent que les réponses à une question concernant le nombre d'heures passées devant la télévision dépendent du choix particulier des durées proposées aux sujets.

Certains ont utilisé ces résultats pour mettre en cause la validité des mesures de bien-être. Mais leurs arguments, qui s'appuient sur l'idée qu'il est impossible, dans nos sociétés, que tellement de personnes soient heureuses, relèvent de critères plus idéologiques que scientifiques. Il n'y a aucune raison sérieuse pour que la distribution des réponses à des items de bien-être doive se conformer à une distribution normale. Comme le note très justement Veenhoven (1984), la prévalence du bien-être sur le mal-être dans les réponses ne reflète pas forcément la présence d'un biais : il est tout aussi plausible de considérer que cela correspond à l'état habituel des individus, tout comme prévaut la santé (physique) sur la maladie.

On peut aussi remarquer que cet effet, que l'on appellera dorénavant "effet de positivité", n'est pas spécifique au bien-être. Il correspond, selon Peeters (1989), à l'effet d'indulgence ou l'effet Pollyanna que l'on constate dans beaucoup de domaines, c'est-à-dire à la tendance à donner la priorité, dans des processus de jugement, aux informations positives par rapport aux informations négatives. Les études portant sur l'asymétrie entre positif et négatif sont très nombreuses, qu'il s'agisse d'auto-attribution de traits de compétence ou de traits de personnalité. Malheureusement, elles ignorent largement les études sur le bien-être et l'inverse est tout aussi vrai.

On peut tout de même trouver deux explications, l'une ponctuelle, l'autre plus générale, pour rendre compte de cet effet de positivité.

L'explication de Headey & Wearing (1988) fait intervenir un mécanisme adaptatif qu'ils ont appelé le "sens de supériorité relative". Ce mécanisme est adaptatif dans la mesure où il permet à l'individu de maintenir son estime de soi. Il pourrait s'expliquer par le fait qu'un individu peut toujours trouver, dans un domaine quelconque, un aspect où il est effectivement supérieur aux autres et maximiser le poids accordé à cet aspect dans son évaluation. Les auteurs considèrent aussi que ce sens de supériorité relative peut trouver sa source dans la manière dont les individus se comparent aux autres : en restreignant leurs comparaisons à des groupes de référence relativement similaires, les individus maximisent leurs chances de s'auto-évaluer plus positivement.

Une autre explication fait référence à la notion d'illusion mentionnée antérieurement. Pour Taylor & Brown (1988), il existe une tendance générale chez les individus à se créer des illusions positives. Celles-ci entraînent une conception positive exagérée du soi, une sur-estimation de son propre contrôle et un optimisme exagéré à l'égard du futur. Dans une publication ultérieure, Taylor (1991) propose un modèle dans lequel les illusions positives sont un cas particulier d'un mécanisme plus général qui rend compte des réactions des individus en face des événements. Ce modèle décrit, avec force détails, la mobilisation très rapide de l'organisme d'un individu suite à un épisode négatif, grâce à des réponses physiologiques, émotionnelles, cognitives ou sociales. A long terme,

l'impact de cet épisode négatif tend à être minimisé, ce qui permet de maintenir les illusions positives. Ainsi, l'effet de positivité est bien plus qu'un biais de réponse : il correspond à un fonctionnement psychologique réel et adapté des individus. Ce qui est biaisé, c'est, au sens statistique, la distribution des réponses ; ce ne sont pas les réponses elles-mêmes.

L'effet de positivité mérite donc d'être étudié en tant que tel et non pas comme un biais de réponse. Il peut permettre, ainsi que le suggèrent Byrne & Eysenck (1993), d'appréhender les différences individuelles en matière de personnalité à partir de différences dans les structures et les processus cognitifs des individus. Eux-mêmes ont montré que le bien-être, envisagé en termes d'affectivité positive, était associé à une tendance à interpréter positivement du matériel verbal ambigu. Dans le même ordre d'idée, Klein & Kunda (1993) ont montré, dans des études expérimentales, que les estimations des individus à propos de la fréquence de certains de leurs comportements passés ont tendance à être biaisées de manière à leur permettre de maintenir leur croyance qu'ils sont supérieurs aux autres.

2.4 Conclusion

Ce chapitre nous a permis d'aborder les aspects les plus généraux de la mesure du bien-être dans la littérature. Ceux-ci ne constituent toutefois que des aspects partiels qui seront complétés dans les prochains chapitres à partir de considérations liées à la structure du bien-être et aux modèles nomologiques qui ont été proposés.

Nous avons vu qu'il existait de nombreux instruments pour mesurer le bien-être. La quasi totalité de ces instruments prend la forme de questionnaires d'auto-évaluations. Mais ces questionnaires n'ont pas tous connu le même succès dans la littérature. Seuls quelques-uns sont fréquemment utilisés.

Cette situation présente des avantages. Elle permet de disposer d'informations complètes et répliquées à propos d'une même échelle. Mais elle comporte aussi des inconvénients, car les questionnaires les plus célèbres n'ont pas forcément que des qualités. Ainsi, l'ABS comporte des items qui diffèrent quant à leur niveau de spécificité et qui font référence tantôt à l'activité et à l'engagement, tantôt au plaisir, ce qui rend sa signification quelque peu ambiguë. Cette ambiguïté est encore accentuée par la confusion qui préside au calcul d'un score général de bien-être à partir de deux scores dont on a montré qu'ils étaient indépendants. On pourrait ajouter aussi, toujours en ce qui concerne l'ABS, que le désintérêt porté par Bradburn aux contenus des items (seul compte, en effet, le nombre d'expériences vécues) est sans doute par trop radical. Parmi les échelles pour

personnes âgées, le LSI reste l'instrument le plus utilisé et son influence est telle que les nouvelles échelles qui sont apparues depuis sa création s'en inspirent souvent largement. Pourtant, le LSI n'a jamais été conçu pour mesurer le bien-être, ainsi que nous l'avons rappelé. Nous avons vu aussi, dans le chapitre précédent, que certains items de cette échelle relevaient d'une position implicite concernant le déclin du bien-être chez les personnes âgées, sans que celui-ci soit aucunement démontré. Ces deux exemples illustrent assez clairement que les questionnaires de bien-être sont encore très perfectibles.

Même si les questionnaires auto-administrés sont largement dominants dans la mesure du bien-être comme dans le cas de beaucoup de construits psychologiques, d'autres techniques ont aussi été utilisées, surtout de manière expérimentale, afin de valider les questionnaires. Si certaines de ces techniques sont assez classiques (évaluations par des pairs, évaluations cliniques), d'autres sont plus originales (niveaux moyens d'humeur, différence entre nombres d'événements positifs et négatifs rappelés, indices comportementaux, choix forcé entre des adjectifs de même niveau de désirabilité sociale). On notera surtout à propos de ces techniques que les auteurs partent souvent du postulat qu'elles mesurent la même chose que les questionnaires et/ou qu'elles fournissent des mesures parfaitement valides (à la différence des mesures issues de questionnaires). Aucun de ces deux postulats ne peut être accepté sans réserve, surtout lorsqu'ils impliquent des conceptions théoriques sur la nature du bien-être. C'est le cas notamment des techniques qui se basent sur le calcul du niveau moyen de l'humeur. Lorsque l'on utilise ces mesures pour valider les réponses à des échelles de bien-être, on considère que le bien-être n'est rien d'autre que l'humeur moyenne. Cela n'est pas impossible mais c'est un raccourci un peu rapide qui tend à sous-estimer l'importance d'autres aspects du bien-être. Quoi qu'il en soit, ces positions théoriques mériteraient d'être explicitées et discutées.

Même si les questionnaires élaborés pour mesurer le bien-être ne sont pas à l'abri de critiques, les vérifications des qualités psychométriques de ces échelles ont, en général, donné de bons résultats. De nombreuses études ont établi que les réponses étaient consistantes et qu'elles avaient une stabilité temporelle satisfaisante. Les résultats sont tout aussi encourageants en ce qui concerne la validité convergente, la validité divergente et la validité de construit estimée à partir du pourcentage de variance due au trait mesuré. On notera toutefois un désintérêt manifeste pour tous les aspects liés à la validité de contenu, qui ne sont jamais pris en compte. Par ailleurs, nous avons aussi pu remarquer que la vérification des qualités psychométriques des échelles était trop souvent envisagée uniquement de manière empirique sans tenir compte des considérations théoriques sous-jacentes. La mesure est un test de théories et lorsque l'on vérifie la stabilité temporelle des réponses, on teste l'hypothèse que le bien-être est un trait ayant une certaine stabilité. On

ne peut donc à la fois arguer d'une bonne stabilité temporelle des mesures et défendre l'idée que le bien-être est un état : il y a une contradiction interne dans cette démarche.

Mises à part les critiques ponctuelles adressées à telle ou telle échelle, les questionnaires de bien-être ont été la cible de critiques plus fondamentales que nous avons essayé de passer en revue. Si on excepte les positions idéologiques contestant la possibilité même de mesurer le bien-être, on peut distinguer quatre grands sujets de polémiques : les influences de l'humeur, de la désirabilité sociale, du contexte et les biais dans les distributions des réponses.

L'influence de l'humeur du moment sur les réponses aux items de bien-être a principalement été mise en évidence dans des études expérimentales. Plusieurs indices laissent penser que ces effets sont beaucoup moins nets dès lors qu'il s'agit de conditions naturelles et/ou lorsque les mesures du bien-être ne se limitent pas à un seul item. Si une telle influence existe, elle peut être le fait des relations entre humeur et mémoire ou des stratégies adoptées par les sujets pour minimiser le coût cognitif associé à une évaluation globale de leur vie. Un autre aspect à prendre en considération est la possibilité de relations réciproques entre humeur et bien-être. Celles-ci commencent seulement à être envisagées dans le contexte plus général des relations entre traits et états. Là encore, sur le plan théorique, une clarification semble nécessaire car si le bien-être est envisagé comme une humeur moyenne, on ne voit pas très bien comment l'humeur pourrait être une source de biais.

Vieux serpent de mer de la psychologie, la désirabilité sociale a donné lieu à des débats qui paraissent un peu datés. Il est certain que les corrélations sont élevées lorsque les échelles de désirabilité sociale ont un contenu similaire à celui des échelles de bien-être. Dans le cas contraire, les corrélations diminuent sans disparaître totalement. Mais ce qui est plus important, c'est que le contrôle de la désirabilité sociale n'augmente pas les différents indicateurs de validité, ce qui suggère qu'elle constitue une composante essentielle du bien-être. Que la désirabilité sociale soit conçue en termes d'auto-duperie ou en termes de gestion de son impression, ses relations avec le bien-être soulèvent des questions qui dépassent de loin le cadre restreint des biais de réponse. L'auto-duperie et les phénomènes d'illusion qui lui sont associés constituent peut-être l'un des mécanismes fondamentaux permettant de maintenir un niveau élevé de bien-être. Quant à la gestion de son impression, elle s'exprime notamment par un besoin d'approbation sociale qui se manifeste dans des comportements variés, autres que celui de répondre à un questionnaire, ainsi que l'ont noté Snyder & Ickes (1985). En contribuant à renvoyer à un individu une image positive de lui-même, l'approbation sociale peut lui permettre, là aussi, de maintenir son niveau de bien-être.

Les effets de contexte semblent pouvoir affecter les réponses aux items de bien-être. Le contexte peut être déterminé par la situation sociale (e.g. présence d'une personne devenant un standard de référence auquel les sujets se comparent lorsqu'ils répondent) ou tout simplement par les items qui ont été retenus dans le questionnaire (qui fournissent aux sujets, selon la cognition sociale, des indices leur permettant de construire leurs réponses). S'il est difficile de contester les résultats concernant les effets de contexte, on peut tout de même relever deux limites assez sérieuses de ces travaux. Tout d'abord, ces effets semblent dépendre assez largement du nombre d'items utilisés pour mesurer le bien-être, ce qui n'étonnera nullement les psychométriciens qui savent que la part de variance d'erreur associée à un item unique est importante. Ensuite, on peut s'interroger quant à l'intérêt que présentent de tels effets. Si ceux-ci sont utilisés pour discréditer les mesures de bien-être (et toutes les mesures issues de questionnaire en général), la démonstration est fallacieuse : ce n'est pas parce que l'on a montré, dans une étude, que les réponses à un item changeaient en fonction du contexte que l'on a montré que les réponses à des échelles composées de plusieurs items sont invalides. En revanche, si ces effets sont étudiés pour eux-mêmes, ils peuvent aider à comprendre les processus cognitifs qui sont à l'oeuvre dans les réponses. Dans les travaux de l'équipe de Schwarz et de ses collaborateurs, on trouve un peu de ces deux objectifs. Seul le second nous semble digne d'intérêt et mérite aussi de retenir l'attention des chercheurs qui utilisent des questionnaires.

Enfin, les biais constatés dans les distributions de réponses ne semblent pas constituer un argument allant à l'encontre de la validité des mesures de bien-être. Il est tout à fait probable que la prévalence du bien-être sur le mal-être soit réelle, tout comme l'est la prévalence de la santé physique sur la maladie. A nouveau, le débat sur la distribution des réponses gagnerait à être considéré dans un cadre plus général, celui des travaux sur les effets de positivité qui essaient d'expliquer ces effets à l'aide de mécanismes adaptatifs. La synthèse et les propositions de Taylor (1991) sont particulièrement stimulantes à cet égard.

En bref, le chercheur qui souhaite utiliser des mesures de bien-être n'a que l'embarras du choix. Il est assuré de disposer de résultats relativement fiables quelle que soit l'échelle qu'il retient étant donné ce que l'on sait de la manière dont les échelles ont été développées et des travaux psychométriques généraux. L'interrogation principale concerne la validité de contenu de ces échelles et, plus généralement, la signification précise des scores qui sont obtenus. C'est à cette question primordiale que nous avons choisi de nous attacher.

CHAPITRE III

LA STRUCTURE DU BIEN-ETRE SUBJECTIF

3.1 Introduction

Même si la plupart des auteurs s'accordent sur le fait qu'il y a un facteur commun à toutes les échelles de bien-être (e.g. Kozma & Stones, 1980 ; Chamberlain, 1988), conçu parfois comme un facteur de second ordre (e.g. Liang, 1984), peu d'entre eux acceptent l'idée d'un bien-être unidimensionnel (Kamman et al., 1984a ou McGreal & Joseph, 1993 constituent des exceptions notables). Les travaux sur la structure du bien-être occupent dès lors une place importante dans la littérature. Il s'agit de dégager, au-delà des communautés de variance, des composantes plus spécifiques qui viennent enrichir la signification du concept.

Les études structurales sur le bien-être peuvent se situer à des niveaux différents. On trouve des analyses, largement exploratoires, menées sur des questionnaires comprenant un grand nombre d'items construits pour la circonstance. On trouve aussi des analyses plus spécifiques portant sur des instruments connus, ayant un nombre réduit d'items. Ces études utilisent souvent les modèles d'équations structurales, dans une démarche que les auteurs qualifient de confirmatoire, même si les structures testées dérivent le plus souvent d'analyses factorielles exploratoires préalables. La spécificité des facteurs dégagés par cette dernière stratégie est un obstacle sérieux à leur réplication : trop dépendants du caractère idiosyncrasique des items composant les échelles analysées, les dimensions ainsi dégagées ont eu, pour la plupart, une existence éphémère malgré l'intérêt théorique potentiel de certaines interprétations. A côté de ces deux grandes stratégies dont nous présenterons quelques études prototypiques, il existe des démarches plus originales, rarement utilisées, mais qui présentent un intérêt indéniable. Nous en verrons deux exemples : une étude structurale du bien-être qui passe par l'intermédiaire des représentations et une application de l'approche en facettes.

Par-delà les spécificités des dimensions mises en évidence dans ces approches, on peut essayer de dégager les dimensions générales les plus solides. Dans la seule revue de question sur la structure du bien-être, celle de Chamberlain (1988), l'auteur recense quatre couples de dimensions : positif et négatif ; affectif et cognitif ; fréquence et intensité ; interne et externe. Parmi ces quatre, nous n'approfondirons que les deux premiers. En effet, nous verrons à l'occasion de la discussion sur le couple positif-négatif que d'une part, la signification des dimensions de fréquence et d'intensité est ambiguë et que, d'autre

part, il semble qu'elle ne fasse que reprendre une autre proposition, beaucoup plus générale celle-là, faite à propos de la structure de l'émotion. En ce qui concerne le couple interne-externe, dû principalement à Lawton (1983 b; Lawton et al., 1984)^[1], Chamberlain (1988) considère qu'il se superpose partiellement à la distinction positif-négatif et que certains arguments permettent de le rapprocher de la distinction affectif-cognitif. De plus, les études empiriques qui supportent ce couple sont largement insuffisantes. Pour toutes ces raisons, ces deux dimensions ne seront pas envisagées, dans la suite du texte, comme des dimensions fondamentales du bien-être. Seuls les couples positif-négatif et cognitif-affectif feront l'objet d'une présentation détaillée.

Nous verrons tout d'abord la distinction positif-négatif car c'est celle qui a marqué en premier lieu l'histoire du bien-être. Nous en avons vu les prémisses à propos de l'échelle de balance affective (ABS) de Bradburn (cf. section 2.1.2), qui montrait une indépendance entre une composante positive et une composante affective du bien-être. Dans un premier temps, seront exposés les arguments empiriques venant étayer cette indépendance. Celle-ci a suscité de nombreux commentaires visant à en expliquer l'origine : certaines explications sont techniques et invoquent divers artefacts potentiels, d'autres reposent sur des considérations plus théoriques. Cette présentation fera l'objet d'un second point. Nous verrons ensuite que la discussion sur les dimensions positive et négative du bien-être renvoie aux travaux sur la structure des émotions dont nous ferons le bilan. Enfin, avant de conclure sur cette question, nous proposerons de passer en revue les arguments théoriques nous paraissant les plus aptes à supporter cette distinction.

Alors que le couple positif-négatif est relativement bien assuré au niveau empirique, la distinction entre composantes affectives et composantes cognitives du bien-être repose surtout sur des considérations théoriques générales. Nous verrons que malgré l'intérêt heuristique de cette distinction, les arguments empiriques font quelque peu défaut, ce qui s'explique probablement par la difficulté à trouver des définitions opérationnelles satisfaisantes pour ces deux dimensions. Il faudra donc s'interroger sur la manière dont cette distinction est traduite dans la littérature au niveau des items.

La conclusion générale sera l'occasion de faire le bilan de ces travaux et de se demander pourquoi, alors que les deux couples que nous venons d'évoquer font l'objet d'un consensus relatif, il n'existe pratiquement pas de travaux ayant tenté d'étudier conjointement ces quatre dimensions. Cette question paraît pourtant primordiale et il n'est pas exclu que ce désintérêt manifeste cache une difficulté théorique fondamentale, au-delà des difficultés de mise en oeuvre pratique.

[1] : On pourra consulter le premier chapitre (section 4) où nous avons brièvement parlé des propositions de Lawton.

3.2 Les études structurales sur le bien-être

Il n'est pas question ici de présenter l'ensemble des analyses structurales, analyses factorielles pour la plupart, ayant été menées à partir d'items censés mesurer le bien-être subjectif. Cela serait très long, vite fastidieux et sans grande valeur informative. Nous préférons aborder cet aspect à partir d'études prototypiques s'inscrivant dans des perspectives différentes et utilisant des stratégies spécifiques. Les limites inhérentes à chacune des stratégies font qu'aucune d'entre elles ne peut être considérée comme entièrement satisfaisante, mais elles sont complémentaires. Dans un premier temps nous verrons une étude dont la particularité est de proposer une analyse structurale portant sur un grand nombre d'items construits de manière à être des indicateurs globaux du bien-être. Nous verrons ensuite des travaux plus spécifiques portant sur l'analyse structurale d'items de questionnaires classiques de bien-être ; ces études ont surtout été effectuées avec des échelles de bien-être pour personnes âgées. Dans une troisième partie, nous présenterons une application de l'approche en facettes à la mesure du bien-être. Enfin, la structure du bien-être a aussi été abordée à partir des représentations d'individus tout-venant. Le choix des études prototypiques a été relativement simple. Pour chacune des deux dernières stratégies, elles se sont imposées d'elles-mêmes : il n'existe en effet, à notre connaissance, qu'un seul exemple d'application au domaine du bien-être subjectif. Il s'agit, pour les représentations, de l'étude de Thomas & Stock (1988) et, pour la théorie en facettes, de celle de Levy & Guttman (1975)^[2]. En ce qui concerne la première stratégie, l'étude de Andrews & Withey (1976) semblait tout indiquée. C'est une des toutes premières conceptions du bien-être distinguant d'une part les composantes affective et cognitive et, d'autre part, les composantes positive et négative^[3]. La seconde stratégie sera illustrée par plusieurs publications complémentaires de Jersey Liang et ses collaborateurs, qui proposent un ensemble d'analyses factorielles contraintes sur deux échelles de bien-être pour personnes âgées, le LSI-A et la PGC-MS^[4].

[2] : En fait, nous avons vu, dans le chapitre consacré à la théorie en facettes, une autre application due à Levy (1981b). Cependant, celle-ci n'est qu'un exemple incomplet et non prototypique puisqu'il s'agissait de construire une phrase en facettes qui pouvait s'appliquer à des items pré-construits.

[3] : La même année, Campbell et al. (1976) ont publié un ouvrage général sur le bien-être de la même ampleur que celui de Andrews & Withey. Les conceptions du bien-être en termes structuraux, défendues dans ces deux ouvrages, ne sont pas fondamentalement différentes, mais la démarche de Andrews & Withey nous semble plus illustrative que celle de Campbell et ses collaborateurs.

[4] : Nous nous sommes déjà expliqué à propos de l'utilisation de l'expression "analyse factorielle contrainte" en lieu et place de la traditionnelle "analyse factorielle confirmatoire".

3.2.1 Une analyse structurale exploratoire sur des indicateurs généraux de bien-être subjectif : l'exemple de Andrews & Withey (1976)

C'est au début des années 70 que Frank Andrews & Stephen Withey ont commencé une vaste entreprise sur la mesure du bien-être. Celle-ci s'inscrit dans le cadre du mouvement des indicateurs sociaux. L'un de ses principaux objectifs est d'établir des indicateurs valides de bien-être. Les résultats présentés par ces auteurs ont été obtenus grâce aux réponses de plus de 5000 individus composant quatre échantillons représentatifs de la population des Etats-Unis à des questionnaires comportant plusieurs centaines de questions. Il s'agit sans doute là d'une des études les plus extensives sur le bien-être, avec celle de Campbell, Converse & Rodgers publiée la même année (1976). Ces deux ouvrages restent encore aujourd'hui des références de base sur la question.

Pour Andrews & Withey (1976), les indicateurs de bien-être peuvent être envisagés à différents niveaux de spécificité. Au niveau le plus général, on trouve ceux qui font référence à la vie dans sa globalité. A un niveau plus spécifique, ce sont encore des évaluations générales, mais celles-ci portent sur des aspects particuliers de la vie que les auteurs appellent des "*concerns*"^[5]. Deux types de "*concerns*" peuvent être distingués : les domaines de vie (les lieux, les activités, les personnes..) et les critères (les valeurs, les normes, les aspirations...)^[6]. Au niveau le plus spécifique, les auteurs considèrent les évaluations d'un domaine de vie en fonction d'un certain critère. L'item "comment évaluez-vous votre indépendance au travail ?" est un exemple d'indicateur de ce niveau, l'indépendance étant un critère, le travail étant un domaine de vie. Nous nous intéresserons ici uniquement aux indicateurs les plus généraux^[7].

La caractéristique principale des indicateurs généraux est qu'ils doivent porter sur une évaluation globale de la vie d'un individu ("*life-as-a-whole*"). C'est à partir de cette définition que les auteurs ont mis au point une liste de 68 indicateurs créés pour la circonstance ou issus d'études antérieures. Certains de ces indicateurs sont obtenus par combinaison de deux items ou plus. Ils peuvent être présentés sous la forme d'une typologie qui croise plusieurs dimensions. Des exemples correspondant à chacun des types utilisés par les auteurs sont regroupés dans le tableau 2.5.

[5] : Il est très difficile de trouver une traduction française satisfaisante de ce terme. Défini par les auteurs (p. 11) comme "les aspects de la vie pour lesquels les gens éprouvent des sentiments", il désigne tout ce qui intéresse les individus dans leur vie, tout ce qui les "concerne".

[6] : Un exemple d'indicateurs pour chacun des deux types est le suivant : "comment évaluez-vous votre travail ?" et "comment évaluez-vous ce que vous avez accompli dans votre vie ?".

[7] : Cette classification des items de bien-être en fonction de leur niveau de spécificité est intéressante, mais elle ne s'est jamais imposée. Il est difficile de savoir exactement pour quelle raison, mais en lisant l'ouvrage de Andrews & Withey, on s'aperçoit que les auteurs restent très évasifs sur la définition des domaines de vie et des critères et qu'ils n'effectuent pas eux-mêmes, de manière explicite, le lien entre les items qu'ils ont construits et ce modèle conceptuel.

Généralité de l'évaluation	Etendue de l'évaluation	Perspective de l'évaluation			
		absolue	relative	longue	courte
Générale	totale	<i>How do you feel about your life as a whole ?</i>	<i>My life is much worse than most other people's ... somewhat worse ... about the same ... somewhat better ... much better</i>	différence entre : <i>"where do you expect your life to be five years from now ?" et "where were your life five years ago ?"</i>	<i>How much feelings about your life as a whole change from month to month ?</i>
	partielle	<i>How do you feel about the good and pleasant parts of your life ?</i>	/	/	/
Spécifique	/	<i>How do you feel about how happy you are ?</i> les items de l'ABS	/	/	/

Tableau 2.5 : Exemples d'indicateurs généraux du bien-être utilisés par Andrews & Withey (1976)

Ainsi qu'on peut le constater à la lecture du tableau, seuls six types d'items ont été construits sur les 16 possibles. Cela n'est pas dû à une impossibilité pratique ; les auteurs estiment que ces six types d'indicateurs sont suffisamment variés et qu'ils sont les plus utiles pour la mesure du bien-être.

Les auteurs rapportent de nombreuses analyses, obtenues avec des techniques différentes (échelonnement multidimensionnel, analyse factorielle et analyse en clusters) sur des échantillons différents de sujets répondant à des sous-ensembles d'items particuliers. Nous ne ferons état ici que des résultats les plus intéressants.

Les indicateurs absolus, généraux, avec une étendue totale constituent, dans toutes les analyses, un groupe d'indicateurs fortement liés entre eux, et corrélés aussi avec l'ensemble des autres items. Ils occupent une position centrale dans les espaces issus des échelonnements multidimensionnels et obtiennent, en général, des saturations élevées sur le premier facteur avant rotation d'une analyse en composantes principales. Tous ces résultats indiquent de manière assez convergente que ce type d'indicateurs est le plus intéressant quand on a besoin d'une évaluation générale du bien-être.

L'autre résultat qui nous intéresse ici est celui que produit un sous-ensemble de 12 indicateurs qui correspondent à une perspective absolue dans la typologie des auteurs. Ces indicateurs comprennent notamment les items de l'échelle de balance affective de Bradburn (cinq items de même valence formant un seul indicateur). Le premier résultat relevé par les auteurs concerne l'indépendance entre les deux indicateurs correspondant à l'ABS de Bradburn. Autrement dit, les évaluations des individus concernant les aspects positifs de leur vie ne sont pas corrélées aux évaluations concernant les aspects négatifs. Une analyse factorielle suivie d'une rotation varimax confirme bien évidemment ce résultat en faisant apparaître un facteur saturant les indicateurs positifs et un facteur saturant les indicateurs négatifs. Ils sont interprétés respectivement comme facteur d'affectivité positive et facteur d'affectivité négative. Mais un troisième facteur émerge de l'analyse factorielle : les auteurs l'interprètent comme un facteur d'évaluation cognitive, tout en reconnaissant que sa signification conceptuelle reste à préciser et ce, notamment, parce que les indicateurs plus affectifs n'en sont pas totalement indépendants. C'est le cas, par exemple, des items censés évaluer le bonheur des individus, que les auteurs avaient tendance à considérer comme des indicateurs affectifs du bien-être : en fait, les résultats montrent qu'ils ont des saturations élevées sur le facteur cognitif, même s'ils sont corrélés de manière non négligeable avec les facteurs affectifs.

On retiendra donc de cette étude la mise en évidence empirique de trois dimensions que les auteurs assimilent à des sources de variation relativement indépendantes des mesures de bien-être. La paternité de cette conception du bien-être est souvent attribuée à Andrews & Withey, même si des formes similaires se retrouvent chez d'autres auteurs (cf., par exemple, Beiser, 1974 ; Campbell et al., 1976). En y faisant référence, on oublie toutefois que ses concepteurs eux-mêmes ont émis des réserves quant à la signification du facteur cognitif. S'ils l'ont appelé ainsi, écrivent-ils (p. 88), c'est parce qu'il fallait bien lui donner un nom mais celui-ci "*may imply a greater conceptual clarity than we believe actually exists*".

3.2.2 Analyses factorielles contraintes sur des échelles de bien-être pour personnes âgées : les travaux de Liang

Alors que le mouvement des indicateurs sociaux a beaucoup utilisé des questionnaires construits pour la circonstance en fonction d'objectifs spécifiques, le courant gérontologique sur le bien-être a largement fait usage d'échelles classiques dont nous avons vu quelques exemples dans le troisième chapitre. Parmi ces échelles, le LSI-A de Neugarten et al. (1961) et la PGC-MS de Lawton (1972, 1975) sont les plus employés et ont servi de base pour l'étude de la structure du bien-être chez les personnes âgées. Comme le souligne à juste titre Ryff (1989c), il y a quelque chose de fondamentalement

paradoxal à utiliser, pour mettre en évidence la structure d'un concept, des instruments qui n'ont pas toujours été développés pour mesurer ce concept. C'est effectivement le cas pour le LSI-A qui a été proposé, nous l'avons vu, comme une mesure de l'adaptation chez les personnes âgées et non pas comme une mesure de bien-être.

Malgré ces limites, plusieurs auteurs travaillant sur le bien-être des personnes âgées se sont posés comme objectifs de clarifier les problèmes liés à la conceptualisation et à la mesure du bien-être sur la base de la dimensionnalité des échelles existantes. Les travaux de Jersey Liang et ses collaborateurs, qui se sont engagés dans un programme d'analyses factorielles de la PGC-MS et du LSI-A, constituent l'une des tentatives les plus notables en ce domaine.

Ces travaux ont la particularité d'utiliser l'analyse factorielle contrainte du modèle LISREL, technique que les auteurs estiment supérieure aux analyses factorielles classiques, notamment quant au choix de la dimensionnalité et à la possibilité de prise en compte des erreurs de mesure. De plus, toujours au niveau méthodologique, les auteurs comparent les solutions obtenues à partir de matrices de corrélations linéaires, de matrices de corrélations tetrachoriques (dans le cas d'items dichotomiques) ou polychoriques (dans le cas d'items trichotomiques) ou encore de matrices de covariance. Les résultats les plus marquants de cette comparaison sont que les estimations des saturations sont plus élevées dans le cas de coefficients tetrachoriques ou polychoriques (et, par conséquent, les estimations des erreurs de mesure sont plus faibles). Ce résultat n'est pas étonnant car on sait que les coefficients tetra- ou polychoriques sont plus élevés que les coefficients de corrélations classiques lorsqu'ils sont appliqués aux mêmes données. S'il est vrai que ces coefficients doivent être préférés lorsque les items sont dichotomiques ou trichotomiques, il faut tout de même noter que les solutions factorielles ne changent pas de manière sensible si l'on excepte la valeur absolue des paramètres. Comme ce sont plus les solutions factorielles qui nous intéressent ici que les valeurs des paramètres en eux-mêmes, nous ne ferons pas état de ces différences plus avant.

Analyses structurales du PGC-MS

Outre l'utilisation de l'analyse factorielle contrainte, la principale originalité des modèles structuraux présentés par Liang est de faire intervenir un facteur de second ordre qui rend compte des corrélations entre les facteurs de premier ordre. Selon les publications et les échelles, ce facteur peut prendre un nom différent, mais dans tous les cas il peut être assimilé à un facteur de bien-être général. Une illustration typique des modèles structuraux testés par Liang est présentée dans la figure 2.10.

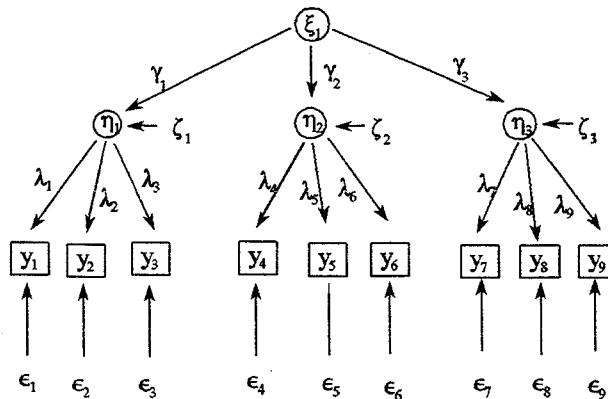


Figure 2.10 : Illustration fictive d'un modèle structural testé par Liang avec 9 variables observées, 3 facteurs de premier ordre et un facteur de second ordre

Dans le cas de la PGC-MS, Liang & Bollen (1983) proposent un modèle à trois facteurs de premier ordre, identique à la solution obtenue par Lawton (1975) et par Morris & Sherwood (1975) avec des analyses factorielles exploratoires. Deux items ne sont pas pris en compte car ils font référence aux relations sociales, que les auteurs considèrent comme relevant d'un domaine conceptuellement différent de celui du bien-être. Plusieurs modèles sont testés et, pour améliorer l'adéquation, les auteurs sont amenés à admettre l'existence de corrélations entre les erreurs de mesure. Celles-ci ne sont pas interprétées, en l'absence de connaissances théoriques suffisantes. Ce modèle sera testé dans de nouvelles études, notamment pour estimer la stabilité des structures factorielles selon le sexe (Liang & Bollen, 1985), selon les cultures (comparaison Etats-Unis - Japon, Liang & al. 1987 a) et selon l'époque (comparaison à six ans d'intervalle, Liang & Lawrence, 1989). Les résultats de ces études sont résumés dans le tableau 2.6. On s'aperçoit que les modèles sont globalement identiques, mais que certains items sont jugés inadéquats dans des solutions ultérieures.

ITEMS	FACTEURS DE 1 ^{er} ORDRE	<i>Agitation</i>	<i>Dissatisfaction</i>	<i>Attitudes toward own aging</i>
	ETUDES	A B C D	A B C D	A B C D
Little things bother me more this year		X X X —		
I sometimes worry so much that I can't sleep		X X X /		
I am afraid of a lot of things		X X — X		
I get mad more than I used to		X X X X		
I take things hard		X X X X		
I get upset easily		X X — /		
I sometimes feel that life isn't worth living			X X X —	
Life is hard for me much of the time			X X X /	
How satisfied are you with your life today			X X — —	
I have a lot to be sad about			X X X —	
Things keep getting worse as I get older				X X X X
I have as much pep as I had last year				X X X X
As you get older you are less useful				X X X X
As I get older, things are better/worse than I thought they would be				X X — —
I am as happy now as when I was younger				X X X X

Légende: A (Liang & Bollen, 1983) ; B (Liang & Bollen, 1985) ; C (Liang et al., 1987) ; D (Liang & Lawrence, 1989)

X : item saturé ; — : item inadéquat à la solution retenue ; / : item non retenu dans le questionnaire de l'étude

Tableau 2.6 : La structure de la PGC-MS d'après différentes études de Liang et ses collaborateurs

Analyses structurales du LSI-A

Les analyses portant sur le LSI-A suivent la même logique que celle que nous venons de relater. Les objectifs de clarification conceptuelle sont toujours affichés et le modèle testé s'appuie partiellement sur les résultats d'analyses factorielles exploratoires préalables effectuées par Adams (1969) et par Hoyt & Creech (1983). De plus, seuls 11 items sont retenus sur les 20 que comporte le LSI-A : les explications de Liang (1984) sur cette sélection font intervenir la validité apparente des items, leur fidélité et leur pattern d'erreurs de mesure corrélées (ces deux derniers indicateurs sont obtenus sur la base de pré-analyses). Les trois facteurs de premier ordre sont ceux qui ont été proposés par Hoyt & Creech (1983) ; le facteur de second ordre est appelé "bien-être subjectif". Dans des publications ultérieures (Liang et al., 1987 b, 1988) l'identité des structures factorielles

est étudiée selon l'origine ethnique et selon l'âge. Quelques petites différences sont à noter mais, globalement, la structure reste la même.

ITEMS	FACTEURS DE 1 ^{er} ORDRE	<i>Mood tone</i>			<i>Zest</i>			<i>Congruence</i>		
	ETUDES	A	B	C	A	B	C	A	B	C
I am just as happy as I was younger		X	X	X						
My life could be happier than it is now		X	X	/						
These are the best years of my life		X	X	X						
I expect some interesting and pleasant things to happen to me in the future					X	X	/			
The things I do are as interesting to me as they ever were					X	X	X			
I feel old and somewhat tired					X	X	/			
Most of the things I do are boring and monotonous					X	X	X			
As I look back on my life, I am fearly well-satisfied								X	X	X
I would not change my past life even if I could								X	X	/
I've gotten pretty much what I expected out of life								X	X	X
I have gotten more of the breaks in life than most of the people I know								X	X	X

Légende : A (Liang, 1984) ; B (Liang et al., 1987 b) ; C (Liang et al., 1988)

X : item saturé ; / : item non retenu dans le questionnaire de l'étude

Tableau 2.7 : La structure du LSI-A d'après différentes études de Liang et ses collaborateurs

Il n'est pas nécessaire d'alourdir davantage la présentation. Ces deux exemples suffisent pour dresser un bilan de ce type de travaux. On remarquera tout d'abord que l'utilisation de l'analyse factorielle contrainte n'apporte guère d'informations supplémentaires par rapport aux résultats des analyses factorielles exploratoires sur lesquelles Liang s'appuie pour tester ses modèles. Tout au plus peut-on mettre en évidence des erreurs de mesure corrélées (mais celles-ci ne sont pas interprétées) ou des différences dans les structures factorielles (que nous n'avons qu'évoquées, mais là n'est pas notre propos). Le résultat le plus important est peut-être celui qui concerne la présence d'un facteur général qui rend compte des corrélations entre les facteurs de premier ordre ; mais en fait de clarification conceptuelle, celui-ci s'appelle tantôt "bien-être subjectif" (pour le LSI-A), tantôt "satisfaction globale avec la vie" (pour la PGC-MS), sans que cette distinction s'appuie sur des justifications théoriques très précises^[8].

[8] : En fait, le facteur "satisfaction globale avec la vie" de la première publication (Liang, 1984) deviendra "bien-être subjectif" dans la seconde (Liang & Bollen, 1985).

Au niveau conceptuel toujours, on notera aussi l'extrême spécificité des facteurs de premier ordre dégagés. Dans les deux échelles que nous venons de prendre comme exemple, on trouve déjà six facteurs différents dont on ne connaît ni les similitudes ni les différences. A cela il faut ajouter certaines réinterprétations de facteurs dans des études qui ne sont pas répertoriées dans les deux tableaux. Ainsi, Liang (1985), à propos du LSI-A, propose de nommer "bonheur" le facteur préalablement intitulé "*mood tone*" et de supprimer définitivement le facteur "*zest*" (et les items qui saturent ce facteur) à cause d'une ambiguïté trop importante^[9].

Plus généralement, le parcours de la littérature gérontologique sur le bien-être, donne très vite l'impression d'être submergé par le nombre des dimensions qui ont été mises en évidence, ces facteurs n'étant ni totalement différents ni totalement semblables. De plus, certains facteurs, saturant les mêmes items, se voient quelquefois interprétés de manière différente. Outre les exemples que nous venons de voir, nous pouvons encore citer l'étude de Hoyt & Creech (1983) qui font correspondre, respectivement, aux facteurs "*congruence*", "*mood tone*" et "*zest*" du LSI-A, les notions de satisfaction avec le passé, avec le présent et avec le futur.

Il est difficile de dresser un bilan global de ces travaux. On peut tout de même remarquer certaines similitudes avec une conception du bien-être qui distingue entre des composantes cognitive et affective. Des facteurs comme "congruence entre les buts atteints et les buts désirés" ou "attitude vis-à-vis de sa vieillesse" semblent correspondre plus ou moins à une composante cognitive alors que des facteurs comme "mood tone", "agitation" ou "bonheur" font plutôt référence à un facteur affectif. La formulation la plus explicite de cette convergence est sans doute celle de Stock et al. (1986) qui proposent une analyse factorielle contrainte des items du LSI-A et de l'ABS : le modèle retenu comporte un facteur cognitif, un facteur d'affectivité positive et un facteur d'affectivité négative. Même si ce modèle n'est pas "pur", puisque les items du LSI-A saturent à la fois le facteur d'affectivité positive et le facteur cognitif, c'est cette conception qui se rapproche le plus de celle d'Andrews & Withey, que nous avons vue plus haut.

Somme toute, les analyses par les modèles d'équations structurales sont un peu décevantes. Il est vrai que l'utilisation du LISREL permet d'obtenir des estimations plus précises des paramètres grâce à l'utilisation de coefficients de corrélation plus adaptés aux échelles de réponses dichotomiques ou trichotomiques. Surtout, le LISREL permet d'obtenir des indicateurs d'adéquation du modèle testé qui ne sont pas disponibles avec des analyses factorielles non contraintes. Sur le plan méthodologique, la supériorité de ce type d'analyses ne fait guère de doute.

[9] : Pourtant, dans deux études ultérieures, les modèles structuraux proposés par Liang (Liang et al., 1987 b, 1988) reprendront à nouveau ce facteur.

Mais les limites de ces études apparaissent surtout sur le plan théorique. Dans l'idéal, ces analyses devraient permettre de confirmer une théorie définitoire. Ce n'est pas le cas ici. Elles confirment simplement un modèle spécifique élaboré pour un questionnaire donné. Malheureusement, ces modèles ne sont pas les mêmes d'un questionnaire à l'autre et dépendent largement des spécificités de contenu des items. De plus, les modèles testés ne sont souvent rien d'autre que les solutions obtenues à la suite d'analyses factorielles non contraintes. On peut remarquer aussi que l'amélioration de l'adéquation est rendue possible en supprimant la contrainte de corrélations nulles entre les erreurs de mesure. Comme ces corrélations indiquent en général qu'une source de variation systématique n'a pas été prise en compte, on peut regretter que les modèles ne soient pas modifiés dans ce sens.

Malgré ces limites, les résultats de ces analyses ne sont pas dépourvus de tout intérêt. Elles permettent tout d'abord d'avoir une confiance accrue dans les structures mises en évidence grâce aux analyses factorielles non contraintes. De plus, au-delà de la spécificité inhérente à chacun des instruments, la distinction des dimensions affectives et cognitives est retrouvée de manière plus ou moins nette ; c'est le cas aussi, dans une moindre mesure, de celle entre les dimensions positives et négatives. Surtout, la principale originalité de ces études est d'avoir su proposer un modèle hiérarchique du bien-être pour rendre compte des corrélations entre les facteurs.

En bref, il nous semble que l'intérêt des modèles d'équations structurales se manifeste surtout lorsqu'il s'agit effectivement de tester des théories définitoires. Dans le cas du bien-être, ces théories sont encore trop peu élaborées pour qu'il soit possible de construire un modèle correctement formalisé. En l'absence d'un tel cadre théorique, l'utilisation de ces méthodes relève donc plus de la sophistication méthodologique que d'une avancée conceptuelle substantielle.

3.2.3 Une analyse structurale des représentations du bien-être : l'étude de Thomas & Stock (1988)

L'objectif de Thomas & Stock (1988, p. 142) est "d'étudier les dimensions du bonheur sans imposer de postulats préalables à propos du construit". L'originalité de ce travail est d'utiliser à cette fin les représentations que se font les individus du bien-être. L'utilisation des représentations comme méthode d'étude d'un concept a été largement critiquée par bon nombre de psychologues du courant social, notamment à travers l'hypothèse de distorsion systématique (e.g. Shweder & D'Andrade, 1980). Pourtant, l'histoire récente des recherches dans le domaine de la structure de la personnalité (cf., par exemple, McCrae & John, 1992 ; Borkenau, 1992 ; Goldberg, 1993) a montré tout

l'intérêt que pouvait présenter l'étude des représentations^[10] ; c'est notamment grâce à la convergence entre les résultats provenant de ce type d'approche et ceux d'approches plus traditionnelles que s'est instauré le large consensus actuel sur un modèle de la personnalité à cinq facteurs, connu sous le nom de "*big-five*". On peut considérer que l'étude de Thomas & Stock procède de la même logique, même si les auteurs n'y font pas allusion.

Cette étude comporte deux phases : la première consiste à demander à des sujets d'évoquer librement des termes à partir des mots inducteurs "bonheur" et "malheur" ; la seconde consiste à faire évaluer la similitude entre les concepts dégagés lors de la phase précédente à l'aide de la technique des comparaisons par paires. Comme les auteurs font l'hypothèse de l'émergence d'une structure composée d'une dimension affective et d'une dimension cognitive, ils demandent aussi aux sujets d'évaluer la similarité entre chacun des concepts dégagés lors de la première phase et les termes "*feeling*" et "*thought*"^[11].

Les auteurs ne retiennent, pour les comparaisons par paires, que les 12 concepts les plus fréquemment cités, plus le terme inducteur positif (bonheur). Ces comparaisons permettent d'obtenir une matrice de similarités entre les concepts qui est soumise à un échelonnement multidimensionnel (EMD). Les résultats de la solution bidimensionnelle retenue par les auteurs sont présentés dans la figure 2.11.

Les résultats de l'EMD font apparaître clairement une opposition, sur la première dimension entre les termes positifs et les termes négatifs. Bien que ce résultat soit cohérent avec la distinction opérée entre affectivité positive et affectivité négative, il n'est pas étonnant que cette opposition apparaisse fortement dans une étude où les sujets doivent juger de la similarité entre des concepts^[12]. La deuxième dimension est interprétée, par les auteurs, comme une dimension d'indépendance personnelle. Celle-ci n'est pas vraiment explicitée, mais il faut sans doute y voir une dimension qui exprime le degré d'indépendance vis-à-vis des autres. On considèrera, sans autre commentaire, que cette interprétation est quelque peu "forcée" au vu des résultats.

[10] : L'intérêt d'étudier les représentations et le débat sur l'hypothèse de distorsion systématique ne sont qu'évoqués ici. Ces questions seront traitées plus en détail au début de la troisième partie.

[11] : L'étude porte sur deux échantillons indépendants : un échantillon tout-venant et un échantillon de religieuses. Rien ne permet de savoir, dans l'article des auteurs, pourquoi cet échantillon si particulier a été interrogé. Comme on peut s'en douter, celui-ci se caractérise par des réponses idiosyncrasiques (Dieu, la prière...). Il ne sera donc fait référence, dans la suite du texte, qu'au premier de ces échantillons.

[12] : Les auteurs considèrent que leurs résultats sont en faveur d'une conception où l'affectivité positive et l'affectivité négative sont des pôles opposés d'un même continuum plutôt qu'en faveur de l'indépendance entre ces deux composantes. Cette conclusion est totalement hors de propos. La technique de collecte des observations et la méthode d'analyse de données ne permettent nullement d'apporter une quelconque réponse à cette question.

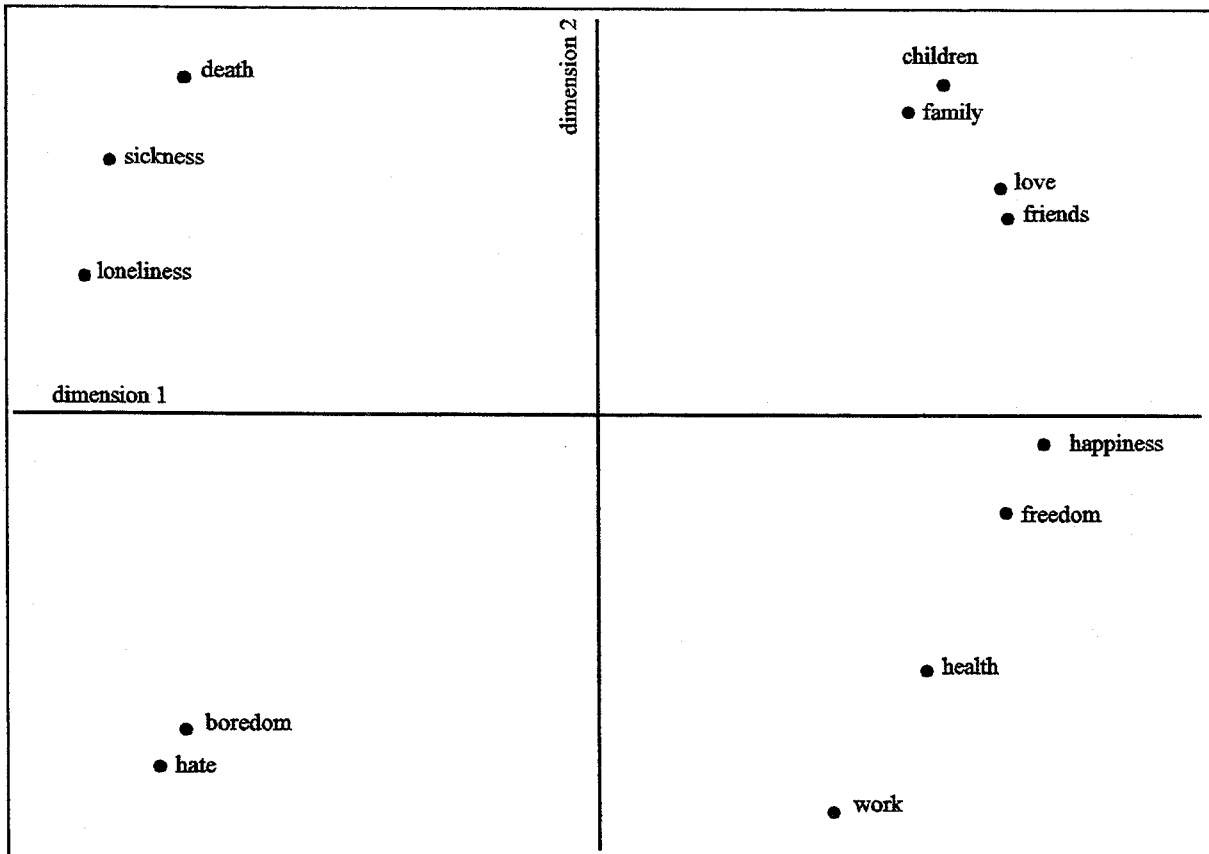


Figure 2.11 : Echelonnement multidimensionnel de la similarité entre 13 concepts issus d'évocations libres liées au bonheur (d'après Thomas & Stock, 1988, p. 146)

Les évaluations des 13 concepts au regard des dimensions cognitives et affectives peuvent permettre de projeter ces dimensions dans la configuration résultante de manière à en faciliter l'interprétation (cf. Tournois & Dickes, 1993). Malheureusement, dans ce cas de figure, cette opération n'apporte aucune information supplémentaire car les évaluations utilisant l'inducteur affectif ("*feeling*") sont corrélées à .98 (!) avec les évaluations utilisant l'inducteur cognitif ("*thought*").

Le seul résultat vraiment assuré de cette étude est l'opposition entre positif et négatif, mais celui-ci est trivial. On regrettera donc certains choix méthodologiques. Il ne semble pas très pertinent d'utiliser, dans une tâche d'évocations libres, deux mots inducteurs opposés. On peut aussi se demander pourquoi les auteurs n'ont retenu que les 12 termes les plus fréquemment évoqués. Il est vrai que la tâche de comparaisons par paire est extrêmement lourde et ne peut s'appliquer qu'à un petit nombre d'items. Cependant, il existe d'autres techniques, plus adaptées pour des échantillons d'items importants, qui permettent d'obtenir des évaluations de similarité ; c'est le cas du tri subjectif.

Cet exemple de recherche utilisant les représentations n'invalide nullement l'intérêt que peut présenter une telle approche. Si nous n'avons pas choisi d'exemple plus illustratif de cet intérêt, c'est que, tout simplement, c'est le seul exemple existant^[13]. D'autres études de ce type sont donc nécessaires, d'autant qu'elles présentent un double avantage : outre les études structurales qu'elles permettent d'effectuer, les tâches d'évocations libres constituent une méthode empirique pour aborder le problème de l'univers des items d'un concept.

3.2.4 Une approche en facettes du bien-être subjectif : l'étude de Levy & Guttman (1975)

Nous avons vu que l'approche en facettes se proposait de définir un concept à partir de l'univers des items en en spécifiant les sources de variation, c'est-à-dire les facettes. La combinaison de ces sources de variation dans une phrase en facettes permet de construire les items qui relèvent du concept et cette phrase peut être validée ou invalidée en fonction des résultats empiriques.

Levy & Guttman (1975) proposent une application de cette approche au domaine du bien-être subjectif. Nous allons résumer leur étude tout en notant que les auteurs restent très vagues sur la manière dont les facettes ont été choisies. Elles relèvent, semble-t-il, plus de l'intuition que d'une réelle démarche théorique.

Dans un premier temps, l'univers des items est défini de manière très générale, ainsi que le montre la figure 2.12.

An item belongs to the universe of well-being items if and only if its domain asks for
 (cognitive) (level)
 (affective) assessment of the () of the state
 (instrumental) (treatment)
 of a social group in some life aera, and the range is ordered from "very satisfactory" to
 "very unsatisfactory" according to the normative criterion of the respondent for
 that aera of life.

Figure 2.12 : Définition de l'univers des items de bien-être d'après Levy & Guttman (1975).

[13] : Du moins à notre connaissance.

Cette définition, beaucoup trop vaste et insuffisamment spécifiée, est remplacée par une phrase dont les facettes sont clairement explicitées. Celle-ci est présentée dans la figure 2.13.

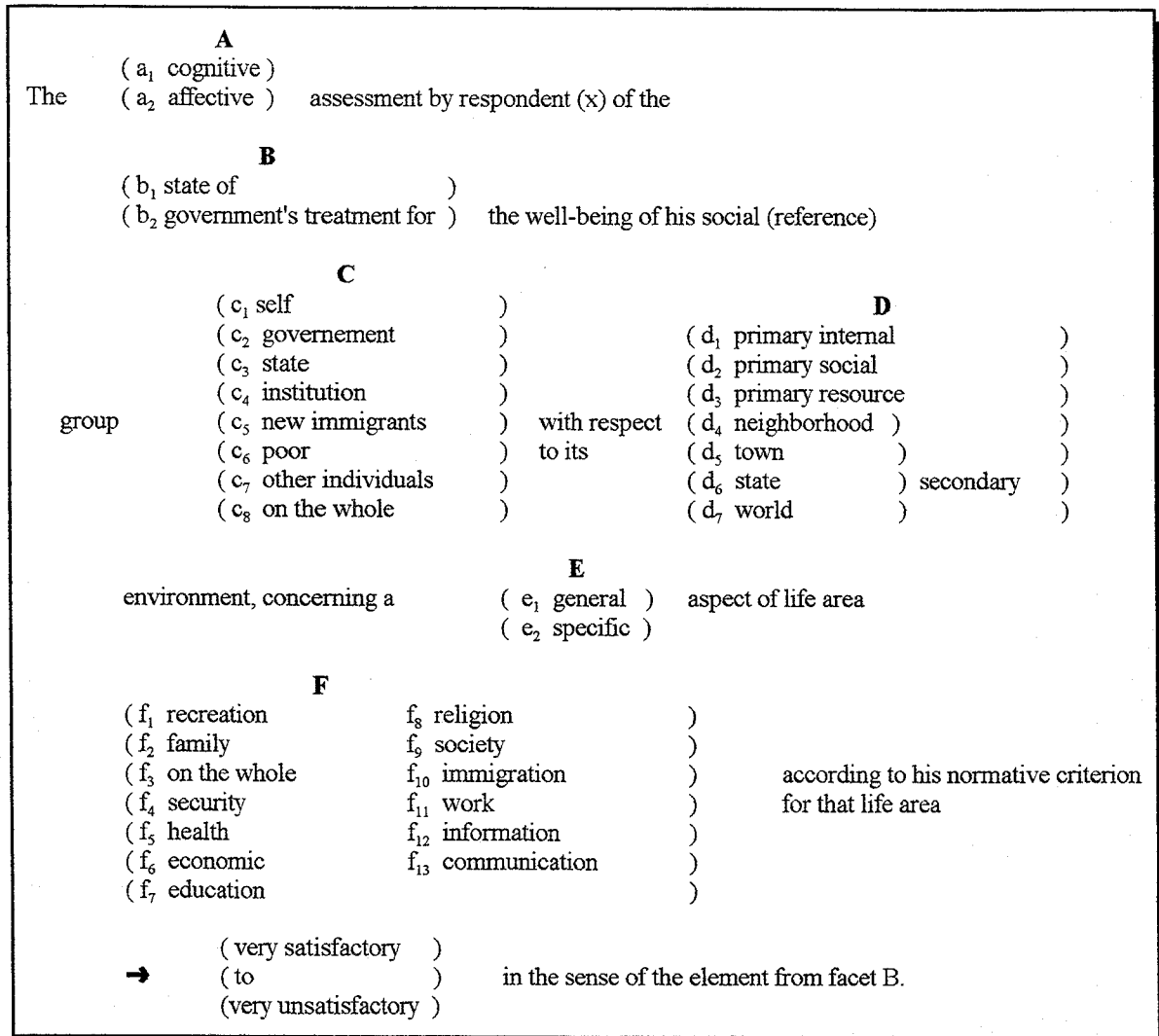


Figure 2.13 : Une phrase en facettes définissant le bien-être subjectif (d'après Levy & Guttman, 1975)

Ainsi qu'on peut le constater, cette phrase comporte six facettes de contenu dont la combinaison permet de définir 5824 structuples ($2 \times 2 \times 8 \times 7 \times 2 \times 13$) à partir desquels peuvent être construits des items. On retrouve dans la facette (A) la distinction entre une composante affective et une composante cognitive du bien-être. Comme Levy & Guttman considèrent que le bien-être est une attitude, sa présence n'est pas étonnante. La seconde facette renvoie à ce qui est évalué par le sujet : la situation présente ou la politique gouvernementale vis-à-vis d'un certain groupe social (facette C) dans un certain

environnement (facette D)^[14]. Les facettes (E) et (F) spécifient les domaines de vie concernés par l'évaluation. Aucune de ces facettes n'est clairement définie dans l'article auquel nous faisons référence et le choix des éléments de chacune des facettes n'y est pas discuté : en l'absence d'information, on a l'impression que ces choix sont largement arbitraires.

L'étude empirique conduite par Levy & Guttman ne comporte pas, bien évidemment, les 5824 items possibles. Seul un échantillon de 24 items est constitué, à partir de critères qui ne sont pas explicités. Sur ces 24 items, nous ne retiendrons, dans la suite du texte que les 12 qui concernent ce que les auteurs appellent le bien-être personnel^[15]. Ce sont ceux qui sont définis par le premier élément de la facette (C), l'évaluation que font les individus du bien-être d'autrui ne nous intéressant pas ici. La liste de ces items, ainsi que le structuple auquel ils correspondent, est présentée dans le tableau 2.8^[16].

A partir de ce tableau, il serait possible d'entrer dans le détail de l'assignation des items à chacune des facettes afin de discuter de la pertinence de certains choix. Nous ne prendrons que deux exemples. Le premier exemple concerne la facette (F) des domaines de vie. Si l'assignation de l'item n° 4 au domaine des loisirs (f_1 : *recreation*) ne pose guère de difficulté, il n'en est pas de même pour les items 8 à 12 qui concernent tous les cinq l'habitat, mais qui sont aussi assignés au domaine des loisirs. Les justifications théoriques de ce choix ne paraissent nullement évidentes. L'item n° 5 concerne la santé, mais il est assigné au domaine de la sécurité. On peut raisonnablement penser qu'il s'agit d'une coquille typographique. En revanche, qu'en est-il pour l'item n° 2 qui est lui aussi assigné au domaine de la sécurité : s'agit-il de la même coquille (auquel cas l'humeur serait assimilée au domaine de la santé) ou d'une autre ? Au vu des résultats présentés ultérieurement, il semble qu'il s'agisse de la même coquille mais ceci ne résoud en rien les problèmes théoriques posés par la mise en correspondance de l'humeur avec la santé.

Le second exemple fait référence à la facette qui distingue les aspects affectif et cognitif. L'intérêt théorique de cette facette est évident mais, comme elle n'est pas définie précisément, il faut se servir du contenu des items pour essayer de mettre en relief les spécificités de ces deux composantes. Les items qui comprennent des termes comme "humeur" ou "heureux" sont considérés comme relevant de la composante affective. Les

[14] : Il ne faut pas s'étonner de trouver les "nouveaux immigrants" comme élément de la facette (C). En effet, l'étude empirique se déroule en Israël.

[15] : En fait, lorsque l'on regarde en détail la liste des items des auteurs, on s'aperçoit qu'il y en a 13 et non pas 12. A nouveau, on est conduit à s'interroger sur cette bizarrerie, sans pouvoir l'expliquer (peut-être est-ce tout simplement un oubli ?).

[16] : Avant de consulter le tableau et de lire les lignes qui suivent, le lecteur curieux et peu pressé peut essayer, à partir du seul libellé des items, de retrouver leur assignation théorique.

items cognitifs sont ceux qui font référence à une évaluation, à un jugement, à un désir conscient. En revanche, ce qui est plus étonnant, c'est que les items concernant la satisfaction des individus font partie de la composante affective. Chez beaucoup d'autres auteurs, nous l'avons vu avec l'exemple de l'étude de Andrews & Withey (1976), les items de satisfaction relèvent explicitement de la composante cognitive car ils obligent le sujet à effectuer une évaluation. Le débat est ouvert et nous verrons plus bas qu'il est difficile de le clore, même aujourd'hui.

n°	structuples	items
1	a ₂ b ₁ c ₁ d ₁ e ₁ f ₃	Generally speaking, are you happy these days ? (very happy ...very unhappy)
2	a ₂ b ₁ c ₁ d ₁ e ₁ f ₄	How is your mood these days ? (very good all the time ... not good almost all the time)
3	a ₁ b ₁ c ₁ d ₂ e ₁ f ₂	In general, how do you evaluate your family life ? (very good ... very bad)
4	a ₂ b ₁ c ₁ d ₂ e ₁ f ₁	In general are you satisfied with the way you spend your leisure time ? (very satisfied ... not at all satisfied)
5	a ₁ b ₁ c ₁ d ₃ e ₁ f ₄	In general, how do you evaluate your health these days ? (very good ... not at all good)
6	a ₁ b ₁ c ₁ d ₃ e ₁ f ₆	Is your family income sufficient ? (definitively sufficient ... insufficient)
7	a ₂ b ₁ c ₁ d ₃ e ₁ f ₇	Are you satisfied with your education level ? (very satisfied ... very unsatisfied)
8	a ₂ b ₁ c ₁ d ₃ e ₁ f ₁	In general, are you satisfied with the apartment you live in ? (very satisfied ... very unsatisfied)
9	a ₁ b ₁ c ₁ d ₄ e ₁ f ₁	In general, how do you evaluate the neighborhood you live in ? (very good ... not at all good)
10	a ₂ b ₁ c ₁ d ₅ e ₁ f ₁	In general, are you satisfied with life in your town these days ? (very satisfied ... not at all satisfied)
11	a ₁ b ₁ c ₁ d ₅ e ₂ f ₁	Do you want very much to continue living in this town ? (definitively yes ... definitively no)
12	a ₁ b ₁ c ₁ d ₅ e ₂ f ₁	Do you want very much to move to another town ? (definitively no ... definitively yes)

Tableau 2.8 : Les items de bien-être "personnel" de Levy & Guttman (1975) et les structuples correspondant (en petits caractères apparaissent les échelons extrêmes de l'échelle de réponses, qui en compte 4 ou 5 selon les items ; en caractères gras sont mis en relief les mots-clés de chaque item, qui serviront ultérieurement pour la présentation des résultats).

Enfin, on peut signaler que certains items comportent une période de référence temporelle explicite (cf. les items n° 1, 2 et 5) alors que les autres n'en comportent pas. Comme cette distinction est aussi une source potentielle de variation des réponses, elle aurait pu faire l'objet d'une facette.

Dans l'approche en facettes, la définition est validée si on réussit à montrer que les facettes constituent des sources de variation des réponses. Le plus souvent, cette opération de validation passe par l'application d'un échelonnement multidimensionnel. Celui-ci permet de tester l'hypothèse selon laquelle les items les plus proches dans l'espace résultant de cette analyse sont ceux qui ont le plus d'éléments de facettes en

commun. C'est une SSA ("*Smallest Space Analysis*" ou "*Similarity Structure Analysis*"), forme particulière d'échelonnement multidimensionnel, qui a été appliquée à la matrice d'association entre les 12 items du tableau 2.8. Les coefficients d'association ne sont pas des corrélations linéaires mais des coefficients monotones. Les résultats de la solution bidimensionnelle sont présentés dans la figure 2.14.

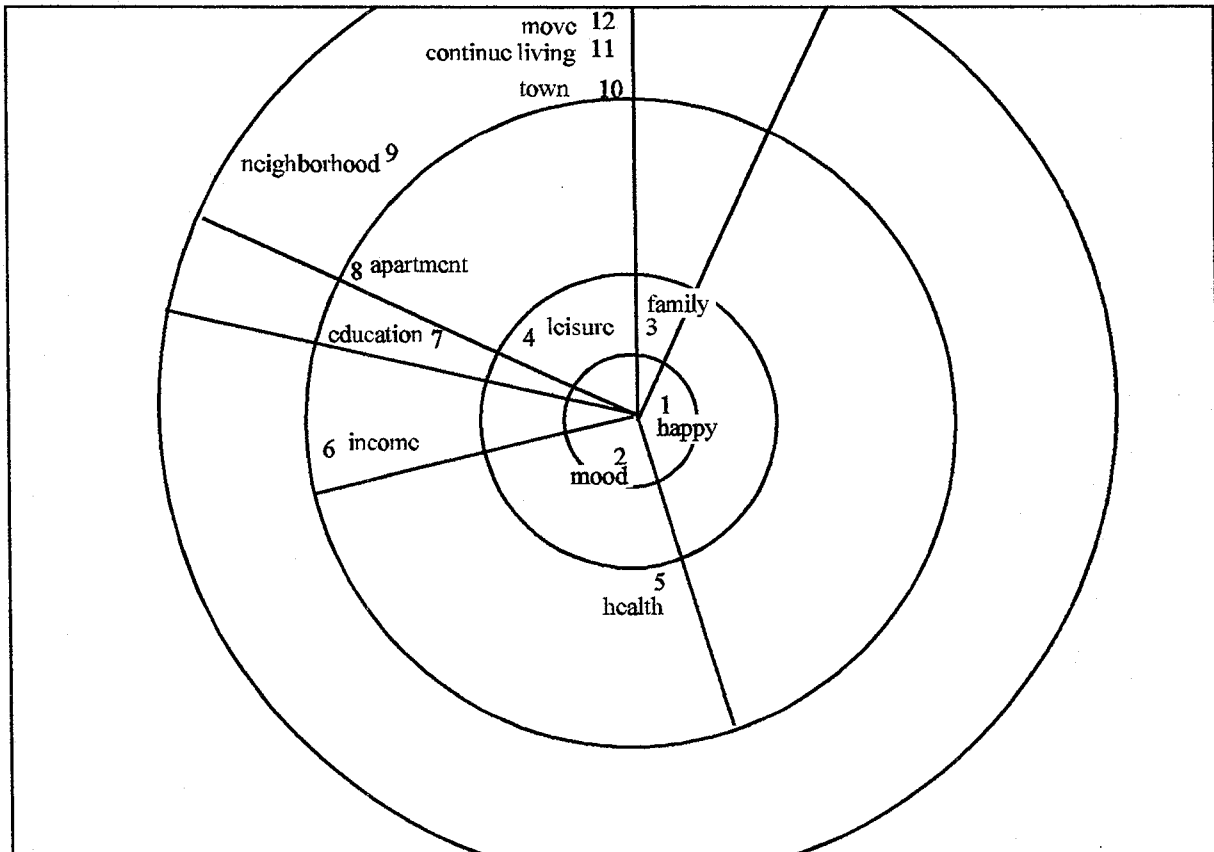


Figure 2.14 : La structure du bien-être personnel d'après Levy & Guttman (1975)

Cette représentation graphique des résultats fait apparaître une figure caractéristique, que l'approche en facettes désigne sous le nom de radex (cf. section 3.2 de la première partie). La facette (F) joue un rôle angulaire : les domaines de vie forment des régions distinctes, symbolisées par les traits qui partent du centre de la configuration. On remarquera que les items concernant l'habitat se situent effectivement dans la région "loisirs" et que l'item relatif à l'humeur se trouve dans la région "santé". Aussi est-on conduit à se demander s'il y a eu confirmation empirique de cette assignation ou si ces items ont été classés après coup, sur la base des résultats.

La facette (D), quant à elle, joue un rôle modulaire : elle découpe la configuration en cercles concentriques. Dans le cercle central, se trouvent les items en rapport avec

l'environnement primaire intime (d_1) ; les items du second cercle font référence à l'environnement social primaire (d_2) ; le troisième comprend les ressources de l'environnement primaire (d_3) ; enfin, dans le quatrième, les items correspondent à l'environnement secondaire (d_4, d_5). Cet éloignement des items à partir du centre reflète en quelque sorte la distance psychologique.

Les autres facettes ne sont pas validées par les résultats. C'est le cas, en particulier, de la facette (A) qui distingue les composantes affectivité et cognitive. Les auteurs ne s'expliquent pas sur ce point.

Cet exemple illustre parfaitement les qualités et les défauts des applications empiriques de l'approche en facettes. La formalisation d'une définition permet de valider ou d'invalidier explicitement les facettes prises en compte. Malheureusement, cette application souffre d'un manque évident d'assises théoriques dans le choix des facettes et de leurs éléments. De plus, la correspondance entre les items et les structuples n'est pas aussi rigoureuse qu'on pourrait le souhaiter. Enfin, l'étude empirique porte sur un nombre d'items ridiculement inférieur à celui des structuples théoriques. Tout comme pour le cas des études structurales qui s'appuient sur les représentations, ces remarques critiques ne remettent pas en cause l'intérêt de l'approche en facettes. Elles ont pour seul objectif d'en relever les points faibles afin de pouvoir améliorer les applications ultérieures.

3.3 Le couple positif-négatif dans l'étude du bien-être subjectif

3.3.1 Bien-être et mal-être : des dimensions différentes ?^[17]

Dans une acception assez large du concept de santé mentale, l'idée qu'il existe un continuum unique permettant d'ordonner les individus selon qu'ils se sentent plutôt "bien dans leur peau" ou qu'ils présentent des symptômes pathologiques semble intuitivement fort plausible. Dans la première revue de synthèse de la littérature, Wilson (1967, p. 303) déclare d'ailleurs que toutes les études réalisées ont en commun le fait qu'elles "explorent le côté positif du continuum de l'adaptation".

Ceci explique que lorsque Bradburn (1969), dans une des premières études de vaste ampleur sur le bien-être, montre que non seulement le bien-être et le mal-être sont des dimensions différentes mais qu'elles sont indépendantes, il en est le premier surpris. Cette surprise s'étendra à l'ensemble de la communauté scientifique^[18]. Depuis cette date, les travaux sur cette question ont été très nombreux et, aujourd'hui encore, ils continuent à agrémenter les pages des revues spécialisées.

Malgré son caractère contre-intuitif, ce résultat semble relativement bien établi, notamment grâce à des études qui utilisent des techniques différentes. De plus, il apparaît que le réseau nomologique caractérisant le bien-être n'est pas le même que celui qui caractérise le mal-être. Au passage, on a pu relever d'autres résultats tout aussi contre-intuitifs tels que, par exemple, une corrélation positive entre le nombre d'événements positifs et le nombre d'événements négatifs vécus par les individus (cf. Headey & Wearing, 1989). Au lieu d'indépendance entre bien-être et mal-être, on parle plutôt aujourd'hui d'indépendance entre affectivité positive et affectivité négative (respectivement PA et NA pour "*positive affect*" et "*negative affect*"). Les raisons de ce changement terminologique apparaîtront plus clairement au fil du texte.

C'est en utilisant l'échelle de balance affective (ABS) que Bradburn (1969) a mis en évidence cette indépendance, en corrélant le score total aux items positifs et le score total aux items négatifs. C'est aussi avec cette échelle que des auteurs comme Beiser (1974) ou Andrews & Withey (1976) ont réussi à répliquer ce résultat. Dans une adaptation française de cette échelle, Dickes (1989 a) aboutit aux mêmes conclusions. D'autres techniques ont aussi été utilisées pour confirmer ce résultat. Ainsi, Gannon et al. (1992) valident un modèle factoriel contraint d'indépendance entre PA et NA. Emmons

[17] : Ce titre reprend celui d'un article de Headey et al. (1984 b).

[18] : En fait, ce résultat était établi dès 1965 (cf. Bradburn & Caplovitz, 1965), mais l'usage veut que l'on retienne l'ouvrage de 1969 qui est plus complet.

& Diener (1985b, 1986) montrent que les scores moyens d'humeur positive mesurée pendant plusieurs jours consécutifs, ne sont pas significativement corrélés aux scores moyens d'humeur négative. Plus récemment, et avec une technique identique, Hepburn & Eysenck (1989) trouvent le même résultat. Zevon & Tellegen (1982) ainsi que Watson (1988 b) ont utilisé une approche quelque peu différente mais qui repose toujours sur des mesures quotidiennes de l'humeur. Zevon & Tellegen (1982) rapportent les résultats d'analyses factorielles en plan P, dans lesquelles les matrices analysées comprennent, en lignes, les occasions temporelles et, en colonnes, les variables. Les résultats de cette étude idiographique font apparaître deux facteurs, l'affectivité positive et l'affectivité négative, pour 21 des 23 sujets composant l'échantillon examiné (chaque sujet a rempli un questionnaire pendant 90 jours consécutifs). Dans l'étude de Watson (1988 b), 80 sujets ont rempli quotidiennement un questionnaire d'humeur pendant une moyenne de 7 semaines. L'auteur calcule tout d'abord, pour chaque sujet, la corrélation entre PA et NA à partir d'une matrice "occasions temporelles × variables", puis la moyenne de ces corrélations pour l'ensemble des individus. Cette corrélation moyenne est égale à -.20, ce qui, bien que statistiquement significatif, n'est pas très élevé.

Les autres arguments utilisés pour conforter l'indépendance entre PA et NA concernent la différence de réseau nomologique. Il s'agit d'établir que certaines variables sont corrélées à l'une des affectivités mais pas à l'autre (et inversement) pour montrer la validité divergente de ces deux dimensions. C'est Bradburn lui-même qui a fourni les premiers éléments allant dans ce sens. Il a montré en effet que NA (mais pas PA) était liée aux difficultés dans le mariage, dans le travail, aux tensions interpersonnelles et aux soucis. A l'inverse, PA (mais pas NA) est corrélée avec le niveau d'implication de l'individu dans son environnement et avec son niveau de contact social. Depuis, des résultats similaires ont été constamment obtenus (cf., parmi de multiples autres, Beiser, 1974 ; Cherlin & Reeder, 1975, Warr, 1978 ; Harding, 1982 ; Kop, 1993 a). On a aussi trouvé de nouvelles variables qui présentaient les mêmes caractéristiques : les plus célèbres sont, pour NA, le névrosisme, la santé et les événements de vie négatifs (cf., respectivement, Costa & McCrae, 1980 ; Watson, 1988 b ; Reich & Zautra, 1981, 1983) et, pour PA, l'extraversion et les événements de vie positifs (Costa & McCrae, 1980 ; Reich & Zautra, 1981, 1983). On ajoutera volontiers à cette liste les travaux expérimentaux sur l'humeur, trop souvent oubliés dans la littérature sur le bien-être, qui montrent que les effets de PA ne sont pas symétriques à ceux de NA (cf. Isen, 1987 et la section 4.2 du chapitre suivant). Avec Chamberlain (1988, p. 584-585), qui suggère aussi que NA serait plus corrélée avec des variables personnelles et PA avec des variables situationnelles, on peut résumer ces résultats comme suit :

- l'affectivité positive est liée à l'extraversion, à l'estime de soi, à la fréquence des contacts sociaux et à la disponibilité d'un soutien social ;
- l'affectivité négative est liée au névrosisme, au contrôle interne, aux événements négatifs, à la dépression, à l'anxiété et à la santé.

Affectivité négative et affectivité positive ont donc été mises en évidence par différentes méthodes. Pourtant, tout le monde ne défend pas l'indépendance totale entre ces deux dimensions, soutenue par Bradburn ou Gannon et al. (1992). Plusieurs auteurs défendent une indépendance partielle sur la base d'analyses factorielles avec rotation oblique (e.g. Headey et al., 1984 b) ou sur la base d'une simple corrélation entre les scores aux items positifs et négatifs (e.g. Kop, 1993 a) : dans les deux cas, on peut estimer que PA et NA partagent 25% de variance commune, ce qui est trop pour que ces deux dimensions soient totalement indépendantes et trop peu pour que ce soient les deux pôles opposés d'un même continuum. Mais il y a aussi quelques rares auteurs qui défendent l'idée d'un bien-être unidimensionnel. Parmi ceux-ci, Kammann et ses collaborateurs (cf. surtout Kammann et al., 1984 a), sont sans doute les plus représentatifs. Selon eux, des concepts comme ceux d'anxiété, de dépression ou de névrosisme forment tous, malgré leurs différences, le pôle négatif du bien-être. Et si les corrélations entre le bien-être et ces variables ne sont pas plus élevées, c'est parce que l'on doit s'attendre à une relation curvilinéaire : comme les mesures d'anxiété ou de névrosisme ne couvrent qu'une partie du continuum, à des scores bas sur ces traits peuvent correspondre des niveaux de bien-être très différents. Eux-mêmes, avec un instrument qu'ils ont mis au point, l'Affectomètre (cf. section 2.1.1), ne trouvent qu'une corrélation de -.66 entre PA et NA, ce qui reste assez éloigné de -1. En explication, ils évoquent l'existence d'un éventuel "effet plafond" pour les réponses aux items négatifs et d'un "effet plancher" pour les réponses aux items positifs.

Pour convaincantes qu'elles soient, ces explications peuvent toutefois difficilement rendre compte de résultats comme celui de l'existence de corrélats différents pour PA et NA. Elles incitent tout de même à examiner avec plus d'attention les critiques et les explications ponctuelles qui ont été adressées à Bradburn et, plus généralement, aux résultats mettant en évidence une indépendance entre PA et NA. C'est ce que nous allons faire à présent, en remettant à plus tard une discussion sur les explications théoriques générales venant conforter l'idée d'une indépendance au moins partielle entre ces deux dimensions.

3.3.2 Critiques et explications ponctuelles de l'indépendance entre affectivité positive et affectivité négative

Comme l'indépendance entre PA et NA a été établie pour la première fois avec l'ABS^[19], il n'est pas étonnant que les auteurs aient cherché à expliquer ce résultat à partir des spécificités propres à cet instrument de mesure. Nous en verrons plusieurs exemples. D'autres pourtant ont proposé des explications moins spécifiques. Certaines sont surtout

[19] : Il est certain que d'autres, avant Bradburn, ont évoqué cette possibilité (cf. Jahoda, 1958 ; Herzberg et al., 1957 ; Nowlis, 1965), mais cela a eu peu ou pas d'impact sur les recherches concernant le bien-être.

techniques et attribuent les résultats à des artefacts d'ordre méthodologique. Mais c'est à Diener et al. (1985 b) que revient le mérite d'avoir proposé une explication fort séduisante faisant intervenir notamment la manière dont les sujets répondent aux questionnaires de bien-être. Ironie de l'histoire, quelques années plus tard, ces mêmes auteurs apportent des résultats qui, sans qu'ils le reconnaissent explicitement, portent le coup de grâce définitif à cette explication.

Explications faisant intervenir des caractéristiques spécifiques de l'ABS

Plusieurs critiques concernent directement le contenu des items de l'ABS. Dès 1967^[20], Wilson fait remarquer que les items négatifs font plutôt référence à des expériences générales alors que les items positifs concernent des expériences plus spécifiques. Cette critique sera reprise par Brenner (1975) et par Cherlin & Reeder (1975). En accordant le même poids à des sentiments très généraux et à des sentiments beaucoup plus spécifiques, on augmenterait la probabilité de ne pas trouver de corrélation entre les deux ensembles d'items.

Plus fondamentale, la remarque de Cherlin & Reeder (1975) fait intervenir la signification pouvant être attribuée au contenu des items. Ils considèrent que les items positifs de l'ABS sont principalement des mesures du niveau d'activation ou d'engagement alors que les items négatifs font référence à des expériences déplaisantes. Ces interprétations différentes renvoient aux travaux sur la structure des émotions que nous évoquerons en détail dans la section suivante. Signalons juste que le niveau d'activation et le caractère plaisant constituent, chez certains auteurs, les dimensions de base des expériences émotionnelles et que ces deux dimensions sont indépendantes. Il ne serait donc pas étonnant de trouver une corrélation nulle entre les items positifs et négatifs de l'ABS.

Outre le contenu des items, la pertinence de l'échelle de réponses a aussi été questionnée. A l'origine, l'ABS proposait une "échelle" dichotomique, les sujets devant indiquer s'ils avaient connu ou non chacune des expériences. Plusieurs auteurs (e.g. Brenner, 1975 ; Cherlin & Reeder, 1975 ; Diener, 1984 ; Kammann et al., 1984 a) ont fait remarquer que l'utilisation de cette échelle de réponses revenait à accorder autant de poids à un sentiment éprouvé une seule fois qu'à une expérience très fréquente. Un tel procédé peut faire baisser artificiellement des corrélations. En remplaçant les réponses dichotomiques par une échelle de réponses formulée en termes de fréquences, Brenner (1975) ou Warr et al. (1983) rapportent des corrélations de l'ordre de -.50 qui semblent remettre en cause l'idée de l'indépendance entre PA et NA. Mais ce résultat n'a pas

[20] : Les items de ce qui allait devenir l'ABS avaient déjà été utilisés par Bradburn & Caplovitz en 1965.

toujours été répliqué : en modifiant aussi l'échelle de réponses, Beiser (1974), Andrews & Withey (1976) ou encore Dickes (1989 a) continuent à trouver des corrélations très faibles, voire non significatives, entre PA et NA^[21].

Enfin, Diener (e.g. Diener, 1984) évoque le fait que les consignes de l'ABS font une référence explicite à une période temporelle (au cours de ces dernières semaines). Or, pour Diener, il est tout à fait possible que PA et NA soient fortement corrélées négativement sur de courtes périodes temporelles (à la limite, à un moment donné, on ne peut éprouver qu'un seul sentiment) et qu'ils deviennent relativement indépendants au fur et à mesure que la durée de la période augmente. Avec Emmons (Diener, & Emmons, 1985), il publiera une série d'études visant à montrer ce phénomène. Souvent utilisé, cet argument donnera lieu pourtant à peu de vérifications empiriques. Watson (1988 a) se trouvera dans l'incapacité de répliquer les résultats de Diener & Emmons (1985) ; il signale même que cet article rapporte des résultats surprenants, comme une corrélation de -.63 entre PA et NA lorsque la période temporelle est de un mois, qui chute à -.10 lorsqu'elle est de trois semaines. Il faut toutefois reconnaître que ces deux résultats ont été obtenus avec des méthodes différentes et qu'il est difficile de les comparer directement. Watson considère que les variations dans le niveau de la corrélation entre PA et NA dépendent beaucoup plus du choix des items. Nous y reviendrons plus bas.

Explications faisant intervenir des artefacts méthodologiques

Nous avons déjà évoqué les remarques de Kammann et al. (1984 a) concernant l'éventuelle existence d'une relation curvilinéaire entre PA et NA ainsi que celle d'un effet plafond pour NA et d'un effet plancher pour PA. Même si ces explications sont plausibles, il n'est guère probable qu'elles puissent rendre compte des multiples résultats, obtenus avec des méthodes différentes, qui confirment le résultat de Bradburn. Tout au plus, de tels effets, s'ils existent, peuvent atténuer légèrement les corrélations, mais non les faire disparaître totalement.

Certains auteurs contestent l'indépendance entre PA et NA sur la base des erreurs de mesure. Lorsque celles-ci ne sont pas prises en compte explicitement, elles font diminuer les corrélations entre les mesures. Ce phénomène d'atténuation due à la fidélité est bien connue dans le cadre de la théorie classique du score vrai (cf. Dickes et al., 1994). A l'appui de cette constatation, Green et al. (1993, 1994) rapportent des données, obtenues grâce à une méthodologie multitraits-multiméthodes, qui permettent de prendre en compte les erreurs de mesure aléatoires et les erreurs de mesure systématiques (ou corrélées). Dans ce cas, il est possible d'estimer la corrélation "vraie" entre des construits.

[21] : Dans une étude empirique rapportée dans la troisième partie, on verra que l'on a trouvé une corrélation de -.21 entre PA et NA, avec une version française de l'ABS, présentant aux sujets une échelle de fréquence en sept points.

Les résultats rapportés par les auteurs les incitent à conclure qu'une structure avec un facteur bipolaire est plus appropriée qu'une structure avec deux facteurs orthogonaux^[22].

Une autre explication, faisant intervenir la qualité des réponses de certains sujets, est proposée par Schmitt & Stults (1985, cf. aussi à ce sujet, Pilotte & Gable, 1990). Dans une étude sur données simulées, ils montrent qu'il suffit que 10% des sujets répondent de manière incorrecte à des items négatifs (i.e. négations grammaticales ou opposés sémantiques), pour qu'un facteur saturant ces items apparaisse en plus du facteur saturant les items positifs. Pour que cette explication soit plausible, il faut toutefois admettre qu'une partie des sujets réponde aux items négatifs de la même manière qu'aux items positifs sans s'apercevoir qu'ils ont un sens opposé.

Enfin, toujours en référence aux travaux utilisant l'analyse factorielle, on ne peut exclure totalement l'hypothèse que les facteurs positif et négatif soient des "facteurs de difficulté". Appelés ainsi en raison d'une longue tradition de recherche dans le domaine des épreuves de performance, ces facteurs apparaissent lorsque des items appartenant à une même dimension diffèrent de manière importante en termes de moyennes. Nous verrons en effet, dans la troisième partie, qu'il y a de sérieux arguments pour penser que les moyennes des items positifs de bien-être sont assez nettement inférieures à celles des items négatifs (après qu'ils soient recodés dans le sens du bien-être, cf. Kop, 1993 b).

Toutes les explications techniques venant d'être avancées sont plausibles. Toutefois, elles ne rendent compte que du possible caractère artefactuel de l'indépendance totale entre les dimensions d'affectivité positive et d'affectivité négative. A notre avis, elles ne remettent pas en cause l'existence d'une indépendance partielle, d'autant que les différences entre les réseaux nomologiques caractérisant ces deux dimensions sont beaucoup plus difficiles à expliquer avec de tels artefacts.

Explications faisant intervenir l'intensité

Au-delà des explications ponctuelles que nous venons de voir, Diener et al. (1985 b) ont fait l'hypothèse, théoriquement plus satisfaisante, que deux dimensions indépendantes intervenaient dans la structure de l'affectivité : "La première dimension est la fréquence d'affectivité positive ou le nombre de fois ("*amount of time*") que l'affectivité positive prédomine sur l'affectivité négative. La seconde est l'intensité ou la force ("*strength*") avec laquelle on éprouve l'affectivité." (Diener et al., 1985 b, p. 1255).

[22] : En fait, les auteurs appuient surtout leurs conclusions sur la comparaison entre les corrélations brutes et les corrélations pour lesquelles les erreurs de mesure sont prises en compte. Utilisant l'analyse factorielle contrainte, ils auraient pu, explicitement, tester un modèle orthogonal et un modèle unifactoriel (complété éventuellement par un modèle à deux facteurs obliques). On notera aussi que leurs résultats diffèrent selon les items utilisés.

Les différentes expériences rapportées dans cet article semblent montrer que l'intensité est une composante importante de l'affectivité (cf. Larsen & Diener, 1987, pour une revue de question sur cette dimension). Un instrument sera mis au point pour la mesurer (*Affect Intensity Measure*, cf. Larsen & Diener, 1985). L'existence de cette dimension s'appuie principalement sur des corrélations *positives* élevées constatées entre l'intensité avec laquelle sont éprouvés des affects positifs et des affects négatifs. Deux autres résultats de cette étude sont importants pour notre propos. Le premier est l'existence de corrélations *negatives* élevées entre la fréquence de l'affectivité positive et négative. Le second est l'indépendance entre l'intensité de l'affectivité et la fréquence (ce n'est pas parce que l'on éprouve souvent des sentiments positifs que ces sentiments sont intenses). Ainsi, si les réponses aux questionnaires d'affectivité dépendent à la fois de l'intensité des sentiments et de leur fréquence, il n'est pas étonnant de trouver des corrélations nulles entre ces réponses, puisque l'association négative en termes de fréquences est compensée par l'association positive en termes d'intensité.

Cette explication théorique n'a pas rencontré grande opposition dans la littérature. Elle a le mérite d'être plausible, globale et heuristique (on a notamment proposé des typologies permettant de distinguer quatre groupes de sujets selon le niveau de fréquence et le niveau d'intensité des affects, cf. Chamberlain, 1988). Pourtant, cette hypothèse soulève à la fois des questions théoriques et méthodologiques.

Au niveau théorique, on peut s'interroger sur la signification de la dimension d'intensité. A y regarder de plus près, l'interprétation de cette dimension n'est pas aussi claire qu'il y paraît au premier abord. Chez Larsen & Diener (1985), l'intensité ne serait rien d'autre que la variabilité de l'humeur, dont le caractère fondamental a été mis en évidence par Wessman & Ricks en 1966. En 1987, les mêmes auteurs estiment pourtant que la variabilité de l'humeur et son intensité ne sont pas des concepts identiques, même s'ils sont corrélés : l'intensité reflète la force typique des états affectifs alors que la variabilité reflète l'étendue et la fréquence des changements de ces états. En 1990, une autre interprétation est proposée : pour Thomas & Diener (1990), l'intensité est identique à la dimension d'activation qui fait partie de la structure de l'émotion dont on parlera plus loin. De manière plus fondamentale encore, on peut se demander si l'intensité est bien une composante du bien-être. Dans les premiers travaux de Diener & Larsen, cela ne semble faire aucun doute : l'intensité et la fréquence sont deux dimensions indépendantes (cf. la vérification empirique à l'aide d'une matrice multitraits-multiméthodes chez Larsen & Diener, 1985) qui sous-tendent les réponses aux items de bien-être. Pourtant, dans l'article de 1987, les auteurs affirment que l'intensité est indépendante du bien-être. Comme on peut difficilement envisager qu'un concept puisse être tout à la fois une composante d'un autre concept et en être indépendant, il faut accepter l'idée d'un changement conceptuel. Celui-ci existe effectivement. Il apparaît en filigrane dans l'article de 1987 et de façon beaucoup plus explicite en 1991 (Diener et al., 1991 b). Le bien-être n'est plus alors

conçu que comme la fréquence, la durée ou encore la prédominance des affects positifs sur les affects négatifs. C'est pourquoi on peut alors parler d'indépendance entre bien-être (c'est-à-dire fréquence des affects) et intensité. Nous allons revenir sur ce point afin d'évaluer les conséquences pour la question de l'indépendance entre PA et NA, juste après avoir évoqué quelques aspects méthodologiques liés à la mesure de l'intensité.

Au niveau méthodologique, deux procédures ont été utilisées pour mesurer l'intensité, chacune d'entre elles appelant des réserves. Dans la première, on demande à des sujets de remplir, pendant plusieurs jours consécutifs, des questionnaires d'affectivité comprenant des items positifs et des items négatifs. L'intensité émotionnelle positive est opérationnalisée comme le niveau moyen d'affectivité *calculé sur l'ensemble des jours où l'affectivité positive est supérieure à l'affectivité négative*. C'est l'inverse pour l'affectivité négative. On se rappelle que l'existence de la dimension d'intensité repose sur le fait que l'on trouve une corrélation positive entre intensité de l'affectivité positive et intensité de l'affectivité négative. Or, Cooper & McConville (1993) ont montré, à partir de simulations fort convaincantes, que l'on pouvait retrouver cette corrélation avec des données totalement aléatoires, pour peu que l'on utilise la même procédure d'opérationnalisation de l'intensité^[23]. La seconde procédure est plus traditionnelle puisqu'elle consiste à utiliser un questionnaire, l'AIM (*Affect Intensity Measure*, Larsen & Diener, 1985). Dans ce questionnaire^[24], les sujets doivent indiquer, *sur une échelle de fréquence*, l'intensité avec laquelle ils éprouvent des émotions. La signification des réponses est donc ambiguë puisqu'elles combinent la fréquence et l'intensité. De plus, certains items sont susceptibles d'une double interprétation. Prenons l'exemple de l'item — traduit littéralement — "mon humeur négative est d'intensité moyenne". Lorsqu'un sujet répond "jamais" à cet item, cela peut indiquer deux réalités différentes. Soit le sujet ressent une humeur négative de manière très peu intense, soit, au contraire, il ressent ce type d'humeur de manière très intense. Or, dans la codification des réponses, seule la deuxième interprétation est prise en compte.

Si nous revenons, maintenant, à la pertinence de l'explication de l'indépendance entre PA et NA à partir des dimensions intensité et fréquence, il faut se souvenir qu'elle repose entièrement sur l'hypothèse que les sujets qui répondent aux items de bien-être utilisent *à la fois* des informations concernant la fréquence de leurs affects et des informations concernant leur intensité. Or, c'est justement cette hypothèse qui est mise à mal dans un article de Diener et de ses collègues paru en 1991.

[23] : Larsen & Diener n'ont pas encore, à notre connaissance, réagi à ces critiques. Dans leur article de 1987 (p. 7), ils évoquent bien quelques doutes quant à cette procédure, mais ils se satisfont finalement de la convergence des mesures.

[24] : On trouvera chez Williams (1989) des arguments montrant que le caractère unidimensionnel de cette échelle est discutable.

Le titre de cet article est éloquent : "*Happiness is the frequency, not the intensity, of positive versus negative affect*". Autrement dit, le bien-être doit être conceptualisé en termes du rapport entre la fréquence des affects positifs et des affects négatifs et non en termes d'intensité. Les arguments avancés en faveur de cette thèse sont au nombre de trois. L'un évoque les conséquences néfastes que peuvent avoir d'intenses expériences positives : elles ont tendance à diminuer l'intensité des expériences positives ultérieures, voire même à être suivies d'expériences négatives tout aussi intenses. Le second argument est que la fréquence des affects est mesurée de manière plus précise et plus valide que leur intensité. Bien que discutable^[25], cet argument n'invalide par l'explication sur l'indépendance. En revanche, le dernier est beaucoup plus embarrassant, puisque les auteurs réussissent à montrer, à travers plusieurs études empiriques, que le postulat sur lequel s'appuyait tout le modèle explicatif est faux. En effet, ils montrent que "(...) la fréquence relative de l'affectivité positive est une condition à la fois **nécessaire et suffisante** pour produire des scores élevés sur différentes mesures de bonheur" (dont l'ABS) (Diener et al., 1991 b, p. 123, ce sont les auteurs qui soulignent). Ainsi, outre les problèmes méthodologiques et théoriques soulevés par la dimension intensité des affects, ces résultats incitent à abandonner cette explication.

L'intensité est aussi évoquée d'une autre manière dans les travaux de Diener, pour rendre compte de l'indépendance entre PA et NA. Selon Diener & Iran-Nejad (1986), il y aurait indépendance lorsque les émotions sont de faible intensité mais, lorsque celles-ci sont intenses, des corrélations négatives assez importantes devraient être constatées. Cette hypothèse est plausible. Néanmoins, les résultats empiriques rapportés en sa faveur ne nous convainquent pas plus que les précédents. Pour comprendre nos réserves, la méthodologie utilisée dans cette étude doit être prise en compte. Les auteurs demandent à des sujets de remplir, quotidiennement, des questionnaires comprenant des adjectifs décrivant des états affectifs positifs et des états affectifs négatifs. Les réponses à chaque adjectif s'effectuent sur une échelle en 7 points, le point 0 signifiant que l'affect n'est pas du tout éprouvé, le point 6 signifiant qu'il est ressenti de manière très intense. Mais, pour montrer que PA et NA sont liées par une relation négative et non linéaire dans le cas d'émotions intenses, ils demandent aux sujets de ne remplir les questionnaires *que lorsqu'ils ressentent une émotion de manière assez forte*. C'est justement cette dernière condition qui fait que leur démonstration n'est pas acceptable. Pour le comprendre il suffit d'examiner la figure 2.15.

[25] : L'argument consistant à dire qu'il vaut mieux définir un construit de telle manière parce que la mesure en est plus facile et d'un niveau de mesure plus élevé ne paraît pas défendable. C'est au chercheur qu'il appartient de mettre au point des mesures valides en fonction de la définition de son objet d'étude et non l'inverse. De plus, l'affirmation (p. 122) selon laquelle la fréquence peut être mesurée à un niveau d'intervalles ou même de rapports, alors que l'intensité ne pourrait être mesurée qu'au niveau ordinal ne repose sur aucune justification sérieuse, en dehors d'arguments du sens commun.

Le graphique A représente un nuage de points entre deux variables indépendantes. Lorsque Diener & Iran-Nejad (1986) demandent à leurs sujets de remplir les questionnaires uniquement lorsqu'ils ressentent des émotions intenses, il se produit un effet de sélection qui conduit inévitablement au nuage de points représenté dans le graphique B : il présente une relation négative non linéaire. Ce que nous voulons montrer à travers cet exemple, c'est que l'hypothèse testée par les auteurs n'est pas réfutable. Il n'est pas certain qu'elle soit fausse, mais les éléments empiriques rapportés dans l'article ne permettent pas de juger de sa validité.

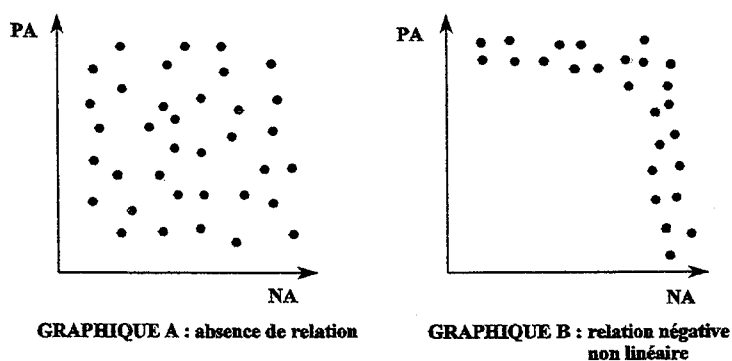


Figure 2.15 : Illustration fictive de l'effet de sélection des valeurs extrêmes sur la forme du nuage de points entre deux variables

Ces longs commentaires sur l'utilisation de l'intensité des émotions pour expliquer l'indépendance entre PA et NA nous ont parfois éloigné de notre sujet. Ils ne sont pas inutiles pourtant car ils sont assez typiques des problèmes conceptuels et/ou méthodologiques que l'on rencontre dans la littérature sur le bien-être subjectif. Nous en verrons d'autres exemples. Avant cela, il faut nous intéresser, toujours dans le cadre du couple positif-négatif aux travaux sur la structure des émotions qui proposent un cadre général permettant de mieux comprendre l'indépendance entre PA et NA.

3.3.3 La structure de l'affectivité

Plutôt que la structure de l'affectivité, on aurait pu intituler ce paragraphe, à l'instar de ce que l'on peut trouver dans la littérature, la structure de l'émotion ou de l'humeur ou des sensations. Sous ces termes différents se cachent, apparemment, des réalités similaires, même s'il est difficile de connaître exactement les nuances conceptuelles utilisées par les auteurs qui sont rarement explicites sur ce point^[26]. Aujourd'hui, on aurait tendance à utiliser de façon privilégiée l'expression "structure des émotions" (cf. Larsen & Diener, 1992). Cependant, suivant en cela Russell (1980), on considèrera qu'une émotion est représentée dans le langage par un affect ; nous préférons donc parler de la structure de l'affectivité, mais comme la terminologie n'est pas fixée, nous emploierons indifféremment l'un ou l'autre des termes ci-dessus.

Les travaux sur les émotions ont une longue histoire et cette histoire se confond avec l'identification des dimensions qui sous-tendent les expériences émotionnelles. La première typologie en ce domaine est sans doute celle de Galien (sanguin, phlegmatique, bilieux, atrabilaire), qui sera ensuite reprise notamment par Kant. Wundt lui donnera une interprétation plus dimensionnelle, sans en changer véritablement la teneur. Le caractère plaisant (vs déplaisant) d'une émotion et son intensité semblent constituer les dimensions de base de l'expérience émotionnelle. Les chercheurs travaillant dans le domaine du bien-être ont longtemps ignoré les travaux sur les émotions, si ce n'est pas l'intermédiaire de citations de circonstance, sans réel impact. C'est notamment le cas de Bradburn lui-même (cf. Cherlin & Reeder, 1975) dont les publications ne font jamais référence à une littérature dont on va pourtant voir les liens profonds avec les études sur le bien-être.

On peut considérer qu'il y a eu deux traditions de recherche différentes sur la structure de l'affectivité^[27], ces traditions conduisant à des résultats longtemps considérés comme antinomiques.

Dans la première, on fait évaluer à des sujets la similarité entre des expressions émotionnelles faciales ou vocales ou, directement, entre des adjectifs censés indiquer des émotions particulières. Une méthode alternative consiste à faire évaluer des indicateurs de l'émotion avec la technique du différenciateur sémantique. Les données obtenues de la sorte sont soumises, en général, à un échelonnement multidimensionnel dont les résultats font apparaître deux dimensions principales : un facteur bipolaire décrivant le caractère plaisant ou déplaisant d'une émotion et un facteur unipolaire indiquant son niveau d'intensité ou son niveau d'activation. Une troisième dimension a quelquefois été

[26] : Cf. notre discussion sur les aspects terminologiques de l'étude des émotions dans la section 1.3.

[27] : Cf., pour une présentation rapide de ces deux traditions et de la bibliographie s'y rapportant, Bush (1973), Russell (1980) et Watson & Tellegen (1985).

proposée, mais sans grand succès. Par exemple, Russell & Mehrabian (1977) proposent d'y ajouter la soumission-dominance, c'est-à-dire le degré de contrôle que l'on peut avoir sur les événements accompagnant les émotions. Cependant, comme cette troisième dimension est corrélée de manière non négligeable avec la première, elle n'a pas été retenue. De manière plus générale, c'est surtout l'absence de réplique qui a provoqué la faillite des solutions tridimensionnelles^[28].

Dans la seconde tradition, on utilise de manière privilégiée des questionnaires auto-administrés qui sont traités ensuite par analyse factorielle. Les résultats de ces études font apparaître plusieurs facteurs, généralement unipolaires, dont le nombre varie, selon les auteurs, entre 5 et 12 (par exemple, tristesse, anxiété, colère, tension, exaltation, etc.). C'est à Watson (cf. Watson & Tellegen, 1985) que revient le mérite d'avoir clarifié les structures obtenues dans ce courant de recherches. Grâce à des analyses secondaires systématiques de toutes les matrices de corrélations publiées sur la question, les auteurs peuvent identifier, au-delà des spécificités de contenu propres à chaque questionnaire, deux facteurs de second-ordre bipolaires et indépendants qu'ils désignent sous le nom d'affectivité positive et d'affectivité négative.

Il restait à réconcilier les résultats obtenus par ces deux approches. Dans les deux cas, on est en présence de solutions bidimensionnelles, mais les facteurs n'ont pas la même interprétation. C'est toujours à Watson & Tellegen (1985) que l'on doit une proposition de réponse séduisante. On sait depuis Schlosberg (1952) que les termes caractérisant les émotions peuvent être représentés sur un cercle dans un espace à deux dimensions. Techniquement, cette structure s'appelle un "circumplex" et dans le langage de l'analyse factorielle, cela indique que la matrice de corrélation n'a pas de structure simple. Autrement dit, comme tous les termes affectifs sont situés sur un cercle dans lequel de petites distances correspondent à de fortes similarités et de grandes distances à des oppositions maximales, il n'est pas possible de trouver un système d'axes de coordonnées qui permettent de rendre compte mieux qu'un autre des relations entre les variables. Cela paraîtra clair à l'examen de la figure 2.16.

Ce cercle de l'affectivité est divisé en huit secteurs à l'intérieur desquels se trouvent des termes décrivant des émotions similaires. Ces secteurs sont appelés des "octants". Il paraît évident que, pour une structure de ce type, n'importe quel système d'axes de coordonnées orthogonaux est à même de rendre aussi bien compte des relations entre les termes affectifs. Ainsi, l'opposition entre les résultats des deux traditions s'explique par

[28] : Récemment, Roberts & Wedell (1994) ont émis l'hypothèse que les différences dans les structures des émotions pouvaient être dues, dans certains cas, à des effets de contexte. En effet, dans des études qui se basent sur des similarités perçues, deux éléments seront évalués d'autant plus proches que les éléments de l'ensemble sont très différents. Pourtant, ces mêmes éléments seront probablement jugés comme étant beaucoup plus dissemblables dans le cas où les éléments à évaluer sont très similaires. De tels effets de contexte seraient surtout perceptibles avec des techniques comme le tri subjectif.

le fait qu'elles utilisent des systèmes d'axes de coordonnées différents. Ils sont représentés en lignes pointillées pour la première et en traits pleins pour la seconde.

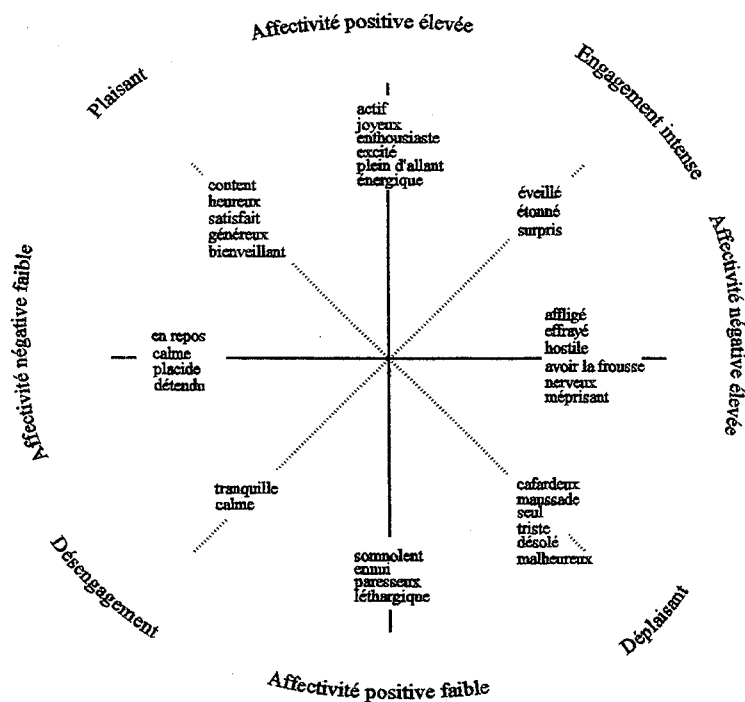


Figure 2.16 : Le circumplex de l'affectivité (adapté d'après Watson & Tellegen, 1985)

Ainsi, les deux structures de l'affectivité sont tout à fait compatibles. On passe de l'une à l'autre par une simple rotation orthogonale des axes de coordonnées. Les débats qui subsistent à ce sujet dans la littérature proviennent principalement du fait que certains auteurs préfèrent utiliser les résultats d'analyses factorielles avant rotation, auquel cas les facteurs s'interprètent comme "plaisant-déplaisant" et "activation" (ou "engagement"), alors que d'autres s'appuient sur les résultats d'une rotation varimax, grâce à laquelle ce sont les facteurs d'affectivité positive et d'affectivité négative qui apparaissent.

Cette structure de l'affectivité est relativement consensuelle^[29]. Le fait qu'on la retrouve avec des méthodes différentes (jugements de similarités ou questionnaires d'auto-évaluation) et dans des pays différents (Etats-Unis, Suède, Japon, Israël, Grèce, Pologne, Estonie, Chine, Croatie, etc.) est un sérieux argument en faveur de sa stabilité (cf. Larsen

[29] : Dans la littérature sur l'émotion, on trouve aussi des travaux qui privilégient une approche taxonomique sur une approche dimensionnelle. L'étude de Storm & Storm (1987) et, surtout, celle de Shaver et al. (1987) en constituent de bons exemples. Les résultats obtenus avec ces deux types d'approches ne semblent pas antithétiques, même s'il est difficile, pour le moment, de les intégrer dans un même cadre formel.

& Diener, 1992, pour les références concernant ces études). Ce qui n'est pas résolu aujourd'hui, c'est le choix des dimensions de base. Comme il est clair qu'au niveau mathématique cette question n'a pas de sens, ce sont des arguments théoriques qui sont utilisés par les chercheurs.

Mais, même au plan théorique, il n'y a pas de réponse univoque à cette question. On peut être tenté, comme l'ont fait Meyer & Shack (1989), de superposer le circumplex de l'affectivité à une structure bifactorielle de la personnalité. Cette idée n'est pas nouvelle, mais elle est intéressante. Larsen & Ketelaar (1991) notent d'ailleurs que l'étude des relations entre émotion et personnalité est en train de devenir une sous-discipline à part entière et que de telles études sont prometteuses. Pour des auteurs comme Tellegen (1985, cité par Watson & Tellegen, 1988), les dimensions "de base" de l'affectivité (affectivité positive et affectivité négative) sont identiques aux facteurs de second ordre qui se dégagent de certaines études de la personnalité et qui sont interprétés comme des dimensions d'émotionnalité positive et d'émotionnalité négative^[30]. Dans leur étude empirique, Meyer & Shack (1989) proposent aux sujets un questionnaire comprenant à la fois des items classiques de personnalité et des items d'affectivité. Leurs analyses permettent de montrer l'émergence d'une solution bifactorielle, dans laquelle l'un des facteurs sature à la fois les items d'affectivité positive et d'extraversion, l'autre, les items d'affectivité négative et de névrosisme. La mise en correspondance de ces deux structures est illustrée par la figure 2.17.

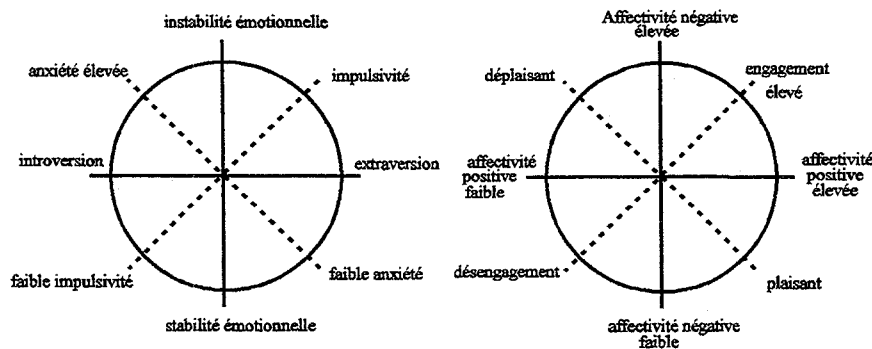


Figure 2.17 : Mise en correspondance du circumplex de la personnalité (à gauche) et du circumplex de l'affectivité (à droite) (adapté d'après Meyer & Shack, 1989)

Cette mise en correspondance montre qu'il peut être heuristique d'envisager, selon un même schéma, la structure de la personnalité et la structure de l'affectivité. Cependant, elle ne permet pas de décider quelles en sont les dimensions de base. En ce qui concerne

[30] : Meyer & Shack (1989) montrent d'ailleurs que la structure de l'affectivité est la même, que les items soient formulés en termes d'état (i.e. en ce moment) ou en termes de traits (i.e. en général).

la structure de la personnalité, certains auteurs (e.g. Gray, 1981) ont en effet proposé d'utiliser l'impulsivité (qui définit notamment les "extravertis névrotiques" sensibles aux signaux de récompense) et l'anxiété (qui définit notamment les individus qui ont une plus grande sensibilité aux signaux de punition, de nouveauté) à la place de l'extraversion et du névrosisme, comme dimensions de base. Or, ces deux dimensions sont tout simplement obtenues par une rotation de 45° des axes du circumplex de la personnalité ainsi que l'illustre la figure ci-dessus. Ainsi, nous sommes d'accord avec Larsen & Diener (1992, p. 41) lorsqu'ils écrivent que la question des dimensions de base du circumplex est une "pseudoquestion" et que les chercheurs devraient plutôt s'employer à établir "*the conditions under which one rotational scheme is more heuristically useful than others, not which scheme is more correct in some absolute sens*".

Les travaux sur la structure de l'affectivité constituent une contribution substantielle par rapport au débat sur l'indépendance entre PA et NA. Il paraît clair en effet que nombre de querelles sur le statut total, partiel ou inexistant de l'indépendance seraient caduques si les auteurs utilisaient une terminologie unique. Larsen & Diener (1992, p. 30) notent, par exemple, que certains auteurs, à l'intérieur du même article, utilisent indifféremment les termes "heureux/malheureux" et "positif/négatif" pour faire référence à la même dimension. Or, le circumplex de l'affectivité montre clairement qu'il ne s'agit pas de la même dimension.

Mais au-delà des choix terminologiques, c'est la sélection des items qui prend une importance primordiale. Si un chercheur construit une échelle avec des items appartenant aux octants "affectivité négative élevée" et "affectivité positive élevée" du circumplex, il ne doit pas être étonné que la corrélation entre les items positifs et négatifs soit nulle^[31]. En revanche, si, comme le font Kammann et Flett (1983), les items sont choisis en fonction de leur corrélation élevée (positive ou négative) avec une question comprenant dans son libellé le mot "heureux", il est logique de trouver une forte corrélation négative entre items positifs et négatifs car l'adjectif "heureux" appartient à la dimension "plaisant-déplaisant"^[32]. Le plus souvent pourtant, les échelles n'ont pas une structure aussi claire : les items sont souvent choisis sans référence au circumplex ce qui rend la signification des échelles imprécise. Ainsi, Lawton et al. (1993) font remarquer que l'échelle d'affectivité positive utilisée par Diener et ses collaborateurs dans leurs nombreuses publications ne contient que des adjectifs de l'octant "plaisant" du circumplex, alors que l'échelle d'affectivité négative contient des adjectifs provenant de deux octants différents :

[31] : Watson et al. (1988) ont proposé des échelles construites selon ce principe.

[32] : Ils obtiennent effectivement une corrélation de .66 entre "PA" et "NA", ce qui les conduit à penser que le bien-être est unidimensionnel. Nous l'avons vu plus haut.

"déplaisant" et "affectivité négative élevée"^[33]. En examinant en détail leurs publications, on trouve d'ailleurs plusieurs résultats cohérents avec cette proposition. Tout d'abord, les corrélations entre items positifs et items négatifs sont faibles mais pas nulles, ce qui est logique. Ensuite, on doit s'attendre à ce que la corrélation moyenne entre les items positifs soit plus élevée que celle entre les items négatifs. C'est effectivement ce que l'on constate à travers, par exemple, les données rapportées par Diener & Iran-Nejad (1986) : elle est égale à .80 dans le premier cas alors qu'elle n'est que de .60 dans le second.

3.3.4 L'indépendance entre affectivité positive et affectivité négative sur le plan théorique

Si l'indépendance entre PA et NA semble bien établie au niveau empirique au sens qu'attribuent Watson & Tellegen (1985) à ces deux dimensions dans le circumplex de l'affectivité, il reste à s'interroger sur la signification théorique de cette indépendance. La littérature sur cette question est trop parcellaire pour qu'il soit possible de se faire une opinion définitive, mais plusieurs travaux méritent d'être cités parce qu'ils offrent, à notre avis, des hypothèses de recherche intéressantes pour l'avenir.

L'idée selon laquelle bonheur et malheur ne seraient pas entièrement symétriques n'est pas très récente. On la trouve déjà dans la littérature philosophique (cf. Tatarkiewicz, 1976) et dans certains travaux scientifiques du début du siècle. C'est ainsi, par exemple, que Casson (1930, cité par Tatarkiewicz, 1976) considère que le corps humain contient beaucoup plus d'organes sensibles à la douleur que d'organes sensibles au plaisir.

Plus généralement, si PA et NA sont indépendants, c'est qu'ils ont des fonctions et des origines différentes. Nous avons déjà vu que ces deux dimensions se caractérisaient par des réseaux nomologiques différents, ce qui, tout en étant compatible avec cette idée, reste toutefois uniquement descriptif. Zautra & Reich (1983) vont plus loin lorsqu'ils mettent en relation l'affectivité avec deux systèmes de besoins indépendants. Le premier concerne l'évitement ou la réduction d'états négatifs pouvant être engendrés par des événements aversifs. Le second concerne la promotion d'états émotionnels positifs tels que la satisfaction et l'épanouissement personnel par l'intermédiaire d'expériences de vie gratifiantes et enrichissantes. Ces deux ensembles fonctionneraient séparément, ce qui signifie que la satisfaction d'un besoin de l'un des ensembles n'entraîne pas la satisfaction d'un besoin de l'autre ensemble. On ne peut s'empêcher de rapprocher cette hypothèse de la théorie psycho-physiologique de Gray (e.g. Gray, 1981) qui distingue deux systèmes neurologiques motivationnels qui seraient responsables de différences individuelles

[33] : Les items utilisés par Diener et ses collaborateurs (e.g. Diener & Emmons, 1985) sont, pour l'affectivité positive, "happy", "joyful", "pleased", "enjoyment/fun" et, pour l'affectivité négative, "depressed/blue", "unhappy", "frustrated", "angry/hostile" et "worried/anxious".

stables. Ces deux systèmes, qui sont indépendants, ont pour fonctions respectives de réguler le comportement en présence de signaux de récompense (système d'activation comportementale) et de punition (système d'inhibition comportementale)^[34].

Le cadre de référence le plus général sur cette question est, semble-t-il, celui que propose Taylor (1991). Dans un article de synthèse très complet, l'auteur passe en revue les travaux, menés dans le cadre de traditions de recherche différentes, ayant montré les effets asymétriques des événements positifs et négatifs. Taylor propose une hypothèse de mobilisation-minimisation afin de rendre compte de ces effets. Selon cette hypothèse, il y aurait mobilisation de l'organisme de manière à minimiser ensuite son impact. Plus spécifiquement, l'auteur montre que les événements négatifs semblent engendrer plus de réponses physiologiques, affectives, cognitives et comportementales que les événements positifs.

La phase de mobilisation de l'organisme s'exprime dans les domaines suivants :

- activité physiologique : les événements négatifs évoquent plus d'activités physiologiques que les événements positifs (mais les travaux sur cette question sont rares, notamment parce qu'ils posent la question de l'équivalence affective des événements positifs et négatifs) ;
- affectivité : les événements négatifs paraissent être de meilleurs déterminants de l'humeur que les événements positifs ;
- attention : les événements négatifs conduisent les individus à restreindre et à focaliser leur attention, particulièrement sur les caractéristiques qui déterminent l'état négatif ;
- pondération de l'information dans les jugements : de nombreuses études ont montré que les aspects négatifs d'un objet, d'un événement sont pondérés plus fortement que les aspects positifs dans les jugements^[35] ;
- attribution : les événements négatifs entraînent plus d'activités d'attribution que les événements positifs (en particulier, plus de causes potentielles des événements sont passées en revue et ils entraînent des attributions plus extrêmes) ;
- activité cognitive : un nombre considérable de recherches dans le domaine de l'humeur a montré que les humeurs positive et négative avaient des effets différents sur l'activité cognitive (par exemple, l'humeur positive est associée à l'utilisation de stratégies de traitement de l'information plus rapides et moins coûteuses) ;

[34] : Pour une application de cette théorie à la relation entre personnalité et émotion, cf. Larsen & Ketelaar (1991).

[35] : Des travaux ont cependant montré le contraire, notamment dans l'évaluation des aptitudes d'autrui. Skowronski & Carlston (1987) expliquent ce résultat par le fait que dans le domaine des aptitudes ce sont les informations positives qui sont pertinentes alors que dans beaucoup d'autres domaines (personnalité, attitudes...), ce sont les informations négatives qui sont les plus pertinentes.

- activité sociale : il y a peu d'études sur la question, mais il semble que les individus ont plus tendance à se tourner vers autrui, à chercher leur soutien ou leur compagnie après un événement négatif.

On trouve, au niveau de la phase de minimisation, des données similaires à celles qui viennent d'être brièvement décrites et qui montrent la résistance de l'organisme aux événements négatifs. On citera notamment les analyses qui établissent que les événements négatifs sont moins accessibles en mémoire, qu'il est plus difficile d'induire une humeur négative qu'une humeur positive ou encore que les informations ambiguës ont tendance à être interprétées positivement.

La difficulté de trouver des explications générales à ces phénomènes tient au fait qu'ils concernent des classes de réponses différentes : physiologiques, affectives, cognitives et sociales. On peut pourtant espérer que celle-ci sera surmontée un jour. Pour en revenir plus spécifiquement à notre propos, nous pensons que de tels travaux établissent de manière quasi irréfutable que l'affectivité positive et l'affectivité négative ne peuvent plus être considérées comme des pôles opposés d'une même dimension. Les travaux synthétisés par Taylor en apportent des preuves qui complètent de manière très élégante les études corrélationnelles que nous avons passées en revue.

3.3.5 Conclusion

Il ressort de manière assez nette de cette discussion que les aspects positifs et négatifs du bien-être doivent être considérés séparément. Nombreux sont les résultats empiriques qui montrent que la corrélation entre affectivité positive et affectivité négative est faible et que ces deux dimensions n'ont pas les mêmes réseaux nomologiques. Plusieurs auteurs ont proposé des explications visant à sous-estimer cette indépendance. Il ne semble guère faire de doute que certains résultats soient le fait d'artefacts d'ordre méthodologique. Cependant, ces artefacts ne parviennent pas à expliquer pourquoi certaines variables sont corrélées avec PA et non avec NA et inversement. Il en est de même d'ailleurs de la proposition théorique de Diener (cf. Diener et al., 1985 b) faisant intervenir dans les réponses aux items de bien-être deux dimensions indépendantes, la fréquence et l'intensité. Nous avons vu que cette hypothèse était peu plausible, en raison de difficultés tant théoriques que méthodologiques, et qu'elle était impropre à rendre compte des différences de réseaux nomologiques.

Les travaux menés à propos de la structure de l'affectivité permettent de mieux comprendre l'indépendance entre PA et NA. La proposition d'un circumplex de l'affectivité rend compte des relations complexes existant entre les adjectifs choisis pour représenter ce domaine. Selon le type de marqueurs utilisés, les corrélations entre items

positifs ou négatifs pourront être nulles, faiblement négatives ou fortement négatives. Cette structure fournit en quelque sorte un modèle de mesure permettant de justifier le choix de tel ou tel item. Il paraît évident, aujourd'hui, que l'indépendance entre PA et NA ne devrait plus être étudiée sans référence à ce circumplex. Pourtant, force est de constater que la littérature sur le bien-être n'a pas encore totalement intégré ces travaux sur les émotions. Trop d'auteurs les ignorent encore, ce qui perpétue des querelles stériles quant à l'existence de l'indépendance. De plus, comme certaines habitudes sont bien ancrées, ce sont presque toujours les mêmes instruments qui sont utilisés. Or, ceux-ci n'ont pas, en général, une structure très claire parce que les items retenus ne sont pas représentatifs du domaine affectif : c'est le cas notamment lorsque le nombre d'octants représentés n'est pas le même pour les items positifs et pour les items négatifs. De plus, ces échelles contiennent souvent des items qui ne sont pas forcément des marqueurs purs des octants du circumplex. On pourrait évidemment effectuer des analyses secondaires des études publiées en tenant compte de ces différences dans la composition des instruments, mais la tâche est considérable et, surtout, on ne dispose que rarement des informations suffisantes pour mener ces analyses à bien. Il semble donc beaucoup plus réaliste d'abandonner dans l'avenir les échelles qui ne seraient pas construites sur la base du circumplex de l'affectivité. Les échelles proposées par Watson et al. (1988) constituent un pas allant dans ce sens.

Dans l'histoire du circumplex de l'affectivité, le débat sur ses dimensions de base occupe une place centrale. Il importe de reconnaître que, dans une telle structure, n'importe quel système d'axes orthogonaux est aussi efficace. Certains préféreront les dimensions d'affectivité négative et d'affectivité positive parce qu'elles semblent se superposer aux dimensions d'extraversion et de névrosisme de la personnalité. D'autres préféreront les dimensions plaisant-déplaisant et activation parce qu'elles correspondent approximativement aux dimensions de l'espace sémantique d'Osgood et al. (1957) ... ou pour de toutes autres raisons. Ces choix sont tout aussi légitimes pourvu que la signification des dimensions soit précisée de façon univoque. En particulier, il ne faut pas oublier que le pôle positif de l'affectivité positive inclut une connotation plaisante accompagnée d'une forte implication émotionnelle^[36]. Il est tout aussi important de remarquer que les termes "heureux" ou "satisfait", qui sont souvent utilisés dans les échelles de bien-être, ne sont pas des marqueurs purs de la dimension d'affectivité positive. Ils sont saturés sur la dimension plaisant-déplaisant et indiquent à la fois une affectivité positive élevée et une affectivité négative faible.

De manière générale, les rapports entre les aspects positifs et négatifs constituent un vieux serpent de mer de la psychologie. On raconte déjà des débats à ce sujet à propos des dimensions de l'espace sémantique d'Osgood et on les retrouve dans d'autres

[36] : Tout comme chez Eysenck, la dimension d'extraversion-introversion se définit à partir de la sociabilité et de l'impulsivité.

domaines (cf., par exemple, l'article de Andrich & Van Schoubroeck, 1989, à propos de la dimensionnalité du questionnaire de santé de Goldberg). On peut considérer que le débat sur la symétrie positif-négatif est maintenant dépassé et qu'il faut proposer des cadres d'explication plus généraux que ceux que l'on connaît. L'asymétrie ne se réduit pas à la présence d'artefacts dus à l'utilisation de questionnaires. Il semble donc plus intéressant de rechercher des explications théoriques qui rendent compte de ce phénomène. Les travaux de Claves & Timmers (e.g. 1993) sur l'asymétrie entre traits positifs et négatifs ou ceux de Taylor (1991) que nous avons mentionnés, constituent, à notre avis, de réelles avancées dans ce sens.

3.4 Le couple affectif-cognitif dans l'étude du bien-être subjectif

Paradoxalement, alors que la distinction positif-négatif a suscité de nombreux débats accompagnés d'une pléthore d'études empiriques, la distinction entre les aspects affectifs et cognitifs du bien-être semble acceptée de manière quasi consensuelle, alors que les éléments empiriques en sa faveur sont plutôt rares et ambigus.

La division tri-partite de l'esprit humain en composantes affective, cognitive et comportementale est bien antérieure à l'avènement de la psychologie scientifique. Hilgard (1980) en présente un historique très détaillé. C'est sans doute dans l'étude des attitudes que ces distinctions ont été les plus utilisées. En suivant Ajzen (1989), on peut les décrire brièvement à partir du type de réponses habituellement utilisé pour les étudier.

	Composante		
	cognitive	affective	conative (comportementale)
mode de réponse			
verbal	expressions de croyances envers l'objet d'attitude	expressions de sentiments envers l'objet	expressions d'intentions comportementales
non verbal	réactions perceptives envers l'objet d'attitude	réactions physiologiques envers l'objet	comportements observés envers l'objet

Figure 2.18 : Les réponses utilisées pour la mesure des différentes composantes de l'attitude (d'après Ajzen, 1989, p. 242)

Malgré l'importance que revêt cette distinction dans les théories sur les attitudes, on est surpris de constater la rareté des travaux empiriques (cf. Breckler & Wiggins, 1989a) et leurs résultats peu concluants. Les corrélations calculées entre les composantes affective et cognitive sont souvent supérieures à .80 (e.g. Bagozzi & Burnkrant, 1985)^[37], ce qui suscite des doutes quant à l'intérêt de cette distinction. Au niveau de l'opérationnalisation de ces composantes, on peut aussi remarquer une certaine confusion. Ajzen (1989, p. 246) note, en effet, que certains auteurs ont utilisé la dimension évaluative du différenciateur sémantique comme une mesure de la composante affective des attitudes alors que d'autres l'ont utilisée comme une mesure de la composante cognitive. En revanche, les discussions théoriques ont presque toujours fait apparaître la nécessité de traiter séparément ces deux composantes. Il serait trop long ici de passer en revue l'ensemble de ces arguments, aussi nous contenterons-nous de citer ceux de

[37] : Il s'agit de corrélations entre variables latentes, c'est-à-dire corrigées pour l'atténuation due à la fidélité.

Greenwald (1968, cité par Bagozzi, 1978). Selon lui, les composantes des attitudes ont des antécédents différents : la composante affective résulterait de l'association de l'objet d'attitude avec un stimulus affectif par l'intermédiaire de mécanismes classiques du conditionnement ; la composante cognitive résulterait de l'association entre un objet d'attitude et du matériel éducatif par l'intermédiaire de processus de traitement de l'information et de processus d'apprentissage.

Ce rapide détour par l'étude des attitudes n'est pas gratuit. On se rappelle on effet (cf. section 1.3) que le bien-être est parfois assimilé à une attitude. Il n'est donc pas étonnant qu'une distinction entre composantes affective et cognitive ait été proposée dans le cas du bien-être. Mais, pas plus que dans le cadre plus général des attitudes, on ne trouve dans la littérature sur le bien-être d'éléments visant à préciser clairement la signification de ces deux composantes.

Lorsque les chercheurs qui travaillent dans le domaine du bien-être font référence à des aspects cognitifs et affectifs, ils se contentent le plus souvent de distinguer entre des items formulés en termes de satisfaction et des items portant sur des expériences plaisantes ou déplaisantes de la vie quotidienne (cf. Campbell, 1981 ; Veenhoven, 1984 ; Chamberlain, 1988). Andrews & McKennell (1980) considèrent en effet que les items de satisfaction sont plutôt cognitifs parce qu'ils incluent, explicitement ou non, une notion de comparaison (avec son passé, avec les autres...)^[38]. Ainsi, dans un cas, les évaluations sont basées sur des indices affectifs, émotionnels, dans l'autre, elles sont basées sur des indices rationnels, comparatifs.

Ces définitions ne sont guère convaincantes. Elles reposent sur des postulats quant aux processus utilisés par les sujets pour répondre aux items des questionnaires de bien-être mais ces postulats n'ont jamais été vérifiés. De plus, toute réponse à un item entraîne nécessairement une activité cognitive de la part du sujet qui y répond : celle-ci serait-elle moins importante dans un cas que dans l'autre ? Ce serait, dans cette perspective, une distinction basée sur une différence de degré, mais il est peu probable que les auteurs l'envisagent effectivement ainsi.

On pourrait aborder le problème sous un autre angle en essayant d'examiner ce qui fait la spécificité des items selon qu'ils sont habituellement considérés comme des items affectifs ou comme des indicateurs de la composante cognitive. Très schématiquement, on peut remarquer que les items affectifs utilisent principalement l'adjectif "heureux", alors que les items cognitifs emploient souvent l'adjectif "satisfait". On constatera tout d'abord que ces deux adjectifs, qui semblent renvoyer à des réalités différentes dans le cas

[38] : Il est vrai que nombre d'items de bien-être font référence explicitement à une comparaison (avec les personnes de son âge, avec soi-même il y a cinq ans, etc.), mais rien n'indique que les items qui ne sont pas formulés de cette façon entraînent chez le sujet une activité de comparaison.

du bien-être, sont des quasi synonymes dans les travaux sur l'émotion, puisqu'ils appartiennent au même octant du circumplex (cf. figure 2.16 ci-dessus). De plus, nous avons vu que dans le cadre de l'approche en facettes (Levy & Guttman, 1975, cf. section 3.2.4, première partie), les items de type "je suis satisfait avec..." renvoient à la composante affective !

Comme il est difficile de préciser la distinction affectif-cognitif au niveau conceptuel, on peut essayer d'analyser les arguments empiriques qui ont été avancés en sa faveur. Nous l'avons dit, les études empiriques sont rares. On peut tout de même dégager trois grands types d'études qui abordent cette question avec des approches différentes.

Dans les études structurales exploratoires utilisant l'analyse factorielle, on analyse un ensemble d'items de bien-être hétérogène, positifs et négatifs. Nous n'avons trouvé, en tout et pour tout, que deux études qui, par cette technique, mettent en évidence l'existence d'un facteur cognitif. L'une, celle de Andrews & Withey (1976), a déjà été présentée au début de ce chapitre, l'autre est due à Beiser (1974). Bien qu'antérieure à la précédente, elle n'a pas connu la même renommée. Dans les deux cas, les auteurs mettent en évidence trois facteurs. Deux facteurs sont interprétés facilement parce qu'ils saturent respectivement les items positifs et négatifs de l'ABS de Bradburn. Le troisième est interprété comme un facteur cognitif par Andrews & Withey (nous avons vu les réserves des auteurs à propos de l'appellation de ce facteur), comme un facteur de satisfaction à long terme par Beiser, mais cette différence d'appellation est mineure.

D'autres travaux ont utilisé l'analyse factorielle contrainte (McKennell, 1978 ; McKennell & Andrews, 1980 ; Andrews & McKennell, 1980 ; cf. aussi Horley & Little, 1985, qui ont répliqué les résultats précédents). En général, les analyses portent sur les items de l'ABS, considérés comme des items affectifs "purs", sur des items de satisfaction dans différents domaines de vie et sur des items d'évaluation générale de la vie. Les auteurs montrent qu'après avoir partialisé la variance due aux facteurs affectifs (PA et NA), une part importante de la variance totale de certains items reste inexplicée. Un facteur supplémentaire est donc introduit dans le modèle afin d'en rendre compte. C'est ce facteur qui est assimilé à la dimension cognitive. La démarche suivie par les auteurs illustre bien le fait que la composante cognitive n'est pas définie explicitement, mais qu'elle correspond à "ce qui reste après avoir enlevé la variance due au facteur affectif". Il ne fait aucun doute que le modèle qui inclut ce facteur supplémentaire est plus adéquat que les modèles précédents. Cependant, les auteurs remarquent que les items autres que ceux de l'ABS sont saturés à la fois sur les facteurs affectif et cognitif, ce qui ne contribue pas à préciser la signification de ce nouveau facteur.

Enfin, certaines études semblent montrer la validité divergente des composantes affective et cognitive. Campbell et al. (1976), par exemple, établissent que l'âge est lié positivement à des mesures cognitives du bien-être mais négativement à des mesures affectives. Headey et al. (1984 a) montrent que les événements de vie vécus sont plus liés aux mesures affectives qu'aux mesures cognitives. C'est un résultat convergent qui est obtenu par De Haes et al. (1987), selon lesquels, chez les femmes atteintes d'un cancer du sein, le niveau de satisfaction (i.e. composante cognitive) reste stable alors que le niveau d'affectivité diminue. Enfin, pour Moorman (1993), la relation entre bien-être et performance dans le domaine du travail est plus importante pour la composante cognitive que pour la composante affective.

Mais d'autres études n'ont pas réussi à retrouver ces deux composantes, y compris lorsque les analyses s'inscrivaient dans une démarche confirmatoire. C'est le cas notamment de l'étude de Levy & Guttman (1975), dans laquelle la facette distinguant les items affectifs et cognitifs n'est pas validée^[39]. C'est le cas aussi d'une étude de Andrews lui-même (Abbey & Andrews, 1985), qui est pourtant un de ceux qui a le plus tenté de formaliser cette distinction. Enfin, rappelons que dans l'étude de Thomas & Stock (1988) les évaluations de concepts liés au bonheur en termes affectif et cognitif corrélaient à .98.

Si on devait tirer une conclusion de ces travaux, on pourrait dire qu'il semble exister des sources de variation qui ne sont pas prises en compte par les facteurs d'affectivité positive et d'affectivité négative. C'est ce que montrent notamment les résultats de certaines analyses factorielles exploratoires ou contraintes. Mais la signification de cette source de variation supplémentaire n'est pas évidente. Lorsque Andrews et ses collègues proposent de l'assimiler à un facteur cognitif parce que le bien-être peut être conceptualisé comme une attitude, ils incitent toutefois à la prudence des interprétations. Ces conseils de prudence ne seront pas toujours suivis. Il est vrai que dans un domaine comme celui du bien-être, où règne un certain flou conceptuel, il est rassurant de pouvoir se rattacher à un modèle définitoire solide. Mais cette solidité n'est qu'apparente au vu des résultats empiriques. C'est le cas du bien-être ; c'est aussi celui des attitudes^[40]. De plus, force est de constater que les rares études qui se proposaient de retrouver ces distinctions à partir d'items construits pour la circonstance se sont soldées par un échec (Levy & Guttman, 1975 ; Abbey & Andrews, 1985).

[39] : De manière générale, les applications de la théorie en facettes ayant utilisé la distinction affectif-cognitif se sont souvent soldées par un échec (cf., par exemple, Shye, 1978 b).

[40] : On sait que la littérature sur les attitudes est considérable. Dawes & Smith (1985) ont recensé 20209 références entre 1970 et 1979. Parmi toutes ces références, il y avait, avant l'étude de Breckler & Wiggins (1989 a) seulement cinq études empiriques s'attachant à montrer la validité divergente des composantes affective et cognitive : Mann, 1959 ; Woodmansee & Cook, 1967 ; Kothandapani, 1971 ; Breckler, 1984 (il y a quelques études plus récentes, mais elles se contentent de proposer des analyses secondaires en utilisant l'analyse factorielle contrainte).

En ce qui concerne les items utilisés pour opérationnaliser ces deux composantes, on a parfois l'impression que la distinction se limite à opposer ceux qui interrogent les individus par rapport à eux-mêmes à ceux qui les interrogent sur des aspects particuliers de leur vie. Dans le premier cas, les sujets doivent, à la manière des questionnaires de personnalité, s'auto-attribuer des traits ("je suis heureux" ; "je suis déprimé", etc.) ; dans le second, ils sont invités à évaluer leur travail, leur famille, leur santé ("je suis satisfait par rapport à mon travail", etc.). Mais cette distinction n'est pas aussi simple. Si on en croit Headey et al. (1984 a), les items de satisfaction seront plutôt cognitifs ou plutôt affectifs selon les domaines. La satisfaction avec son niveau de vie serait plus cognitive alors que la satisfaction avec son mariage, sa vie sexuelle ou sa santé serait plus affective. Il n'est donc pas exclu de penser que les items de satisfaction sont cognitifs lorsque les sujets peuvent comparer explicitement leur situation avec celle des autres. Pour des domaines plus intimes, en revanche, les comparaisons seraient plus difficiles et les sujets répondraient en fonction d'indices plus affectifs. Cette hypothèse est plausible, mais elle demande à être étayée par des informations sur la façon dont les sujets répondent aux items de bien-être. Malheureusement de telles informations n'existent pas. Avant d'en savoir davantage sur les processus de réponse, une autre explication, qui reste à un niveau plus descriptif, peut être proposée. La distinction habituellement opérée entre items cognitifs et items affectifs reflète surtout une différence de degré qui exprime la distance psychologique par rapport à l'environnement interne d'un individu. Nous ne prétendons pas que l'affectif, c'est ce qui est interne à l'individu et que le cognitif, c'est ce qui lui est externe ! Nous essayons simplement de rendre compte de certaines régularités constatées dans l'opérationnalisation des items de bien-être et suggérons que la distinction entre composantes cognitive et affective pourrait être abandonnée au profit d'une dimension exprimant la distance psychologique^[41].

Cette idée n'est pas seulement intuitive. Les résultats de Levy & Guttman (1975) font clairement apparaître une facette dont le rôle modulaire indique l'éloignement des items par rapport à l'environnement primaire de l'individu. Dickes (1989 a) met en évidence une satisfaction dans l'environnement primaire et une satisfaction dans l'environnement secondaire. On se rappelle peut-être aussi que Lawton (1983 a, 1983 b) distingue entre un bien-être interne et un bien-être externe. Cette distinction pourrait recouper une distinction entre des items ayant une focalisation interne et des items centrés sur des domaines de vie "externes". Enfin, toujours à l'appui de cette idée, McKennell (1978) fait l'hypothèse que la composante affective du bien-être serait plus sensible aux événements de vie appartenant à la sphère proche de l'individu, alors que la composante cognitive le serait aux événements de l'environnement social ou politique (sphère publique). Il n'est pas certain que la distance psychologique soit cette source de variation

[41] : Du moins dans l'état actuel des connaissances.

supplémentaire. Mais cette proposition a le mérite d'avoir une signification beaucoup plus claire que la distinction entre affectif et cognitif.

3.5 Conclusion

Dans ce chapitre, nous avons tenté de faire le point sur les travaux s'intéressant à la structure du bien-être. Parmi ceux-ci, ont d'abord été dégagés quatre grands types d'approches de cette question. Les analyses exploratoires, comme celles de Andrews & Withey (1976), portent sur de grands ensembles d'items et ont permis de dégager un modèle conceptuel du bien-être à trois composantes : l'affectivité positive, l'affectivité négative et la satisfaction (cognitive). D'autres études se sont concentrées sur les analyses structurales d'instruments classiques du bien-être, en particulier ceux qui ont été élaborés pour les personnes âgées. Malgré l'utilisation d'analyses factorielles contraintes, qui présentent plusieurs avantages par rapport aux analyses factorielles menées dans une perspective exploratoire, ces études n'échappent pas aux critiques émises depuis longtemps à l'encontre des analyses factorielles, à savoir qu'elles ne peuvent révéler que des dimensions prises en compte dans l'élaboration des items. Ainsi, il n'est pas étonnant que leurs résultats fassent apparaître des facteurs très spécifiques, dont l'interprétation largement idiosyncrasique ne favorise pas l'émergence d'une synthèse constructive. Dans une autre approche, plus originale, Thomas & Stock (1988) ont eu recours aux représentations. Malheureusement, leur étude est quelque peu décevante et seule une opposition positif-négatif, triviale avec cette méthode, apparaît de manière nette. Enfin, Levy & Guttman (1975) ont proposé une application de l'approche en facettes au bien-être subjectif. Même si les facettes retenues sont peu fondées sur le plan théorique, il ressort de cette étude un effet assez net des facettes concernant les domaines de vie et le type d'environnement qui se présente comme une facette ordonnée exprimant la distance psychologique.

De l'ensemble de cette littérature, les spécialistes du bien-être (e.g. Diener, 1984 ; Chamberlain, 1988) ne retiennent que trois grandes dimensions : l'affectivité positive, l'affectivité négative et la satisfaction cognitive.

Nous avons longuement discuté du couple "positif-négatif" dans le cas du bien-être. L'accumulation de données empiriques ne laisse guère de doute sur le fait que l'affectivité positive et l'affectivité négative ne constituent pas les pôles opposés d'une même dimension. Mais la question de l'indépendance partielle ou totale ne prend sens qu'en référence à la structure de l'affectivité étudiée dans le cadre des études sur les émotions. Cette structure se présente sous la forme d'un circumplex dont les dimensions de base sont, selon les auteurs, l'affectivité positive et l'affectivité négative ou le caractère

plaisant-déplaisant et l'activation (ou engagement). Cette clarification des relations entre les états affectifs habituellement utilisés pour mesurer le bien-être est encore insuffisamment reconnue, ce qui fait que subsistent des débats, qui devraient être dépassés, à propos de l'indépendance entre affectivité positive et affectivité négative. Malheureusement, il n'y a aucune raison pour que ces débats cessent dans un avenir proche étant donné que l'on continue à utiliser des instruments dont les items échantillonnent mal les octants du circumplex. Au-delà de ces considérations structurales, l'asymétrie entre les aspects positifs et négatifs donne lieu à des investigations théoriques plus fondamentales, dont l'importance dépasse largement le cadre du bien-être.

Autant la distinction des composantes positives et négatives du bien-être a donné lieu à une littérature pléthorique, remplie de prises de position souvent virulentes, autant celle des composantes affectives et cognitives est acceptée de manière quasi consensuelle alors qu'elle n'est pas assurée, ni empiriquement ni conceptuellement. Ce n'est là qu'un paradoxe de plus des études sur le bien-être, mais celui-ci n'est pas étonnant au vu de la littérature sur les attitudes dans laquelle cette distinction n'est pas mieux établie. Nous avons proposé de remplacer cette distinction, au moins provisoirement et en attendant des études complémentaires, par une dimension exprimant la distance psychologique.

Globalement, les travaux sur la structure du bien-être laissent de nombreuses questions en suspens. Outre la distinction entre composantes affectives et cognitives, on peut se demander, si celle-ci est avérée, comment ce couple de dimensions se combine avec le précédent. Autrement dit, peut-on envisager une distinction entre les aspects positif et négatif pour la composante affective et pour la composante cognitive ? C'est un fait que, dans la littérature, les items censés mesurer la satisfaction (i.e. la composante cognitive) sont presque toujours des items positifs spécifiant des domaines de vie particuliers (à l'inverse des items censés mesurer la composante affective). De rares auteurs le regrettent (e.g. Organ & Near, 1985). Certains ont tout de même essayé de mesurer les aspects positifs et négatifs de la satisfaction dans les domaines (Campbell et al., 1976 ; Carp & Carp, 1981 ; Emmons & Diener, 1985 a), mais les résultats sont peu concluants : en ce qui concerne l'indépendance, par exemple, les conclusions varient selon les auteurs^[42]. D'autres ont proposé une typologie des individus qui se base sur une telle combinaison (cf. Campbell et al., 1976 ; McKennell, 1978 ; Michalos, 1980) ; malheureusement, cette typologie adopte implicitement l'idée que les dimensions affective et cognitive sont des continus qui opposent un pôle positif et un pôle négatif, ce qui n'est guère satisfaisant au vu de ce que l'on sait. Il n'est pas exclu que ce soit à nouveau les travaux sur les émotions qui parviennent à éclairer cette question. Par exemple, l'étude de la typologie du lexique affectif de Clore et al. (1987) distingue, au niveau le plus général, les conditions internes et externes. Les conditions internes, qui seules nous

[42] : Mais à nouveau, il faudrait mener de telles études en faisant explicitement référence au circumplex de l'affectivité, ce qui n'a pas été fait.

intéressent ici, se divisent à leur tour en conditions mentales (états physiques et corporels) et non-mentales. Enfin, ces dernières peuvent être distinguées selon qu'elles correspondent à l'une ou l'autre des composantes affective, cognitive ou comportementale. L'originalité de cette approche réside dans la possibilité de combiner ces composantes. Ainsi, on va trouver des états affectifs (e.g. heureux, découragé, désir ardent), des conditions affectivo-comportementales (e.g. de bonne humeur, maussade, mélancolique), des conditions affectivo-cognitives (e.g. encouragé, rancunier, désespéré, inquiet), des conditions cognitives (e.g. certain, désorienté, surpris) et des conditions cognitivo-comportementales (e.g. attentif, avide, vertueux). Pour l'heure, nous ne pouvons que souligner le caractère incontestablement plus fin de cette approche, sans préjuger de sa pertinence à venir.

Enfin, on peut se demander, en admettant qu'un modèle du bien-être à trois composantes soit acceptable, comment ces composantes se situent les unes par rapport aux autres. Implicitement, il semble que soit admise l'idée selon laquelle l'affectivité positive, l'affectivité négative et la satisfaction se situent à un même niveau, au-dessus duquel se trouve une dimension générale de bien-être subjectif censée rendre compte de la variance commune entre ces trois composantes. En fait, comme cet aspect a été peu étudié, on ne sait pas du tout si ce modèle est approprié. On pourrait tout aussi bien imaginer un modèle dans lequel l'affectivité, par le biais de son influence sur les processus cognitifs, influence la composante cognitive. Des travaux comme ceux de Byrne & Eysenck (1993), montrant des différences de processus cognitifs selon le niveau d'affectivité, pourraient corroborer une telle hypothèse qui reste bien évidemment hautement spéculative.

Quoiqu'il en soit, si l'étude de la structure du bien-être nécessite des investigations complémentaires, il n'en reste pas moins vrai, comme le souligne Emmons (1986), que les modèles explicatifs oublient trop souvent que le bien-être n'est pas un concept unidimensionnel. Le peu que l'on sait sur sa structure mériterait d'être pris en compte de manière systématique dans les élaborations théoriques ainsi que dans l'étude des corrélats du bien-être. Malheureusement, nous allons voir dans les chapitres qui suivent que cela reste l'exception.

CHAPITRE IV

LE RESEAU NOMOLOGIQUE OU LES CORRELATS DU BIEN-ETRE

4.1 Introduction

Dans la tradition psychométrique, la construction du réseau nomologique d'un concept permet d'en préciser la signification. Savoir qu'un concept est corrélé avec telle variable mais ne l'est pas avec telle autre apporte des informations utiles qui lui confèrent une plus grande légitimité dans la construction scientifique et assurent une certaine validité à sa mesure, pourvu que ces informations s'intègrent dans une représentation théorique cohérente.

Or, dans le cas du bien-être, il n'est pas certain que l'on puisse mettre en évidence un réseau théorique dont la cohérence soit à même de rendre compte de toutes les relations qui interviennent. Cela tient sans doute en partie au nombre considérable de variables qui ont été étudiées en relation avec le bien-être. Veenhoven (1984) en recense 3500 dans son ouvrage sur la question, qui reste sans doute encore, à l'heure actuelle, la synthèse la plus complète des corrélats du bien-être.

L'objectif de ce chapitre n'est pas d'analyser en détail chacune des relations faisant intervenir le bien-être : il faudrait pour cela plusieurs volumes ! Il s'agit plutôt de proposer une introduction au chapitre suivant, consacré aux théories du bien-être, introduction nécessaire pour plusieurs raisons. Une vue d'ensemble sur les corrélats et les non-corrélats du bien-être fera apparaître toute la difficulté d'intégrer, dans un même schéma, des relations nombreuses, complexes et quelquefois contradictoires ; c'est pourtant l'enjeu d'une théorie du bien-être que d'en rendre compte. De plus, plusieurs théories s'inscrivent dans le prolongement direct de résultats qui ont profondément marqué l'histoire des recherches concernant ce concept. L'exemple le plus spectaculaire est sans doute celui de l'indépendance partielle entre le bien-être perçu par les individus et des indicateurs objectifs de conditions de vie. Outre les polémiques et les explications techniques suscitées par ce résultat, celui-ci est à la base du développement de modèles, le plus souvent très élaborés, dont nous verrons plusieurs exemples dans le chapitre suivant. Enfin, au-delà de ces considérations historiques, la recension des corrélats du bien-être permettra d'en mieux comprendre la signification et la dynamique.

Mais l'étude des corrélats du bien-être soulève plusieurs problèmes qui peuvent être abordés à un niveau général. C'est par un rapide tour d'horizon de ces questions que

nous commencerons ce chapitre, avant de présenter le panorama le plus large possible des corrélats, en limitant la discussion à certaines relations particulièrement significatives.

4.2 Problèmes théorico-méthodologiques dans l'étude des corrélats du bien-être

Il existe bien peu de variables étudiées de manière traditionnelle en psychologie qui n'ont été mises en relation, sous une forme ou sous une autre, avec le bien-être. Les corrélations calculées étant souvent faibles, ce que l'on sait au moins depuis Watson (1930), il devenait tentant d'essayer de trouver de nouvelles variables, dérivées de considérations théoriques ou d'approches totalement empiriques, qui fassent apparaître des relations plus substantielles. Malgré la diversité de ces variables et des approches qui sous-tendent leur utilisation, on peut reconnaître des aspects communs dans les démarches. Ils illustrent des pratiques et mettent aussi en évidence, dans certains cas, des conceptions théoriques implicites, justifiant que l'on s'y arrête.

On notera tout d'abord qu'il y a eu énormément d'études concernant les relations entre le bien-être et ce que l'on peut appeler les indicateurs objectifs de la qualité de la vie, c'est-à-dire des indicateurs non soumis à une évaluation de la part du sujet. Ce point évoque bien évidemment les débats relatifs aux conceptions objectives et subjectives du bien-être dont il a été question plus haut et fait référence à toute une tradition de recherches socio-économiques sur les indicateurs de la qualité de la vie. Pour Campbell (1976), le flot d'indicateurs statistiques monétaires a longtemps été considéré comme suffisant pour caractériser la qualité de la vie. Toutefois, l'évolution rapide de ces indicateurs au cours des 40 dernières années (augmentation du revenu des ménages, régression de la pauvreté, etc.) ne s'est pas accompagnée de l'évolution d'autres indicateurs de la qualité de la vie concernant, par exemple, la criminalité ou la consommation de stupéfiants. C'est pourquoi les chercheurs ont été amenés à proposer comme alternative des indicateurs non monétaires, qui se sont développés dans un courant connu sous le nom de mouvement des indicateurs sociaux. Ces indicateurs sociaux, même non monétaires, restent pourtant toujours des indicateurs objectifs ne prenant pas en compte l'évaluation subjective que font les individus de leur vie. Ainsi, il n'est pas étonnant que nombre de chercheurs se soient demandés si les *indicateurs objectifs et subjectifs* du bien-être étaient bien des mesures identiques d'un même phénomène. Nous verrons que cette question, largement débattue, a reçu des réponses empiriques qui laissent planer un doute sérieux sur cette identité.

Les relations entre le bien-être et les autres variables ont été étudiées, dans l'immense majorité des cas, à partir de données transversales, analysées par des méthodes corrélationnelles. Or, tous les étudiants savent qu'une corrélation entre deux variables ne signifie nullement que l'une influence l'autre ni, *a fortiori*, que l'une est la cause de l'autre.

Pourtant, dans le domaine du bien-être, les chercheurs semblent avoir oublié cette règle méthodologique de base ou, tout du moins, avoir considéré qu'elle ne s'appliquait pas à leur objet d'étude. Il suffit pour s'en convaincre de lire deux célèbres revues de la littérature sur le bien-être : Diener (1984) intitule son chapitre sur les corrélats "les influences sur le bien-être subjectif" ; McNeil et al. (1986a) intitulent le leur "les prédicteurs du bien-être subjectif". Ce serait bien évidemment une méprise que de considérer que ces intitulés reflètent une erreur méthodologique aussi grossière. Ils reflètent plutôt une conception implicite du bien-être très répandue. En effet, jusqu'à aujourd'hui, la motivation principale des recherches sur le bien-être a été sa prédiction. Comprendre les mécanismes qui font qu'une personne est heureuse représente un enjeu fondamental, notamment dans le cadre des recherches socio-économiques. Ce faisant, on a trop souvent et trop rapidement assimilé la description des caractéristiques des personnes présentant un niveau élevé de bien-être avec l'explication ou la prédiction de ce niveau sans autre regard critique. Il serait souhaitable qu'un historien des sciences se penche un jour sur cette question pour comprendre pourquoi tant de chercheurs ont adopté la position implicite visant à faire du bien-être une variable endogène ultime. Nous n'avons pas cette ambition ici. Tout au plus pouvons-nous remarquer que le fait de penser que le bien-être dépende d'un certain nombre de variables pouvant être manipulées a des côtés rassurants. De plus, cela correspond à certaines conceptions populaires^[1] et à une appréhension intuitive qui peut sembler, à l'occasion, raisonnable. Par exemple, peu de personnes contesteraient le fait que si les individus mariés sont plus heureux que les autres, c'est parce qu'ils sont mariés, ou que si les individus qui ont connu des événements positifs récents dans leur vie sont plus heureux que ceux qui ont plutôt subi des événements négatifs, c'est parce que ces événements positifs leur sont arrivés (ou parce que les événements négatifs sont arrivés aux autres). Pourtant, même dans des cas *a priori* aussi évidents, la situation n'est pas si simple qu'il y paraît. Ce n'est que récemment que ces questions commencent à recevoir des réponses empiriques sérieuses ; celles-ci sont édifiantes : Masteskaasa (1992), par exemple, montre, grâce à une vaste enquête longitudinale portant sur plus de 9000 individus, que si les personnes mariées sont plus heureuses que les autres, ce n'est pas parce le mariage amène (forcément) le bonheur, mais parce que les personnes heureuses ont une plus grande probabilité de se marier que les autres. C'est aussi grâce à une étude longitudinale que Headey et ses collègues (cf. Headey et al., 1989 ; Headey et Wearing, 1991) ont montré que le bien-être pouvait être, en partie du moins, la "cause" des événements de vie ultérieurs que connaissent les individus^[2]. En plus de ces résultats empiriques, il y a peu de corrélats qui résistent à un examen critique remettant en cause leur statut de prédicteur du bien-être. Même si on est

[1] : Même si d'autres conceptions populaires soutiennent le contraire !

[2] : Nous reviendrons plus en détail sur le travail de Headey et ses collègues, qui est particulièrement important et intéressant. Contentons-nous de noter pour l'instant que plusieurs précautions méthodologiques ont été prises afin d'éviter que ce résultat ne puisse être le fait d'artefacts.

porté à croire que certaines variables statutaires classiques, dont l'antériorité par rapport au bien-être ne semble faire aucun doute, sont à l'abri de telles critiques, il n'en est rien. Que l'on pense, par exemple, au niveau d'éducation ou au revenu, pour lesquels certaines études ont cru déceler une (faible) relation avec le bien-être ; il est tout aussi concevable de considérer que l'éducation ou le revenu influencent le bien-être que de considérer la relation inverse, à savoir que les personnes heureuses, étant mieux adaptées, sont celles qui réussissent le mieux dans leurs études et qui ont les possibilités d'avancement les plus rapides dans leur carrière.

Il n'est pas question ici de défendre l'idée que le bien-être est la cause de tout ! Il est important, toutefois, d'envisager de telles possibilités, ce qui a rarement été fait jusqu'à présent, même si la situation commence à évoluer. Cette évolution réclame un changement total de perspective et d'habitudes, ce qui n'est pas chose aisée. Cela explique peut-être qu'il y ait si peu de recherches dans lesquelles le bien-être joue un rôle explicite de variable indépendante, de prédicteur. Veenhoven, un des rares auteurs sensibilisé très tôt à cet aspect de la recherche sur le bien-être, a tenté, en 1988, de recenser les études menées dans cette perspective. Cette revue d'une littérature peu nombreuse fait surtout apparaître une influence du bien-être sur les contacts sociaux, sur l'implication active, sur la santé, voire sur l'espérance de vie. Toutefois, si on accepte une vision plus large du bien-être, on peut faire référence à une littérature plus importante portant sur les effets de l'humeur et, principalement, de l'affectivité positive. Ces travaux, essentiellement expérimentaux — l'humeur y est induite par différentes manipulations situationnelles —, s'inscrivent dans le cadre de la cognition sociale ou de la psychologie sociale cognitive, ce qui explique peut-être qu'ils sont rarement cités dans la littérature classique sur le bien-être. Dans une revue de synthèse de ces travaux, Isen (1987) recense les principaux effets mis en évidence. Il semble que l'affectivité positive influence la coopérativité dans les négociations, le comportement d'aide, la sociabilité, la gentillesse ou encore la prise de risque. Des effets concernant les processus cognitifs (plus grande intégration du matériel, plus grande flexibilité dans la résolution de problèmes...) ont aussi été relevés, ceux-ci pouvant avoir un effet indirect sur le comportement social (la faculté d'intégrer les informations peut faciliter l'intégration de ses propres idées avec celles des autres, ce qui favorise, par exemple, la perception de soi comme faisant partie d'un groupe ; la plus grande flexibilité associée à l'affectivité positive peut faciliter la résolution des problèmes inter-personnels...). Ces résultats ne sont pas incompatibles — loin s'en faut — avec ceux que produisent les études sur le bien-être. Ils suggèrent d'ailleurs des hypothèses intéressantes quant à sa dynamique. Si le bien-être favorise l'émergence d'états d'affectivité positive et si ces états influencent, notamment, les relations sociales, les bénéfices tirés d'une bonne qualité des contacts sociaux peuvent, en retour, contribuer à maintenir un niveau élevé de bien-être. L'intégration de ces travaux dans une théorie générale reste donc, à terme, un des enjeux les plus importants de la recherche sur le bien-être.

Les autres remarques sur les corrélats du bien-être concernent des aspects plus spécifiques. Nous en discuterons trois : les aspects liés au contrôle des variables dans l'étude de relations bivariées, ceux qui concernent la différenciation des composantes du bien-être et ceux qui touchent aux éventuelles différences de signification du bien-être pour des groupes différents.

Lorsqu'une relation entre deux variables est mise en évidence, on sait qu'elle peut être due à l'influence d'une troisième variable qui n'aurait pas été prise en compte dans l'étude. Cette situation est très courante et elle incite à contrôler, par des méthodes appropriées, ces variables "parasites" ou confondues. Un exemple simple peut illustrer cette situation. De nombreuses études nord-américaines ont montré que le niveau de bien-être des noirs était (légèrement) inférieur à celui des blancs. Mais est-ce l'origine ethnique qui est liée au bien-être ou les variables socio-économiques associées à cette origine ? On sait que les noirs ont des revenus plus faibles que les blancs, ont un niveau d'éducation moyen moins élevé, habitent dans des quartiers moins agréables, etc. Pour retenir une relation entre bien-être et origine ethnique, il faudrait donc s'assurer qu'une telle relation subsiste lorsque toutes ces variables sont contrôlées ; autrement dit, il faudrait montrer que les noirs restent moins heureux que les blancs même lorsqu'ils ont les mêmes revenus, le même niveau d'éducation, habitent dans les mêmes quartiers, etc. Il en va de même pour beaucoup d'autres variables. Malheureusement, ce contrôle n'est pas toujours facile à opérer, notamment parce que le nombre des variables confondues est potentiellement très important. Ainsi, à quelques rares exceptions près, les corrélats du bien-être ont été étudiés sans contrôler les éventuelles variables parasites. Et il est difficile d'évaluer, *a posteriori*, dans quelle mesure certaines relations sont dues totalement ou en partie à l'influence d'autres variables.

Une autre difficulté rencontrée dans la recension des corrélats du bien-être est liée à son opérationnalisation. Nous avons vu qu'il existait de nombreux instruments pour mesurer le bien-être, qui, même s'ils corrélaient entre eux de manière substantielle, peuvent mesurer chacun des aspects plus spécifiques. On peut donc constater des différences dans les corrélats selon les instruments utilisés et, plus généralement, selon les aspects particuliers du bien-être sur lesquels est mis l'accent. La littérature en fournit plusieurs exemples, dont l'âge est sans doute le plus connu. Alors qu'il ne semble pas y avoir de relation entre l'âge et le bien-être quand on utilise des indicateurs généraux, une relation significative apparaît dans le cas d'indicateurs plus spécifiques. En fait, Campbell et al. (1976), par exemple, ont montré que les aspects cognitifs du bien-être sont liés positivement à l'âge alors que les composantes affectives le sont négativement. Le sexe en offre un autre exemple. De leur méta-analyse, Haring et al. (1984) concluent que le niveau de bien-être moyen des hommes est légèrement supérieur à celui des femmes. Mais dans une méta-analyse plus récente, Wood et al. (1989) montrent que la relation entre sexe et bien-être dépend de la manière dont ce dernier a été opérationnalisé, cette

relation pouvant même s'inverser. Au total, il reste, là aussi, difficile d'évaluer l'impact des différences d'opérationnalisation dans le cas de certaines variables, soit qu'elles aient été trop peu étudiées, soit que les mesures classiques du bien-être ne renvoient pas de manière univoque à des composantes bien définies^[3]. On regrettera toutefois que ces différences d'opérationnalisation soient trop rarement prises en compte, notamment dans les recherches de type meta-analytique.

Enfin, certains auteurs ont souligné qu'avant de comparer les niveaux de bien-être de différentes études, il fallait s'assurer que sa signification était identique dans les différents groupes impliqués (e.g. Liang et al., 1987 a, 1987 b). Cette question est généralement abordée par la comparaison des structures factorielles des mesures dans les différents groupes. Celle-ci n'a guère été étudiée que pour des variables telles que le sexe, l'âge ou l'origine ethnique (cf. Herzog & Rodgers, 1981 ; Carp & Carp, 1983 ; Hoyt & Creech, 1983 ; Wilson et al., 1985 ; Liang et Bollen, 1985 ; Liang et al., 1987 ; Weingarten & Bryant, 1987 ; Liang & Lawrence, 1989 ; Andrews, 1989 ; McCulloch, 1991). De telles études font apparaître des résultats mitigés. Il est rare que les auteurs concluent à une invariance factorielle totale, même si les différences sont rarement très importantes. Ils conseillent toutefois de ne pas les négliger et d'en tenir compte dans l'interprétation des résultats des études comparatives.

On retiendra tout d'abord de ces remarques générales que certaines relations mises en évidence sont sujettes à caution étant donné les problèmes liés aux variables parasites, aux spécificités de l'opérationnalisation du bien-être ou encore aux différences dans les structures factorielles. Mais le point le plus important reste l'interprétation de ces relations. Il faut absolument se garder de toute interprétation en termes d'influence causale, même si cela peut sembler, dans certains cas, intuitivement correct. Des interprétations pouvant paraître contre-intuitives dans un premier temps sont pourtant cohérentes avec une nouvelle conception du bien-être que viennent étayer plusieurs résultats empiriques, ainsi qu'on le verra dans le prochain chapitre. Ce faisant, le bien-être change de statut et donc de signification.

4.3 Les corrélats du bien-être

La classification des corrélats présentée reprend essentiellement celle de Veenhoven (1984), qui a finalement regroupé en plus de 200 catégories les 3500 relations

[3] : Il est intéressant de noter au passage que les différences constatées dans les relations selon les composantes du bien-être qui sont mesurées peuvent servir d'arguments pour assurer la validité divergente de ces composantes. Nous en avons vu plusieurs exemples dans le chapitre précédent.

recensées dans son ouvrage. Comme toute classification, celle-ci repose sur des choix dont la pertinence peut être discutée. Toutefois, comme celle-ci n'est rien d'autre qu'un moyen pratique d'organiser une présentation qui reste surtout descriptive, elle est largement suffisante : les aspects plus fondamentaux seront abordés ailleurs (cf. chapitre 5). Comme l'ouvrage de Veenhoven est encore la synthèse la plus actuelle sur la question, nous l'avons pris comme base dans ce chapitre. Cette synthèse est complétée par d'autres revues générales de la littérature (Larson, 1978 ; Diener, 1984 ; McNeil et al., 1986a) ou des revues plus spécifiques portant sur la relation entre le bien-être et une seule autre variable : c'est notamment le cas des revues quantitatives de la littérature (méta-analyses). Elle est aussi augmentée de quelques variables dont les relations avec le bien-être n'avaient pas été étudiées à l'époque où Veenhoven publiait son ouvrage.

Cette section prendra surtout la forme de tableaux complétés de quelques commentaires visant à préciser certains points ayant suscité des polémiques ou correspondant à des résultats particulièrement importants dans l'histoire du bien-être. Le bien-être y est successivement envisagé en relation avec les conditions de vie, les caractéristiques individuelles et ses "antécédents". Chaque tableau ventile les relations étudiées^[4] selon qu'elles sont positives, négatives, qu'il y a absence de relation ou que les résultats sont inconsistants (le plus souvent par contradiction entre différentes études). Lorsqu'aucune source bibliographique n'est précisée, les résultats présentés proviennent des revues de la littérature de Larson (1978), de Veenhoven (1984), de Diener (1984) ou de McNeil et al. (1986a). Afin de ne pas alourdir outre mesure la présentation, ces sources ne seront pas systématiquement répétées, sauf lorsque nous ferons référence à des commentaires ou à des prises de position spécifiques à certains auteurs. Malgré l'absence de systématisation et le caractère diffus que caractérise une telle démarche, elle offre au lecteur une image assez fidèle de l'ensemble des recherches corrélationnelles sur le bien-être.

[4] : Sans préjuger du type de coefficient d'association utilisé, même si, dans la grande majorité des cas, il s'agit d'un coefficient de corrélation linéaire.

4.3.1 Bien-être et conditions de vie

Variables	$r > 0$	$r < 0$	$r = 0$	$r > 0$ ou $r = 0$
A. Bien-être et société				
Conditions économiques	* prospérité économique (produit national brut, niveau de pauvreté, suffisance alimentaire) (<i>la relation est assez importante mais non linéaire</i>)	* inégalité économique	* sécurité économique (inflation, taux de chômage)	* croissance économique (<i>concerne plus les changements de conditions économiques ayant eu lieu dans un passé proche que les conditions actuelles</i>)
Conditions politiques	* liberté (coercition gouvernementale, liberté de la presse) * démocratie (séparation des pouvoirs, élections libres, etc.) * calme politique (violence politique, grèves, manifestations, etc.)			
Différences régionales	* prospérité économique locale		* habitat urbain (vs rural) <i>Les résultats sont différents selon les pays (Veenhoven, 1984) et selon l'âge des individus (McNeil et al., 1986a)</i> * taille de la ville <i>Les résultats sont différents selon les pays</i>	
B. Bien-être et place de l'individu dans la société				
Sexe			Sexe	
Age			Age <i>mais $r > 0$ avec les composantes plus cognitives du bien-être et $r < 0$ avec les composantes plus affectives</i>	
Statut socio-économique	* revenu * confort du logement * prestige professionnel (perçu par les autres, perçu par soi...) * niveau social global (<i>indicateurs composites</i>)			* niveau d'éducation

Variables	r > 0	r < 0	r = 0	r > 0 ou r = 0
Statut minoritaire		* migrants * homosexualité (il est possible que cette relation évolue avec le temps) * minorité ethnique	* religion (pas de résultats concludants)	
C. Bien-être et travail				
Avoir un emploi	* avoir un travail (vs chômage)		* femme au foyer (vs ayant une activité professionnelle) * travail à temps partiel (pas de résultats concludants)	* retraite (résultats contradictoires)
Type d'occupation	(emploi qualifié, avec responsabilités, etc. vs emploi non qualifié)			
Travail bénévole	* membre d'une association * activité (quantité d') dans une association			
D. Bien-être et relations interpersonnelles				
Mariage	* avoir un conjoint * statut marital des personnes non mariées (célibataires > veuves > divorcées) * bien-être du conjoint			
Enfants				* avoir des enfants * nombre d'enfants
Amis et proches	* avoir des amis ou des proches (nombre, fréquence des contacts) * avoir un confident * contacts sociaux * activités sociales			

Tableau 2.9 : Bien-être et conditions de vie

A. Bien-être et société

Les travaux recensés dans le cadre des relations entre le bien-être des individus et les caractéristiques des sociétés dans lesquelles ils vivent sont surtout le fait d'études sur des indicateurs agrégés au niveau d'un pays ou de régions. En général, lorsque des relations sont mises en évidence, elles sont relativement élevées mais il est nécessaire de noter que les corrélations calculées entre des données agrégées sont, toutes choses égales par ailleurs, supérieures à celles obtenues sur des données non agrégées.

Les relations entre le niveau de bien-être moyen des individus d'un pays et sa richesse ont fait l'objet de débats. Lorsque Easterlin propose, en 1974, une conception relativiste du bien-être qui insiste sur le fait que les individus s'auto-évaluent en fonction de critères de comparaison relatifs et non pas absolus, il s'appuie sur des résultats qui montrent une indépendance entre la richesse d'un pays et le niveau moyen de bien-être de ses habitants. Pourtant, Veenhoven (1991 b, mais voir aussi Veenhoven & Ouweneel, 1989 et Ouweneel & Veenhoven, 1991) conteste cette conclusion. Reprenant les mêmes données, il parvient à établir de manière très convaincante qu'il existe bien une relation, relativement importante, entre ces variables et que les interprétations d'Easterlin reposent sur une représentation graphique très discutable des résultats. Il vérifie par ailleurs que cette relation subsiste sur des données plus récentes. De plus, cette relation n'est pas linéaire. Pour les pays les plus pauvres, le bien-être moyen augmente assez régulièrement en fonction de leur richesse ; en revanche, pour les pays industrialisés, cette relation disparaît presque complètement, les différences de richesse entre les pays ne se reflétant plus au niveau du bien-être moyen de leurs habitants. L'analyse de Veenhoven vient mettre à mal une conception encore largement partagée sur l'indépendance entre le niveau de bien-être et les conditions économiques car l'article d'Easterlin a longtemps constitué une référence incontournable. Il importe toutefois de remarquer que cette relation ne concerne que les pays en deçà d'un certain seuil de richesse et qu'un phénomène identique semble se produire aussi au niveau individuel. Larson (1978) note en effet que les relations entre le bien-être et les revenus des individus sont plus importantes pour les individus ayant les revenus les plus bas. Un tel résultat prend sens si on considère que le bien-être est altéré lorsque les besoins les plus élémentaires ne sont pas satisfaits. Lorsque les conditions économiques sont telles qu'elles permettent de vivre décemment, le surplus de richesse semble n'avoir plus d'influence sur le niveau de bien-être.

B. Bien-être et place de l'individu dans la société

Sous l'appellation générale de "place de l'individu dans la société", Veenhoven regroupe des variables qui caractérisent le statut social des individus : sexe, âge, appartenance à des groupes minoritaires et différents indicateurs de statut socio-

économique. Comme ces informations font partie de celles qui sont habituellement recueillies dans les enquêtes, on dispose de nombreuses informations ayant donné lieu à plusieurs revues de synthèse. Il en ressort que la part de variance commune entre ces variables et le bien-être est extrêmement faible, à la surprise de beaucoup de chercheurs. Andrews & Withey (1976) rapportent que ces variables, envisagées conjointement, "expliquent"^[5] moins de 10% de la variance du bien-être. De tels résultats ont beaucoup surpris les spécialistes des enquêtes sociales qui ont l'habitude de trouver des pourcentages beaucoup plus élevés, y compris pour des phénomènes subjectifs. Inglehart et Rabier (1986) ont comparé les pourcentages de variance commune entre différents indicateurs subjectifs et un ensemble de cinq variables de statut social (sexe, âge, éducation, revenu et profession) sur des données obtenues auprès de vastes échantillons représentatifs des Etats-Unis et de pays européens. Cet ensemble de variables "explique" 33% de la variance du positionnement sur une échelle politique droite-gauche, 30% de l'identification aux partis politiques, 26% des intentions de vote, 23% du type de valeurs... et seulement 6% du bien-être. On se doute que ces résultats ont suscité beaucoup d'interrogations et de tentatives d'explication. Les explications générales seront abordées dans le cadre du chapitre sur les théories. Nous nous contenterons ici d'illustrer quelques explications plus spécifiques proposées pour l'âge, le revenu et le niveau d'éducation.

Les relations entre l'âge et le bien-être ont fait l'objet de nombreuses discussions. Les résultats de la méta-analyse de Stock et al. (1983) montrent une quasi absence de relation. Pourtant, depuis Campbell et al. (1976), on a régulièrement trouvé une relation positive entre l'âge et les composantes du bien-être faisant référence à la satisfaction (c'est-à-dire à des évaluations plus cognitives) et une relation négative pour des aspects plus affectifs, faisant référence à l'enthousiasme ou à la joie de vivre. Il est donc possible que les coefficients positifs et négatifs se compensent dans les moyennes calculées dans les méta-analyses. Toutefois, les relations sont de faible intensité, mais répliquables (Larson, 1978, estime que l'âge explique entre 0 et 1% de la variance du bien-être).

La relation positive entre l'âge et les composantes cognitives du bien-être a beaucoup intrigué les auteurs qui partaient du postulat que les personnes âgées, étant objectivement dans des situations plus difficiles, devaient être moins satisfaites que les plus jeunes (cf. Carp & Carp, 1981 ; Felton 1987). Campbell et al. (1976) ont passé en revue des explications pouvant rendre compte de ces résultats : l'augmentation des niveaux de satisfaction avec l'âge peut s'expliquer par la mort sélective des plus insatisfaits ou par des effets de cohorte. Les auteurs considèrent que ces deux

[5] : Nous nous sommes expliqué plus haut sur l'erreur qui consiste à croire que les corrélats du bien-être en sont les causes. Il faut donc comprendre ici le terme expliquer au sens statistique, c'est-à-dire "partager un certain pourcentage de variance commune".

explications, bien que plausibles, sont relativement peu probables^[6]. Les autres explications avancées par les auteurs à titre d'hypothèses ne s'excluent pas et sont difficiles à départager au niveau empirique. On avance que l'augmentation de la satisfaction est peut-être liée à des changements objectifs dans les conditions de vie (augmentation des salaires, etc.), à des effets subjectifs d'accommodation à la situation ou à une meilleure adéquation personne-environnement^[7]. Cette dernière explication fait référence à l'hypothèse selon laquelle au fur et à mesure que l'on vieillit, on trouve des "niches" qui correspondent davantage à ce que l'on souhaite. En ce qui concerne la relation négative entre l'âge et les aspects affectifs du bien-être, Campbell et al. (1976) font référence à la diminution, avec l'âge, de l'intensité des plaisirs ; ils évoquent aussi un espoir d'évolution des conditions de vie plus grand chez les plus jeunes, qui les conduirait à une évaluation plus critique de leur situation actuelle, en termes cognitifs. Il n'y a donc pas de réponses définitives permettant d'expliquer les relations complexes entre l'âge et le bien-être. Quoiqu'il en soit, il ne faut pas oublier que malgré son intérêt théorique, cette relation est peu importante en termes d'intensité.

Ce qui est étonnant dans les relations entre ces différents indicateurs de *statut socio-économique*, ce n'est pas qu'elles soient positives, mais qu'elles soient beaucoup plus faibles que ce à quoi on pouvait s'attendre : Larson (1978) rapporte en effet que ces variables n'"expliquent" qu'entre 1 et 9% de la variance du bien-être. Ces résultats étonnent et embarrassent depuis longtemps les chercheurs sur le bien-être, parce que les variables de statut socio-économique sont considérées comme des indicateurs objectifs des conditions de vie, à la fois accessibles et évidents. Certains (e.g. Zapf, 1987 ; Olson & Schober, 1993) parlent même du "paradoxe de la satisfaction"^[8], pour faire état de leur désarroi. Pourtant, il paraît clair, aujourd'hui, que les indicateurs objectifs des conditions de vie et les indicateurs subjectifs de bien-être ne mesurent pas la même réalité ; il serait naïf de croire que des conditions de vie satisfaisantes sont suffisantes pour rendre une personne heureuse. Pourtant, l'attribution de cette absence de relation à des artefacts techniques est encore avancée en explication. C'est ainsi que certains (e.g. Douthitt et al., 1992 ; Mullis, 1992) considèrent qu'il faut améliorer l'opérationnalisation des indicateurs économiques alors que d'autres (e.g. Moum, 1988) incriminent les biais de réponses aux items des questionnaires de bien-être. D'autres encore font référence au caractère inapproprié des méthodes linéaires traditionnellement utilisées, en arguant du fait que les

[6] : Mais cf. Felton (1987) à propos des effets de cohorte et, plus bas, le paragraphe sur la santé à propos de la relation entre le bien-être et l'espérance de vie.

[7] : Dans la théorie de l'adéquation personne-environnement ("*person-environment fit*", cf. French et al., 1974), les besoins individuels diffèrent en fonction de niveaux optima sur plusieurs variables dans des interactions complexes : certains ont besoin de plus d'amis que d'autres, mais dans certaines situations seulement, etc. Il s'agit d'une théorie clairement interactionniste.

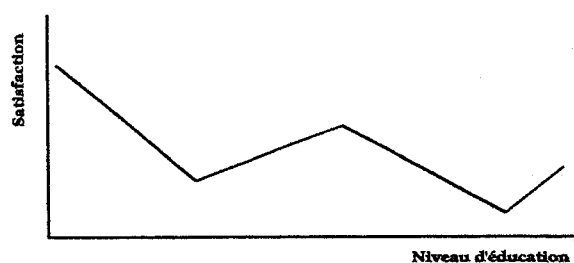
[8] : Inglehart & Rabier (1986) préfèrent parler, quant à eux, de "paradoxe des petites variations inter-groupes".

relations entre indicateurs objectifs des conditions de vie et bien-être seraient non linéaires (cf. Diener et al., 1993).

Cet argument a souvent été avancé à propos du niveau d'éducation (cf. Larson, 1978) pour lequel on dispose de résultats contradictoires. Alors que certains concluent à une relation positive qui subsiste lorsque des variables parasites sont contrôlées (e.g. Glenn & Weaver, 1981), Campbell et al. (1976) trouvent, à plusieurs reprises, que le niveau d'éducation est corrélé négativement avec le bien-être. Cette étude est d'ailleurs intéressante parce qu'elle montre, à partir de données recueillies sur un échantillon représentatif de la population des Etats-Unis, que la relation entre le niveau d'éducation et la satisfaction dans différents domaines de vie est non linéaire. Plus spécifiquement, elle a souvent une forme en W (cf. figure 2.19), que les auteurs ont essayé d'analyser en détail.

Le second pic inférieur du W correspond à des personnes qui ont commencé des études universitaires sans les achever. Les auteurs pensent qu'en plus d'un sentiment de frustration qui pourrait influencer sur leur bien-être, ces individus adoptent les références du groupe social de ceux qui ont effectué (et terminé) leurs études universitaires. Or, ces derniers ont, objectivement, des conditions de vie supérieures aux leurs.

Ainsi, s'ils expriment moins de satisfaction, c'est qu'ils adoptent des standards de référence trop élevés par rapport à leur situation^[9]. Le premier pic supérieur du W exprime le fait que ce sont les personnes les moins éduquées qui ont les niveaux de satisfaction les plus élevés. Cela se comprend, expliquent les auteurs, si on considère que les individus sont d'autant plus modérés dans l'évaluation de leurs conditions de vie qu'ils ont connu d'autres situations. Or, cette "conscience des alternatives", comme la nomment Campbell et al., est acquise soit par l'éducation soit par des expériences extra-scolaires (voyages, déménagements, etc.), ces deux facteurs étant bien évidemment liés. Comme les personnes les moins éduquées sont celles qui ont connu peu de situations alternatives, ils relativisent moins leur niveau de satisfaction. Malgré son caractère subtil, cette explication n'est pas très convaincante : on ne comprend pas très bien, en effet, pourquoi on observe un tel phénomène dans le cas du niveau d'éducation et pas dans celui du revenu ; il est tout aussi plausible de penser



[9] : On verra, dans le prochain chapitre, que de nombreux modèles du bien-être se sont construits autour de l'idée que les individus s'auto-évaluent en fonction de la perception qu'ils ont de leur situation par rapport aux situations des individus faisant partie du même groupe social de référence.

que les personnes ayant des niveaux élevés de revenus ont aussi cette "conscience des alternatives".

Les résultats relatifs au *sexe* diffèrent quelque peu selon les auteurs : Veenhoven (1984) conclut à une absence de relation ; dans leur méta-analyse, Haring et al. (1984) aboutissent à la conclusion que le niveau de bien-être des hommes est supérieur à celui des femmes ; la méta-analyse de Wood et al. (1989) conduit au résultat inverse. Nous avons vu que cette divergence pouvait s'expliquer en partie par la non prise en compte des différences d'opérationnalisation du bien-être. Mais en tout état de cause, les auteurs sont d'accord pour admettre que si cette relation existe, elle est très faible ; Larson (1978) estime que le sexe explique entre 0 et 1% de la variance du bien-être.

En ce qui concerne le *statut minoritaire*, c'est sans doute l'origine ethnique qui a été la plus étudiée, notamment aux Etats-Unis, à travers la comparaison du niveau de bien-être des noirs et des blancs. Les résultats de la méta-analyse de Stock et al. (1985) font apparaître que les blancs ont un niveau de bien-être moyen supérieur à celui des noirs, mais que l'origine ethnique n'explique que 2% de la variance du bien-être (les estimations de Larson, 1978, sont encore plus faibles). De plus, ainsi que nous l'avons signalé en introduction, la prise en compte de plusieurs variables parasites (revenu, niveau d'éducation, etc.) pourrait encore faire baisser ce pourcentage.

C. Bien-être et travail

La relation entre le chômage et le bien-être préoccupe de nombreux chercheurs. Il paraît relativement bien établi que les personnes qui ont un *emploi* ont un niveau de bien-être moyen supérieur à celles qui en recherchent un mais, à nouveau, l'intensité de la relation est faible : Larson (1978) l'estime à moins de 1%. Les effets de la retraite ont été aussi très étudiés, y compris en France (cf. Poitrenaud et al., 1987), mais il ne s'en dégage pas de conclusion très claire. Il est probable que de nombreuses variables devraient être prises en compte, telles que la possibilité pour les personnes de ne pas prendre leur retraite.

On peut concevoir que toutes les variables liées au *bénévolat* et à l'implication dans des associations renvoient plus aux liens et aux contacts sociaux qu'au travail proprement dit. Ces variables sont, de plus, des exemples typiques d'une causalité douteuse : les personnes sont-elles heureuses parce qu'elles participent à des associations ou sont-ce les personnes les plus heureuses qui s'impliquent le plus dans de telles activités ?

D. Bien-être et relations interpersonnelles

Selon Larson (1978), le *mariage* "explique" entre 1 et 4% de la variance du bien-être. Cette relation semble bien établie mais nous avons vu que Masteskaasa (1992) apportait des éléments empiriques laissant entrevoir que les personnes heureuses avaient une probabilité plus importante de se marier. Ce résultat est à mettre en parallèle avec tous ceux qui montrent que les activités et les contacts sociaux sont plus fréquents chez les individus dont le niveau de bien-être est élevé. On se rappelle aussi à ce sujet les études expérimentales sur l'humeur (cf. Isen, 1987), qui mettent en évidence l'influence de l'affectivité positive sur différents indicateurs de la qualité des relations interpersonnelles.

Sous le terme assez vague d'*activités sociales* sont regroupées toutes les études portant sur les interactions sociales, qu'elles portent sur la fréquence de ces activités ou sur la taille du réseau social de l'individu. Les estimations de Larson (1978) font apparaître entre 1 et 3% de variance commune entre bien-être et activités sociales ; la méta-analyse de Okun et al. (1984) permet d'élever ces estimations jusqu'à 8%. Il faut aussi signaler que certains résultats laissent penser que, plus généralement, le niveau global d'activité d'un individu est lié au bien-être (cf. Diener, 1984 ; McNeil et al., 1986a ; Stones & Kozma, 1986 a).

4.3.2 Bien-être et caractéristiques individuelles

Variables	$r > 0$	$r < 0$	$r = 0$	r inconsistant
A. Ressources personnelles				
Santé	Santé			
Efficacité mentale générale (être capable de gérer correctement les problèmes de sa vie)	* efficacité mentale positive (intégrité de l'identité, force de l'ego, résolution...) * maturité psycho-sociale (stades de Erikson, actualisation de soi...)	* désordres mentaux		
Apitudes spécifiques	* aptitudes sociales (empathie, chaleur humaine, ...)		* intelligence (confirmé par Diener, 1984)	* planifier ses buts personnels
Richesse de la vie mentale			* diversité des affects éprouvés * variabilité de l'humeur * complexité cognitive (différenciation conceptuelle, dépendance à l'égard du champ, rigidité mentale) <i>(les résultats sont plutôt inconsistants)</i>	* intensité des affects
Adaptation	* adaptation			
Influences génétiques	influences génétiques			
B. Traits de personnalité				
Traits de personnalité	* extraversion * sens de la vie * estime de soi * internalité * androgynie * recherche de sensations	* névrosisme		

Variables	$r > 0$	$r < 0$	$r = 0$	$r > 0$ ou $r = 0$
Mécanismes de défense	* intellectualisation, revirement d'opinion	* se retourner contre soi, contre les autres, projection	* refoulement	
Orientation temporelle				* orientation temporelle
C. Style de vie				
Temps de loisirs	* niveau d'activité, changements de loisirs, ...			
Habitudes de consommation			* alimentation, alcool, tabac	
Sommeil			* sommeil (quantité, heure précoce)	
Vie ascétique			* vie ascétique	
D. Aspirations, buts, valeurs				
Aspirations	* aspirations liées aux valeurs, au caractère, à la santé, à la famille et à la situation internationale	* aspirations liées au domaine économique * volonté de changer sa vie	* aspirations liées au travail, à la société	
Buts dans la vie	* plaisirs, préserver sa santé	* s'améliorer soi-même psycho-logiquement, améliorer ses relations intimes, focaliser son énergie sur les études et le travail	* buts matériels, idéaux	
Motivations sous-jacentes	* tendance à essayer d'exceller dans les compétitions * tendance à rechercher l'approbation sociale (mais les résultats sont divergents) * tendance à rechercher des situations excitantes			

Variables	$r > 0$	$r < 0$	$r = 0$	$r > 0$ ou $r = 0$
Valeurs	* intérêt pour les valeurs, les principes moraux * importance accordée aux valeurs hédonistes, sociales, religieuses	* importance accordée aux valeurs intellectuelles, à l'individualisme, à la liberté, au contrôle de soi	* importance accordée aux valeurs politiques (conservatisme...), aux valeurs liées à la personnalité (ambition, honnêteté...), aux valeurs esthétiques	
Religion	* religiosité (membre d'une église, importance de la religion, participation religieuse)		* type de religion	
Conventionnalisme	* conservatisme politique		* attitude libérale envers la sexualité	
Représentations du bien-être	* valorisation du bien-être			
E. Evaluations de différents aspects de la vie				
Evaluations de soi	* caractérisations positives de soi (stabilité de l'image de soi, congruence entre soi et idéal du soi) * auto-évaluations (estime de soi, confiance en soi, satisfaction générale avec soi-même)			
Evaluations impliquant autrui	* évaluation de son mariage * évaluation de ses enfants * évaluation de sa famille (corrélation moyenne = .29 selon la méta-analyse de Okun & Stock, 1987) * évaluation de ses amis * évaluation des gens en général (confiance...)			

Variables	r > 0	r < 0	r = 0	r > 0 ou r = 0
Evaluations de la société	* liberté perçue * attitudes envers les institutions (gouvernement, école, mariage, système de santé, religion) * évaluation de son pays			* acceptation du système social (anomie) (cf. Dobson et al., 1979 qui concluent à une relation positive)
Evaluations de sa position sociale	* évaluation de son revenu (suffisance, tranquillité financière, satisfaction globale) * évaluation de son niveau d'éducation * satisfaction globale avec sa position sociale			
Evaluations de son travail	* caractéristiques (positives) perçues du travail * satisfaction avec des aspects spécifiques du travail * satisfaction globale avec le travail (relation importante) ⁽¹⁰⁾			
Evaluations de ses loisirs	* évaluation de son temps de loisirs * satisfaction globale avec ses loisirs			
Evaluations de son environnement de vie	* satisfaction avec son logement * satisfaction avec son voisinage * satisfaction avec différents aspects de son environnement (coût de la vie, attractivité de la ville...)			
Evaluations de sa santé	* évaluations de sa santé (pour mémoire, nous en avons déjà parlé)			

Tableau 2.10 : Bien-être et caractéristiques individuelles

[10] : Cette relation est en général importante, ainsi que le montre la méta-analyse de Okun & Stock (1987), produisant un coefficient moyen de .33 et celle, un peu plus récente, de Tait et al. (1989, cités par Judge & Watanabe, 1993), qui fournit un coefficient de .44.

A. Ressources personnelles

On se doute que les relations entre le bien-être et la *santé* ont été beaucoup étudiées. Les auteurs admettent généralement une relation assez importante entre ces deux variables : le coefficient moyen de la méta-analyse de Okun et al. (1984) est de .32 ; les estimations de Larson (1978) varient entre 4 et 16% de variance commune. La santé peut être évaluée de différentes manières. On peut recueillir des informations directes en demandant aux sujets d'évaluer leur état de santé (voire leur satisfaction vis-à-vis de la santé) ou procéder de manière plus indirecte en faisant estimer aux sujets le nombre de visites chez le médecin, le nombre d'hospitalisations ou encore la fréquence de prises de médicaments. On peut aussi recueillir une évaluation externe faite par un médecin. Bien évidemment, les relations sont plus fortes dans le premier cas, mais elles subsistent dans le second^[11]. Toutefois, aucun des indicateurs utilisés n'est entièrement satisfaisant : les auto-évaluations de la santé peuvent très bien être contaminées par le niveau de bien-être ; la quantité de soins ne reflète pas nécessairement l'état de santé réel d'un individu (elle peut indiquer son niveau d'inquiétude par rapport à sa propre santé) ; enfin, comme le note Larson (1978), les évaluations des médecins, même si ce sont les plus objectives, ne sont pas forcément aptes à rendre compte de l'état de santé d'une personne (qui peut ressentir des douleurs ou d'autres symptômes qui ne sont pas pris en compte par ce type d'évaluation).

Au vu de ces résultats, il est difficile de nier que la santé et le bien-être partagent une part non négligeable de variance commune, la santé apparaissant, dans les revues de la littérature, comme l'un des corrélats les plus importants du bien-être. Il est plus difficile cependant de préciser laquelle des deux variables influence l'autre, si une relation causale existe. Pourtant, quelques études longitudinales semblent montrer que le niveau de bien-être est un prédicteur non négligeable de l'espérance de vie. C'est le cas de l'étude de Paltore (1969, cité par Veenhoven, 1984) réalisée aux Etats-Unis et de celle de Grand & Boquet (1989), en France. L'étude de Grand & Boquet tend à montrer que le bien-être, mesuré de manière très fruste avec seulement trois items, tend à être un meilleur prédicteur de la mortalité que des indicateurs *a priori* plus importants, comme la santé.

Quant à l'*efficacité mentale générale*^[12], elle a été relativement peu étudiée. On reconnaîtra tout de même sous cette appellation des concepts que certains considèrent comme des composantes mêmes du bien-être (cf. le modèle de Ryff, 1989a, 1989c, décrit dans le premier chapitre).

[11] : Même si George & Landerman (1984) contestent cette relation sur la base d'analyses secondaires qu'ils ont menées en contrôlant plusieurs variables parasites.

[12] : Cf. tableau 2.10 pour la signification que Veenhoven donne à cette expression.

En ce qui concerne les *aptitudes spécifiques* (intelligence, etc.), seules les aptitudes sociales semblent liées au bien-être. Ce résultat doit être mis en relation avec l'ensemble de ceux relatifs aux relations interpersonnelles qui vont globalement tous dans le même sens.

Les travaux relatifs à la *richesse de la vie mentale* (diversité des affects, complexité des affects, etc.) ne montrent généralement pas de liaison. En ce qui concerne l'intensité des affects, les résultats sont plutôt contradictoires. Les études recensées par Veenhoven (1984) font pratiquement toutes apparaître des relations négatives. Depuis, le concept d'intensité affective a connu un succès important dans la littérature, notamment suite aux travaux de Diener et de ses collègues (cf. Diener & Emmons, 1985 ; Diener et al., 1985 b) sur l'indépendance entre affectivité positive et affectivité négative (cf chapitre précédent). Or, il ressort de ces études une relative indépendance entre les mesures d'intensité des affects et les mesures classiques de bien-être.

Le bien-être a aussi été associé à la notion d'*adaptation*, concept vague s'il en est. La corrélation moyenne calculée entre des mesures de bien-être et des mesures d'adaptation est estimée à .38 dans l'étude méta-analytique de Okun & Stock (1987).

Les travaux qui étudient les *influences génétiques* sont rares dans le domaine du bien-être. Pourtant, on peut trouver quelques résultats, chez Tellegen et al. (1988) ou chez Hershberger (1994), montrant que ces influences génétiques ne sont pas à négliger.

B. Traits de personnalité

L'étude la plus systématique sur les relations entre *traits de personnalité* et bien-être est due à Emmons & Diener (1985 b), mais de nombreux autres travaux, plus spécifiques, ont été menés dans cette perspective. On retiendra surtout de ces résultats les relations importantes qui sont mises en évidence entre le bien-être et les deux grands facteurs de la personnalité de Eysenck, à savoir l'extraversion et le névrosisme^[13]. De manière plus générale, on peut noter que le niveau des relations entre les traits de personnalité et le bien-être est plus important que celui de la plupart des variables qui ont été envisagées jusqu'à présent.

Deux points sont à noter concernant la relation entre le bien-être et l'*extraversion*. Tout d'abord, l'extraversion est principalement liée à l'affectivité positive, très peu à l'affectivité négative. Ensuite, Emmons & Diener (1986) ont montré que c'était la

[13] : Plus récemment, McCrae & Costa (1991) ont montré que le bien-être était lié à tous les facteurs du "big-five".

composante de sociabilité de l'extraversion qui rendait compte des relations avec le bien-être et non pas la composante d'impulsivité.

Les relations entre le *névrosisme* et le bien-être sont relativement importantes. Dans leur méta-analyse, Okun & Stock (1987) rapportent un coefficient moyen de -.33, ce qui est comparable à ce que l'on obtient pour la santé. De plus, il faut noter que ce résultat est obtenu avec des mesures globales du bien-être, alors que le névrosisme semble être fortement lié à l'affectivité négative et relativement indépendant de l'affectivité positive.

La relation entre le *sens de la vie* et le bien-être est importante. Nous avons d'ailleurs souligné à plusieurs reprises que ces deux concepts étaient difficilement distinguables au niveau conceptuel ainsi que l'attestent des travaux comme ceux de Chamberlain & Zika (1988), de Zika & Chamberlain (1992) ou de Korotkov (1993).

L'*internalité* est aussi liée au bien-être. La tendance à se percevoir comme la cause des événements est une des relations les plus solidement établies dans ce domaine. Certains auteurs (e.g. Zautra & Reich, 1980, 1983) notent pourtant que les personnes heureuses ont tendance à s'attribuer la cause des événements positifs et à attribuer à des causes externes les événements négatifs.

Dans sa revue de travaux, Diener (1984), considère que les résultats relatifs à l'*androgynie* sont plutôt inconsistants. Mais dans une publication plus récente, Hunt (1993) a fourni plusieurs arguments qui laissent penser que cette relation est tout de même positive.

C. Style de vie

Il ne ressort pas grand chose des variables que Veenhoven (1984) considère comme des indicateurs de style de vie si on excepte la relation avec les *loisirs*. Il nous semble que ce résultat mérite d'être rattaché à ceux qui concernent le niveau global d'activité, les loisirs ne constituant qu'un aspect particulier de l'ensemble des activités entreprises par un individu.

D. Aspirations, buts, valeurs

Veenhoven (1984) propose, sous la rubrique des aspirations, buts et valeurs, un ensemble de variables relativement ignorées des autres chercheurs, qu'il discute de manière détaillée. Les études empiriques sur lesquelles s'appuient les résultats présentés dans le tableau 2.10 sont donc peu nombreuses et il y a donc tout lieu de les considérer avec prudence. En lisant le compte-rendu de ces investigations, on ne peut pas s'empêcher

de penser qu'elles font état, dans une certaine mesure, des représentations que se font les individus du bien-être. Lorsque des personnes sont interrogées sur leurs aspirations, leurs buts dans la vie, sur leurs valeurs ou encore sur leurs soucis, il est plus que probable que les individus expriment ce qui leur paraît central pour leur propre bonheur. Lorsque, de plus, ces questions sont posées dans le contexte d'une recherche sur le bien-être ou sur la qualité de la vie, comme c'est souvent le cas, c'est encore plus évident.

Dans cette énumération, on sera plus sensible aux thèmes évoqués qu'aux relations en elles-mêmes. En effet, lorsque des thèmes sont fréquemment évoqués, on peut penser qu'ils sont pris en compte, d'une manière ou d'une autre, dans l'évaluation que fait un individu de son bien-être général. Ils peuvent constituer, à ce titre, des facettes importantes dans la conceptualisation du bien-être. En revanche, les relations présentent moins d'intérêt, principalement parce que l'on ne dispose pas de leur intensité^[14]. Comme plusieurs de ces résultats sont issus d'enquêtes portant sur de vastes échantillons, la significativité des différences ne peut refléter leur importance.

Mentionnons, en ce qui concerne les représentations du bien-être, que les résultats relatifs aux *causes perçues du bien-être* sont similaires à ceux qui relèvent des aspirations. La seule différence concerne la santé : ce sont les personnes les moins heureuses qui mentionnent le plus la santé. Il faut toutefois noter que ces résultats rapportés par Veenhoven (1984) ne sont pas confirmés par l'étude de Felton (1987), dans laquelle il n'apparaît pratiquement pas de relation entre représentation et niveau de bien-être.

E. Evaluations de différents aspects de la vie

On peut s'étonner de voir figurer, dans un chapitre sur les corrélats du bien-être, des variables concernant l'évaluation de différents aspects de la vie. En effet, certaines conceptions du bien-être considèrent que ces variables font partie intégrante du concept. Pour d'autres, dont Veenhoven fait partie, il faut distinguer entre évaluation globale de sa vie et évaluations d'aspects plus spécifiques. Quelle que soit la position adoptée, il est intéressant de pouvoir disposer d'une liste des domaines de vie dont l'évaluation a été mise en relation avec le bien-être général. Parmi ces domaines, celui du travail se distingue par le nombre considérable d'études auxquelles il a donné lieu. La satisfaction avec le travail a été examinée en relation avec le bien-être, bien sûr, mais aussi pour elle-même. Dans ce dernier cas, des auteurs ont notamment essayé de mettre en relation des indicateurs subjectifs de satisfaction avec le travail et des indicateurs objectifs (salaire, conditions de travail, etc.) : tout comme dans le cas du bien-être, la relation entre ces deux types d'indicateurs se révèle très faible.

[14] : En ce qui concerne les *valeurs*, Veenhoven note que les relations sont assez faibles.

Toutes les relations rapportées sous cette rubrique des *évaluations* sont relativement importantes. On ne s'en étonnera pas outre mesure puisque les variables dont il est question sont des évaluations subjectives, de même que le bien-être. De plus, comme les instruments utilisés pour mesurer la satisfaction dans les domaines de vie ont la même forme que ceux destinés à la mesure du bien-être, ils partagent une part d'erreur systématique due à la méthode, ce qui contribue à hausser les corrélations. Ce sont souvent aussi des mesures effectuées avec un seul item, comportant donc une part importante d'erreur de mesure, ce qui contribue, en revanche, à baisser les corrélations. Ces deux effets ont tendance à se compenser, ainsi que le montrent Lance et Sloan (1993), à l'aide d'un modèle d'équations structurales appliqué à une matrice multitraits-multiméthodes. Il s'ensuit que la taille des quelques coefficients moyens rapportés peut être assez fiable.

4.3.3. Les "antécédents" du bien-être

Etant donné les problèmes liés au sens de la causalité présentés par l'immense majorité des corrélats du bien-être que nous venons de voir, il est tentant d'essayer de recueillir des informations sur des variables dont on sait qu'elles sont temporellement antérieures à la mesure du bien-être. C'est dans cette optique que Veenhoven (1984) fait référence aux "antécédents" du bien-être. Malheureusement, ces études sont rares et, ce qui est plus grave, ce sont le plus souvent des études rétrospectives interrogeant les sujets sur leurs conditions de vie passées. Cette procédure ne permet évidemment pas d'assurer la validité interne (au sens de Cook & Campbell, 1979) des relations mises en évidence. Autrement dit, des antécédents rétrospectifs ne sont pas de vrais antécédents : on ne sait pas, en général, évaluer dans quelle mesure les informations rétrospectives fournies par un sujet sont modifiées par des reconstructions cognitives. Tout ceci fait que les résultats présentés par Veenhoven sont, pour le moins, douteux.

Pour résumer les résultats de ces études, nous dirons que le bien-être semble lié négativement avec certaines conditions défavorables dans l'enfance (divorce des parents, maladies ou décès familiaux, etc.) mais pas avec toutes (de mauvaises conditions socio-économiques ne semblent pas avoir une influence sur le bien-être ultérieur). Plusieurs recherches ont eu aussi comme objectifs d'étudier les relations entre le bien-être et les conditions de vie directement antérieures. Dans ce cadre, on a surtout fait référence aux événements de vie vécus par les individus. Selon Veenhoven (1984), les résultats de ce type d'études sont inconsistants, alors que selon Diener (1984), on peut mettre en évidence un effet positif sur le bien-être des événements de vie positifs et un effet négatif des événements négatifs.

La relation entre événements de vie et bien-être pose aussi des problèmes théoriques fondamentaux : elle sera donc abordée plus en détail avec les résultats de recherches plus récentes, dans le prochain chapitre. Enfin, il n'existe pratiquement pas d'études sur les caractéristiques personnelles antérieures. A titre anecdotique, signalons que la beauté physique évaluée par des juges indépendants sur des photographies datant de la fin de l'adolescence semble liée au bien-être ultérieur. Plus sérieusement, le bien-être actuel est lié au bien-être antérieur, ce que nous avons déjà vu dans le chapitre consacré aux instruments de mesure (cf. section 2.2.1).

Enfin, depuis la synthèse de Veenhoven, d'autres études ont été menées. L'une d'elles est particulièrement informative puisqu'il s'agit d'une enquête longitudinale focalisée sur le bien-être et menée sur un échantillon représentatif de la population australienne. Etant donné l'impact des résultats de ce panel sur les conceptions théoriques du bien-être, ceux-ci seront présentés dans le prochain chapitre.

4.4 Conclusion

Que pouvons-nous retenir de cette énumération de variables corrélées avec le bien-être ? Il semble tout d'abord que le résultat le plus important concerne les relations entre le bien-être subjectif et les conditions de vie objectives. Nous avons vu que ces relations étaient relativement importantes lorsqu'elles portent sur des données agrégées. En revanche, elles s'atténuent fortement au niveau individuel et vont même jusqu'à disparaître complètement. Ces résultats ne sont pas forcément contradictoires. La relation curvilinéaire mise en évidence entre les indicateurs économiques d'un pays et le niveau moyen du bien-être de ses habitants indique que cette relation est rendue possible par la présence, dans l'échantillon, de quelques pays très pauvres dans lesquels les conditions de vie ne permettent pas de remplir les besoins les plus élémentaires au sens de la hiérarchie de Maslow. Dès lors que l'on se limite aux pays industrialisés, le niveau de bien-être moyen d'un pays est pratiquement indépendant de sa richesse. On ne sera pas surpris d'apprendre, qu'au niveau individuel, la plupart des études ont été réalisées dans les pays les plus riches. Or, dans ces derniers, on peut penser que les besoins de base sont satisfaits pour l'immense majorité des habitants. Il est dès lors possible que les conditions de vie n'aient un impact sur le bien-être qu'en dessous d'un certain niveau : au-delà, celles-ci ne peuvent plus rendre compte de la variance du bien-être.

Le bien-être n'est pas seulement relativement indépendant des variables de type monétaire. Il l'est aussi, de façon plus générale, de toutes les variables qui caractérisent le statut social d'un individu. Il serait faux d'affirmer que ces relations n'existent pas du tout ; elles sont toutefois tellement faibles que leur pouvoir informatif est quasi nul. C'est

pourquoi nous nous montrerons réservé quant aux hypothèses interprétatives avancées pour rendre compte de communautés de variance dépassant rarement 5%.

En revanche, des relations assez importantes, et non triviales, sont constatées entre le bien-être et les relations interpersonnelles ou les activités sociales. Ces résultats sont obtenus tant dans des études corrélationnelles que dans des études où l'humeur est induite expérimentalement. On peut sans doute rapprocher cette relation de celle qui a été mise en évidence entre le bien-être et la composante de sociabilité de l'extraversion. De manière plus générale d'ailleurs, les relations entre le bien-être et les traits de personnalité apparaissent significatives et d'une ampleur non négligeable. Parmi les caractéristiques personnelles ne faisant pas référence au statut social, la santé occupe une place de choix. On considère souvent en effet que la santé est un des corrélats "objectifs" du bien-être les plus solides. La forme de cette relation n'est pas forcément celle que suggère le sens commun. Des résultats obtenus grâce à des études longitudinales permettent de montrer que le bien-être est un bon prédicteur de l'espérance de vie, ce qui laisse entendre qu'il peut influencer sur la santé. En ce qui concerne les relations entre les buts, les aspirations ou les valeurs des individus, les résultats sont plus difficiles à interpréter, tout d'abord en raison du faible nombre d'études qui se sont intéressées à cette question, ensuite en raison de la faiblesse des relations mentionnées. On trouve tout de même, dans cette liste de variables, des convergences avec des résultats antérieurs, notamment en ce qui concerne les relations entre le bien-être et une tendance assez générale pouvant être assimilée à la désirabilité sociale, lorsque celle-ci n'est pas envisagée comme un biais de réponse. Il nous semble que des relations entre le bien-être et des variables comme la tendance à rechercher l'approbation sociale, le conservatisme politique ou même la coopérativité dans les négociations interpersonnelles ou la gentillesse participent toutes à cette convergence. Enfin, le bien-être est lié de manière importante à des évaluations positives concernant différents aspects de sa vie. Quand le bien-être est conçu comme une évaluation générale de sa vie, comme c'est le cas chez Veenhoven (1984), ces résultats ne sauraient surprendre ; ils incitent plutôt à penser que ces variables mériteraient d'être étudiées comme des composantes du bien-être et non comme des corrélats.

A force de vouloir à tout prix identifier les variables susceptibles d'influencer le niveau de bien-être des individus, on a eu trop tendance à oublier, pendant presque deux décennies de recherche intensive, que le bien-être présentait des caractéristiques similaires à celles de dispositions individuelles relativement stables. Il n'est pas très difficile de comprendre les motivations sous-jacentes à ce parti pris sur la nature du bien-être, largement dominant pendant cette époque. Malheureusement, en se focalisant sur la prédiction du bien-être, les modèles théoriques se sont en partie sclérosés dans la quête de variables intermédiaires pouvant expliquer l'absence de relation entre conditions de vie et bien-être. Il est possible que de telles variables intermédiaires existent, mais il est surtout probable, comme on va le voir à présent, que ces variables ne soient que des conséquences du bien-être et non des causes.

CHAPITRE V

THEORIES NOMOLOGIQUES ET MODELES EXPLICATIFS DU BIEN-ETRE SUBJECTIF

5.1 Introduction

Les philosophes d'abord, les scientifiques ensuite ont toujours essayé de déterminer quelles sont les raisons qui font qu'une personne est heureuse. Dans son ouvrage, Tatariewicz (1976) recense une liste impressionnante de théories, plus ou moins élaborées, plus ou moins fantaisistes, conçues par des philosophes venant d'époques et d'horizons différents, ayant pour principal objectif de déterminer quelles sont les causes du bonheur. La présentation qu'il en donne est empreinte d'une ironie à peine dissimulée. Il montre en effet qu'à chacune de ces théories on peut en opposer une autre, tout aussi plausible, qui soutient exactement le contraire. C'est ainsi que pour certains le bonheur se trouve dans l'amour alors que pour d'autres il est dans la réduction de tous les attachements ; pour être heureux, certains estiment qu'il faut connaître la vérité, d'autres conseillent de vivre dans l'illusion ; le bien-être peut se trouver dans le combat, dans l'adversité, il peut aussi se trouver dans la tranquillité... Ces quelques exemples suffisent à illustrer notre propos. Le lecteur en trouvera plusieurs dizaines d'autres dans l'ouvrage cité.

Contrairement à une idée répandue, émise par Dodge en 1930 et reprise par de nombreux auteurs (e.g. Wilson, 1967 ; Veenhoven, 1984), les théories sur le bien-être *ne sont pas* restées au niveau auquel les philosophes grecs les ont laissées. Quiconque feuillette la littérature scientifique sur le bien-être pourra s'en rendre compte et constatera qu'il est très vite submergé par la quantité des modèles explicatifs proposés. Ce foisonnement de théories, ou plutôt de micro-théories, s'étend à des concepts voisins comme, en particulier, la satisfaction avec le travail (cf. Furnham, 1991)^[1]. Cependant, malgré cette profusion, il est vrai que ces modèles restent souvent partiels, ne rendant compte que de certaines régularités empiriques, et que tous peuvent faire l'objet de critiques, que ce soit au niveau de leurs postulats ou de la méthodologie utilisée pour les valider.

[1] : Les modèles concernant des composantes spécifiques du bien-être, comme la satisfaction dans certains domaines de vie, ne seront pas évoqués dans ce chapitre pour plusieurs raisons. La principale est qu'ils sont très spécifiques et qu'il n'est donc pas évident de les transposer à un niveau plus général. Une autre est que l'exercice qui consisterait à passer en revue l'ensemble de ces travaux est pratiquement insurmontable.

Les raisons qui ont conduit les auteurs à proposer autant de modèles ne sont sans doute pas étrangères aux résultats qui montrent une indépendance partielle ou totale entre le bien-être et les conditions de vie objectives, présentés dans le chapitre précédent. Ces résultats, qui expriment ce que certains ont appelé le paradoxe de la satisfaction, ont suscité de nombreux commentaires et engendré des explications diverses. Ainsi, Olson & Schober (1993) distinguent des explications techniques et des explications comportementales. Parmi les premières, il a été fait référence au fait : que les mesures des conditions objectives n'étaient pas assez précises ; que les individus devaient évaluer des contenus auxquels ils n'avaient jamais pensé ; qu'il n'y avait pas de relation linéaire entre indicateurs objectifs et indicateurs subjectifs ; qu'il y avait des aspects affectifs dans les réponses qui n'étaient pas liés aux conditions objectives. Les secondes sont encore plus nombreuses. Elles partagent toutes le postulat selon lequel l'insatisfaction est un état psychologique qui ne peut être enduré longtemps par un individu à cause de la tension qui lui est associée. Citons rapidement :

- la désirabilité sociale contamine les réponses aux indicateurs subjectifs ;
- l'expression de son insatisfaction fait partie de la socialisation culturelle, celle-ci étant partiellement indépendante de l'expérience personnelle ;
- la satisfaction d'un individu ne s'élève que lorsque ses conditions s'améliorent significativement par rapport à celles d'un groupe de référence pertinent ;
- les personnes vivant dans des conditions privilégiées sont plus souples dans les standards de référence qu'elles adoptent ; elles sont donc plus critiques et plus insatisfaites ;
- le revenu n'est pas un indicateur valide du bien-être dans le groupe des individus qui vivent dans les meilleures conditions parce que la satisfaction des besoins élémentaires entraîne l'apparition de nouveaux besoins ;
- les individus qui vivent dans les plus mauvaises conditions se résignent et adaptent leur standard de référence à leur situation ;
- les standards de référence individuels ne sont pas déterminés par la situation : la perception de conditions objectives identiques ne se sera donc pas la même pour des individus différents ;
- des processus cognitifs comme la réduction de la dissonance cognitive peuvent influencer sur l'évaluation que font les individus de leur situation ;
- la satisfaction augmente avec l'âge, quelles que soient les conditions de vie.

Il n'y a pas de correspondance terme à terme entre les explications de cette liste et les modèles un tant soit peu formalisés que propose la littérature. Certaines d'entre elles n'ont même jamais donné lieu à vérification empirique, du moins à notre connaissance. Le seul intérêt de cette liste est de faire apparaître ce qui sera, nous allons le voir, l'un des points d'ancrage de nombreux modèles : les référents internes utilisés par les individus pour évaluer leur bien-être. Ainsi Zapf (1987) considère que le bien-être est fortement

déterminé par ses expériences personnelles et par les références à sa propre histoire ainsi qu'aux groupes sociaux auxquels les individus se comparent.

Etant donné la diversité des modèles, il serait souhaitable d'essayer de les évaluer en les confrontant systématiquement aux mêmes données empiriques. Malheureusement, de telles tentatives sont pratiquement inexistantes, comme le souligne Wright (1985) et, lorsqu'elles existent (cf. Emmons & Diener, 1985 a ; Diener et al., 1993), elles restent à un niveau beaucoup trop superficiel pour qu'il soit possible d'en tirer des conclusions sérieuses. Cette situation n'est sans doute pas due à la mauvaise volonté des chercheurs. Pour pouvoir confronter des modèles, il faut que ceux-ci soient suffisamment formalisés de manière à mettre en lumière leurs convergences et, surtout, leurs divergences. Nous n'étonnerons sans doute personne en disant que ce niveau de formalisation n'est presque jamais atteint. De plus, il n'est pas du tout exclu que certains modèles doivent être envisagés comme complémentaires plutôt que concurrents, étant donné qu'ils ne concernent pas exactement les mêmes aspects.

Afin de pouvoir préciser les spécificités de tous ces modèles, il semble indispensable d'en proposer une classification opérationnelle. C'est ce que nous tenterons de faire dans la section qui suit. Cette classification se démarque assez nettement de la manière traditionnelle d'aborder les théories du bien-être mais nous essayerons de nous en expliquer.

La troisième section sera consacrée à la présentation détaillée des modèles proprement dit. Nous avons abandonné l'idée de présenter de manière exhaustive toutes les études qui, souvent, se contentent de tester des sous-modèles. Aussi, seuls les modèles les plus généraux et les plus prototypiques seront analysés et critiqués. A cette occasion, nous serons amené à entreprendre des simulations pour tester le bien-fondé des méthodes mises en oeuvre. Il va sans dire qu'une théorie générale du bien-être n'a pas encore vu le jour. Pourtant, le modèle d'équilibre dynamique de Headey & Wearing (1989, 1991), s'il est encore bien incomplet, nous semble aller dans ce sens. Il fera l'objet de la cinquième section. Avant cela, il nous faudra examiner la validité d'un postulat largement accepté dans la plupart des modèles du bien-être que nous passerons en revue. Ce postulat consiste à assimiler le bien-être à une variable endogène. Cette question renvoie à un débat entre les partisans d'une approche "*top-down*" du bien-être et les partisans d'une approche "*bottom-up*"^[2]. Ce débat correspond approximativement à ce que nous avons déjà évoqué à propos des corrélats du bien-être, à savoir l'interprétation de corrélations comme des causes alors qu'elles pourraient être interprétées comme des effets. C'est parce que le modèle d'équilibre dynamique s'appuie notamment sur des résultats remettant en

[2] : On pourrait traduire respectivement ces deux expressions par approche descendante et approche ascendante. Cependant, comme les expressions américaines sont beaucoup plus évocatrices, nous continuerons à les utiliser.

cause la suprématie quasi absolue des modèles *bottom-up* que nous avons choisi de le présenter après avoir discuté de cette question.

5.2 Une proposition de classification des modèles explicatifs du bien-être

Peu d'auteurs se sont aventurés à proposer une classification des théories du bien-être. Il est souvent fait référence à telle ou telle famille de modèles dans la littérature, la plupart du temps à titre d'introduction théorique, mais ces présentations restent partielles et partiales notamment parce que les auteurs ne s'appuient, en général, que sur les travaux menés dans une perspective similaire à la leur^[3]. La classification que nous proposons n'échappe pas entièrement à ces limites. Les modèles que nous avons retenus sont exclusivement des modèles psychologiques (ou socio-psychologiques) et il n'est pas certain que telle ou telle variante spécifique y soit toujours parfaitement représentée. Toutefois, celle-ci peut être conçue comme une grille générale de lecture pouvant être enrichie et complétée par la suite.

Avant d'exposer nos critères de classification, il n'est pas inutile d'illustrer, par l'exemple, les distinctions entre modèles proposés dans la littérature. Deux exemples permettront de montrer que ces distinctions sont insuffisantes et d'introduire les caractéristiques de base des familles les plus répandues de modèles.

Au niveau des théories philosophiques, von Wright (1963, cité par Shin & Jonhson, 1978) distingue trois grandes classes de modèles :

- les modèles des idéaux épicuriens, dans lesquels le bonheur consiste en la possession de certaines choses qui procurent un plaisir passif. On reconnaîtra sans difficulté ici les travaux visant à mettre en relation le bien-être avec des conditions de vie matérielles, objectives, abordés dans le chapitre précédent.
- Les modèles des utilitaristes insistent sur la satisfaction des désirs et des besoins : la satisfaction des besoins est la cause du bonheur alors que la persistance de besoins non satisfaits entraîne le mal-être. Nous verrons que cette classe de modèles subsiste encore aujourd'hui à travers de nombreuses variantes. Alors que les théories précédentes sont plutôt normatives, celles-ci font place à la subjectivité des individus par l'intermédiaire de leurs désirs ou de leurs besoins.

[3] : Ceci tient surtout au fait, à notre avis, que le bien-être a été étudié, quelquefois sous d'autres noms, tant par différentes sous-disciplines de la psychologie que par différentes disciplines des sciences humaines. Nous avons déjà eu l'occasion de voir que même les travaux psychologiques (e.g. ceux sur les émotions et la structure de l'affectivité) n'avaient pas eu le retentissement que l'on aurait pu attendre dans le milieu de la recherche sur le bien-être.

- Une dernière catégorie regroupe tous les modèles qui voient, à la suite d'Aristote, la source du bonheur dans l'activité qui accompagne l'accomplissement de ses propres capacités. Plus dynamique que les précédentes, cette catégorie est aussi plus difficile à appréhender. On dirait aujourd'hui que ce n'est pas la satisfaction d'un besoin en elle-même qui est la cause du bonheur, mais l'activité qui a été déployée pour y parvenir, à condition que cette activité ait permis à l'individu d'exprimer pleinement ses compétences et ses habiletés.

Bien que largement antérieure à la plupart des travaux scientifiques sur le bien-être, cette classification fait apparaître des catégories dans lesquelles s'inscrivent encore nombre de modèles actuels. Ceux-ci se sont précisés, bien sûr, mais nous verrons que l'idée sous-jacente est la même. Cette classification est pourtant incapable de rendre compte de développements plus récents qui, à partir de bases identiques, se sont enrichis en faisant appel à différentes variables intermédiaires qui revêtent une importance théorique considérable. En faisant intervenir beaucoup plus directement l'individu, à travers son activité cognitive, ces modèles sont en rupture avec les théories classiques. Ils constituent d'ailleurs aujourd'hui les modèles les plus répandus, sous des formes assez diversifiées.

Le complément le plus notable apporté à cette classification des théories philosophiques est celui que propose Diener (1984), dans sa revue de la littérature. Cette nouvelle classification essaie d'intégrer les études théoriques et empiriques menées dans le cadre des sciences sociales. Elle est résumée dans le tableau 2.11.

Théories	Description
Téléologiques	Il y a bien-être lorsqu'un but est atteint ou lorsqu'un besoin est satisfait
Plaisir et Douleur	Bonheur et malheur sont liés. Des phases de bien-être sont forcément précédées de phases de mal-être correspondant à des besoins non satisfaits
de l'Activité (ou auto-téléologiques)	Le bien-être dérive de l'activité déployée pour parvenir à certains états finaux (buts ou besoins)
<i>bottom-up vs top-down</i>	Le bonheur est la somme de multiples petits plaisirs ou de petites satisfactions vs il existe une propension générale à appréhender la vie de manière positive
Associationnistes	Ensemble de modèles ayant en commun le fait "qu'ils cherchent à expliquer pourquoi certains individus ont un tempérament qui les prédispose au bonheur" à partir de travaux sur la mémoire, le conditionnement ou les processus cognitifs (p. 566)
du Jugement	Le bonheur résulte d'une comparaison entre les conditions actuelles et un standard de comparaison. Les modèles se distinguent en fonction du standard utilisé.

Tableau 2.11 : Les théories du bien-être d'après Diener (1984)

Ainsi qu'on peut s'en apercevoir à la simple lecture du tableau, cette classification n'est guère satisfaisante. Elle fait apparaître des distinctions qui n'en sont pas vraiment et confond des niveaux différents. Ainsi en est-il, par exemple, des théories "plaisir-douleur" dont les différences avec les théories téléologiques ne sont pas claires. A la limite, le fait que le malheur précède le bonheur peut être conçu comme une hypothèse spécifique qui dérive de théories téléologiques, mais de là à expliquer, avec cette hypothèse, pourquoi les gens sont heureux, il semble qu'il y ait un fossé difficile à combler. Le second exemple concerne la distinction entre théories *bottom-up* et *top-down*. Il est manifeste, au vu de la description qui en est donnée, que cette catégorie n'est nullement homogène et qu'elle renvoie à des caractéristiques (ou des postulats) susceptibles de s'appliquer à tout modèle plutôt qu'à des hypothèses sur l'origine du bonheur. Sans entrer dans les détails, il paraît clair en effet que les théories téléologiques peuvent être qualifiées de théories *bottom-up* alors que les théories dites associationnistes, telles qu'elles sont décrites dans le tableau, sont des théories *top-down*. En bref, il nous semble que la classification proposée par Diener n'a pas les qualités que l'on pourrait en attendre. Au lieu de classer les modèles dans des catégories homogènes exclusives, elle s'appuie sur des distinctions qui correspondent à des niveaux de spécificité très différents selon les catégories. Si le but premier d'une classification est de fournir une grille de lecture simple et fonctionnelle, celle de Diener ne remplit pas son rôle.

La classification alternative sur laquelle nous allons nous baser est élaborée à partir de cinq grandes familles de modèles, d'inégale importance, dont la plupart peuvent être subdivisées en plusieurs sous-modèles. Il s'agit des modèles :

- basés sur l'influence des conditions de vie ;
- basés sur l'influence des changements dans les conditions de vie ;
- basés sur l'influence des traits de personnalité ;
- constructivistes ;
- spécifiques, regroupés dans une catégorie résiduelle.

La représentation graphique présentée met en relief les principales caractéristiques de chacune des familles de modèles (figure 2.20).

Les modèles basés sur l'influence des **conditions de vie** se distinguent selon qu'ils font l'hypothèse d'une influence directe ou selon qu'ils font appel à des variables intermédiaires. Les premiers sont pratiquement abandonnés. En effet, les multiples résultats que nous avons résumés dans le chapitre précédent montrent de façon consistante que le bien-être n'est pas ou que fort peu lié aux conditions de vie objectives dans lesquelles se trouvent les individus. En réponse à cette indépendance, les auteurs ont essayé d'imaginer des variables intermédiaires venant médiatiser l'influence des conditions de vie sur le bien-être. Ces variables intermédiaires sont assimilées à des

référénts internes qui permettent aux sujets d'évaluer leurs conditions de vie. Certains modèles, inspirés des travaux sur la comparaison sociale, font appel à des référents interpersonnels. D'autres utilisent des référents intrapersonnels qui découlent de manière plus ou moins directe des notions de soi réel et de soi idéal. Dans tous les modèles, ces référents servent de standards par rapport auxquels les individus confrontent leur propre situation et c'est de cette comparaison que découle un certain niveau de bien-être. Parmi les modèles de cette famille, les plus célèbres sont celui de Campbell et al. (1976), généralement désigné sous le nom de modèle subjectif de la qualité de la vie (cf. Olson & Schober, 1993) et celui des discordances multiples^[4] de Michalos (1985). Tous deux utilisent plusieurs référents.

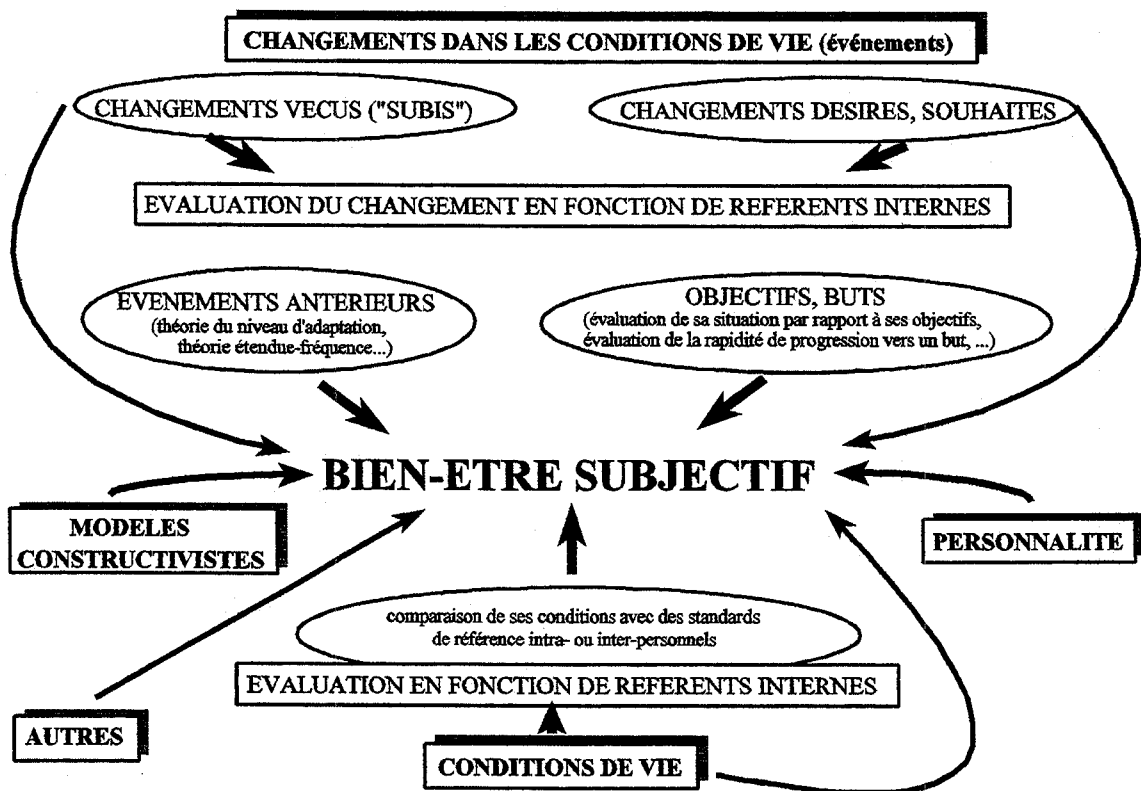


Figure 2.20 : Représentation schématique des différents modèles explicatifs du bien-être subjectif

Alors que les modèles précédents sont essentiellement statiques, la deuxième famille comprend des modèles plus dynamiques puisqu'ils sont basés sur les **changements dans les conditions de vie**. Certes, il y a une part d'arbitraire dans cette distinction étant donné que les conditions de vie ne sont jamais fixées définitivement. Il

[4] : C'est une traduction littérale de "multiple discrepancies model".

ne s'agit donc pas d'opposer des chercheurs selon qu'ils acceptent ou non l'idée de modifications des conditions de vie, mais plutôt de mettre en lumière des différences portant sur l'importance accordée à ces changements. Tous les modèles de cette classe peuvent être conçus comme une autre forme de réponse à l'indépendance entre bien-être et conditions de vie : si celles-ci sont incapables de rendre compte du niveau de bien-être des individus, il est possible que les changements le soient.

Une première distinction peut être opérée, au sein de cette vaste famille, entre ce que nous avons appelé les *changements vécus* et les *changements désirés*. Les premiers renvoient à une tradition de recherche, particulièrement développée dans le milieu psychiatrique, concernant l'influence des événements de vie sur le stress et, plus généralement, sur la santé mentale. Dans cette tradition, les événements de vie sont principalement conçus comme subis par les individus : classiquement, on comptabilise, à l'aide de questionnaires rétrospectifs, les événements de vie positifs et/ou négatifs vécus et on étudie leur influence sur des indicateurs de santé mentale. Les changements désirés renvoient à ce que Diener (cf. tableau 2.11) regroupe sous les appellations de théories téléologiques et de théories de l'activité (ou auto-téléologiques). Ces modèles envisagent le bien-être comme résultant de l'atteinte de buts ou de la satisfaction de besoins. Ils concernent donc bien des changements dans les conditions de vie puisque le bien-être est sous la dépendance de modifications de la situation d'un individu par rapport à ses objectifs ; de plus, ces changements sont souhaités, que ce désir soit conscient, comme dans le cas des buts, ou inconscients, comme dans le cas des besoins.

Les modèles de la famille "changements dans les conditions de vie" peuvent aussi être distingués selon que l'influence est directe ou médiatisée par des variables intermédiaires. En ce qui concerne les changements vécus, les variables intermédiaires font principalement intervenir *l'évaluation du changement* en fonction d'*événements antérieurs*. Le nouvel événement peut être évalué en fonction de la moyenne des événements antérieurs, comme le supposent les partisans de la théorie du niveau d'adaptation, ou en fonction d'une valeur située à mi-chemin entre la moyenne et la médiane des événements précédents pour la théorie "étendue-fréquence". En ce qui concerne les changements désirés, les auteurs ont surtout fait référence aux évaluations que font les individus des *objectifs* qu'ils se sont fixés et à la rapidité avec laquelle les individus progressent vers un but.

C'est au sein de ces deux premières familles de modèles que l'on peut situer l'essentiel des travaux théoriques et empiriques sur le bien-être. Mais pour être complet, il faut ajouter de rares études dans lesquelles ce sont les traits de **personnalité** qui deviennent les principales causes du bien-être. Les modèles que nous avons appelés **constructivistes** sont un peu à part. Ils font référence à des travaux, dont nous avons déjà vu plusieurs exemples (cf. section 2.3), menés dans le cadre de la cognition sociale.

Comme ils font largement appel à des influences transitoires contextuelles, ils répondent plus, à notre avis, à la question "pourquoi les individus répondent-ils qu'ils sont heureux (ou malheureux) à des items de bien-être ?" qu'à la question "pourquoi les individus sont-ils heureux (ou malheureux) ?". C'est pourquoi ces modèles, qui se limitent principalement aux propositions de Schwarz & Strack (1991), seront envisagés séparément. Enfin, nous avons introduit une **catégorie résiduelle** pour recevoir des micro-théories relativement spécifiques ou des considérations théoriques générales qui n'ont jamais vraiment donné lieu à des vérifications empiriques. Citons quelques exemples :

- L'adéquation entre les besoins, les valeurs d'une personne et les opportunités fournies par l'environnement pour les satisfaire est l'une des formes des théories de l'adéquation personne-environnement (Caplan, 1983 cité par Furnham, 1991). Cette théorie est souvent évoquée dans les travaux sur le bien-être et elle pourrait constituer une réelle approche interactionniste, mais il n'y a pas, à notre connaissance, d'études empiriques qui s'en inspirent directement.
- L'analyse transactionnelle de Berne (1964) a été utilisée par Mills (1986) pour montrer l'importance des contacts avec autrui pour le bien-être.
- Morawetz (1977) s'appuie sur les théories économiques de l'utilité pour montrer que, toutes choses égales par ailleurs, la satisfaction avec le revenu est d'autant plus faible que la distribution des revenus à l'intérieur d'une communauté est inégale.
- Beaucoup de travaux menés sur les personnes âgées font référence, explicitement ou non, aux théories de l'activité sociale (cf. Lemon et al., 1972, cités par Okun et al., 1984).

Toute classification recèle une part d'arbitraire. Celle-ci n'échappe pas à la règle. Pour qu'elle réponde à ses objectifs, il est nécessaire de caricaturer certains travaux, alors que la réalité est souvent beaucoup plus complexe, car il y a des postulats partagés, des influences mutuelles et donc des modèles mixtes.

Par exemple, certains auteurs ont proposé d'inclure des variables de personnalité dans des modèles basés sur l'influence des conditions de vie. Mais ces variables n'ont en général qu'un rôle accessoire. Elles contribuent à améliorer le pourcentage de variance expliquée, sans avoir de bases théoriques solides. Aussi, ces modèles restent-ils fondamentalement attachés à des explications du bien-être en termes de conditions de vie médiatisées par des référents internes. C'est la raison pour laquelle nous les rattachons à cette famille. Des remarques du même ordre peuvent être faites à propos des modèles

basés sur les traits de personnalité : si certains testent des modèles qui comportent d'autres types de variables, leur influence reste annexe dans la théorie adoptée par les auteurs.

Une similitude nette entre les familles dégagées concerne les variables intermédiaires. Il n'y a pas de différences fondamentales dans les variables utilisées dans les modèles qui se basent respectivement sur les conditions de vie et sur les changements dans ces conditions. Certains auteurs ne font d'ailleurs pas la distinction entre ces deux familles. Quand ils parlent de référents internes (Guttek et al., 1983), de théories du jugement (Diener, 1984 ; Meadow et al., 1992) ou de théories relativistes (Veenhoven, 1991 b), tous considèrent des modèles dans lesquels il y a une opération cognitive de comparaison entre deux éléments. L'un des éléments est toujours un standard de référence, même s'il diffère selon les théories ; l'autre est constitué soit par la situation, soit par le changement de situations, selon la famille de modèles. On comprend donc l'origine de ces appellations : ce sont des théories du jugement parce qu'elles impliquent une évaluation cognitive ; ce sont des théories relativistes parce qu'une situation ou un événement n'est pas évalué dans l'absolu, mais relativisé dans son impact en fonction de référents internes.

A travers cette classification, une première impression qui se dégage est que les théories du bien-être sont, en grande majorité, situationnistes plutôt que personnologiques. C'est le plus évident dans le cas des modèles basés sur une influence directe des conditions de vie ou des changements vécus. Pourtant, il est remarquable que, malgré la saillance de ce constat, les théories du bien-être n'aient jamais été envisagées sous cet angle. Ce secteur de recherche a toujours été à l'écart des controverses qui ont agité les travaux faisant intervenir les traits de personnalité ou les attitudes^[5]. Pourtant, tout comme les explications exclusivement personnologiques du comportement ont montré leurs limites, nous verrons que les explications exclusivement situationnistes du bien-être ont aussi les leurs. On peut toutefois se demander dans quelle mesure les modèles indirects du bien-être, qui font intervenir des référents internes, ne pourraient pas jouer le même rôle que les théories interactionnistes dans les modèles du comportement. Cela n'est pas exclu, mais il nous semble tout de même qu'il y a une différence de taille entre ces deux approches théoriques. Alors que les théories interactionnistes font appel à des différences interindividuelles stables qui viennent interagir avec les caractéristiques de la situation, les modèles indirects du bien-être se contentent de faire intervenir des variables ponctuelles, pour lesquelles la question de la stabilité temporelle et

[5] : Ce rapprochement entre l'histoire des recherches sur le bien-être, la personnalité et les attitudes pourra peut-être étonner. Il est vrai que les controverses qui ont agité le milieu de la psychologie pendant les années 70 concernaient l'explication de *comportements* à partir de différences individuelles stables comme les traits de personnalité ou les attitudes. Que le bien-être ne soit pas un comportement ne fait guère de doute. Mais qu'il soit envisagé comme un état, comme une sensation, comme une expérience de la conscience, on ne voit pas pourquoi la forme des théories devant être développées pour en rendre compte serait fondamentalement différente de celle des théories censées expliquer le comportement.

situationnelle ne se pose pas ; de plus et surtout, ces variables sont envisagées en superposition avec les conditions ou les événements de vie plutôt qu'en interaction. La mise au point de modèles réellement interactionnistes devrait constituer, à notre avis, l'une des priorités de la recherche sur le bien-être pour les années futures. Ainsi pourraient-êtré testées des conceptions déjà exprimées par des philosophes comme Tatarkiewicz (1976, p. 176), pour qui il est essentiel que les sources du bonheur rencontrent des dispositions qui permettent à la personne d'en profiter.

5.3 Les modèles du bien-être

Cette section sur les modèles du bien-être comprendra surtout une discussion sur les deux plus grandes familles : ceux qui se basent sur les conditions de vie et ceux qui s'appuient sur les changements. Les modèles qui s'intéressent à l'influence des traits de personnalité sont peu nombreux, à l'inverse des précédents. Ce n'est que depuis ces dernières années que, le renouveau des recherches sur la personnalité aidant, on commence à voir des travaux qui dépassent de simples constats corrélationnels. Les modèles constructivistes seront présentés à partir de l'article de Schwarz & Strack (1991) dans lequel les auteurs proposent une synthèse de leurs propres travaux. Enfin, les modèles de la catégorie résiduelle, trop spécifiques, ne seront pas décrits plus en détail.

5.3.1 Bien-être général et satisfactions spécifiques

Avant d'entrer dans le détail des modèles explicatifs du bien-être, il paraît utile de relever et d'examiner un postulat — explicite chez certains, implicite chez d'autres — que partagent presque toutes les théories. Le bien-être général serait fonction de satisfactions spécifiques à propos des différents domaines de la vie : on est heureux en général à condition d'être heureux dans son travail, dans sa famille, d'être satisfait de ses revenus, de sa santé, etc... Les différences entre les modèles ne s'expriment que quant à la manière dont sont dérivées les satisfactions plus spécifiques.

Campbell et al. (1976) sont parmi les premiers à défendre l'idée selon laquelle le bien-être général est fonction de satisfactions dans différents domaines de vie. Ils sont finalement assez surpris de constater qu'une simple régression linéaire additive permet de prédire 54% de la variance du bien-être général à partir de satisfactions dans 17 domaines de vie. D'autres modèles plus complexes, tenant compte notamment des relations non linéaires ne permettent d'améliorer que de manière insignifiante ce pourcentage. C'est à des conclusions identiques qu'aboutissent Andrews & Withey (1976). Un ensemble de 12 items de satisfactions spécifiques permet d'expliquer 45% de la variance totale d'un indicateur de bien-être général. Les auteurs vérifient aussi que la valeur prédictive du modèle n'augmente pas par la prise en compte de relations non linéaires, d'interactions entre variables ou par la mise en oeuvre de différents modes de pondération des satisfactions spécifiques.

Pourtant, dans ces deux études, un résultat intrigue les chercheurs. Les associations bivariées entre les satisfactions spécifiques et le bien-être général font apparaître un pattern assez caractéristique : les corrélations les plus élevées sont obtenues pour les domaines de vie les plus proches des sujets. La hiérarchie constatée entre les domaines

de vie est d'ailleurs sensiblement identique à celle que l'on obtient par le calcul des distances à partir du centre de la configuration, dans des analyses structurales de ces items par échelonnement multidimensionnel. Ces résultats amènent les auteurs à se demander si les modèles ne pourraient pas être améliorés en tenant compte de l'importance que les individus accordent à chacun des domaines de vie^[6]. L'idée est simple : la satisfaction dans un domaine devrait contribuer d'autant plus au bien-être général que celui-ci est évalué comme important. On passe ainsi de modèles dits additifs à des modèles dits multiplicatifs (puisque dans les régressions multiples, on utilise souvent le *produit* des évaluations d'importance et de satisfaction concernant un domaine ; il s'agit en fait de modèles additifs avec pondération). Les observations recueillies par Campbell et al. (1976) permettent de tester cette hypothèse^[7]. Contrairement aux attentes, les modèles multiplicatifs ne s'avèrent pas meilleurs que les simples modèles additifs.

Malgré ces résultats décourageants, d'autres continueront à se référer à ces modèles, que ce soit pour le bien-être ou pour des concepts plus spécifiques. Ainsi, en appliquant le même raisonnement, on peut essayer de prédire la satisfaction avec le travail à partir de satisfactions concernant des aspects du travail, pondérées par leur importance. Les résultats ne se sont pas avérés plus concluants que pour la prédiction de la satisfaction générale (cf. Quinn & Mangione, 1973, cités par Organ & Near, 1985). Une application originale de cette hypothèse a été proposée par Near & Rechner (1993) dans une étude portant sur des indicateurs agrégés. Selon ces auteurs, si le modèle multiplicatif est correct, la corrélation entre la satisfaction avec le travail et le bien-être général devrait être plus importante dans les pays où le travail est fortement valorisé. Les résultats rapportés par les auteurs ne sont pas meilleurs que ceux qui ont été obtenus au niveau individuel.

Soupçonnant d'éventuelles imperfections méthodologiques qui viendraient infirmer une hypothèse pourtant intuitivement fort plausible, Mastekaasa (1984) entreprend une évaluation systématique des modèles additifs et multiplicatifs à l'aide d'une méthodologie plus rigoureuse que celles qui sont employées habituellement (méthodes de régression plus adéquates, mesures ipsatives de l'importance pour contrôler l'acquiescement et la désirabilité sociale, etc.). Les résultats sont mitigés : le modèle multiplicatif semble meilleur pour la satisfaction avec le travail, mais pas pour la satisfaction générale. Lorsque l'on examine les résultats dans le détail, on s'aperçoit que

[6] : Voir aussi Bharadwaj & Wilkening (1977), pour qui les prédicteurs de la satisfaction générale ne sont pas les mêmes pour tous car ils varient selon les intérêts et les besoins des individus.

[7] : Classés par ordre décroissant d'importance moyenne, on trouve les domaines suivants : santé, mariage, vie de famille, gouvernement, amis, logement, travail, ville, croyance religieuse, loisirs, argent, associations. A quelques exceptions près, cette hiérarchie est parallèle à celle obtenue à partir des études structurales et des analyses bivariées : elle exprime quelque chose de l'ordre de la distance psychologique.

le modèle multiplicatif, dans le cas de la satisfaction avec le travail, n'améliore que très faiblement le pourcentage de variance expliquée par un simple modèle additif.

Récemment, Marsh (1993) a proposé un article de synthèse sur cette question, appliqué non pas directement au bien-être, mais à l'estime de soi. Toutefois, il ne semble pas que les conclusions puissent être très différentes entre ces deux concepts. Selon l'auteur, William James proposait déjà en 1890 que l'évaluation qu'une personne fait d'elle-même reflète une moyenne pondérée de domaines spécifiques, les pondérations étant données par l'importance, la valeur attachée à chaque domaine. Cette idée a été largement reprise, surtout dans les travaux sur l'estime de soi. L'article auquel nous faisons allusion vise à recenser et à confronter les différents modèles proposés dans ce cadre. Il est frappant tout d'abord de constater la richesse de ces modèles : alors que les recherches sur le bien-être se sont contentées de modèles à pondérations simples, les modèles d'estime de soi ont été beaucoup plus élaborés. Cela n'est sans doute pas étranger au fait que les pondérations simples n'ont donné que des résultats médiocres. Pour se faire ne serait-ce qu'une idée de la diversité et de la complexité des modèles, on peut citer :

- les modèles sans pondération (l'évaluation globale est la simple somme d'évaluations plus spécifiques) ;
- les modèles avec pondérations constantes (les pondérations sont les mêmes pour tous les sujets) ;
- les modèles avec pondérations individuelles (une pondération différente est associée à chaque domaine pour chaque individu), parmi lesquels on peut distinguer les modèles basés sur :
 - * l'importance accordée par chaque sujet à chaque domaine ;
 - * la différence entre perception du soi réel et du soi idéal ;
 - * la combinaison de l'importance d'un domaine et de la différence entre soi réel et soi idéal ;
 - * la différence entre l'évaluation du domaine et l'importance accordée à ce même domaine ;
 - * le degré de confiance accordée à l'évaluation du domaine ;
 - * etc.

La conclusion des confrontations empiriques opérées par Marsh (1993) est claire : les modèles à pondérations constantes sont nettement meilleurs que les modèles à pondérations individuelles (même lorsque celles-ci sont très compliquées). Mais les modèles à pondérations constantes sont équivalents, voire moins bons, que les modèles sans pondération du tout. Dans tous les cas, les différences dans les pourcentages de variance expliquée sont très faibles.

Il nous semble qu'une première leçon peut être tirée de cette analyse critique. Il paraît inutile que les chercheurs dans le domaine du bien-être cherchent à développer des

modèles à pondérations individuelles plus complexes que l'importance du domaine. Plutôt que d'essayer de complexifier les modèles, on pourrait s'interroger sur leurs échecs. L'une des premières explications qui vient à l'esprit est que ce ne sont pas les satisfactions spécifiques qui influencent le bien-être général, mais le niveau de bien-être qui influe sur la satisfaction dans chaque domaine de vie. Une telle explication renvoie directement au débat entre les approches *top-down* et *bottom-up*. Nous examinerons donc sa plausibilité empirique plus loin dans ce chapitre.

5.3.2 Les modèles du bien-être basés sur les conditions de vie

Nous avons distingué, plus haut, les modèles basés sur les conditions de vie selon qu'ils font référence à une influence directe ou indirecte. Plus personne ne défend aujourd'hui l'idée d'une influence directe des conditions de vie^[8]. Nous ne discuterons donc pas ici les modèles qui se rattachent à cette conception et renvoyons le lecteur au chapitre précédent qui fait état de nombreux résultats montrant que les conditions de vie sont relativement indépendantes du niveau de bien-être subjectif.

Dans tous les modèles à influence indirecte, l'idée générale est que les conditions sont évaluées non pas dans l'absolu, mais en fonction de standards de référence qui sont propres aux sujets. Ces standards de référence sont arbitraires (cf. Veenhoven, 1991 b), c'est-à-dire qu'ils ne correspondent pas aux conditions objectives. Pour évaluer leur satisfaction, les individus effectuent une différence (le plus souvent) ou un rapport (rarement) entre les conditions telles qu'ils les perçoivent et un ou plusieurs standards de référence. Ceux-ci peuvent être inter-personnels (les conditions de vie sont comparées avec celles d'autres personnes) ou intra-personnels (les conditions de vie actuelles sont comparées avec celles déjà connues dans le passé ou attendues pour le futur). La plupart des auteurs essaient de rattacher de manière explicite, lorsque cela est possible, ces standards de références à des théories psychologiques bien établies. Le tableau 2.12 présente les standards les plus utilisés, des exemples d'auteurs qui en ont fait usage et les références théoriques auxquelles ils renvoient.

Nombre de ces standards ont été utilisés de manière isolée, dans des études spécifiques. Les résultats sont en général en faveur de l'hypothèse qui stipule que les individus sont d'autant plus satisfaits que la différence entre l'évaluation de leurs conditions et le niveau auquel elles sont comparées est faiblement négative ou positive selon les cas. Ces modèles seront illustrés par deux études prototypiques qui s'appuient toutes deux sur plusieurs standards de référence : le modèle subjectif de la qualité de la

[8] : On trouve encore dans certains travaux empiriques des modèles dans lesquels les conditions de vie sont considérées comme des variables externes, mais elles ne jouent qu'un rôle accessoire.

vie de Campbell et al. (1976) et le modèle des discordances multiples de Michalos (1985).

Standards de référence Différences entre sa situation actuelle et...	Exemples d'utilisation	Référents théoriques
STANDARDS DE REFERENCE INTRAPERSONNELS		
ce que l'on désire (aspiration)	Andrews & Whitey (1976) Campbell et al. (1976) Michalos (1985)	Théorie de l'aspiration (Lewin et al., 1944) Distinction entre soi réel et soi idéal
ce que l'on a eu de meilleur dans le passé	Campbell et al. (1976) Michalos (1985)	
ce que l'on attendait	Campbell et al. (1976) Michalos (1985)	Théorie de la dissonance cognitive de Festinger (1957)
ce dont on a besoin	Michalos (1985)	Théorie de l'adéquation personne- environnement (Caplan, 1983)
STANDARDS DE REFERENCE INTERPERSONNELS		
ce que les autres ont (groupe pertinent pour l'individu)	Andrews & Whitey (1976) Campbell et al. (1976) Michalos (1985)	Théorie du groupe de référence (Merton & Kitt, 1950) Théorie de la comparaison sociale (Wills, 1981)
ce que l'on mérite ^[9]	Carp & Carp (1982) ; Carp et al. (1982) ; Michalos (1985)	Théorie de l'équité (Adams, 1965)
ce que des autres "plus mal lotis" ont ^[10]	Heidrich & Ryff (1993a, 1993b)	

Tableau 2.12 : Exemples de standards de référence utilisés dans la littérature

A. Le modèle subjectif de la qualité de la vie de Campbell et al. (1976)

L'étude de Campbell et al. (1976) a eu un impact très important dans les recherches sur le bien-être. Plusieurs raisons y ont sans doute contribué : l'étude porte sur un échantillon représentatif de la population américaine, des données très nombreuses ont été recueillies, les analyses présentées sont très fines, etc. Mais la raison majeure est sans doute l'originalité du modèle proposé qui a inspiré, directement ou non, bon nombre de travaux plus récents. Sous sa forme la plus simple, le modèle peut s'exprimer de la

[9] : Ce standard de référence peut être conçu à la fois comme un standard interpersonnel et intrapersonnel.

[10] : "Downward comparisons".

manière suivante : la satisfaction dans un domaine dépend de la différence entre la perception de sa situation actuelle et un certain standard de référence auquel elle est comparée. Il est représenté dans la figure 2.21.

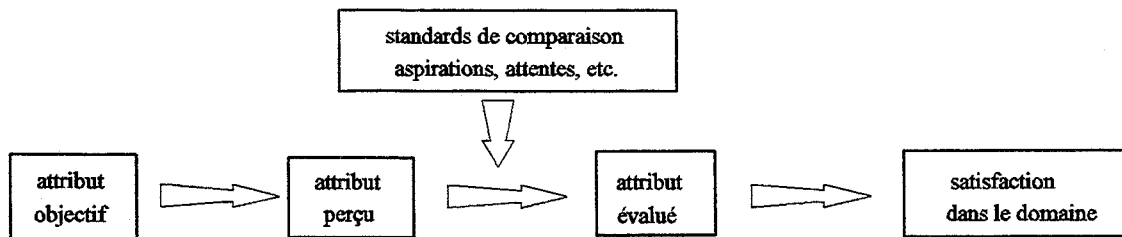


Figure 2.21 : Le modèle subjectif de la qualité de la vie de Campbell et al. (1976, p. 13)

Ce modèle a été testé dans deux domaines de vie, le logement et le voisinage, mais les auteurs estiment qu'il peut être étendu aux autres domaines. Une échelle de 0 à 100 est présentée aux sujets : le point 0 est censé représenter, dans l'exemple du logement, le pire logement qu'ils puissent imaginer et le point 100, le logement idéal, parfait. Les sujets sont invités à évaluer, sur cette échelle, leur logement actuel ainsi que des logements hypothétiques, en fonction de différents référents rapportés ci-dessous (la liste étant ordonnée en fonction de la moyenne des réponses, indiquée entre parenthèses) :

- le meilleur logement qu'ils peuvent espérer (aspiration) (90.7) ;
- le logement qu'ils pensent avoir dans cinq ans (expectation, attente) (82.9) ;
- celui qu'ils ont le plus apprécié dans leur vie (82.3) ;
- celui de leurs proches (77.6) ;
- celui de leurs amis (76.7) ;
- le leur actuellement (75.1) ;
- celui de l'américain moyen (71.7) ;
- celui qu'ils avaient il y a cinq ans (71.6) ;
- celui qu'ils avaient juste avant l'actuel (60.1) ;
- celui qu'ils ont le moins aimé dans leur vie (43.5) ;

Les auteurs calculent ensuite les différences entre l'évaluation des conditions actuelles et les autres évaluations et examinent les corrélations entre ces différences et la satisfaction dans les domaines^[11]. C'est, de manière assez nette, la différence entre l'évaluation de sa situation actuelle et son niveau d'aspiration qui est la plus fortement corrélée avec la satisfaction dans le domaine, même si les autres différences contribuent

[11] : Les auteurs effectuent aussi les mêmes opérations en se basant sur des rapports au lieu des différences. Comme ces deux estimations aboutissent à des résultats quasi identiques, il ne sera plus fait référence, dans la suite, qu'au modèle en termes de différences.

à améliorer le pourcentage de variance expliquée (qui est égal à 47% dans une régression pour un maximum possible de 60% à cause de la fidélité de la variable dépendante, dont 33% pour le niveau d'aspiration à lui seul). Ce résultat conduit les auteurs, après plusieurs analyses intermédiaires dont nous ne parlerons pas ici, à proposer une modification de leur modèle initial. Dans ce nouveau modèle, la satisfaction dépend de la comparaison entre sa situation actuelle et son niveau d'aspiration, ce dernier dépendant lui-même des évaluations effectuées en référence aux autres standards. Il est illustré par la figure 2.22.

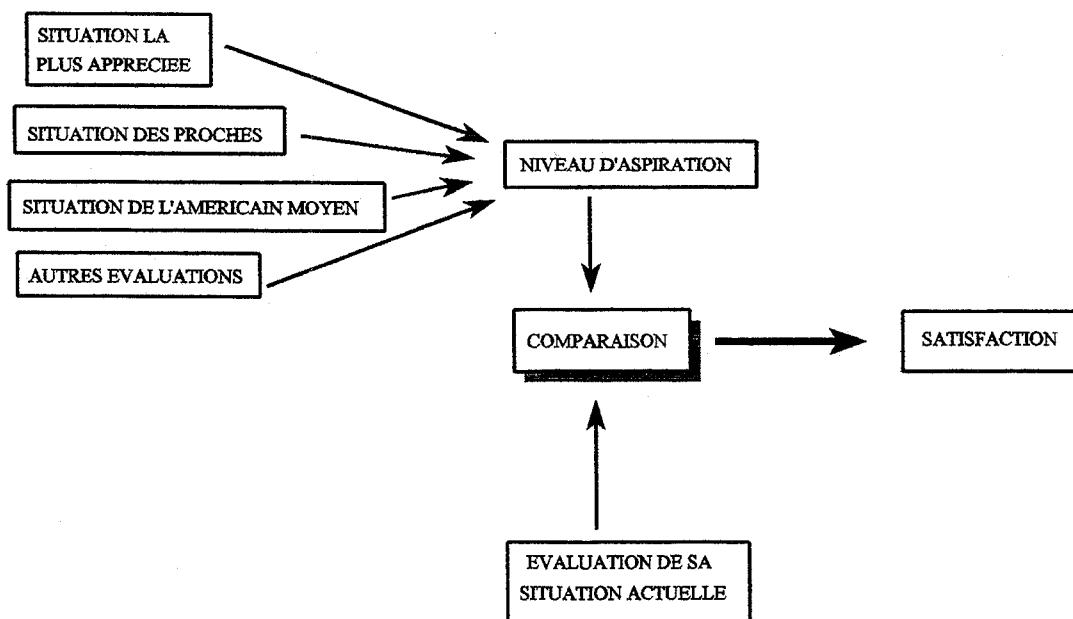


Figure 2.22 : Le modèle modifié de Campbell et al. (1976, p. 193).

Le meilleur prédicteur du niveau d'aspiration est la situation connue la plus appréciée. On trouve ensuite l'évaluation de la situation des proches et celle de l'américain typique. Les autres variables ne contribuent que très faiblement à la prédiction du niveau d'aspiration.

Le modèle de Campbell et al. (1976) a séduit beaucoup d'auteurs, même s'il a été critiqué. Par exemple, Carp & Carp (1982), en se basant sur la théorie de l'équité, estiment que la différence entre l'évaluation de sa situation actuelle et ce que l'on mérite est un meilleur prédicteur de la satisfaction que celle qui implique le niveau d'aspiration. Dans leur étude empirique, ils confirment cette hypothèse en procédant par des régressions dans lesquelles toutes les variables indépendantes sont entrées en même temps

plutôt que pas à pas, comme l'ont fait Campbell et al. (1976). Ce résultat sera affiné dans une étude ultérieure (Carp et al., 1982) qui montre que la différence "situation actuelle - situation méritée" est le meilleur prédicteur de la satisfaction chez les personnes âgées, alors que la différence "situation actuelle - situation aspirée" l'emporte chez les plus jeunes. D'autres critiques plus fondamentales ont été adressées au modèle de Campbell et al. (1976) ; nous les examinerons après avoir présenté le modèle de Michalos (1985).

B. Le modèle des discordances multiples de Michalos (1985)

Le modèle des discordances multiples proposé par Michalos (1985) est dans la lignée directe de celui de Campbell et al. (1976). Il est représenté sous forme schématique dans la figure 2.23. On peut voir que les différences perçues entre sa situation et différents standards de référence constituent également le noyau central de ce modèle.

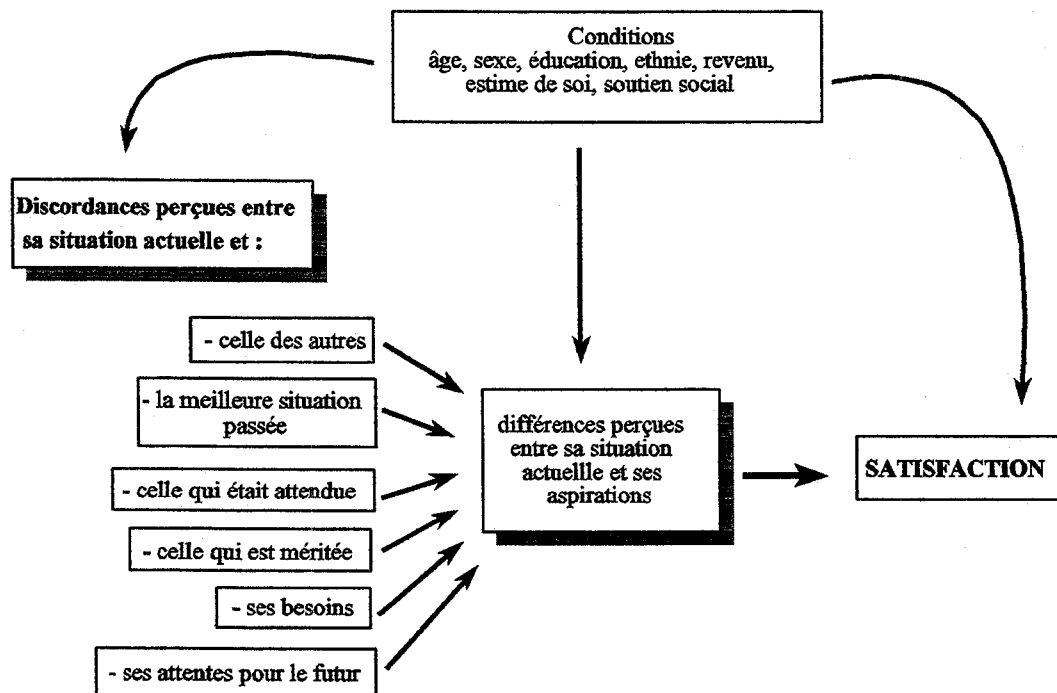


Figure 2.23 : Le modèle des discordances multiples de Michalos (1985, p. 357).

Nous ne discuterons pas ici des différentes "conditions" introduites par Michalos dans son modèle. Signalons seulement qu'en l'absence de justification théorique à ces choix, un tout autre ensemble de conditions, notamment en ce qui concerne les variables de personnalité, aurait aussi bien pu être introduit.

La différence théorique la plus importante entre les modèles de Michalos et de Campbell et al. concerne les relations entre les différentes comparaisons. Chez Campbell et al., les évaluations en fonction de différents standards de référence influencent directement le niveau d'aspiration et la comparaison de ce niveau avec sa situation actuelle prédit la satisfaction. Chez Michalos, la différence entre le niveau d'aspiration et la situation actuelle n'est plus influencée directement par les différentes évaluations, mais par les *discordances* entre la situation actuelle et les différents standards de référence. Une autre différence tient aux standards de référence utilisés. Cependant, ces changements nous paraissent mineurs par rapport à l'idée générale commune aux deux modèles.

En revanche, il y a quelques nouveautés intéressantes sur les plans empirique et méthodologique. Tout d'abord, le modèle de Michalos a été testé systématiquement pour de nombreux domaines de vie^[12], alors que Campbell et al. (1976) s'étaient limités aux domaines du logement et du voisinage. La validité du modèle s'en trouve donc améliorée. Mais la différence la plus notable concerne la manière dont les différences sont établies. Chez Campbell et al. (1976), les discordances étaient calculées à partir des réponses des sujets. Chez Michalos (1985), elles ne sont plus calculées, les sujets y répondent directement. Autrement dit, pour une discordance donnée, les sujets ne répondent plus à deux questions, mais à une seule^[13]. La première remarque qui vient à l'esprit en prenant connaissance de cette procédure est que les items de discordance (qui constituent les variables indépendantes) et les items de satisfaction (les variables dépendantes) mesurent pour ainsi dire la même chose^[14] ! Dans la discussion de ses résultats, Michalos reconnaît que cette critique lui a été formulée par McKennell, dans un échange de correspondance. Mais celle-ci ne lui paraît pas justifiée. Pour en évaluer la portée, il propose d'examiner les relations entre ces deux types de variables et une troisième variable à laquelle l'une des premières est théoriquement liée. Si les items de satisfaction et de discordances mesurent la même chose, écrit Michalos, on devrait trouver le même réseau de relations pour les deux types de variables. Le corpus de données recueillies lui permet de trouver 952 cas dans lesquels il peut vérifier cette hypothèse, en comptant le nombre de fois que les items de satisfaction et de discordance sont reliés significativement à une troisième

[12] : Dans son article de 1985, Michalos note que 2184 coefficients de régression ont été estimés !

[13] : Cela donne, pour la discordance "situation actuelle - aspiration" dans le cas de la vie en général, l'item suivant : "*Consider your life as a whole. How does it measure up your general aspirations or what you want ?*". Les réponses se font sur une échelle en sept points. Les variables dépendantes sont mesurées avec des items du type : "*How do you feel about your life as a whole right now ?*".

[14] : On peut aussi soupçonner des problèmes de fidélité étant entendu que chaque mesure n'est obtenue qu'avec un seul item.

variable. Comme cela ne se produit que dans 32% des cas, Michalos conclut que ces deux types de variables ne mesurent pas la même chose^[15].

La justification apportée par Michalos à cette critique nous semble peu fondée. En effet, on ne voit pas très bien pourquoi c'est un seuil de significativité totalement arbitraire qui est choisi pour décider si deux relations sont identiques ou non. Cette procédure permet, par exemple, de rejeter l'identité de deux relations qui seraient significatives respectivement à $p = .04$ et à $p = .06$. Dans ce cas, l'intensité des relations est pratiquement identique (à condition que le nombre de sujets soit le même dans les deux cas), mais la fixation d'un seuil arbitraire conduit à les considérer comme différentes.

De plus, même si on admet que les items de satisfaction et de discordances ne mesurent pas la même chose, on peut se demander quel est le sens de la relation. Il est tout à fait plausible en effet de faire l'hypothèse que c'est le niveau de satisfaction qui fait que l'on perçoit de faibles différences entre sa situation actuelle et son niveau d'aspiration, plutôt que l'inverse. Cette hypothèse sera examinée en détail, plus bas, dans le cadre des approches *top-down* et *bottom-up*. Nous allons maintenant présenter, de manière plus générale, les critiques qui ont été adressées aux modèles d'influence indirecte des conditions de vie en commençant par celui de Campbell et al., qui est plus fondamental.

[15] : Il faut tout de même noter qu'il procède de la même manière pour tester l'ensemble de son modèle et qu'il aboutit à une conclusion opposée alors que seulement 771 relations sont significatives sur les 2184 estimées, soit 35%.

C. Critiques adressées aux modèles d'influence indirecte des conditions de vie

Le modèle de Campbell et al. (1976) a fait l'objet de critiques méthodologiques. L'une d'entre elles retiendra particulièrement notre attention^[16]. Elle consiste à dire (cf. Carp & Carp, 1982 ; Carp, Carp & Millsap, 1982) qu'il est fallacieux de tester un modèle dans lequel une même variable figure à la fois comme variable dépendante et comme variable indépendante (sous forme d'une différence avec une autre variable). En effet, la corrélation entre les évaluations de sa satisfaction dans un domaine et de sa situation actuelle est très élevée (aux alentours de .80), ce qui laisse penser, là aussi, que ces deux variables mesurent la même chose. Ce résultat n'a d'ailleurs pas échappé aux auteurs qui testent indifféremment des modèles dans lesquels la variable dépendante est le niveau de satisfaction ou l'évaluation de sa situation actuelle. Les résultats étant quasi identiques dans les deux cas, nous limiterons notre discussion au cas simple où la variable dépendante est l'évaluation de sa situation actuelle^[17].

Si cette critique semble légitime, elle reste hypothétique : on ne sait pas quel peut être son impact réel sur les résultats. Pour évaluer cet impact, nous avons eu recours à des simulations sur données aléatoires. Ce qui nous préoccupe est de savoir si les résultats obtenus par Campbell et al. peuvent être dus à des artefacts liés à la méthode utilisée.

Dans un premier temps, nous avons travaillé sur deux variables obtenues par un programme informatique de génération aléatoire de nombres. Les contraintes suivantes ont été fixées : les distributions des deux variables sont toutes deux gaussiennes, ont une même moyenne (100) et un même écart-type (15). Ces données ont été générées pour 1000 individus hypothétiques afin d'assurer une certaine stabilité aux calculs et les simulations ont été répétées 20 fois. Nous avons ensuite calculé la corrélation entre ces deux variables (appelons-les x et y) puis la corrélation entre l'une d'elles et la différence entre les deux (entre y et y - x). Les résultats de ces simulations sont présentés dans le tableau 2.13.

On remarque tout d'abord que les coefficients obtenus lors de ces 20 simulations sont très proches les uns des autres. Cette stabilité permet de se limiter à un petit nombre de simulations. Les corrélations entre x et y oscillent autour de zéro, ce qui était attendu étant donné la nature aléatoire des variables. En revanche, les corrélations entre y et

[16] : L'autre critique méthodologique avancée concerne la fidélité des scores de différence. Bien que justifiée, elle ne peut pourtant remettre en cause les conclusions positives tirées par Campbell et ses collaborateurs. En effet, le manque de fidélité ne peut qu'atténuer la corrélation entre deux variables et non pas l'augmenter.

[17] : Formellement, le modèle s'écrit, dans le premier cas : $y' = \beta (y - x) + \alpha + \epsilon$ et, dans le second cas, $y = \beta (y - x) + \alpha + \epsilon$, avec $r_{yy'} \approx .80$. Les simulations qui sont présentées par la suite portent uniquement sur le second cas de figure, mais les résultats sont tout à fait similaires lorsque l'on se place dans le premier cas. La démonstration ne perd donc pas son caractère de généralité.

(y - x) sont très élevées. Autrement dit, on peut toujours mettre en évidence une relation entre une variable et la différence entre cette même variable et une autre variable, même quand ces variables ne sont aucunement liées entre elles^[18].

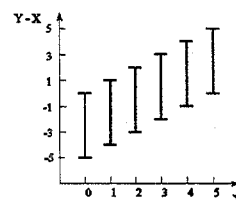
simulation	$r_{y,x}$	$r_{y,y-x}$	simulation	$r_{y,x}$	$r_{y,y-x}$
1.	-.02	.71	11.	.03	.70
2.	-.03	.70	12.	-.01	.69
3.	-.03	.73	13.	-.02	.71
4.	-.00	.71	14.	.01	.73
5.	-.03	.71	15.	-.03	.71
6.	.04	.70	16.	-.04	.72
7.	-.03	.71	17.	-.01	.69
8.	.06	.71	18.	-.02	.72
9.	.03	.68	19.	-.05	.71
10.	.08	.69	20.	.02	.69

Tableau 2.13 : Corrélations entre y et x et entre y et (y - x) sur des données simulées (y et x ont même moyenne, même écart-type et se distribuent normalement).

Cependant, la situation qui vient d'être décrite est quelque peu artificielle : les données générées ne correspondent pas à celles que l'on obtient avec des dispositifs réels. En particulier, il est très rare que les variables aient même moyenne et même écart-type. Pour tenir compte de cette objection éventuelle, nous avons effectué une nouvelle série de simulations qui, tout en conservant la génération aléatoire d'observations, obéit à un certain nombre de contraintes. Quatre variables sont utilisées, correspondant aux spécifications suivantes :

variable	moyenne	écart-type
y	75	15
x_1	90	10
x_2	60	15
x_3	60	25

[18] : Cela se comprend facilement si on regarde le graphique ci-contre. Admettons que deux variables Y et X soient mesurées sur une même échelle variant de 0 à 5. Les valeurs extrêmes de la différence (Y - X) dépendent forcément de la valeur de Y. Ainsi, si Y est égal à 0, la différence (Y - X) ne peut varier qu'entre -5 (i.e. 0 - 5) et 0 (i.e. 0 - 0) et ainsi de suite. On obtient donc un graphique qui montre qu'il y a nécessairement une relation linéaire forte entre les variables (cf., à ce sujet, Wall & Payne, 1973, cités par Wright, 1985).



La génération de ces données se fait toujours à partir d'une distribution gaussienne, mais elle répond à une contrainte supplémentaire : les données doivent se situer dans l'intervalle [0 ; 100], pour tenir compte de l'échelle de réponses utilisée par les auteurs (elles sont, de plus, arrondies à la valeur entière la plus proche). Concrètement, cela s'obtient en générant une nouvelle donnée aléatoire tant que la valeur obtenue ne se trouve pas dans l'intervalle désiré. Les moyennes choisies correspondent approximativement aux valeurs observées par Campbell et al., respectivement pour l'évaluation de sa situation actuelle ($m = 75$), le niveau d'aspiration ($m = 90$) et l'évaluation de la situation précédente ($m = 60$). Les écarts-types sont choisis de manière à différer les uns des autres, sans correspondre aux données réelles car les auteurs n'indiquent pas leur valeur. Les résultats des 20 simulations, chacune portant sur 1000 observations, figurent dans le tableau 2.14.

simulation	$r_{y,x1}$	$r_{y,y-x1}$	$r_{y,x2}$	$r_{y,y-x2}$	$r_{y,x3}$	$r_{y,y-x3}$
1.	.04	.86	.01	.67	-.00	.53
2.	.01	.86	.01	.67	.06	.49
3.	.01	.86	-.01	.68	.00	.52
4.	-.01	.86	.04	.64	-.03	.55
5.	-.04	.86	.01	.68	.06	.50
6.	-.04	.87	-.00	.69	-.03	.56
7.	.05	.85	-.03	.68	.04	.51
8.	.03	.86	-.03	.70	-.04	.56
9.	.01	.85	-.06	.69	-.02	.53
10.	-.05	.87	-.03	.68	.05	.51
11.	.03	.86	.03	.65	.02	.51
12.	-.03	.87	.05	.65	.06	.50
13.	-.05	.87	.01	.68	-.03	.54
14.	-.02	.86	-.00	.68	-.04	.56
15.	-.00	.86	-.03	.70	-.08	.58
16.	.01	.87	.04	.67	.03	.51
17.	.02	.87	.01	.69	-.08	.59
18.	-.04	.88	-.03	.70	.02	.53
19.	.02	.85	.00	.67	.00	.52
20.	-.04	.86	.07	.65	.03	.51

Tableau 2.14 : Corrélations entre y et x_i et entre y et $(y - x_i)$ sur des données simulées (les variables ont respectivement, pour moyenne, 75, 90, 60, 60 et, pour écart-type, 15, 10, 15, 25. Elles se distribuent normalement et trouvent leurs valeurs dans l'intervalle [0 ; 100]).

Les corrélations entre y et les différentes variables x_i sont toujours proches de zéro, ce qui ne fait qu'illustrer la génération aléatoire des observations. Les corrélations entre y et les scores de différences $(y - x_i)$ sont toujours importantes, mais elles ne sont pas

uniformes. Les coefficients sont d'autant plus élevés que l'écart-type de la variable x_i est faible. Ainsi, les résultats de cette seconde série de simulations ne remettent pas en cause les résultats de la première. Ils montrent à nouveau, dans une situation plus proche de la situation empirique réelle, que l'on peut toujours trouver des corrélations importantes entre une variable et le score de différence entre cette variable et une autre variable. De plus, ils les affinent en montrant que l'intensité de cette corrélation dépend de la variance de la seconde variable. On peut donc, même avec des données totalement aléatoires, mettre en évidence des variables ayant un meilleur pouvoir prédictif que d'autres.

Les simulations que nous venons de mener ne peuvent rendre compte de tous les résultats obtenus par Campbell et ses collaborateurs. En particulier, elles n'expliquent pas les corrélations constatées directement entre, par exemple, l'évaluation de sa situation actuelle et son niveau d'aspiration ($r \approx .35$). Pour mieux comprendre ce résultat, il faut tenir compte d'une nouvelle contrainte, plus subtile que les précédentes, puisqu'elle est implicite dans le dispositif d'observation adopté. Rappelons-le brièvement. Les sujets doivent évaluer leur situation actuelle dans un certain domaine et, sur une même échelle de réponse, leur niveau d'aspiration vis-à-vis de ce domaine (i.e. la meilleure situation qu'ils puissent espérer), la situation de leurs proches, la situation d'un américain moyen, la pire situation qu'ils aient connue, etc. Ce dispositif d'observation est habituellement considéré comme étant à stimulus unique et à réponse ordonnée, comme le sont toutes les échelles d'évaluation (cf. Dickes et al. 1994). C'est le principe du dispositif à stimulus unique qui fait problème ici. En effet, on y postule que la réponse à un item se fait indépendamment des réponses aux autres items. Il est facile de comprendre que ce n'est pas le cas ici, au moins pour certains items. Prenons par exemple l'évaluation de sa situation actuelle et de son niveau d'aspiration. Quel que soit l'ordre dans lequel ces deux items sont présentés aux sujets, il est très peu probable qu'une personne déclare que sa situation actuelle est supérieure à la meilleure situation qu'elle puisse espérer. Ainsi, ce dispositif d'observation contient une contrainte implicite qui introduit une dépendance forcée entre certaines réponses^[19]. Pour évaluer l'effet de cette nouvelle contrainte, nous avons effectué une nouvelle série de simulations avec des variables spécifiées de la même manière que dans la série précédente. Concrètement, les variables x_i ont été générées de telle sorte que leurs différences avec y soient toujours positives ou toujours négatives ; autrement dit, pour chaque valeur de x_i , le processus de génération aléatoire se répète jusqu'à ce que la valeur se trouve dans l'intervalle $[0 ; 100]$ et que cette valeur soit supérieure à y (dans le cas où on souhaite que les différences entre y et x soient toujours

[19] : Les auteurs ne fournissent pas les données permettant de vérifier directement cette hypothèse. On sait que les moyennes sont assez nettement différentes (par exemple, autour de 75 pour l'évaluation de sa situation actuelle et autour de 90 pour le niveau d'aspiration), ce qui ne signifie pas, bien évidemment, que les différences entre les scores individuels soient toutes du même signe. Néanmoins, cette hypothèse nous paraît plausible pour certains items, comme le niveau d'aspiration ou l'évaluation de la pire situation jamais connue. En revanche, elle ne s'applique pas aux réponses à d'autres items comme l'évaluation de la situation de l'américain moyen.

positives). Les résultats de cette dernière série de simulations sont présentés dans le tableau 2.15.

L'introduction de cette contrainte supplémentaire change nettement la physionomie des résultats. Les corrélations entre y et les x_i ne sont plus nulles, mais modérément élevées. Il est d'ailleurs intéressant de noter que ces valeurs oscillent autour de .35, ce qui correspond approximativement aux corrélations que trouvent Campbell et al. entre la satisfaction et le niveau d'aspiration. De plus, contrairement aux simulations précédentes, les corrélations entre y et les différences entre y et x_i ne sont plus forcément supérieures aux corrélations simples entre y et les x_i . En particulier, elles sont quasi identiques dans le cas d'une variable ayant une variance élevée (cas de la variable x_3), mais restent nettement supérieures lorsque la variance est faible (cas de la variable x_1). Ici aussi, il est intéressant de constater que les corrélations calculées par Campbell et al. entre la satisfaction et la différence entre la satisfaction et le niveau d'aspiration se situent autour de .85.

simulation	$r_{y,x1}$	$r_{y,y-x1}$	$r_{y,x2}$	$r_{y,y-x2}$	$r_{y,x3}$	$r_{y,y-x3}$
1.	.42	.84	.44	.53	.37	.37
2.	.42	.85	.40	.58	.27	.46
3.	.38	.85	.38	.58	.38	.36
4.	.41	.86	.46	.54	.37	.39
5.	.42	.85	.45	.57	.39	.36
6.	.42	.86	.43	.57	.39	.38
7.	.39	.85	.48	.52	.36	.39
8.	.38	.85	.44	.56	.37	.35
9.	.44	.85	.41	.58	.32	.40
10.	.41	.87	.45	.57	.36	.37
11.	.37	.86	.43	.52	.35	.38
12.	.36	.87	.44	.55	.40	.37
13.	.42	.85	.44	.54	.35	.39
14.	.36	.86	.37	.59	.38	.34
15.	.41	.85	.45	.57	.39	.36
16.	.43	.85	.39	.58	.37	.38
17.	.39	.85	.41	.56	.41	.35
18.	.37	.86	.43	.55	.36	.39
19.	.44	.85	.45	.53	.32	.42
20.	.39	.85	.42	.54	.39	.35

Tableau 2.15 : Corrélations entre y et x_i et entre y et $(y - x_i)$ sur des données simulées (les variables ont respectivement, pour moyenne, 75, 90, 60, 60 et, pour écart-type, 15, 10, 15, 25. Elles se distribuent normalement et trouvent leurs valeurs dans l'intervalle [0 ; 100]. Les différences entre $y - x_i$ sont toutes positives pour x_1 et toutes négatives pour x_2 et x_3).

Quelles conclusions pouvons-nous tirer des résultats de ces simulations ? Ils ne viennent pas remettre en cause fondamentalement le modèle général de Campbell et al. ; ils montrent cependant que celui-ci est infalsifiable avec la méthodologie utilisée, à la fois du fait de la méthode d'analyse et du fait du dispositif d'observation. Autrement dit, ces résultats auraient pu être obtenus quelles que soient les variables utilisées, y compris si les sujets avaient répondu de manière aléatoire. Quelle que soit la portée heuristique de ce modèle, il n'est pas vérifié, contrairement à ce que pensent beaucoup d'auteurs. Pour qu'une vérification empirique soit possible, il faut trouver une autre méthodologie.

L'étude de Wright (1985) remet aussi sérieusement en cause le modèle de Campbell et al. (1976). Après avoir signalé certains problèmes associés à l'utilisation de scores de différences, l'auteur propose une étude empirique dans laquelle, tout comme dans celle de Campbell et al., des discordances calculées sont mises en relation avec des évaluations de satisfaction^[20]. La grande différence est que l'auteur prend la précaution de contrôler, dans les analyses, l'effet de la variable "évaluation de sa situation actuelle". On ne sera pas surpris, suite aux résultats des simulations précédentes, que les corrélations partielles entre les discordances et la satisfaction diminuent de manière importante pour se rapprocher de zéro. Ainsi, c'est l'évaluation de sa situation actuelle qui rend compte de l'essentiel de la variance commune entre les items de satisfaction et les discordances calculées, ce qui, somme toute, n'est pas vraiment étonnant.

Comme Michalos n'utilise pas de discordances calculées, il échappe à ces critiques. Cependant, le problème évoqué plus haut à propos de la signification des items destinés à mesurer les discordances reste entier. Il faudrait des arguments bien plus décisifs que ceux de Michalos pour accepter l'idée que les items de discordance et de satisfaction ne sont pas des mesures alternatives du même concept. On a malheureusement l'impression que ces modèles ont tendance à s'auto-valider dans une démarche essentiellement tautologique.

D'autres critiques ont été reformulées à l'encontre de ces modèles. Plus générales, elles s'adressent à l'ensemble des thèses relativistes, c'est-à-dire à tous les modèles indirects qui font intervenir des mécanismes cognitifs de comparaison. Dans tous les modèles qui font intervenir des comparaisons interpersonnelles, l'individu compare sa situation à celle d'un groupe de référence qui est relativement proche de lui dans la hiérarchie sociale. Pour Veenhoven (1991 b), cette hypothèse entraîne un corollaire évident, l'égalité des niveaux de bien-être moyen entre des pays riches et des pays pauvres, du fait que l'évaluation de son niveau de bien-être est un processus relatif et que les standards de comparaison ne dépendent pas des conditions de vie objectives. Cette

[20] : En fait, dans l'introduction de l'article, les critiques semblent plutôt dirigées vers Michalos que vers Campbell et al. Toutefois, la méthodologie utilisée par l'auteur (différences calculées) le rapproche beaucoup plus du modèle des seconds que de celui du premier.

hypothèse dérivée n'est pas vérifiée empiriquement, ainsi que nous l'avons vu dans le chapitre précédent : il y a une relation (non linéaire) entre la richesse d'un pays et le niveau moyen de bien-être de ses habitants.

Dans une étude consacrée aux relations entre bien-être et revenu, Diener et al. (1993) utilisent une démarche identique. Ils essayent de tester des hypothèses opérationnelles dérivées des modèles relativistes. La première série d'hypothèses concerne la comparaison sociale. Si les individus évaluent leur satisfaction en fonction de personnes qui sont proches et similaires, les niveaux de bien-être devraient être supérieurs, à revenus égaux, dans des régions pauvres (des Etats-Unis) que dans des régions plus riches. Une vérification de la même hypothèse peut s'appliquer à l'origine ethnique : à niveaux égaux de revenus, les noirs, dont on peut penser qu'ils se comparent préférentiellement à des individus appartenant à la même origine ethnique, devraient être plus heureux que les blancs car le groupe de référence des noirs a des conditions de vie plus défavorables. Les auteurs testent ces deux hypothèses par analyse de variance appliquée à des données recueillies sur un échantillon très important, représentatif de la population américaine. Si ces hypothèses sont correctes, on s'attend à trouver un effet significatif de l'interaction entre le revenu et l'origine ethnique (ou la région géographique) sur le bien-être. Malgré la taille des échantillons, l'interaction n'approche la significativité statistique ni dans un cas ni dans l'autre. Une autre série d'hypothèses renvoie à des standards de référence intrapersonnels et, plus spécifiquement, aux attentes (i.e. le bien-être sera d'autant plus élevé que les individus ont des conditions de vie proches de leurs attentes). Ces hypothèses peuvent être testées, selon Diener et ses collaborateurs, pour l'origine ethnique et pour le niveau d'éducation. Pour le niveau d'éducation, elle peut être formulée ainsi : les personnes les moins éduquées devraient être plus heureuses, à revenus égaux que les personnes plus éduquées car elles auraient pu s'attendre à moins. Les résultats sont aussi négatifs que dans la première série d'hypothèses.

Les critiques adressées aux modèles basés sur l'influence indirecte des conditions de vie apparaissent assez sérieuses et les partisans de cette conception n'y ont jamais vraiment répondu. Il faut tout de même noter un point important. Dans l'étude de Diener et al. (1993) rapportée ci-dessus, les hypothèses sont testées à partir de différences de revenu *réelles, objectives*, alors que dans toutes les applications habituelles, ce sont des discordances *perçues, subjectives* qui sont utilisées. Il ne fait guère de doute, qu'à nouveau, on ne mesure pas la même chose dans les deux cas. Nous renvoyons à plus tard dans ce chapitre pour une interprétation alternative de ces résultats.

5.3.3 Les modèles basés sur les changements de conditions de vie

Rappelons brièvement que les modèles de cette famille peuvent être distingués selon deux facettes. La première opère une différenciation entre les changements qui sont plutôt envisagés comme vécus ou subis par les sujets et ceux qui sont conçus comme désirés, souhaités et qui requièrent une activité délibérée de leur part. La seconde distingue les modèles dans lesquels c'est l'influence directe des changements qui est envisagée et ceux qui font intervenir des variables intermédiaires.

Ces différents types de modèles ne sont pas équitablement représentés dans la littérature. Certains ont suscité beaucoup plus de recherches que d'autres. La présentation qui suit rendra compte de ce déséquilibre.

A. Les modèles basés sur l'influence directe des changements vécus

De nombreux travaux ont été menés à propos de l'influence des événements de vie vécus par les individus sur différents indicateurs de bien-être lorsque celui-ci est envisagé dans une acception très générale. Cependant, la plupart de ces études ont porté sur des indicateurs de bien-être négatif (dépression, anxiété, symptômes physiques, tentatives de suicide, consommation d'alcool, etc.) dans des traditions de recherche assez indépendantes de celles qui nous occupent ici. Les travaux utilisant des indicateurs positifs du bien-être sont beaucoup plus rares. Dans leur revue de la littérature sur la question, Zautra & Reich (1983) n'en recensent que sept, dont quatre sont à mettre à l'actif de l'un ou de l'autre des auteurs !

Etant donné ce que l'on sait de l'indépendance entre les aspects positifs et négatifs du bien-être, il n'est pas envisageable d'utiliser les résultats obtenus impliquant le bien-être négatif pour tirer quelque conclusion que ce soit à propos du bien-être positif. Aussi les études qui mettent directement en relation les événements de vie vécus avec le bien-être sont-elles moins nombreuses qu'on pourrait l'imaginer. Les résultats de la revue de Zautra & Reich (1983) font état d'influences partielles : les événements positifs auraient un effet sur le bien-être positif, mais pas sur le bien-être négatif, l'inverse s'appliquant aux événements négatifs.

Mais l'influence directe des événements de vie est loin d'être établie. Zautra & Reich (1983) ne donnent pas d'indication quant à la taille des corrélations entre événements de vie et bien-être positif. Plus récemment, Valiant (1993) rapporte des résultats qui montrent que si événements négatifs et dépression semblent effectivement liés, la relation entre événements positifs et bien-être est sujette à caution.

A l'inverse de ce qui s'est passé pour l'effet des conditions de vie, il semble que les auteurs aient compris très tôt que l'influence des événements de vie pouvait dépendre de la manière dont ces derniers sont appréhendés par les sujets, ce qui explique sans doute en partie pourquoi les recherches ne sont pas plus nombreuses. Si certaines explications sont plutôt techniques — c'est le cas de celle de Russell & Davey (1993) qui demandent que soit distingué le niveau de l'événement (micro, meso ou macroscopique) — la plupart font intervenir des caractéristiques des événements qui sont médiatisées par le cadre de référence interprétatif du sujet.

B. Les modèles du bien-être basés sur l'influence indirecte des événements vécus

La médiation de l'influence des événements vécus par des référents internes peut prendre des formes diverses. Dans les modèles les plus connus, la théorie du niveau d'adaptation et la théorie étendue-fréquence, le cadre de référence interprétatif est fourni par les événements antérieurs et l'évaluation prend surtout la forme d'une comparaison intrapersonnelle. D'autres modèles plus spécifiques font appel à telle ou telle caractéristique des événements, utilisée par les individus pour donner une valeur à l'événement. Ces derniers seront tout d'abord brièvement décrits avant d'examiner en détail les précédents, qui ont eu beaucoup plus d'importance dans les recherches sur le bien-être.

L'idée selon laquelle l'influence d'un événement dépend de l'évaluation de certaines de ses caractéristiques se trouve déjà dans les travaux sur les thérapies cognitives (cf. Lichter et al., 1980) ; ces constatations ont d'ailleurs été mises à profit dans des études qui rapportent des tentatives pour améliorer le bien-être des individus (e.g. Fordyce, 1977 ; Lichter et al., 1980 ; Staats, 1991). Plusieurs caractéristiques à prendre en compte ont été mentionnées dans la littérature. Csikszentmihalyi & Wong (1991) insistent sur le fait que ce sont les opportunités de mettre en oeuvre ses propres compétences (que l'événement fournit à l'individu) qui permettent d'attribuer une valeur à cet événement. Ainsi, un événement objectivement négatif pourra être évalué positivement dans le cas où l'individu a pu exprimer ses capacités personnelles pour y faire face. Vinokur & Caplan (1986) distinguent, quant à eux, une composante cognitive et une composante émotionnelle dans les événements de vie. La composante cognitive concerne le caractère désirable d'un événement, alors que la composante émotionnelle concerne son caractère plaisant. Un même événement peut très bien être désirable et déplaisant (e.g. le décès d'un parent si celui-ci est atteint d'une maladie incurable et douloureuse). Zautra & Reich (1980, 1983) font appel au degré de contrôle des événements tel qu'il est perçu par les sujets. Ils parviennent à montrer que ce sont surtout les événements positifs impliquant un contrôle personnel qui ont une influence sur le bien-être (positif) : les événements positifs qui ne

sont pas sous le contrôle des individus ne semblent pas augmenter le bien-être^[21]. Les auteurs concluent aussi à l'existence d'effets croisés entre polarité des événements et polarité du bien-être. En particulier, les événements négatifs attribués à des causes internes et stables semblent pouvoir affecter le bien-être positif ; les événements positifs qui ne sont pas attribués à des causes personnelles peuvent augmenter le bien-être négatif (à cause de l'augmentation du stress causée par un événement, même positif). Ces résultats ne sont pas sans rappeler les travaux sur les liens entre style attributif et dépression. La méta-analyse de Sweeney et al. (1986) montre que la dépression est liée à l'attribution des événements négatifs à des causes internes, stables et globales et l'attribution des événements positifs à des causes externes, instables et spécifiques. Mais la question qui se pose, dans le cas de la dépression comme dans celui du bien-être, est la même : le type d'attribution doit-il être considéré comme une cause ou comme une conséquence ? L'intérêt de ces variables intermédiaires pour l'explication du bien-être dépend beaucoup de la réponse à cette question.

La théorie étendue-fréquence ("range-frequency theory")

Parmi les deux principaux modèles concernant l'influence indirecte des changements dans les conditions de vie, la théorie étendue-fréquence est sans doute la moins connue, même si elle est souvent citée dans la littérature. Cette théorie s'inscrit dans une tradition expérimentale et psychophysique^[22], ce qui constitue à la fois un avantage et un inconvénient. C'est un avantage dans la mesure où, nous allons le voir, il s'agit sans doute du modèle explicatif du bien-être le plus formalisé qui ait jamais été proposé. Mais cette rigueur dans la formalisation le rend difficilement transposable à des situations naturelles, ce qui questionne sérieusement sa validité écologique.

La théorie étendue-fréquence a été élaborée par Parducci (e.g. 1984, mais les premiers travaux remontent au début des années 60). Nous nous appuyerons dans la suite du texte sur la présentation qui en a été faite par Smith et al. (1989). Dans cette théorie, comme dans d'autres modèles relativistes, le bien-être engendré par un événement ne dépend pas de la valeur absolue de ce dernier mais de la valeur qu'il prend en relation avec d'autres événements contextuels. Plus spécifiquement, la satisfaction engendrée par

[21] : Mais les événements qui sont sous le contrôle des individus sont en fait presque toujours des événements qui nécessitent une action de la part des sujets, dans les études empiriques de Zautra & Reich. Il y a donc une confusion possible avec des concepts comme celui d'activité dont on a vu qu'il était lié au bien-être.

[22] : Au sens le plus général, la psychophysique vise à expliciter les relations entre des continus objectifs et des continus subjectifs du jugement. C'est bien ce dont il est question dans ce modèle, comme on le verra par la suite.

un événement dépend de la forme de la distribution des autres événements : elle est supérieure lorsque la distribution est biaisée négativement^[23].

Dès ses premiers travaux, Parducci remarque que le point neutre d'une distribution d'événements (c'est-à-dire celui qui ne provoque ni satisfaction ni insatisfaction) ne correspond pas à la moyenne des autres événements mais se situe à mi-chemin entre la valeur médiane et la valeur centrale de la distribution. Il en découle deux principes : un principe d'étendue et un principe de fréquence. Pour expliquer ces principes, il est plus simple de s'appuyer sur un exemple. Considérons que l'événement auquel on s'intéresse est une note obtenue par un étudiant à un examen. Les événements contextuels qui vont servir à l'évaluation de ce nouvel événement seront, dans ce cas, les notes obtenues lors d'examens antérieurs. Ces événements peuvent être situés sur une échelle graduée de 0 à 20^[24], leur distribution pouvant être de n'importe quelle forme. Un rang de classement peut être attribué au nouvel événement sur cette échelle.

Le principe d'étendue reflète la tendance à évaluer un événement en fonction de la distance entre cet événement et la valeur minimale de la distribution^[25]. Autrement dit, un événement sera jugé d'autant plus satisfaisant qu'il se situe à proximité de la valeur maximale de la distribution. Ces valeurs extrêmes sont des valeurs subjectives dans le sens où elles vont dépendre du contexte (i.e. des autres événements). Dans notre exemple, pour un sujet donné, elles correspondraient à la note la plus basse et à la note la plus élevée obtenue. Le principe d'étendue s'exprime de la manière suivante :

$$R_{ic} = \frac{(S_i - S_{\min})}{(S_{\max} - S_{\min})} \quad (1)$$

avec : R_{ic} évaluation selon l'étendue de l'événement i dans le contexte c

S_i valeur de l'événement i dans la distribution

S_{\min} valeur minimale de la distribution

S_{\max} valeur maximale de la distribution

[23] : Ceci est facile à comprendre : une distribution biaisée négativement indique que la majorité des événements ont une valence positive. Ainsi, une personne est d'autant plus satisfaite que les événements qu'elle connaît sont positifs.

[24] : Plus généralement, il s'agit de situer les événements sur une dimension qui permette, au minimum, de les ordonner.

[25] : Ainsi, une note de 10 dans une distribution allant de 0 à 12 serait considérée comme très satisfaisante alors que la même note sera jugée peu satisfaisante dans une distribution variant de 7 à 18.

Le principe de fréquence reflète la tendance à évaluer un événement en fonction du nombre total d'événements qui se situent en dessous de cet événement^[26]. Ce principe s'exprime de la manière suivante :

$$F_{ic} = \frac{(r_{ic} - 1)}{(N_c - 1)} \quad (2)$$

avec : F_{ic} évaluation selon la fréquence de l'événement i dans le contexte c
 r_{ic} rang de l'événement i dans le contexte c
 N_c nombre total d'événements dans le contexte c

L'évaluation globale d'un événement est fonction de ces deux principes. Elle s'exprime ainsi :

$$J_{ic} = wR_{ic} + (1 - w)F_{ic} \quad (3)$$

avec : J_{ic} évaluation de l'événement i dans le contexte c
 w poids relatif accordé au principe d'étendue

La valeur empirique de w , telle qu'elle ressort de plusieurs études psychophysiques est proche de 0.5.

Ces deux principes permettent de comprendre comment est évalué un événement particulier. En ce qui concerne la satisfaction générale par rapport à un ensemble d'événements, la théorie étendue-fréquence considère qu'elle est proportionnelle au biais de la distribution. Si les événements sont concentrés près de la valeur maximale de la distribution (biais négatif), le bien-être général sera plus élevé. Ainsi, la forme de la distribution a des effets différents sur la satisfaction générale et sur la satisfaction avec un événement particulier. Soit deux distributions ayant même étendue mais dont l'une est biaisée positivement, l'autre négativement, la satisfaction avec un événement particulier sera la plus importante lorsque la distribution est positivement biaisée ; en revanche, la satisfaction générale sera plus importante lorsque la distribution des événements est négativement biaisée^[27].

Dans l'étude empirique rapportée par Smith et al. (1989), l'idée générale est de présenter à des sujets des distributions hypothétiques de 32 événements (notes obtenues à des examens, d'une part et pourboires reçus lors de services rendus, d'autre part) et de leur demander d'évaluer la satisfaction engendrée par différents événements se situant dans ces distributions. Trois types de distribution sont utilisés : une distribution biaisée

[26] : Ainsi, une note de 7 pourra être jugée satisfaisante dans une distribution allant de 3 à 15, à condition que la majorité des notes se situe entre 3 et 6.

[27] : Car, dans ce cas, il y a beaucoup d'événements positifs et peu d'événements négatifs.

positivement avec une moyenne faible (étendue de 25 à 45 pour les notes, $m = 33$) ; une distribution biaisée négativement avec une moyenne élevée (étendue de 25 à 45, $m = 37$) ; une distribution biaisée négativement avec une moyenne faible (étendue de 21 à 41, $m = 33$). Un exemple de distribution biaisée positivement est présenté dans la figure 2.24.

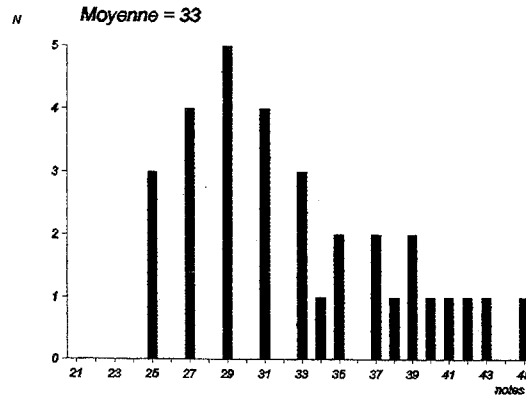


Figure 2.24 : Exemple de distribution biaisée positivement utilisée par Smith et al. (1989) (les événements sont les notes reçues à des examens)

Les résultats de cette étude montrent qu'effectivement, la satisfaction générale, mesurée soit par la moyenne des événements individuels, soit par une évaluation directe^[28], est supérieure lorsque les distributions sont biaisées négativement, alors que la satisfaction avec un événement particulier est supérieure lorsque la distribution des événements est biaisée positivement.

La théorie étendue-fréquence a été critiquée car elle semblait incompatible avec les comportements. Nous avons vu que la satisfaction est d'autant plus importante que la distribution des événements est biaisée négativement. Pourtant, on constate que les individus ont tendance à construire des distributions biaisées positivement. Ainsi, Brickman (1975, cité par Smith et al., 1989) demande à des sujets de répartir un budget de 120 dollars sur une période de 24 jours. Il constate que la majorité des sujets choisit de dépenser de faibles sommes un grand nombre de jours et des sommes importantes à de rares occasions. Ceci signifie que la distribution des événements est biaisée positivement et donc, que les sujets ne choisissent pas la stratégie qui leur procurerait la plus grande satisfaction^[29]. En fait la contradiction entre ce résultat et la théorie étendue-

[28] : On demande aux sujets d'évaluer leur satisfaction avec tous les événements présentés.

[29] : C'est-à-dire dépenser rarement de faibles sommes et fréquemment des sommes plus importantes.

fréquence n'est qu'apparente grâce à la notion d'étendue subjective de la distribution. Souvent, dans les études empiriques, l'étendue réelle de la distribution est assimilée à son étendue subjective. Or, il semble que des événements non observés réellement sont imaginés par le sujet lorsqu'il se représente la distribution des événements. Si tel est le cas, l'étendue subjective est plus grande que l'étendue réelle de la distribution et une distribution biaisée positivement peut devenir biaisée négativement à condition d'étendre l'étendue subjective de la distribution vers le bas. Ceci est illustré par la figure 2.25.

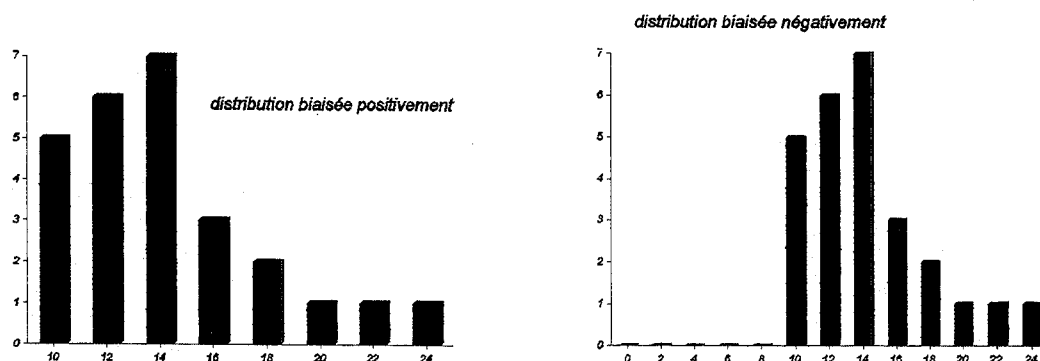


Figure 2.25 : Exemple hypothétique d'une même distribution sous deux étendues différentes.

Ce processus n'est pas sans rappeler certaines études dans lesquelles les individus effectuent des comparaisons avec des personnes "plus mal loties" (e.g. Heidrich & Ryff, 1993a, 1993b). Il indique que les principes d'étendue et de fréquence s'appliqueraient tant aux comparaisons intrapersonnelles qu'aux comparaisons interpersonnelles.

La théorie étendue-fréquence est très séduisante dans sa construction. Elle permet de comprendre pourquoi les événements n'ont pas forcément un effet correspondant à leur valence et souligne l'importance des événements contextuels. Son point faible réside dans son caractère artificiel. Son applicabilité nous semble douteuse dans le cas d'événements pour lesquels il est difficile de construire une distribution (événements rares ou uniques par exemple). Par ailleurs, le principal postulat que nous croyons déceler dans cette théorie semble discutable. Il paraît en effet difficilement concevable qu'à l'intérieur de n'importe quelle classe d'événements^[30] on puisse trouver une dimension unique, continue et quantifiable sur laquelle un individu pourrait situer tous les événements quels qu'ils soient. Enfin, la réalité de l'existence des processus cognitifs complexes qui seraient utilisés par les individus reste à démontrer.

[30] : La notion de classe d'événements est nécessaire à la cohérence de cette théorie. Il n'est pas concevable en effet de construire des distributions avec des événements qui ne partagent pas une même signification. Pour montrer la généralité de cette théorie il faudrait pouvoir préciser à la fois la notion de classes d'événements et celle de dimension d'évaluation. Ceci semble une tâche bien difficile dès lors que l'on ne se contente pas d'événements élémentaires comme des notes scolaires ou des rentrées d'argent.

La théorie du niveau d'adaptation

La théorie du niveau d'adaptation partage avec la théorie étendue-fréquence l'idée que la satisfaction engendrée par un nouvel événement est fonction des événements antérieurs. Mais le processus d'évaluation est plus simple puisque le nouvel événement est simplement confronté à la moyenne des événements précédents, celle-ci étant continuellement réadaptée en fonction des dernières expériences.

La théorie de l'adaptation ne fait que reprendre l'intuition selon laquelle les individus s'adaptent aux circonstances. La simplicité de cette idée a sans doute beaucoup contribué au succès de cette théorie qui a été développée par Helson en 1964 et appliquée au bien-être par Brickman & Campbell en 1971. Mais l'article le plus célèbre de cette tradition de recherche est celui que Brickman et ses collaborateurs ont publié en 1978. Le principe de base de cette théorie est exprimé simplement dans cet article (p. 918) : "*(...) people's judgments of current levels of stimulation depend upon whether this stimulation exceeds or falls short of the level of stimulation to which their previous history has accustomed them*". Autrement dit, un nouvel événement est confronté à une valeur, appelée niveau d'adaptation, qui est la moyenne des évaluations attachées aux événements précédents. Si l'événement dépasse cette valeur, il est jugé satisfaisant et contribue à augmenter le bien-être général ; dans le cas contraire, il est jugé insatisfaisant. Mais un même événement, jugé satisfaisant à un certain moment pourra très bien être jugé insatisfaisant à un autre moment car le niveau d'adaptation ne reste pas figé : il est réactualisé en fonction des nouveaux événements. Les comparaisons sont donc intrapersonnelles et le standard de référence s'ajuste continuellement.

Comme les références à cette théorie s'appuient principalement sur l'étude empirique de Brickman et al. (1978), il semble utile de la décrire de manière assez détaillée, d'autant que des points importants sont souvent oubliés dans la littérature qui la mentionne et qu'elle pose des problèmes d'ordre méthodologique.

L'idée originale de Brickman et al. (1978) consiste à comparer des groupes de sujets ayant connu un événement très positif et un événement très négatif. L'événement positif choisi est un gain important à la loterie (entre 50000 et un million de dollars) ; l'événement négatif est un accident ayant entraîné une paralysie. Un groupe contrôle complète l'échantillon. Si la théorie du niveau d'adaptation est correcte, les sujets ayant connu un événement positif et ceux ayant connu un événement négatif doivent être aussi heureux que les sujets du groupe contrôle parce leur niveau d'adaptation a changé.

Le groupe d'accidentés comporte 29 sujets (para- ou tétraplégiques) fréquentant un centre de réadaptation. Le taux de participation est de 85%, mais on ne sait pas

pourquoi cinq sujets n'ont pas participé à l'étude. Le groupe de gagnants à la loterie a été sélectionné sur la base de l'importance du gain, ainsi que sur des critères de proximité géographique. Sur les 42 sujets contactés, seuls 20 ont été effectivement interrogés, ce qui représente un taux de participation de 52% (les autres n'étant pas joignables ou ayant refusé de participer à l'étude). Le groupe contrôle a été choisi au hasard dans les localités où habitent les gagnants à la loterie, à partir de l'annuaire téléphonique : il comprend 22 sujets pour un taux de participation de 41%. La seule différence significative entre ces trois groupes par rapport aux variables socio-démographiques classiques (sexe, éducation, origine ethnique, catégorie socio-professionnelle...) concerne l'âge, et celle-ci est de taille. L'âge moyen des gagnants est de 44 ans, celui du groupe contrôle de 46 ans mais il est seulement de 23 ans pour les accidentés.

Les accidentés sont tous interrogés en face à face. La majorité des sujets des autres groupes ont été interrogés par téléphone (18 sur 22 pour le groupe contrôle et 16 sur 22 pour le groupe des gagnants), les autres en face-à-face. Différentes questions sont posées aux sujets. Celles qui nous intéressent ici concernent le bien-être. Le bonheur général est mesuré avec trois questions : l'une porte sur leur bien-être actuel, les autres sur leur bien-être passé (6 mois auparavant pour le groupe contrôle ; avant l'événement pour les autres) et attendu dans le futur (dans quelques années). Les auteurs souhaitent aussi obtenir une estimation de ce qu'ils assimilent à une mesure du plaisir quotidien : pour cela, ils demandent aux sujets d'indiquer le plaisir qu'ils prennent à chacune des activités suivantes : bavarder avec un ami ; regarder la télévision ; prendre son petit déjeuner ; entendre une blague drôle ; recevoir un compliment ; lire un magazine ; acheter des vêtements (sauf pour le groupe des accidentés). Les résultats moyens sont présentés dans le tableau 2.16.

Groupes	Bien-être général			Plaisirs quotidiens
	Passé	Présent	Futur	
Gagnants (1)	3.77	4.00	4.20	3.33
Contrôle (2)	3.32	3.82	4.14	3.82
Accidentés (3)	4.41	2.96	4.32	3.48
p < .05 pour les groupes...	2 et 3	2 et 3	/	1 et 2

Tableau 2.16 : Moyennes des trois groupes expérimentaux aux différentes variables dépendantes de l'étude de Brickman et al. (1978) [5 accidentés n'ont pas répondu à toutes les questions concernant les plaisirs quotidiens ; 10 accidentés, 3 gagnants et 1 sujet du groupe contrôle n'ont pas répondu à la question sur le bien-être futur]

Les résultats du groupe des gagnants à la loterie sont ceux qui supportent le mieux la théorie du niveau d'adaptation. Le fait qu'ils aient gagné une somme d'argent importante ne les rend pas plus heureux que les individus du groupe contrôle. De plus, ils rapportent significativement moins de plaisir dans les activités de la vie quotidienne, ce qui est compatible avec l'hypothèse selon laquelle leur niveau d'adaptation a augmenté et qu'il faut des événements beaucoup plus satisfaisants pour qu'ils en retirent du plaisir^[31].

En revanche, contrairement à ce que sous-entendent nombre d'auteurs qui citent les résultats de cette expérience, les accidentés ne sont pas aussi heureux que les individus du groupe contrôle et ne rapportent pas plus de plaisirs dans les activités quotidiennes, ce que reconnaissent d'ailleurs tout à fait explicitement Brickman et al. (1978, p. 924). Ces derniers voient cependant une confirmation indirecte de la théorie du niveau d'adaptation dans le fait que les accidentés rapportent un niveau de bien-être supérieur dans le passé. Cela signifierait que les événements qui deviennent plus positifs ne sont pas ceux du présent, mais ceux du passé. Ainsi, selon les auteurs, il n'est pas étonnant que les événements présents ne soient pas plus satisfaisants car ils sont évalués par rapport à un niveau d'adaptation élevé dans le passé.

L'utilisation qui est faite de cette étude dans la littérature est assez déroutante. Alors que les résultats sont somme toute assez mitigés et que les auteurs ne cachent pas certains problèmes méthodologiques, ils sont pratiquement toujours cités, sans regard critique, en faveur de la théorie du niveau d'adaptation. Ces problèmes méthodologiques ont d'ailleurs été soulignés par Veenhoven (1991 b) pour qui il n'est pas indifférent que les accidentés aient été interrogés en face à face alors que la majorité des individus du groupe contrôle l'ont été par téléphone. Certains résultats montrent, en effet, que les niveaux de bien-être moyen issus de situations en face à face sont en général supérieurs à ceux qui proviennent d'entretiens téléphoniques (cf. Smith, 1979). En plus de cette remarque, on peut s'interroger sur d'autres aspects méthodologiques dont il est difficile d'évaluer l'impact réel sur les résultats :

- les pourcentages de refus de participation à l'étude sont assez importants, surtout pour le groupe des gagnants et le groupe contrôle ;
- la différence d'âge entre les accidentés et les individus des autres groupes est particulièrement marquée ;
- la mesure du bien-être général est obtenue avec un seul item ;
- pour certaines questions, les refus de réponse sont élevés.

[31] : Les auteurs vérifient, dans une seconde étude, que ces résultats ne peuvent pas s'expliquer par le fait que les personnes qui achètent des tickets de loterie ont tendance à prendre moins de plaisir dans leurs activités quotidiennes que ceux qui n'en achètent pas. Il est intéressant de noter au passage que cette étude est l'une des rares qui envisage sérieusement que la relation mise en évidence puisse être interprétée dans un sens causal inverse.

Au-delà de ces remarques méthodologiques, la théorie du niveau d'adaptation a donné lieu à des critiques plus fondamentales, sur les plans empirique et théorique :

- Dans son article sur l'évaluation de la validité des théories relativistes, Veenhoven (1991 b) recense les études qui montrent, de manière cohérente avec les hypothèses dérivées du modèle adaptatif, que le niveau de bien-être des personnes après un événement négatif sérieux n'est pas inférieur à celui de personnes de groupes contrôles. Après avoir relevé les biais méthodologiques de ces travaux, il propose plusieurs autres études, moins souvent citées dans la littérature écrit-il, qui montrent exactement l'inverse.

- La théorie du niveau d'adaptation a aussi été critiquée parce que, de manière quelque peu paradoxale, elle traite les individus comme des récepteurs passifs en considérant que tout événement entraîne mécaniquement un ajustement du niveau d'adaptation (Zautra & Reich, 1980). Si les événements de vie positifs augmentent le niveau d'adaptation et si les événements négatifs le diminuent, on devrait constater une évolution parallèle du niveau d'aspiration des individus. Or, Headey et al. (1989) montrent, dans une étude longitudinale, que le niveau d'aspiration ne semble pas lié aux événements de vie.

- La question de la stabilité du bien-être dans le cadre de la théorie du niveau d'adaptation a été controversée. Certains (e.g. Costa & McCrae, 1980) considèrent en effet que ce modèle implique que le bien-être doit être relativement stable. Pour d'autres (Veenhoven, 1991 b), au contraire, on doit s'attendre à des corrélations test-retest négatives puisque si on est heureux en fonction de certaines circonstances à un moment donné, on devrait l'être moins plus tard une fois que l'adaptation a joué (et inversement). L'hypothèse de corrélations négatives est bien évidemment contredite par les faits, ainsi que nous l'avons vu dans le second chapitre. Pourtant, pour les partisans de la théorie du niveau d'adaptation (e.g. Chamberlain & Zika, 1992), cela n'a pas grand sens d'étudier la stabilité du bien-être sur de longues périodes : il faut l'étudier sur de courtes périodes avant qu'il y ait eu adaptation. Dans une étude déjà citée de Diener et al. (1993), qui vise à tester empiriquement les thèses relativistes à partir de la relation entre le bien-être et le revenu, une corrélation de .001 est trouvée entre les *changements* dans les revenus et le bien-être. Il est vrai que, dans cette étude, la période séparant les deux enquêtes est longue et il est donc possible que des phénomènes d'adaptation se soient produits. Les auteurs notent toutefois que parmi l'ensemble des changements de revenu répertoriés, certains ont dû se produire récemment, ce qui devrait suffire à augmenter la corrélation. On restera cependant réservé quant aux résultats de cette étude dont les résultats pourraient tout aussi bien être utilisés en faveur de la théorie de l'adaptation.

Il semble donc que les arguments empiriques accumulés en faveur de la théorie du niveau d'adaptation ne soient pas à la hauteur de la confiance aveugle qui lui est parfois accordée dans la littérature. Il serait exagéré pourtant de n'accorder aucun crédit aux mécanismes adaptatifs qui semblent pouvoir être mis en évidence dans certaines études. Veenhoven (1991 b) reconnaît bien volontiers que s'il est faux d'affirmer que les personnes qui ont connu de sérieux revers dans leur vie ne sont pas plus malheureuses que les autres, il est manifeste que leur niveau de bien-être n'est pas aussi bas que celui auquel on pourrait s'attendre. Mais même dans le cas le plus favorable, la théorie du niveau d'adaptation ne peut rendre compte que de manière partielle des résultats des recherches sur le bien-être. Si elle permet effectivement d'expliquer pourquoi le bien-être est stable sur de longues périodes, elle est incapable, en revanche, de rendre compte des différences interindividuelles qui existent de manière incontestable (cf. Costa & McCrae, 1980).

En plus des modèles qui considèrent les influences directes et les influences indirectes des changements sur le bien-être, la proposition de Tversky & Griffin (1991) mérite d'être signalée. Elle peut être assimilée à un modèle dans lequel on tient compte à la fois des effets directs et des effets indirects des événements.

Selon Tversky & Griffin (1991), un même événement a deux effets contraires sur le bien-être : un effet de dotation (une expérience positive contribue à augmenter, de manière directe, le bien-être) et un effet de contraste (une expérience positive va servir de nouveau standard de comparaison et rendre d'autres expériences moins positives). Le bien-être est maximum lorsque les expériences positives sont uniquement traitées comme des dotations et les expériences négatives comme des contrastes.

Alors que les effets de contraste ont souvent été mis en évidence dans la littérature sur le bien-être, il n'en est pas de même pour les effets de dotation. Les auteurs expliquent ce résultat par le fait que les mesures du bien-être se basent sur des jugements ; on sait que les jugements sont sensibles aux effets de contraste, alors qu'à l'inverse, les mesures basées sur les choix sont sensibles aux dotations. Ainsi, quelqu'un choisira de gagner 100 francs plutôt que 90 (effet de dotation), même si ses pairs gagnent 120 francs, dans le premier cas, 70 francs dans le second (d'après les théories relativistes, la satisfaction devrait être supérieure si l'individu est dans une situation plus enviable que celle de ses pairs). On peut donc se demander s'il ne vaut pas mieux appréhender le bien-être avec des tâches de choix plutôt qu'avec des tâches de jugement. Il n'en est rien pour les auteurs qui considèrent que les deux approches sont pertinentes, même si aucune n'est, à elle seule, satisfaisante.

Il est difficile d'évaluer le modèle de Tversky & Griffin, qui a suscité, pour l'instant, peu de travaux. On notera tout de même que, de même que la théorie étendue-

fréquence, ce modèle s'appuie sur des études expérimentales dans lesquelles les sujets doivent évaluer leur satisfaction en fonction de situations fictives et quelque peu artificielles. La validité écologique des résultats est donc sujette à caution et il reste à montrer que ces modèles peuvent s'appliquer dans des situations réelles où le sujet qui répond est confronté à une masse d'informations sur ses conditions de vie, sur ses expériences passées et actuelles et sur sa position vis-à-vis de groupes sociaux de référence. De tels modèles restent muets quant à la manière dont l'ensemble de ces informations est intégré par les individus.

C. Les modèles du bien-être basés sur l'influence directe des changements désirés

Il n'est pas toujours facile de distinguer entre les changements vécus et les changements désirés. Cette différenciation a surtout une valeur opérationnelle afin de mettre en lumière les différences entre les modèles qui portent sur ce qu'il est convenu d'appeler les événements de vie et ceux qui sont basés plutôt sur les buts ou les besoins.

La majorité des modèles du bien-être partage le postulat selon lequel les individus sont fondamentalement motivés pour maximiser leur niveau de bien-être. Ce postulat est particulièrement saillant dans les modèles basés sur les buts et les besoins, qui sont d'ailleurs parfois assimilés à des théories motivationnelles (cf. Campbell, 1981). Le point commun de ces modèles est qu'ils considèrent que le bien-être dépend de la satisfaction de ses besoins.

Cette idée n'est pas récente. Déjà présente dans la littérature philosophique, elle a été reprise par Wilson, en 1967, dans la première revue de la littérature scientifique sur le bien-être. Pourtant, force est de constater que les études empiriques s'inspirant directement de l'hypothèse de la satisfaction des besoins ne sont pas nombreuses. Ce modèle jouit d'une certaine acceptation tacite par les auteurs — le fait qu'il soit intuitivement acceptable y est sans doute pour quelque chose — qui l'évoquent lorsque certains de leurs résultats sont compatibles avec une interprétation en termes de besoins^[32]. Mais les études empiriques élaborées de manière à tester directement des hypothèses dérivées de la théorie des besoins sont rares, pour ne pas dire inexistantes.

[32] : Nous avons vu, dans le précédent chapitre, que la relation curvilinéaire entre le bien-être et les revenus était quelquefois interprétée en fonction d'une théorie des besoins : dès lors que le revenu est suffisant pour satisfaire les besoins élémentaires, celui-ci n'a plus d'influence.

Lorsqu'il est fait référence à ces modèles, la hiérarchie des besoins de Maslow est souvent citée^[33]. Pour ne prendre qu'un seul exemple, Campbell et al. (1976, p.12) rapportent qu'ils ont essayé d'élaborer leur étude empirique en adoptant ce cadre théorique. S'ils y ont finalement renoncé, c'est parce qu'il s'avérait difficile de traduire certaines abstractions du système de Maslow dans des items adaptés pour une vaste enquête nationale. Ces auteurs choisiront finalement un moyen terme en décidant de s'intéresser à la satisfaction dans différents domaines de vie.

La théorie de Maslow elle-même, en dehors du contexte des études sur le bien-être, s'est assez peu prêtée aux vérifications empiriques et, lorsque les chercheurs ont tenté de le faire, les résultats ont été assez décevants (cf., à ce sujet, la revue de la littérature de Wahba & Bridwell, 1976). C'est le cas, par exemple de Lawler & Suttle (1972) pour qui, si le système de Maslow est correct, la satisfaction d'un besoin devrait entraîner une diminution de son importance pour un individu. Les résultats ne confirment pas cette hypothèse et conduisent les auteurs à proposer une hiérarchie très simplifiée qui distingue seulement les besoins biologiques et l'ensemble des autres.

Ce sont à nouveau Diener et al. (1993) qui proposent des données empiriques venant tester directement une hypothèse dérivée d'un modèle de satisfaction des besoins, mais cela reste très fruste. Selon ces auteurs, les revenus des individus leur permettent de satisfaire leurs besoins de base. Une fois que ceux-ci sont satisfaits, les individus ont d'autres besoins pour lesquels l'argent ne peut intervenir. On doit donc s'attendre à une relation entre le bien-être et le revenu pour les plus pauvres et une absence de relation pour les plus riches. Cette hypothèse est testée à la fois avec des données recueillies au niveau individuel et des données agrégées au niveau de pays. Aucune des deux sources de données ne confirme cette hypothèse.

Mais les besoins correspondent peut-être à un concept motivationnel trop général pour rendre compte des relations avec le bien-être. Lorsque l'on s'inspire de la théorie de Maslow, on spécifie un petit nombre de besoins, très généraux et identiques pour tous les individus. Conscients de ces limites, certains auteurs ont proposé de recourir à des concepts plus spécifiques à travers les notions de projets personnels (Little, 1983) ou de buts personnels^[34] (Emmons, 1986).

[33] : Rappelons simplement ici que Maslow (e.g. 1968/1972) a proposé un système dans lequel les besoins sont hiérarchisés de sorte que la satisfaction des besoins correspondant à un niveau inférieur fasse émerger des besoins du niveau supérieur. Cette hiérarchie se présente comme suit : besoins physiologiques de base (faim, soif...) ; besoins de sûreté, de sécurité ; besoins d'appartenance (famille, groupe, ...) ; besoins d'autonomie, d'indépendance, d'estime de soi, d'approbation ; besoins de réalisation de soi.

[34] : Nous proposons "buts personnels" comme traduction de "*personal strivings*" qui n'a pas d'équivalent direct en français. Le verbe "*to strive*" signifie "s'efforcer de" et Emmons (1986) définit ces projets personnels comme des objectifs qu'une personne essaie (s'efforce) d'atteindre de manière caractéristique.

Ces concepts semblent appelés à un avenir prometteur dans le domaine de la personnalité. Ils constituent, pour Angleitner (1991), de nouvelles unités de base dans l'étude de la personnalité et ils reflètent, pour Little (1983), tout à la fois les aspects affectifs, cognitifs et comportementaux de la conduite humaine. Pour Brunstein (1993), les concepts de buts personnels peuvent être assimilés à des mesures mixtes idiographiques-nomothétiques de la motivation. Ces mesures sont idiographiques parce que les buts sont spécifiques aux individus ; elles sont nomothétiques parce que les buts sont évalués sur des dimensions communes à l'ensemble des individus.

Pour l'instant, les travaux qui utilisent les buts personnels s'efforcent surtout de mettre en relation le bien-être avec des évaluations qui sont faites de ces buts. En cela, ces études doivent être envisagées comme relevant des modèles à influence indirecte des changements désirés. Nous les verrons donc plus en détail dans le prochain paragraphe.

Mais nous avons besoin du concept de buts personnels pour évoquer certains travaux de Reich & Zautra qui sont habituellement cités en relation avec les modèles sur l'influence des événements de vie. Pourtant, à la lecture de ces travaux, il apparaît que ceux-ci sont plus proches du concept de buts personnels que de celui d'événements vécus. Dans une étude de 1981, Reich & Zautra se proposent de considérer les "événements de vie" à un niveau plus fin que celui auquel ils sont habituellement envisagés. Dans leur étude empirique, qui adopte une démarche expérimentale, des questionnaires de bien-être sont administrés, lors d'un pré-test, à l'ensemble des individus de l'échantillon. Le post-test, qui a lieu un mois plus tard, donne lieu à l'administration des mêmes instruments. Entre les deux, trois groupes de sujets sont constitués et affectés à trois conditions expérimentales. Dans la première, il est demandé aux sujets de s'engager à accomplir 12 activités qu'ils ont jugées plaisantes lors du pré-test. Dans la seconde, les sujets s'engagent à accomplir 2 activités de ce type. Le troisième groupe est un groupe contrôle. Les résultats sont bien évidemment consistants avec les hypothèses ; les sujets les plus heureux lors du post-test sont ceux qui se sont engagés dans de nombreuses activités plaisantes.

Ce ne sont pas tant les résultats de cette étude qui nous intéressent ici — il n'est pas exclu d'ailleurs qu'ils soient dus à des biais méthodologiques comme le reconnaissent les auteurs eux-mêmes — que l'utilisation du concept d'événement de vie. Pour les auteurs, il s'agit surtout de montrer que ce sont principalement les événements de vie qui sont perçus comme contrôlés par les individus qui ont une influence sur le bien-être. Pour nous, la notion d'activités correspond plutôt à celle de buts personnels telle qu'on peut la trouver chez des auteurs comme Emmons. Les exemples d'activités cités par Reich & Zautra (1981 : "aller nager", "participer à une discussion en classe", "visiter un musée", "appeler un ami pour bavarder") ne se différencient pas des exemples de buts cités par

Emmons (1992 : "prendre un bain chaud", "nettoyer la maison", "aller à l'église tous les dimanches").

Reich & Zautra préciseront, en 1983, leur conception. Dans cet article, ils distinguent, parmi les événements de vie quotidiens, les désirs et les demandes. Les désirs correspondent à une motivation intrinsèque ; c'est le fait de faire une action qui cause du plaisir. Les demandes correspondent plus à une motivation extrinsèque ; il s'agit d'événements auxquels on répond plus pour relâcher la pression qu'ils engendrent que par plaisir. Les résultats de l'étude empirique — corrélationnelle cette fois — montrent que c'est la satisfaction engendrée par le résultat des activités issues des demandes qui est lié au bien-être négatif, alors que c'est simplement le nombre d'activités issues des désirs qui est lié au bien-être positif (et pas la satisfaction engendrée par ces activités).

Cette distinction entre désirs et demandes ne fait que reprendre une vieille différenciation entre les théories téléologiques et les théories auto-téléologiques. Dans les premières, il y a bien-être lorsque des états finaux (des besoins, des buts...) sont atteints. Dans les secondes, le bien-être réside plus dans la progression vers ces états finaux que dans leur accomplissement. On peut y superposer aussi certains types de plaisir distingués par Murray (1938/1953) et notamment, le plaisir de l'effet accompagnant la satisfaction de la tension créée par un besoin et le plaisir de l'activité.

Plus récemment, le principal promoteur des théories auto-téléologiques est Csikszentmihalyi (e.g. Csikszentmihalyi, 1975 ; Csikszentmihalyi & Wong, 1991) qui a cependant ajouté une condition essentielle pour que l'activité soit source de bonheur : il faut que celle-ci requiert un niveau de compétence qui correspond approximativement aux aptitudes maximales de l'individu. C'est dans des situations de ce type, appelées "*flow experience*" par Csikszentmihalyi, que le bien-être est le plus élevé^[35].

Au vu de ce qui précède, rien ne permet d'opposer les théories téléologiques et les théories auto-téléologiques. Ainsi que le montre l'étude de Reich & Zautra (1983), elles peuvent même être conçues comme complémentaires. C'est également l'avis de Omodei & Wearing (1990) qui proposent un modèle intégrant ces deux grandes approches : "la satisfaction des besoins et l'engagement (dans des activités) sont des sources conceptuellement distinctes de bien-être, mais elles se chevauchent empiriquement parce qu'elles partagent une source commune : la perception des opportunités pour la satisfaction des besoins." (p. 763). Autrement dit, la perception d'opportunités pour satisfaire ses besoins conduit d'une part, à s'engager dans des activités, ce qui fournit des récompenses intrinsèques et d'autre part, à adopter des comportements qui ont pour résultat la satisfaction des besoins, ce qui conduit à des récompenses extrinsèques. Autant

[35] : Si l'activité requiert des compétences supérieures aux capacités de l'individu, celui-ci a tendance à être anxieux ; à l'inverse, si l'activité requiert des compétences nettement inférieures à ses potentialités, il va s'ennuyer.

l'intention d'intégration théorique est louable, autant la démonstration empirique est peu convaincante. En particulier, l'opérationnalisation des variables dépendantes est tellement proche de celles des variables indépendantes que les relations mises en évidence sont triviales. Nous allons voir un exemple de ce type de difficultés dans le paragraphe suivant.

D. Les modèles du bien-être basés sur l'influence indirecte des changements désirés

Dès 1967, Wilson notait que si le bien-être dépend de la satisfaction des besoins, un besoin n'est jamais satisfait dans l'absolu, car lui-même dépend du niveau d'aspiration des individus.

Actuellement, l'influence indirecte des changements désirés est surtout étudiée à partir du concept de buts personnels. Les travaux comme ceux d'Emmons (e.g. 1986, 1989, 1991, 1992) insistent sur le fait que ce sont les évaluations des buts qui sont liées au bien-être. Dans une étude typique (e.g. Emmons, 1986), on demande aux sujets de donner une liste de 15 objectifs qu'ils essaient d'accomplir ou d'atteindre. Chacun de ces objectifs est ensuite évalué par les sujets en fonction de différentes caractéristiques (e.g. ambivalence, importance, valeur, probabilité de succès, etc.). Ces réponses sont agrégées sur l'ensemble des objectifs et corrélées avec différents indicateurs de bien-être subjectif. Les résultats les plus importants sont les suivants (Emmons, 1986) :

- l'affectivité positive est corrélée avec la valeur accordée aux objectifs, avec la réussite passée et avec l'effort que l'objectif nécessite ;
- l'affectivité négative est corrélée avec la perception d'une faible probabilité de succès, avec les conflits entre les objectifs et avec leur ambivalence ; la probabilité de succès a une relation curvilinéaire (courbe en U) avec l'affectivité négative ;
- la satisfaction (composante plus cognitive du bien-être) est associée à la valeur accordée aux objectifs, à leur importance, à l'absence de conflit entre eux, à la réussite passée et à la probabilité de succès.

Malheureusement, les relations dégagées sont en partie triviales. Par exemple, pour mesurer l'affectivité positive, on demande aux sujets d'indiquer s'ils sont heureux, joyeux et pour mesurer la valeur accordée aux objectifs (qui est une caractéristique corrélée avec l'affectivité positive), on leur demande d'indiquer s'ils seront heureux ou joyeux s'ils atteignent leurs objectifs. Plus généralement, ce type d'études pose un problème crucial quant au sens de la causalité. Peut-on dire qu'une personne est heureuse parce qu'elle évalue positivement (pour simplifier) les objectifs qu'elle poursuit ou doit-on dire l'inverse ?

On regrettera aussi que ces études, qui se veulent idiographiques, n'accordent finalement que très peu d'importance à la nature des buts personnels. Ceux-ci ne sont qu'un prétexte pour déterminer comment les individus perçoivent et évaluent leurs objectifs. Ce n'est que récemment que Emmons (1992) s'est intéressé au contenu de ces buts en les distinguant selon qu'ils sont concrets ou abstraits. Plus intéressante encore est l'étude de Rapkin et Fischer (1992) qui montre que le type d'objectifs poursuivis par des personnes âgées diffère selon le niveau de bien-être. Dans les deux cas, c'est plus la démarche suivie qui mérite un intérêt que les résultats qui restent relativement spécifiques et qui demandent à être répliqués.

Enfin, il existe certains modèles dans lesquels ce n'est pas tant l'atteinte d'un but qui est source de bien-être que la rapidité avec laquelle l'individu estime qu'il progresse vers ce but. L'idée selon laquelle le plaisir est liée à la vitesse de la réduction de la tension créée par un besoin se trouve déjà chez Freud (cf. Murray, 1938/1953) : elle connaît aujourd'hui un regain d'intérêt à travers des travaux théoriques comme celui de Carver & Scheier (1990) ou des travaux plus expérimentaux comme celui de Hsee & Abelson (1991).

5.3.4 Les modèles du bien-être basés sur la personnalité

Autant les modèles présentés ci-dessus sont nombreux et diversifiés, autant les théories du bien-être qui reposent sur l'influence de la personnalité sont rares et peu variés. Costa et al. (1987 a) notent avec humour que s'il existe une "erreur fondamentale d'attribution" qui consiste à surestimer les facteurs dispositionnels dans l'explication du comportement, c'est l'inverse qui semble se produire pour l'explication d'expériences subjectives comme le bien-être.

Deux dimensions se dégagent très nettement des travaux portant sur les liens entre personnalité et bien-être : l'extraversion et le névrosisme. Plus spécifiquement, ces travaux (e.g. Costa & McCrae, 1980 ; Warr et al., 1983 ; Emmons & Diener, 1985 b ; Headey & Wearing, 1989 ; Hepburn & Eysenck, 1989) mettent en évidence une relation entre l'extraversion et l'affectivité positive, d'une part et entre le névrosisme et l'affectivité négative, d'autre part. Les effets croisés, lorsqu'ils sont mis en évidence, sont nettement plus faibles.

Nous avons vu que la stabilité temporelle des mesures de bien-être était bien établie. C'est cette stabilité qui incite certains chercheurs à évoquer l'influence de caractéristiques stables de la personnalité sur le bien-être plutôt que des influences environnementales. L'étude la plus connue est sans doute celle de Costa & McCrae

(1980), qui montre que l'extraversion et le névrosisme peuvent prédire, respectivement, les scores aux items positifs de l'ABS de Bradburn et les scores aux items négatifs de cette échelle après un intervalle de 10 ans. Les corrélations ne sont pas très élevées (respectivement .23 et .39) mais il faut noter que l'écart entre les deux enquêtes est très important, que les coefficients de fidélité de l'ABS sont quelquefois tout juste moyens et que ces valeurs sont, somme toute, bien supérieures à celles qui sont observées avec des indicateurs de conditions de vie. Les modèles basés sur la personnalité sont donc très simples. Ils prennent la forme représentée dans la figure 2.26.

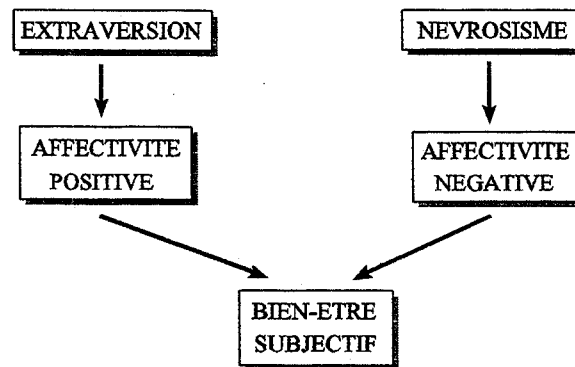


Figure 2.26 : Les modèles du bien-être basés sur l'influence de la personnalité

Les partisans de modèles de la personnalité ne refusent pas l'idée que le bien-être puisse être sensible, de manière momentanée, à certaines conditions de l'environnement. Cela est possible si on considère que le bien-être est composé à la fois d'états affectifs à court terme et d'états affectifs à long terme (Beiser, 1974). Kozma et al. (1990) montrent qu'il est possible de trouver des paradigmes qui permettent de distinguer ces deux aspects.

Si la simplicité des modèles de bien-être basés sur l'influence de la personnalité peut être mise à leur actif, elle constitue aussi leur principal inconvénient. Ces modèles décrivent plutôt qu'ils n'expliquent et on ne comprend pas très bien comment s'opèrent les relations entre bien-être et personnalité. Stones & Kozma (1986 a) ont proposé de concevoir le bien-être comme une "propension". Une propension se définit comme une variable latente, plus large qu'un trait de personnalité, qui contribue à la régularisation du comportement dans des situations différentes et à des moments différents. Cette conception a le mérite de changer quelque peu la perspective classique avec laquelle le bien-être est envisagé. Elle permet d'intégrer des résultats obtenus dans des approches différentes mentionnées antérieurement (stabilité des mesures de bien-être, influence de l'affectivité positive sur les comportements sociaux, sur l'interprétation de stimuli ambigus, etc.), mais elle reste encore trop générale et très descriptive.

Les travaux les plus prometteurs dans ce domaine sont sans doute ceux qui cherchent à clarifier les relations entre personnalité et émotions. Nous les avons déjà évoqués et ils prennent tout leur sens dans la perspective qui nous occupe ici. L'article de Larsen & Ketelaar (1991) est tout à fait exemplaire à cet égard. Selon ces auteurs (p. 132), le postulat des recherches sur les relations entre personnalité et émotions est que l'extraversion représente une *susceptibilité* aux états d'affectivité positive alors que le névrosisme représente une susceptibilité aux états d'affectivité négative. On ne sait pas très bien comment cette susceptibilité se manifeste. McCrae & Costa (1991, cités par Larsen & Ketelaar, 1991) suggèrent deux hypothèses. Dans une conception en termes de tempérament, des traits de personnalité pourraient représenter des différences endogènes dans la sensibilité aux stimuli positifs et affectifs^[36] qui produiraient, à long terme et de manière directe, des différences dans les niveaux d'affectivité positive et d'affectivité négative. Dans une conception plus instrumentale, certaines dimensions de la personnalité favoriseraient la création de conditions de vie spécifiques qui influenceraient, de manière indirecte cette fois, les niveaux de bien-être à long terme. Pour McCrae & Costa, ces deux conceptions ne s'excluent pas. Il est possible que certains traits de personnalité aient une influence directe (e.g. l'extraversion et le névrosisme), alors que d'autres aient une influence indirecte (les autres facteurs du "*big-five*"). Dans leur étude empirique, Larsen & Ketelaar (1991) vérifient que des sujets extravertis sont plus sensibles à une induction expérimentale d'humeur positive alors que les sujets instables émotionnellement sont plus sensibles à une induction expérimentale d'humeur négative^[37].

Longtemps négligés, notamment à cause des attaques virulentes dont a fait l'objet la psychologie de la personnalité dans les années 70, les modèles du bien-être basés sur l'influence de traits de personnalité méritent d'être reconsidérés avec attention. Nul doute que le renouveau actuel des recherches sur la personnalité y contribue de manière substantielle.

[36] : On se rappellera les travaux de Gray (1981) qui distingue un système d'activation comportemental lié aux signaux de récompense et un système d'inhibition comportemental lié aux signaux de punition.

[37] : Nous pouvons noter au passage une conséquence inattendue de ce résultat. Les études expérimentales sur l'humeur qui ne prennent pas la peine de contrôler le niveau de bien-être initial des individus ont toutes les chances de se priver de résultats très intéressants. Il est possible aussi que des phénomènes de ce genre puissent expliquer des divergences entre les résultats d'études où l'humeur est induite expérimentalement et les résultats de celles qui ont lieu en situations naturelles.

5.3.5 Les modèles constructivistes du bien-être

Le titre de ce paragraphe pourrait tout aussi bien être formulé au singulier puisqu'il est vrai que ces modèles se limitent pratiquement à celui de Schwarz & Strack (1991), qui fait la synthèse de plusieurs années de travaux expérimentaux.

Même si le modèle de Schwarz & Strack (1991) a certains points communs avec d'autres modèles relativistes décrits dans ce chapitre, nous le considérons à part parce qu'il n'est pas destiné à expliquer pourquoi les personnes sont heureuses, mais pourquoi elles répondent qu'elles sont heureuses (p. 28). Au vu des nombreuses publications des auteurs et de leurs collègues, il est difficile de savoir s'ils pensent qu'il existe des personnes plutôt heureuses et des personnes plutôt malheureuses. Ce qui est sûr, c'est qu'ils refusent de considérer les réponses aux items de bien-être comme des indicateurs valides de l'état interne des sujets.

Ce modèle s'inscrit dans le cadre de la cognition sociale dont l'un des principes de base est que dans des tâches de jugement (comme les réponses à des items de bien-être), c'est l'information la plus accessible qui est utilisée par les individus. Comme l'information la plus accessible dépend largement d'influences transitoires dues à la situation, ces modèles sont appelés constructivistes pour mettre en relief le fait que les individus "construisent" leurs réponses en fonction de ces informations fournies par la situation.

Le modèle de Schwarz & Strack (1991) est représenté dans la figure 2.27. Nous en expliquerons l'esprit sans entrer dans le détail des nombreux résultats expérimentaux que le lecteur pourra trouver dans l'article de synthèse de 1991.

Lorsque la question posée aux sujets est générale (e.g. "êtes-vous heureux ?" ; "êtes-vous satisfait de votre vie ?"), la tâche demandée est très difficile. Les sujets ont alors tendance à utiliser leur état affectif du moment, leur humeur pour répondre. Plusieurs expériences ont montré que les sujets se déclarent plus heureux après qu'une humeur positive ait été induite. Mais, pour des raisons diverses, l'humeur peut ne pas constituer une source d'information pertinente pour répondre. C'est le cas, par exemple, quand elle est peu prononcée, quand elle est attribuée à des causes externes (e.g. le temps qu'il fait) ou encore lorsque d'autres sources d'informations sont plus directement accessibles. Des manipulations expérimentales ont effectivement montré que lorsque les sujets étaient incités à considérer que le temps était responsable de leur humeur, il n'y avait plus de différences entre les sujets les jours de pluie et les jours de beau temps.

La sélection d'un standard de comparaison dépend, lui aussi, du principe d'accessibilité cognitive. Les expériences qui montrent que les sujets utilisent ces

standards de comparaison sont construites de manière à fournir aux sujets des informations pertinentes que l'on obtient par une question préalable. Par exemple, on demandera au sujet de penser à des événements heureux (vs malheureux) qui lui sont arrivés récemment (vs dans le passé) puis on l'interrogera sur sa satisfaction générale. Dans ce cas, on induit une stratégie de comparaison intrapersonnelle mais des standards de référence interpersonnels peuvent tout aussi bien être utilisés. Les résultats de l'expérience sur le rappel d'événements fournissent des résultats assez curieux qui illustrent les étapes d'évaluation et d'intégration des informations. Les individus répondent qu'ils sont plus heureux lorsqu'ils doivent se rappeler des événements positifs du présent. Ils sont aussi plus heureux quand on leur a demandé de se remémorer des événements négatifs du passé. Dans le premier cas (événements présents), les auteurs estiment qu'il y a un effet d'assimilation, c'est-à-dire que les sujets utilisent de l'information directement accessible par rapport à leur état actuel. Dans le second cas (événements passés), il y a un effet de contraste ; les événements malheureux du passé servent de standard de comparaison qui conduisent les sujets à se déclarer plus heureux actuellement.

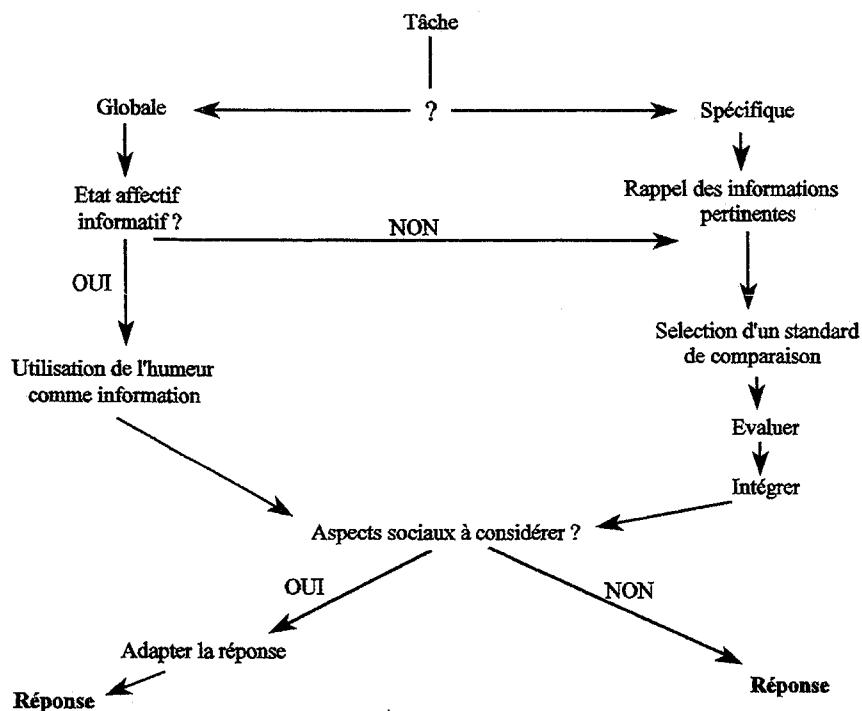


Figure 2.27 : Un modèle constructiviste du bien-être (d'après Schwarz & Strack, 1991, p. 43)

Après que les sujets aient formé un jugement, ils considèrent encore, selon les auteurs, les aspects sociaux de la situation dans laquelle ils se trouvent lorsqu'ils répondent aux items de bien-être. Certains résultats montrent, en effet, que les personnes

s'estiment plus heureuses dans des entretiens en face à face que dans des entretiens téléphoniques. D'autres établissent que les sujets rapportent un niveau de bien-être moins élevé lorsqu'ils sont en présence d'un expérimentateur handicapé. Les auteurs estiment que ce n'est pas tant un effet de la désirabilité sociale, conçue comme une source de différences interindividuelles, qu'un effet de la situation. Ceci leur permet d'expliquer que les corrélations entre échelles de bien-être et échelles de désirabilité sociale ne sont pas très élevées.

Le modèle décrit dans la figure 2.27 reste un modèle théorique hypothétique. Les résultats rapportés en faveur du modèle sont partiels : ils portent chaque fois sur un aspect particulier. Il est en fait assez difficile d'imaginer que ce modèle puisse être testé dans sa globalité. Si la multitude d'effets transitoires mise en évidence par Schwarz et ses collègues existe réellement, les chercheurs de ce courant deviennent en quelque sorte victimes du succès de leurs manipulations expérimentales. Comme tout (ou presque) peut influencer sur les réponses (l'ordre des items, la formulation des items, le nombre de points de l'échelle de réponses, le libellé des échelons, le temps qu'il fait, le dernier résultat de l'équipe nationale de football...), il faudrait pouvoir tenir compte de l'ensemble de ces variables pour tester un modèle global avec des effets interactifs sans doute très complexes. Les auteurs eux-mêmes semblent se perdre un peu dans la multitude de ces résultats puisque l'article de 1991 comporte plusieurs erreurs de cohérence entre les commentaires et les chiffres des tableaux.

Il semble en fait que ces travaux étudient ce que d'autres considèrent comme de l'erreur de mesure. L'existence d'influences transitoires sur les réponses n'est nullement une surprise ; c'est d'ailleurs pour cela que les psychométriciens conseillent de construire des instruments comportant plusieurs items afin que ces influences aient tendance à s'annuler.

Il y a donc deux façons d'envisager les modèles constructivistes. La plus radicale consiste à considérer que les réponses aux questionnaires de bien-être sont totalement déterminées par les circonstances et ne peuvent être considérées comme des indicateurs valides de l'état interne du sujet. Cette conception radicale n'a pas grand intérêt et elle est peu défendable. La grande majorité des résultats obtenus par Schwarz et ses collaborateurs reposent sur des expériences dans lesquelles la variable dépendante (le bien-être) est mesurée avec un seul item, ce qui favorise, bien évidemment, l'apparition d'effets transitoires. Depuis l'étude de Pavot & Diener (1993 b), on sait que ceux-ci disparaissent ou sont fortement atténués lorsque le bien-être est mesuré avec plusieurs items. Sur le plan théorique, cette version radicale n'est pas plus satisfaisante. Contrairement à ce qu'affirment Schwarz & Strack (1991, p. 28), les mesures de bien-être ont une stabilité temporelle assez impressionnante sur de longues périodes (à condition, bien évidemment, qu'elles comportent plus d'un item). Or, les modèles constructivistes

sont incapables de rendre compte de cette stabilité (même partielle) : si les réponses étaient uniquement sous l'influence d'effets transitoires, les corrélations test-retest devraient être proches de zéro.

Une approche plus modérée consiste à reconnaître l'existence de ces effets et l'intérêt qu'il y a à les étudier pour eux-mêmes, sans nier pour autant l'existence de la stabilité des réponses. Les études menées dans le cadre de la cognition sociale sont précieuses pour mieux comprendre la manière dont les individus répondent aux items de bien-être. Mais ces études doivent venir en complément, non en opposition, des modèles plus traditionnels. Il serait en effet dommage que les querelles stériles qui ont longtemps agité les recherches sur la personnalité gagnent aujourd'hui le domaine du bien-être.

5.4 Approches *top-down* et *bottom-up* du bien-être

5.4.1 Introduction

La distinction entre approches *top-down* et *bottom-up* du bien-être apparaît en filigrane de toute la revue de la littérature que nous avons proposée. Nous pouvons à présent nous y intéresser de manière spécifique, disposant des informations suffisantes pour en faire un examen raisonné.

Nous serons plus précis, plus bas, sur la différenciation entre approches *top-down* et *bottom-up*. Pour l'instant contentons-nous d'énoncer qu'elle renvoie au sens de la causalité entre le bien-être et d'autres variables. Dans les approches *top-down*, le bien-être est conçu comme étant la cause des relations avec d'autres variables, alors que dans les approches *bottom-up*, largement majoritaires, le bien-être est conçu comme étant la résultante d'autres variables.

Le choix entre une approche *top-down* ou une approche *bottom-up* est fondamental et il en recoupe bien d'autres. Il renvoie en particulier au choix d'une théorie définitoire pour le bien-être et à la construction de modèles nomologiques personnologiques ou situationnistes. Surtout, il renvoie à l'étude du bien-être en tant que trait ou en tant qu'état. Pour Veenhoven (1991 a), parler de l'opposition entre un "bien-être trait" et un "bien-être état", c'est exactement la même chose que confronter approches *top-down* et *bottom-up*. En effet, dans une approche *bottom-up*, on mettra l'accent sur la variabilité du bien-être, sur sa sensibilité aux conditions externes, sur la possibilité de le manipuler expérimentalement, bref sur le fait qu'il s'agit d'un état psychologique. En revanche, dans une approche *top-down*, on insistera sur sa stabilité temporelle, sur son potentiel prédictif ou sur son caractère diffus, ce qui constituent les caractéristiques d'un trait psychologique.

Autant il est difficile de nier qu'il existe des états de bien-être qui varient en fonction des circonstances extérieures et qui justifient que l'on s'intéresse à l'émergence d'états d'affectivité positive avec des démarches *bottom-up*, autant il est difficile de ne considérer le bien-être que comme un état. Nous avons déjà vu plusieurs arguments venant appuyer une conception du bien-être comme trait. Ils sont rappelés brièvement ci-dessous :

- les réponses aux questionnaires de bien-être sont relativement stables, y compris sur de longues périodes (cf. Stones & Kozma, 1986 a; McNeil et al., 1986 b ; Costa et al., 1987 a, 1987 b) ;
- le bien-être montre non seulement une certaine stabilité temporelle, mais aussi une constance inter-situationnelle (Diener & Larsen, 1984) ;

- des composantes à court-terme du bien-être ne sont pas suffisantes pour rendre compte de certains résultats, il faut faire l'hypothèse de composantes à plus long terme (cf. Kammann & Flett, 1983 ; Kozma et al., 1990) ;
- il semble y avoir une composante stable dans les questionnaires d'humeur momentanée ainsi que le montrent les résultats des études utilisant les modèles traits-états latents de Steyer (cf. Eid et al., 1994) ;
- il y a une convergence entre la structure des états affectifs et la structure des traits affectifs (cf. Tellegen, 1985 ; Watson & Tellegen, 1985 ; Meyer & Schack, 1989).

Nous garderons à l'esprit ces résultats, mais comme ils sont relativement bien établis et qu'ils ont déjà été discutés, nous examinerons essentiellement, dans cette section, la différenciation entre approches *top-down* et *bottom-up* à partir de la question du sens de la causalité.

5.4.2 Le sens de la causalité : premiers éléments

Etymologiquement, le bonheur, qu'il faut comprendre comme bonne fortune, suggère plus une approche *top-down* qu'une approche *bottom-up*. Pourtant, la littérature scientifique a constamment privilégié la seconde. Pour Stones & Kozma (1980 a, 1986 b), l'opposition entre ces deux approches rejoint la distinction philosophique entre l'idéalisme de Kant et le réalisme réductionniste de Locke. Chez Kant, la compréhension procède des éléments d'ordre supérieur vers des éléments d'ordre inférieur. A l'inverse, dans le réalisme réductionniste, la compréhension doit être recherchée dans les éléments les moins abstraits, ceux-ci se combinant de manière simple pour former une composante d'ordre supérieur. C'est en référence à cette seconde conception qu'il faut comprendre toutes les études dans lesquelles le bien-être est réduit à la satisfaction dans différents domaines de vie ou conçu comme étant dépendant des conditions ou des événements de vie. Ceci est particulièrement net dans les théories que Heyink (1993) appelle hédonistes, dans lesquelles le bien-être est considéré comme la somme d'expériences plaisantes. C'est aussi le cas des études où le bien-être est défini comme le niveau moyen d'affectivité positive sur une longue période (e.g. Seidlitz & Diener, 1993).

L'opposition entre approches *top-down* et *bottom-up* ne se retrouve pas qu'en philosophie ou dans les recherches sur le bien-être. La question du sens de la causalité a été posée à plusieurs reprises dans les travaux sur la dépression, par exemple. Alors que Lewinsohn (1974) suggère que la dépression provient d'un manque de renforcements positifs, Sweeney et al. (1982) se demandent si ce ne serait pas plutôt la dépression qui serait la cause d'une faible quantité de renforcements positifs. Il en est de même dans les modèles qui font dépendre la dépression de l'attribution des événements négatifs à des

causes internes, stables et incontrôlables. Les résultats de l'étude longitudinale de Cochran & Hammen (1985) semblent indiquer qu'il serait plus approprié d'envisager en sens inverse la relation entre dépression et style attributif (cf. aussi la méta-analyse de Sweeney et al., 1986).

Il serait nécessaire de s'interroger un jour sur les caractéristiques du contexte historico-culturel qui ont fait de l'approche *bottom-up* la conception dominante. Elles restent difficiles à analyser. Il y a plusieurs exemples presque caricaturaux qui illustrent à quel point cette conception a imposé son point de vue. Le plus frappant est celui que l'on peut suivre dans l'évolution des publications de Bruce Headey et de ses collaborateurs. Ces derniers sont à l'origine de l'enquête longitudinale la plus intéressante qui ait été menée sur le bien-être. Dans leurs premières publications, essentiellement théoriques, ces auteurs affichent clairement une position *bottom-up* (e.g. Headey et al., 1985 b). Pourtant, dès 1984, Headey et al. (1984 a) ont à leur disposition les données d'une pré-enquête portant sur un petit nombre d'individus. Il est manifeste, à la lecture de cet article, que certains résultats les laissent perplexes tant ils viennent remettre en cause leurs postulats théoriques. Ils trouvent notamment que le bien-être influence l'occurrence d'événements de vie ultérieurs et que le sens de la relation entre le bien-être général et la satisfaction dans un domaine de vie dépend du domaine. Mais ces résultats ne suffisent pas à les convaincre : "*On balance, and purely on logical grounds rather than statistical grounds, we believe it reasonable to continue to treat well-being as our dependent variable.*" (Headey et al., 1984 a, p. 209). Ce n'est que quelques années plus tard, après trois vagues de panel successives, que les auteurs se rendront à l'évidence : les conceptions (exclusivement) *bottom-up* du bien-être ne sont pas tenables au niveau empirique (Headey et al., 1989)^[38]. Nous reviendrons plus tard en détail sur cette étude qui peut être considérée comme l'une des contributions les plus originales aux recherches sur le bien-être.

Headey et ses collaborateurs ne sont pas les seuls à s'être interrogés sur la manière dont il fallait concevoir les relations entre le bien-être et les autres variables. De nombreux auteurs évoquent, dans les commentaires de leurs résultats, la possibilité que les relations mises en évidence puissent être attribuées à un effet du bien-être, avant de se rallier finalement, pour la plupart, à la conception dominante. Prenons quelques exemples :

- Bradburn (1969) reconnaît que l'on peut s'interroger sur le sens de la causalité entre l'affectivité positive et des variables comme le niveau de participation sociale ou le caractère varié des expériences ;

[38] : Pour l'anecdote, on notera que le nom de Ruut Veenhoven (e.g. Veenhoven, 1984, 1988) figure dans la liste des auteurs de l'article de 1989 et que le fait que celui-ci se soit interrogé très tôt sur le sens des relations entre le bien-être et ses corrélats n'est sans doute pas totalement étranger à ce revirement de position.

- Emmons (1986), dans ses travaux sur les relations entre bien-être et buts personnels, admet qu'aucune causalité ne peut être déterminée à partir d'une corrélation ; pourtant, toute sa discussion des résultats envisage l'atteinte des buts comme la cause du bien-être ;
- Rapkin & Fischer (1992) font une constatation identique lorsqu'ils montrent que le niveau de bien-être est lié à la nature des buts poursuivis ;
- Kammann et al. (1984 a) émettent des réserves quant à la manière dont est habituellement envisagée la relation entre le bien-être et la santé ;
- Andrews & Withey (1976, p. 115) signalent, dans une note, que leurs analyses sont basées sur l'*intuition* que les évaluations générales du bien-être sont causées par des évaluations plus spécifiques, même si rien dans leurs données ne permet de tester ce postulat.

Ces interrogations légitimes n'ont pas suffi pour que s'impose une autre conception. Les rares auteurs qui ont défendu une approche *top-down* sont ceux qui ont proposé des modèles du bien-être basés sur l'influence de la personnalité. Mais leurs contributions empiriques se sont souvent limitées à montrer que des traits de personnalité pouvaient être des prédicteurs du bien-être, plutôt qu'à essayer d'établir que les corrélations mises en évidence par les tenants d'une approche *bottom-up* pouvaient être interprétées dans un sens inverse. L'une des rares exceptions est l'étude de Stones & Kozma (1986 b)^[39] qui tentent de montrer, à partir de données longitudinales, que le bien-être général influence la satisfaction dans différents domaines de vie plutôt que l'inverse. A part cela, les arguments utilisés sont peu étayés au niveau empirique (cf., par exemple, ceux sur lesquels s'appuient McNeil et al., 1986 a). Il en est de même de certaines analyses que l'on peut trouver dans la psychologie humaniste. Pour Maslow (1968/1972), par exemple, le bien-être est à la fois la conséquence et la cause de la croissance personnelle vers l'épanouissement, l'actualisation de soi ; le bien-être pourrait aussi faciliter les contacts sociaux et stimuler l'activité. Lorsque Veenhoven (1988) fait la synthèse des études ayant pris le parti de montrer les effets du bien-être, deux éléments apparaissent marquants : la rareté des études mais aussi et surtout, les résultats positifs qui semblent indiquer que le bien-être peut influencer la santé, l'espérance de vie, l'implication ou encore les contacts sociaux (cf. chapitre 4).

Mais ces résultats, obtenus quelquefois avec des méthodologies discutables, ne sont pas suffisants pour affirmer l'idée d'une conception *top-down*. Avant d'examiner quelques nouveaux éléments dans ce débat, la distinction entre approches *top-down* et *bottom-up* mérite d'être clarifiée. Certains auteurs l'utilisent dans une acception étroite dans laquelle ne sont considérées que les relations entre le bien-être général et les

[39] : Le titre de leur article illustre d'ailleurs tout à fait la thèse défendue dans l'approche *top-down* : "*Happy are they who are happy...*".

satisfactions plus spécifiques, les autres variables (conditions de vie, événements de vie, etc.), étant toujours conçues comme des variables externes. Dans une acception plus large, elle renvoie à une opposition entre les théories qui conçoivent le bien-être comme la résultante de l'influence de variables exogènes et celles qui le conçoivent comme la cause de ces mêmes variables. Si on utilise la terminologie développée dans la première partie, on dira que l'acception étroite renvoie aux théories définitoires et l'acception large aux théories nomologiques. C'est à l'aide de cette distinction que nous organiserons la suite de ce paragraphe.

5.4.3 Les approches *top-down* et *bottom-up* sur le plan des théories définitoires

Sous l'angle des théories définitoires, les approches *top-down* et *bottom-up* se distinguent quant au sens de la relation entre bien-être général et satisfactions spécifiques. Nous avons vu (section 5.3.1) qu'un postulat largement partagé par les chercheurs faisait dépendre le bien-être de satisfactions plus spécifiques dans différents domaines de vie.

Ce postulat renvoie à une vieille question soulevée par Curtis & Jackson (1962) sur les indicateurs-effets et les indicateurs-causes. Elle a été remise au goût du jour par Bollen, en 1984, à la lumière des modèles d'équations structurales (cf. aussi Cohen et al., 1990 ; Bollen & Lennox, 1991).

Dans leur article de 1962, Curtis et Jackson font remarquer que des indicateurs d'un même concept ne sont pas forcément corrélés. Autrement dit, deux indicateurs corrélés *positivement* à un même concept peuvent avoir entre eux une corrélation nulle voire négative. C'est le cas par exemple de la satisfaction maritale et de la durée du mariage, toutes deux liées positivement à la stabilité maritale, alors qu'elles sont corrélées négativement entre elles.

Cette conception contredit de manière fondamentale la perspective classique de construction d'échelles qui requiert une consistance interne élevée, donc des corrélations positives entre indicateurs.

Pour clarifier cette opposition, il est nécessaire de distinguer entre les indicateurs qui sont les *causes* du concept mesuré et les indicateurs qui en sont les *effets*. Bollen (1984) propose une manière très simple d'envisager ces deux types d'indicateurs.

Dans le cas d'"indicateurs-effets" (il faut lire "indicateurs conçus comme étant les effets de la variable latente"), le concept sous-jacent est la cause des indicateurs. En utilisant la logique et les notations des modèles d'équations structurales (cf. Jöreskog &

Sörbom, 1993), on peut représenter graphiquement un cas simple dans lequel un concept (ξ) est mesuré à partir de deux indicateurs-effets (x_1 et x_2) (figure 2.28).

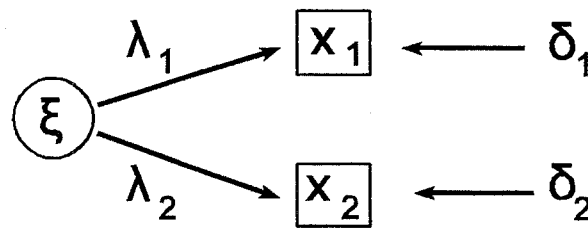


Figure 2.28 : Relations entre un concept et deux indicateurs-effets (d'après Bollen, 1984)

Les équations correspondant à ce modèle sont les suivantes :

$$\begin{aligned} x_1 &= \lambda_1 \xi + \delta_1 \\ x_2 &= \lambda_2 \xi + \delta_2 \end{aligned}$$

Sans que cela ne nuise à la généralité du raisonnement, ce modèle suppose que les erreurs ne sont ni corrélées entre elles ni avec la variable latente et que les saturations des variables observées sur la variable latente sont positives. Ces postulats s'expriment mathématiquement de la manière suivante :

$$\begin{aligned} \text{cov}(\xi, \delta_i) &= \text{cov}(\delta_1, \delta_2) = 0 \\ \lambda_1 &> 0 \\ \lambda_2 &> 0 \end{aligned}$$

La covariance entre x_1 et x_2 est de même signe que leur corrélation. Or :

$$\text{cov}(x_1, x_2) = \text{cov}(\lambda_1 \xi + \delta_1, \lambda_2 \xi + \delta_2) = \lambda_1 \lambda_2 \sigma_\xi^2$$

Il en découle, comme σ_ξ^2 est nécessairement positif, que la corrélation entre x_1 et x_2 est positive. Ce résultat a une conséquence importante. Il indique que le modèle de la consistance interne est approprié dans le cas où les indicateurs sont conçus comme des effets du trait latent. Une corrélation positive entre deux indicateurs est donc une condition nécessaire (mais non suffisante) pour que ces indicateurs mesurent un même concept dans le cas d'indicateurs-effets.

Dans le cas d'"indicateurs-causes" (il faut lire "indicateurs conçus comme ayant une influence causale sur la variable latente"), la situation est différente ainsi que le montre la figure 2.29. Un exemple classique d'indicateurs-causes est le revenu et le

niveau d'éducation pour le statut socio-économique^[40]. Dans ce cas, la variable latente est une variable émergente selon Cohen et al. (1990).

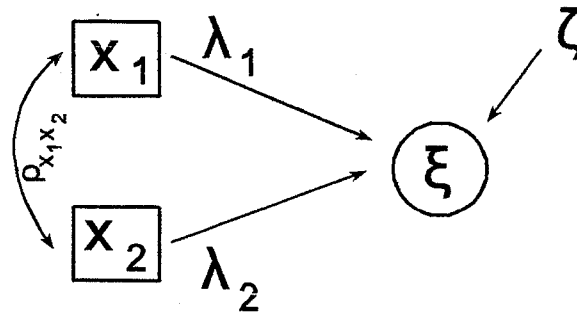


Figure 2.29 : Relation entre un concept et deux "indicateurs-causes" (d'après Bollen, 1984)

L'équation correspondant à ce modèle est la suivante :

$$\xi = \lambda_1 x_1 + \lambda_2 x_2 + \zeta$$

avec :

$$\text{cov}(x_1, \zeta) = \text{cov}(x_2, \zeta) = 0$$

$$\lambda_1 > 0$$

$$\lambda_2 > 0$$

On s'aperçoit que ce modèle n'impose plus aucune contrainte sur la corrélation entre x_1 et x_2 : elle peut être positive, nulle ou négative. Autrement dit, dans le cas d'"indicateurs-causes", le modèle de la consistance interne n'est pas approprié.

Dans sa conclusion, Bollen souligne l'importance de spécifier, avant toute analyse, un modèle de mesure explicite. De plus, il est incohérent de faire quelque postulat que ce soit sur les relations entre les indicateurs avant de déterminer si ce sont des causes ou des effets du concept que l'on cherche à mesurer. Une conséquence importante de ce raisonnement est qu'il fait peu sens de vérifier la consistance interne ou d'effectuer une analyse factorielle dans le cas où les variables sont conçues comme des indicateurs-causes du concept.

Dans les approches *bottom-up*, les satisfactions spécifiques sont conçues comme des indicateurs-causes du bien-être. Cela n'empêche nullement les auteurs d'adopter, implicitement, le modèle de la consistance interne lorsqu'ils effectuent des analyses

[40] : D'autres exemples proposés par Cohen et al. (1990) sont les suivants. Le nombre d'enfants, la maladie de la mère et le nombre d'heures de travail ("indicateurs-causes") pour la disponibilité de la mère pour l'éducation des enfants (variable latente ou émergente). Les niveaux de cholestérol, de pression sanguine et l'histoire familiale des maladies cardiaques ("indicateurs-causes") pour le risque d'accident cardio-vasculaire (variable émergente).

factorielles sur ces variables ou lorsqu'ils évaluent la fidélité d'instrument à partir d'un coefficient α de Cronbach. Autrement dit, les auteurs font comme si les différents indicateurs devaient être corrélés positivement entre eux, ce qui n'est nullement une nécessité dans cette conception.

On pourrait objecter que, de toute façon, les résultats empiriques ayant montré que ces corrélations étaient effectivement positives sont suffisamment nombreux pour que se justifie le modèle de la consistance interne. Mais il se pose dans ce cas un problème théorique délicat. Si des "indicateurs-causes" sont corrélés entre eux, il est raisonnable de penser qu'ils sont eux-mêmes causés par une (ou plusieurs) variable(s) latente(s). Dans une approche *bottom-up*, cette variable latente commune *ne peut être* le bien-être puisque celui-ci est conçu comme étant l'effet de ces indicateurs. Quelle est alors cette variable latente ? Quel est le statut des facteurs dégagés d'analyses factorielles dans les approches *bottom-up* ? Il n'y a, pour l'heure, aucune réponse à ces questions qui ne sont d'ailleurs pas posées dans la littérature. Nul doute, pourtant, qu'il faille le faire un jour.

Mises à part ces considérations générales, certains résultats classiques de la littérature s'accommodent assez mal d'une approche *bottom-up*. C'est le cas, déjà évoqué ci-dessus, des corrélations positives entre des satisfactions dans des domaines très différents. Pourquoi, par exemple, une personne qui se déclare satisfaite avec sa santé a-t-elle également tendance à se déclarer satisfaite avec le gouvernement de son pays ? Dans une approche *top-down*, ce résultat se comprend très bien alors qu'un modèle *bottom-up* nécessiterait des explications fort complexes.

Un autre résultat gênant pour les partisans d'une approche *bottom-up* est la faillite des modèles multiplicatifs du bien-être, c'est-à-dire des modèles dans lesquels on essaie d'améliorer la prédiction du bien-être général en pondérant les satisfactions dans les domaines par l'importance que les individus accordent à chaque domaine. Nous avons vu (section 5.3.1) que ces résultats négatifs n'étaient probablement pas dus à des biais méthodologiques. Les seuls qui, à notre connaissance, ont tenté d'expliquer ce phénomène sont Campbell et al. (1976) qui considèrent que si la pondération par l'importance n'améliore pas la prédiction du bien-être général, c'est parce qu'une personne insatisfaite dans un domaine aura tendance à le juger peu important. Cette explication est plausible, mais il faudrait la vérifier. Il est de toute façon plus économique de penser que la satisfaction dans un domaine est fonction du bien-être général d'un individu, auquel cas il n'y a aucune raison pour que l'importance accordée au domaine joue un rôle dans la relation.

Outre les résultats qu'il est difficile d'expliquer dans le cadre d'une approche *bottom-up*, il y a ceux qui mériteraient une interprétation toute différente de celle que font les auteurs. Par exemple, Near et al. (1983) testent les relations entre satisfaction

générale, satisfaction avec le travail et satisfaction avec des domaines autres que le travail. Dans leur modèle, la satisfaction générale est la variable dépendante, les satisfactions spécifiques constituent des variables intermédiaires et diverses variables exogènes sont censées prédire ces satisfactions spécifiques. Les résultats qu'ils obtiennent sont reproduits dans la figure 2.30^[41].

Il importe de remarquer tout d'abord que ces coefficients ne sont pas directement comparables entre eux car la flèche courbe indique une simple corrélation linéaire alors que toutes les autres flèches indiquent des coefficients de régression. Toutefois, comme les relations sont (très) faibles, les coefficients de régression ne doivent pas différer beaucoup de simples coefficients de corrélations. Ce qui est intéressant pour notre propos, c'est qu'il y a une corrélation assez élevée entre la satisfaction avec le travail et la satisfaction avec les autres domaines (.46). En revanche, les coefficients de régression multiple ne sont que de .003 (!) entre satisfaction avec le travail et satisfaction générale et de .21 entre satisfaction avec les autres domaines et satisfaction générale. Ce résultat incite à se demander si la covariation entre les satisfactions spécifiques ne serait pas due à une cause commune qui serait justement la satisfaction générale. Autrement dit, il aurait été tout à fait justifié, au vu de ces résultats, de tester un modèle avec une causalité inversée. Inutile de dire que les auteurs n'en ont rien fait.

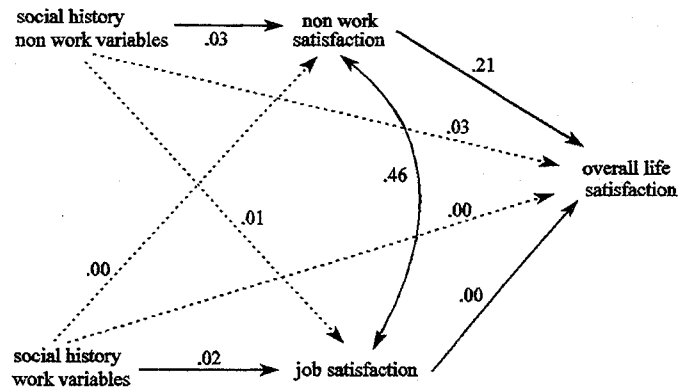


Figure 2.30 : Les résultats des régressions multiples du modèle de Near et al. (1983, p. 140).

Enfin, dans d'autres travaux plus récents, la question de la validité d'une approche *bottom-up* semble explicitement posée. Ainsi, par exemple, Judge & Watanabe (1993), étudient, grâce à des données longitudinales obtenues à cinq ans d'intervalle, la relation entre la satisfaction avec le travail et la satisfaction générale. Les auteurs concluent à des relations bidirectionnelles entre ces deux variables, mais les résultats, là encore, sont plutôt compatibles avec une approche *top-down*. Ainsi, le coefficient de piste entre la

[41] : Notons que les données ont été recueillies dans une étude transversale sur un échantillon représentatif de la population des Etats-Unis (N = 1297). Les résultats de la figure 2.30 ont été obtenus à partir de plusieurs analyses de régression et les coefficients indiqués représentent l'intensité des relations après que les autres variables aient été introduites dans les équations.

satisfaction générale au temps 1 et la satisfaction avec le travail au temps 2 est égal à .26 alors qu'il n'est que de .07 entre la satisfaction avec le travail au temps 1 et la satisfaction générale au temps 2. Citons, pour terminer, l'étude de Lance et al. (1989) qui montre, avec des données transversales cette fois, que le sens de la relation entre satisfaction générale et satisfactions spécifiques dépend des domaines. Selon ces auteurs, la relation serait bidirectionnelle dans le cas du travail, *top-down* dans le cas des activités sociales et *bottom-up* dans celui du mariage. Ces différences pourraient s'expliquer en fonction des caractères spécifique, critique (domaines pour lesquels aucun substitut n'est disponible) et central des domaines. Les domaines spécifiques, critiques et centraux, comme le mariage, contribueraient à la satisfaction générale. Les domaines plus larges, moins critiques et moins centraux seraient influencés par le bien-être général (e.g. les activités sociales). Les domaines larges, critiques et centraux influenceraient et seraient influencés par la satisfaction générale.

5.4.4 Les approches *top-down* et *bottom-up* sur le plan des théories nomologiques

De tous les modèles du bien-être que nous avons présentés dans ce chapitre, il n'y en a pratiquement aucun qui résisterait à une analyse logique montrant que les relations mises en évidence peuvent être considérées comme des effets du bien-être et non l'inverse. C'est un peu plus difficile à envisager pour les modèles basés sur l'influence directe des conditions de vie^[42], mais nous avons vu qu'ils étaient plus ou moins abandonnés. Quant aux modèles basés sur l'influence des événements de vie, nous en reparlerons plus tard.

Cette analyse est d'autant plus facile que les modèles font intervenir des référents internes d'évaluation subjective de tel ou tel aspect. En effet, ces évaluations soit-disant intermédiaires peuvent très bien être "colorées" par le bien-être général d'un individu. De plus, rien n'indique que les individus effectuent réellement les démarches complexes qui leur sont prêtées. On peut d'ailleurs se demander, comme l'ont fait Wortman & Dintzer à propos des modèles attributifs de la dépression (1978, cités par Sweeney et al., 1986), si les individus effectuent ces évaluations de façon spontanée ou s'ils le font uniquement en réponses aux sollicitations des chercheurs ! En se servant *a contrario* des principes de la cognition sociale stipulant que ce sont les informations les plus accessibles qui sont utilisées par les sujets dans des tâches de jugement, Fox & Kahneman (1992) montrent que, dans certains cas, les individus ont recours à leur niveau de satisfaction dans un

[42] : Si c'est plus difficile à envisager, ce n'est tout de même pas impossible, pour certaines variables du moins. On peut penser, par exemple, que ce sont les individus les plus heureux qui seront les mieux à même de recevoir de l'avancement dans leur travail et donc d'augmenter leurs revenus.

domaine pour répondre à une question sur la comparaison entre leur situation et celle des autres !

Nous n'entreprendrons pas une analyse détaillée de tous les modèles. Comme d'habitude, nous nous servirons de quelques exemples qui montreront qu'appliquée aux mêmes résultats, une conception *top-down* est parfois plus plausible ou plus économique qu'une approche *bottom-up*.

Prenons le cas des relations entre le bien-être et la santé. On sait que la santé subjective est plus fortement liée au bien-être (en fait l'affectivité positive) que la santé objective. Les partisans d'une approche *bottom-up* proposent un modèle dans lequel le bien-être est la variable dépendante, la santé une variable exogène et la satisfaction avec la santé une variable intermédiaire. Pour expliquer que l'auto-évaluation de la santé ne correspond pas à l'évaluation objective, il faut faire appel à des mécanismes cognitifs complexes se basant sur des comparaisons intra et/ou interindividuelles qui permettent de construire un jugement relatif de la santé en lieu et place d'une évaluation absolue. En adoptant une conception *top-down*, une explication beaucoup plus simple d'un tel résultat peut être proposée : l'affectivité négative n'est pas nécessairement liée à l'existence de problèmes de santé, mais les personnes qui ont un niveau élevé d'affectivité négative sont plus susceptibles de se plaindre de symptômes physiques que d'autres personnes qui les considèrent comme mineurs.

Il en est de même pour les modèles faisant intervenir des comparaisons. Nous avons vu, avec Diener et al. (1993), que le bien-être était lié aux comparaisons subjectives de revenu, et non aux comparaisons objectives. Dans une approche *bottom-up*, il faut faire intervenir des variables intermédiaires — que l'on ne connaît pas, d'ailleurs — pour expliquer pourquoi les comparaisons subjectives ne correspondent pas aux comparaisons objectives. Il est beaucoup plus économique de faire l'hypothèse d'une relation *top-down* à savoir que c'est le bien-être qui influence la perception des différences entre sa situation et celle des autres.

Au-delà de ces spéculations qui restent, en l'absence de données empiriques, sans grand fondement, on peut constater dans la littérature une tendance, déjà soulignée à propos des théories définitoires, à ne percevoir que les aspects des résultats compatibles avec une approche *bottom-up*. Ainsi Emmons & Diener (1985 a) rapportent une étude des relations entre la satisfaction dans plusieurs domaines et les comparaisons interpersonnelles dans ces mêmes domaines. Les résultats font apparaître des corrélations élevées entre ces deux séries de variables et cet article est souvent cité à l'appui des modèles relativistes. Ce qui nous intéresse dans ces résultats, c'est que les corrélations entre satisfaction et comparaisons interpersonnelles diffèrent de manière sensible selon les domaines. Dans le cadre d'une approche *bottom-up*, cela peut se comprendre. On peut

faire l'hypothèse que l'"influence" des comparaisons interpersonnelles sur la satisfaction est d'autant plus élevée que les individus disposent d'informations sur la situation des individus auxquels ils se comparent. Or, c'est exactement l'inverse qui se produit. Les corrélations sont les plus faibles pour les domaines d'ordre public comme le logement, les notes, les cours suivis — les sujets sont tous des étudiants — alors qu'elles sont particulièrement élevées dans le cas de domaines intimes comme la vie amoureuse ou les amis^[43]. Ces résultats contredisent les hypothèses dérivées d'une approche *bottom-up*. En revanche, ils prennent tout leur sens dans une conception *top-down*. En effet, dans ce cas, il est logique que le niveau de satisfaction influence les comparaisons interpersonnelles dans les domaines pour lesquels les individus disposent de peu d'informations sur la situation d'autrui ; inversement, la satisfaction influence peu les réponses lorsque les individus peuvent réellement comparer leurs conditions à celles de leurs proches^[44].

Dans les modèles d'événements de vie, certains auteurs utilisent des variables intermédiaires faisant intervenir les attributions faites aux événements. Celles-ci peuvent tout aussi bien être la cause du bien-être que son effet. Des résultats obtenus dans le domaine des attitudes sont compatibles avec une telle hypothèse. Ainsi, Regan et al. (1974, cités par Pratkanis, 1989) ont montré que les individus effectuent des attributions internes lorsqu'une personne se comporte de manière consistante avec l'attitude que l'on a vis-à-vis d'elle (e.g. des comportements positifs manifestés par une personne appréciée), mais des attributions externes lorsque le comportement n'est pas consistant avec l'attitude (e.g. des comportements positifs manifestés par une personne dépréciée). Si on considère — comme le font beaucoup d'auteurs — le bien-être comme une attitude, on peut, en se basant sur ces résultats, proposer une hypothèse qui rend compte des relations constatées entre bien-être et attributions. Si une personne a une attitude positive vis-à-vis d'elle-même, elle attribue les événements positifs (qui sont consistants avec son attitude) à des causes internes et des événements négatifs (inconsistants par rapport à son attitude) à des causes externes ; inversement, si une personne a une attitude négative vis-à-vis d'elle-même, elle attribuera les événements positifs à des causes externes et les événements négatifs à des causes internes. Une telle hypothèse rend compte des relations entre bien-être et attributions dans une conception *top-down*.

La situation n'est pas plus claire dans les modèles qui font appel à l'évaluation des buts personnels que poursuivent les individus. On se rappelle que nous avons émis des doutes quant au sens de la causalité entre bien-être et buts à propos des travaux de

[43] : Dans les exemples que nous venons de citer, les corrélations sont respectivement de : .15 ; .25 ; .07 et de .57 et .46. Toutes les mesures sont obtenues avec un seul item.

[44] : Pour être exact, il faut reconnaître que Emmons & Diener (1985 a) évoquent l'hypothèse générale selon laquelle ce serait la satisfaction qui aurait un effet sur les comparaisons et non l'inverse. Mais, comme le font remarquer Fox & Kahneman (1992), cette réserve disparaît dans les citations ultérieures de cette étude qui est habituellement considérée comme l'un des meilleurs arguments en faveur des thèses relativistes du bien-être.

Emmons (1986). L'étude de Brunstein (1993) contribue à répondre en partie à ces critiques en utilisant des données longitudinales recueillies sur une courte période temporelle. Il montre que le bien-être est influencé par le niveau de bien-être précédent et par certaines évaluations des objectifs poursuivis, notamment l'interaction entre l'engagement dans un objectif et son caractère accessible (i.e. si on s'estime fortement engagé dans un but et si on le considère inaccessible, le niveau de bien-être va diminuer). Malheureusement cette étude adopte un point de vue partial. A aucun moment l'auteur n'envisage la possibilité que l'évaluation de l'engagement et/ou du caractère accessible du but soit déterminé, au moins partiellement, par le niveau de bien-être antérieur. Cette hypothèse, issue d'une conception *top-down* aurait dû être testée au même titre que l'hypothèse alternative, issue d'une conception *bottom-up*.

Au vu des lignes qui précèdent, on comprendra que les preuves empiriques en faveur d'une conception *top-down* ne soient pas légion. En effet, pour obtenir des résultats en faveur d'hypothèses, encore faut-il les tester ou, tout du moins, les avoir à l'esprit lors de la conception de l'étude et dans l'interprétation des résultats. Pourtant, si de tels arguments ne sont pas nombreux, ils existent tout de même. Nous avons déjà vu que Mastekaasa (1992) avait réussi à établir, grâce à des données longitudinales, que la relation entre le bien-être et le mariage devait plutôt être comprise comme un effet du bien-être sur la probabilité de se marier. Toujours en rapport avec le mariage, Gilford (1984, cité par McNeil et al., 1986 a) montre que c'est le bien-être qui prédit une interaction maritale positive plutôt que des relations positives entre conjoints qui influencent le bien-être. Citons encore les résultats d'une étude très intéressante de Filipp & Klauer (1991) sur les stratégies de *coping* adoptées en face de l'apparition d'événements critiques comme le cancer. Ces auteurs montrent que certaines stratégies sont plus efficaces que d'autres pour le maintien du niveau de bien-être, mais surtout, que le bien-être influence l'adoption de certaines stratégies et que, dans plusieurs cas, la relation entre bien-être et stratégies de coping est bidirectionnelle. Mais ces arguments restent parcellaires.

Un examen empirique bien plus approfondi des approches *bottom-up* et *top-down* a été proposé par Headey et al. (1989). Cette étude, que nous allons résumer, reste la démonstration la plus magistrale de la faiblesse des modèles *bottom-up*.

Headey et ses collègues se fixent comme objectif de proposer un modèle statistique permettant de résoudre des questions liées au sens de la causalité entre des variables, dès lors que l'on dispose d'au moins trois vagues d'enquêtes dans une étude longitudinale. Trois grandes questions sont abordées grâce à ce modèle, dont seules les deux premières nous intéressent :

- quel est le sens de la relation entre le bien-être général et les satisfactions spécifiques ? (i.e. confrontation des approches *top-down* et *bottom-up* dans la perspective des théories définitoires) ;
- quel est le sens de la relation entre le bien-être et les standards de référence dérivés de la théorie du niveau d'adaptation et du modèle des discordances multiples de Michalos ? (i.e. confrontation des approches *top-down* et *bottom-up* dans la perspective des théories nomologiques).

Les données utilisées proviennent des quatre premières vagues du panel australien sur la qualité de la vie. Les enquêtes ont eu lieu tous les deux ans à partir de 1981. Le questionnaire inclut notamment des items de satisfaction générale, de satisfactions dans différents domaines de vie (mariage, travail, niveau de vie, loisirs, amis, santé), d'attentes, d'aspirations ainsi que l'EPI d'Eysenck (qui mesure l'extraversion et le névrosisme) et des questions socio-démographiques classiques.

Ce qui intéresse les auteurs, c'est de trouver une méthode permettant d'étudier le sens de la relation entre la satisfaction générale (SG) et la satisfaction dans les domaines (SD). Mais cette question n'est pas une question simple parce que pour pouvoir montrer que SG influence SD (ou le contraire), il faudrait établir que des changements sur SG ont précédé des changements sur SD. Les spécialistes de ces problèmes considèrent habituellement que si la relation entre SG_1 et SD_2 (respectivement satisfaction générale au temps 1 et satisfaction dans un domaine au temps 2) est supérieure à la relation entre SD_1 et SG_2 , c'est que la direction de la causalité est probablement de SG vers SD.

D'après la revue de travaux sur cette question de Kessler & Greenberg (1981, cités par Headey et al., 1989), il est possible d'estimer de manière fiable, avec peu de postulats, ces coefficients, si on dispose de trois vagues de panel successives, avec le modèle représenté dans la figure 2.31.

Ainsi qu'on peut le constater en regardant cette figure, le nombre de paramètres à estimer est très important. Sans contraintes supplémentaires, ce modèle n'est pas estimable : il n'est pas identifié. Cependant, Kessler & Greenberg (1981) ont démontré qu'il le devenait si deux contraintes d'égalité étaient imposées. Ces contraintes sont représentées par des flèches affectées du même symbole dans la figure 2.31. Sur le plan théorique, ces contraintes semblent justifiées : elles indiquent, pour la satisfaction générale, que la relation entre SG_1 et SG_2 est égale à celle entre SG_2 et SG_3 .

Malheureusement, il y a une difficulté supplémentaire due à la présence, dans le modèle, à la fois de liens croisés ("*cross-lagged links*" comme, par exemple, $SG_1 \rightarrow SD_2$) et de liens non-croisés entre les mêmes variables (e.g. $SG_2 \rightarrow SD_2$). Comme il est fort probable que les estimations de ces deux relations soient fortement corrélées, les risques

de multicollinéarité deviennent importants, ce qui conduit, une nouvelle fois, à une sous-identification du modèle.

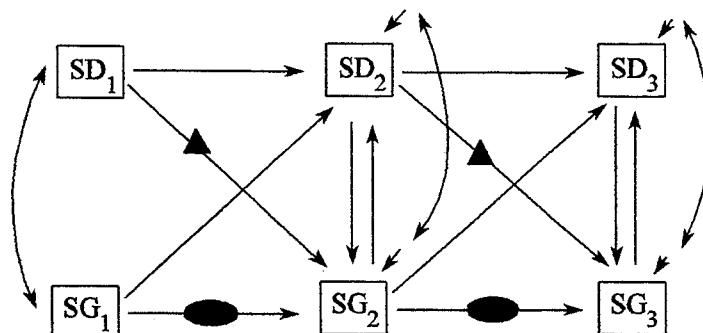


Figure 2.31 : Le modèle de Kessler & Greenberg (1981) pour évaluer le sens des relations causales à partir de données de trois vagues de panel (d'après Headey et al., 1989) (une contrainte d'égalité est imposée aux flèches affectées du même symbole, cf. texte).

Les auteurs proposent donc d'estimer *séparément* les modèles avec les corrélations croisées et les modèles avec les corrélations non-croisées et de comparer les estimations obtenues dans les deux cas. Les résultats des modèles avec des liens non-croisés sont toutefois préférés au niveau théorique, car on peut s'attendre à ce que, si la satisfaction générale influence la satisfaction dans un domaine (ou l'inverse), les effets soient relativement rapides plutôt que décalés de deux ans (période séparant deux enquêtes successives).

Lorsque deux variables sont corrélées, il n'est pas certain que l'une soit la cause de l'autre. La relation peut être artificielle ("*spurious*"), c'est-à-dire due à l'influence d'une autre variable. Il n'est pas exclu que dans le cas des relations entre satisfaction générale et satisfaction dans les domaines, de tels effets existent. Pour évaluer la possibilité de tels effets, les auteurs incluent deux autres spécifications dans les modèles testés. D'une part, cinq variables sont introduites en tant que variables exogènes (âge, sexe [1 = homme ; 2 = femme], statut socio-économique, névrosisme et extraversion). Comme ces variables sont supposées stables, seuls leurs effets sur les variables de satisfaction au temps 1 sont pris en compte. Mais, bien évidemment, il est tout à fait possible que, si des relations artificielles existent, elles soient dues à d'autres variables que celles qui sont incluses dans le modèle. Il existe une possibilité pour tester les effets de variables qui ne sont pas présentes dans le modèle : il suffit d'estimer les corrélations entre les termes d'erreurs pour les équations correspondant aux indicateurs de satisfaction.

En effet, si les termes d'erreurs sont corrélés, ils indiquent qu'une variable supplémentaire devrait être introduite dans le modèle afin de rendre compte de la part de variance commune. On ne sait pas de quelle variable il s'agit, mais on sait qu'elle existe.

Le modèle estimé par les auteurs est illustré dans la figure 2.32, en prenant l'exemple de la relation entre la satisfaction générale et la satisfaction avec le mariage. Ce modèle ne teste que les effets non-croisés. Les résultats des modèles avec effets croisés sont rapportés plus loin. Le modèle a été testé avec LISREL IV à partir de matrices de corrélations ; les coefficients sont des coefficients standardisés^[45].

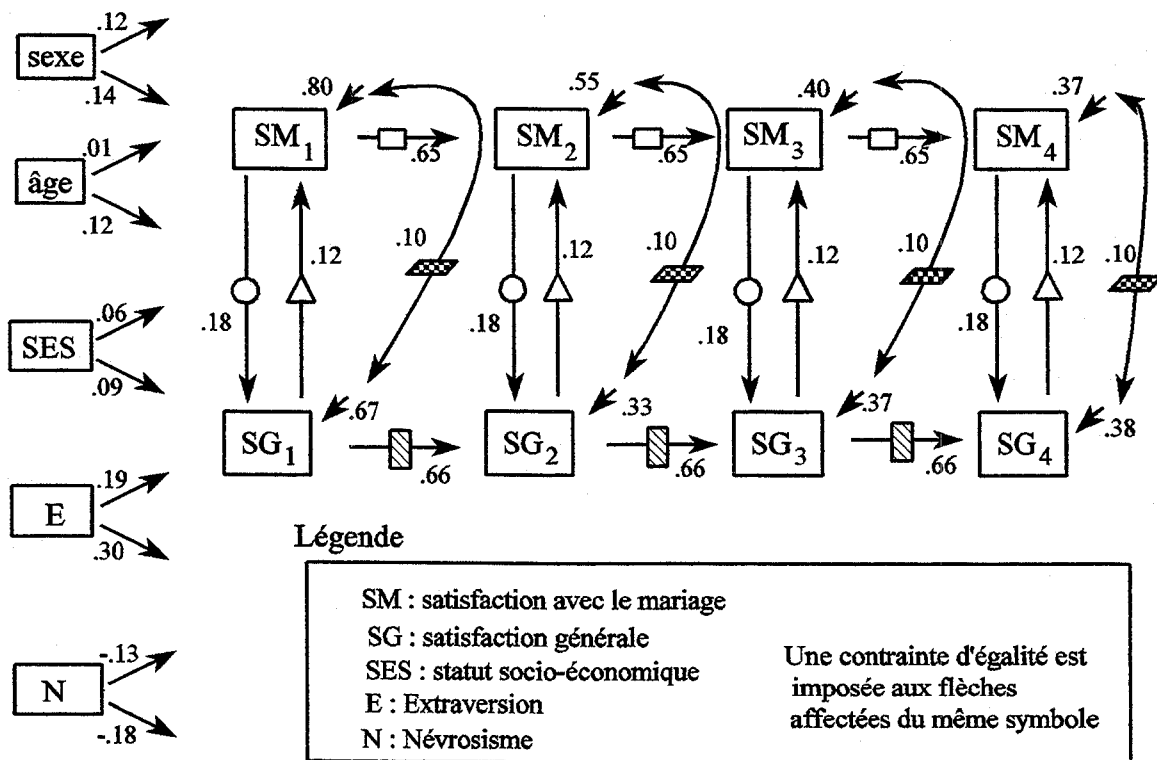


Figure 2.32 : Les relations entre la satisfaction avec le mariage et la satisfaction générale dans le modèle avec effets non-croisés d'après Headey et al. (1989) (pour les variables exogènes, la flèche supérieure indique l'effet sur SM₁, la flèche inférieure, l'effet sur SG₁ ; les flèches courbes indiquent les corrélations entre les erreurs de mesure)

[45] : Les résultats ne sont généralement pas les mêmes si on analyse des corrélations ou des covariances lorsque l'on impose des contraintes d'égalité (cf. Bacher, 1994). Les auteurs indiquent toutefois que l'interprétation des résultats est identique, que l'on utilise des corrélations ou des covariances. Il n'est fait état, dans cette présentation, que du modèle structural. Comme les variables de satisfactions sont mesurées avec plusieurs indicateurs, un modèle de mesure a aussi été estimé permettant de tenir compte des erreurs de mesure auto-corrélées.

Les résultats^[46] obtenus pour la satisfaction avec le mariage indiquent clairement la présence d'une relation bi-directionnelle : la relation *bottom-up* est égale à .18, la relation *top-down* à .12. Ces deux coefficients sont significatifs à un risque de 5%. Les résultats pour l'ensemble des domaines sont résumés dans le tableau 2.17 qui fournit à la fois les estimations des modèles avec effets croisés et celles avec effets non-croisés.

Domaine	Lien	coefficients non-croisés	coefficients croisés
mariage	<i>bottom-up</i>	.18	.08
	<i>top-down</i>	.12	.07
travail	<i>bottom-up</i>	-.07 (ns)	-.03 (ns)
	<i>top-down</i>	.24	.19
niveau de vie	<i>bottom-up</i>	-.04 (ns)	-.02 (ns)
	<i>top-down</i>	.18	.12
loisirs	<i>bottom-up</i>	-.06 (ns)	-.05 (ns)
	<i>top-down</i>	.17	.12
amis	<i>bottom-up</i>	-.01 (ns)	-.01 (ns)
	<i>top-down</i>	.06 (ns)	.04 (ns)
santé	<i>bottom-up</i>	.02 (ns)	.00 (ns)
	<i>top-down</i>	.05 (ns)	.05 (ns)

Tableau 2.17 : Estimation des liens *top-down* et *bottom-up* entre la satisfaction dans les domaines et la satisfaction générale d'après Headey et al. (1989)

On remarquera tout d'abord que les estimations obtenues pour les effets croisés et les effets non-croisés, si elles sont différentes en valeur absolue, vont exactement dans le même sens. Le fait qu'elles soient inférieures dans le cas de liens croisés est tout à fait normal étant donné que dans un cas, les mesures sont obtenues au même moment, alors que dans l'autre elles sont séparées par un intervalle de deux ans. Ainsi, pour le mariage, les résultats indiquent qu'il y a un coefficient de piste de .08 entre la satisfaction avec le mariage au temps t et la satisfaction générale au temps $t + 1$ (effet *bottom-up*) ; ce coefficient est égal à .07 entre la satisfaction générale au temps t et la satisfaction avec le mariage au temps $t + 1$ (effet *top-down*) (ces coefficients sont à comparer à ceux qui sont rapportés dans la figure 2.32).

[46] : Les indicateurs d'adéquation sont satisfaisants pour l'ensemble des modèles. Par exemple, pour le modèle représenté dans la figure 2.32, le χ^2 divisé par le nombre de degrés de liberté est égal à 1.6 ($\chi^2 = 395.94$; ddl = 248 ; les auteurs n'indiquent pas le seuil de probabilité correspondant).

Ce qui est remarquable dans ces résultats, c'est qu'ils indiquent de manière très nette que, pour trois domaines de vie (le travail, le niveau de vie et les loisirs), la relation entre satisfaction générale et satisfaction dans le domaine doit être envisagée avec une approche *top-down* et non avec une approche *bottom-up*. Dans les trois cas, les coefficients *top-down* sont significatifs alors que les coefficients *bottom-up* ne le sont pas. Il reste toutefois qu'aucun de ces effets n'est très important et que le meilleur prédicteur du niveau de satisfaction à un moment donné est le niveau de satisfaction précédent.

En revanche, pour les domaines de la santé et des amis, aucun des effets n'est significatif. Cela ne signifie nullement que la satisfaction avec ces domaines n'est pas liée à la satisfaction générale, bien au contraire. On se trouve dans le cas de relations artefactuelles. Dans le cas des amis, la relation s'explique par l'extraversion qui prédit à la fois les deux variables. Dans le cas de la santé, c'est aussi l'extraversion et, surtout, le névrosisme qui rend compte de la relation. Nous avons dit plus haut que la relation entre satisfaction avec les loisirs et satisfaction générale relevait d'une causalité *top-down*. Il faut en fait y ajouter l'influence d'une autre variable, non incluse dans le modèle, car la corrélation entre les erreurs est assez élevée (.24).

Nous ne détaillerons pas plus avant le test des modèles du niveau d'adaptation et des discordances multiples qui se basent sur une méthodologie identique. Ces résultats indiquent que :

- contrairement à l'hypothèse de la théorie du niveau d'adaptation, les événements ne changent ni les aspirations ni les attentes (des événements négatifs ne les diminuent pas et des événements positifs ne les augmentent pas) ;
- les standards de référence (aspiration, attentes, niveau d'équité) ne sont pas les causes du bien-être, mais en sont ses effets ;
- les discordances entre l'évaluation de sa situation actuelle constituent, de la même manière, des conséquences du bien-être et non des causes.

Au-delà de l'indéniable rigueur méthodologique qui caractérise cette étude, nous soulignerons l'investissement théorique des auteurs qui ont su revenir sur les postulats *bottom-up* qu'ils avaient adoptés auparavant (cf. Headey et al., 1984 a, 1985 b). Encore peu connus, ces résultats incitent à changer totalement de perspective dans la construction de modèles théoriques du bien-être. Il semble évident, en effet, que les modèles exclusivement *bottom-up* doivent être abandonnés, mais la solution ne réside pas dans le recours à des modèles exclusivement *top-down*. Ce qu'il faut comprendre, c'est pourquoi telle variable a une causalité *top-down* avec le bien-être, une autre une causalité *bottom-up*, une autre encore une causalité réciproque, alors que pour d'autres, la relation avec le bien-être est entièrement expliquée par d'autres variables.

Au vu de l'étude que nous venons de présenter en détail, on ne s'étonnera pas que les auteurs qui en sont à l'origine aient cherché à proposer un nouveau modèle du bien-être. Celui-ci, ainsi que nous allons le voir, reste encore sommaire, même s'il essaie d'intégrer et d'expliquer ce que l'on sait des relations entre bien-être, conditions de vie, événements de vie et personnalité. Même si de nombreuses questions sont encore sans réponse, il adopte moins de postulats discutables que la plupart des modèles que nous avons envisagés. Nous allons donc décrire les caractéristiques principales de ce modèle avant de conclure ce chapitre.

5.5 Le modèle d'équilibre dynamique de Headey et Wearing

Pour présenter le modèle d'équilibre dynamique de Headey & Wearing (1989, 1991), le plus simple est de suivre leur raisonnement qui s'appuie sur les nombreux résultats obtenus grâce aux données longitudinales recueillies dans le cadre du panel australien de la qualité de la vie (quatre vagues entre 1981 et 1987). C'est en effet pour expliquer certains résultats surprenants que le modèle a été élaboré. Cette présentation sera aussi l'occasion de retrouver certains résultats discutés antérieurement. Globalement, ils représentent assez bien l'ensemble de la littérature, ce qui ne veut pas dire qu'il n'y ait pas certaines spécificités. Il suffit de lire le compte-rendu d'une méta-analyse pour se rendre compte que les coefficients rapportés peuvent varier beaucoup selon les études. Toutefois, les résultats obtenus dans le cadre du panel australien ont des atouts : ils proviennent d'une enquête longitudinale menée auprès d'un échantillon représentatif de la population d'un état australien ce qui, somme toute, n'est pas si courant.

La première observation de Headey & Wearing (1989, 1991) concerne la stabilité temporelle des mesures de bien-être (cf. tableau 2.18). Le bien-être est mesuré à partir de l'échelle d'affectivité positive (PA) et de l'échelle négative (NA) de l'ABS de Bradburn ainsi qu'avec une échelle destinée, selon les auteurs, à mesurer la composante cognitive du bien-être (satisfaction)^[47].

	1981-83	1981-85	1981-87	α de Cronbach
satisfaction	.64	.51	.52	.92
PA	.39	.31	.35	.64
NA	.42	.36	.40	.65

Tableau 2.18 : Stabilité temporelle et consistance interne des mesures de bien-être utilisées dans le panel australien de la qualité de la vie (Headey & Wearing, 1989, p. 733)

Deux conclusions émergent de ces résultats. Les mesures de bien-être montrent une certaine stabilité. Mais cette stabilité est inférieure à celle à laquelle on pourrait s'attendre si le bien-être était conceptualisé comme un trait car les coefficients de stabilité temporelle sont assez nettement inférieurs aux coefficients de consistance interne. Ainsi, si certaines personnes gardent un niveau de bien-être stable, il est indéniable que d'autres ont des niveaux plus fluctuants.

[47] : Cette échelle comprend les items suivants : "*How do you feel about your life-as-a-whole ?*" (posée une fois au début et une fois à la fin de l'entretien) ; "*How do you feel about ...*" : "*...the sense of purpose and meaning in your life ?*" ; "*...what are you accomplishing in life ?*" ; "*...how exciting your life is ?*" ; "*...the extent to which you are succeeding and getting ahead in life ?*".

Pour expliquer ces fluctuations, les auteurs font appel aux événements de vie. Cette hypothèse est plausible car certaines critiques de la théorie du niveau d'adaptation ont montré qu'il n'y avait pas toujours habitude aux conditions créées par des événements (positifs ou négatifs). Mais les événements de vie ont toujours été traités comme des variables exogènes dans la littérature sans que ce postulat ait été vérifié. L'utilisation de données longitudinales permet d'apporter un éclairage particulier sur cet aspect. Les auteurs proposent de regarder si les événements de vie ont tendance à se répéter au cours du temps. S'ils sont vraiment exogènes, il n'y a aucune raison pour que les individus ayant connu de nombreux événements négatifs (ou positifs) au cours d'une certaine période en connaissent aussi de nombreux au cours d'une période ultérieure.

Les auteurs ont recours à une liste de 94 événements de vie inspirée des listes classiques utilisées dans la littérature sur le stress. Les sujets y répondent en indiquant si ces événements leur sont arrivés pendant la période séparant deux enquêtes (i.e. dans l'enquête de 1985, ils indiquent les événements qui leur sont arrivés entre 1983 et 1985). Deux scores sont obtenus en comptabilisant simplement le nombre d'événements positifs et le nombre d'événements négatifs^[48]. Les corrélations entre ces scores sont rapportées dans le tableau 2.19 (en-dessous de la diagonale).

Ces résultats sont particulièrement intéressants car ils viennent remettre en cause la manière traditionnelle dont on envisage les événements de vie. Ceux-ci ne sont pas totalement exogènes, comme on pourrait le penser. En effet, les corrélations entre le nombre d'événements de même valence au cours du temps sont toutes positives et relativement élevées. L'histoire a tendance à se répéter, écrivent les auteurs dans une formule élégante pour décrire ces résultats. Un autre aspect inattendu réside dans les corrélations entre les événements de valence opposée. Ces corrélations sont positives. Bien qu'inférieures aux valeurs trouvées entre les événements de même valence, elles indiquent qu'une personne qui connaît de nombreux événements positifs a tendance à connaître aussi de nombreux événements négatifs. Sur la base de ces résultats, il est donc difficile d'envisager un effet direct simple des événements de vie sur le bien-être.

Ces résultats ne sont pourtant pas totalement convaincants. En effet, les données sont obtenues de manière rétrospective sur une période de temps relativement longue (deux ans). Il est possible que la personnalité ou l'humeur des individus au moment où ils répondent au questionnaire les conduisent à "oublier" de rapporter certains événements positifs (ou négatifs). Pour répondre à cette objection, les auteurs ne retiennent, dans la

[48] : Cette procédure requiert de déterminer, *a priori*, la valence des événements alors qu'un même événement peut être positif pour un sujet et négatif pour un autre (e.g. changer de logement). A partir de 1985, les auteurs ont demandé aux sujets d'indiquer, pour chaque événement leur étant arrivé, si celui-ci a été positif ou négatif (sur une échelle en 10 points). Comme les corrélations obtenues entre les simples sommes et les scores pondérés par l'évaluation des sujets sont élevées, les résultats présentés seront ceux qui portent sur le nombre d'événements.

liste d'événements, que ceux qu'ils considèrent comme objectifs, c'est-à-dire ceux dont on peut raisonnablement penser qu'il n'y a aucune chance que les individus les oublient, (e.g. se marier, décès d'un enfant)^[49]. Les résultats obtenus en se limitant à ces événements se trouvent au-dessus de la diagonale du tableau 2.19.

	Evénements positifs			Evénements négatifs		
	1981-83	1983-85	1985-87	1981-83	1983-85	1985-87
Positifs						
1981-83	/	.37	.32	.10	.11	.14
1983-85	.52	/	.35	.05	.10	.13
1985-87	.43	.50	/	.04	.07	.07
Négatifs						
1981-83	.15	.18	.20	/	.34	.18
1983-85	.21	.28	.32	.49	/	.33
1985-87	.24	.31	.44	.33	.49	/

Tableau 2.19 : Corrélations entre le nombre d'événements positifs et négatifs selon la période temporelle (corrélations sur tous les événements, au-dessous de la diagonale ; uniquement sur les événements "objectifs", au-dessus de la diagonale ; les coefficients supérieurs à .06 sont significatifs à $p < .05$) (d'après Headey & Wearing, 1991, p. 60)

Il est manifeste que les corrélations concernant les seuls événements objectifs sont plus faibles que celles qui sont obtenues avec l'ensemble des événements. Mais deux points sont à noter. Même plus faibles, ces corrélations permettent d'aboutir aux mêmes conclusions. De plus, les événements retenus comme "objectifs" sont très peu nombreux (6 événements positifs et 20 événements négatifs contre 94 dans la liste initiale). Ceci contribue à diminuer de manière importante la variance des scores et, par conséquent, les corrélations.

Il est assez étonnant finalement d'arriver à une conclusion identique pour le bien-être et pour les événements de vie. Dans les deux cas, il y a une stabilité temporelle substantielle qui laisse pourtant la place pour des changements. Ainsi, on peut se demander ce qui explique cette stabilité. Les auteurs font l'hypothèse qu'elle peut être expliquée par des caractéristiques individuelles stables. Celles-ci incluent à la fois des variables sociales (sexe, âge, statut socio-économique), des traits de personnalité

[49] : Des exemples d'événements plus "subjectifs" sont : "se faire de nouveaux amis" ou "avoir des problèmes avec ses enfants".

(extraversion, névrosisme, ouverture à l'expérience^[50]), et des variables de soutien social (disponibilité de liens intimes et de liens amicaux). Les résultats sont présentés dans le tableau 2.20. Il ne s'agit pas de coefficients de corrélation, mais de coefficients de régression standardisés. Les variables dépendantes sont le bien-être et les événements de vie mesurés en 1985 ; les variables indépendantes sont mesurées en 1983 (sauf l'ouverture à l'expérience, cf. note 50).

	satisfaction	PA	NA	événements positifs	événements négatifs
Sexe (1 = m ; 2 = f)	.08	.02 ^{ns}	-.02 ^{ns}	.01 ^{ns}	-.07 ^{ns}
Age	-.02 ^{ns}	-.15	-.09	-.24	-.22
SES	.04 ^{ns}	.17	-.10	.16	-.02 ^{ns}
Extraversion	.12	.14	.07 ^{ns}	.16 ^[51]	-.03 ^{ns}
Névrosisme	-.24	-.03 ^{ns}	.42	.07 ^{ns}	.16
Ouverture	.16	.12	.09	.17	.14
Liens intimes	.20	.20	-.11	.02 ^{ns}	-.09
Liens amicaux	.13	.02 ^{ns}	-.10	.01 ^{ns}	-.02 ^{ns}
Variance expliquée	20%	16%	26%	21%	13%

Tableau 2.20 : Résultats des régressions multiples sur les variables de bien-être et les événements de vie (d'après Headey & Wearing, 1991, pp. 58 et 61) (SES : statut socio-économique ; ns : non significatif à $p < .05$)

Les résultats du tableau 2.20 confirment, pour le bien-être, des conclusions déjà énoncées. L'extraversion est liée à l'affectivité positive, mais pas à l'affectivité négative. Inversement, le névrosisme permet de prédire l'affectivité négative, mais pas l'affectivité positive. En revanche, l'ouverture est liée positivement à toutes les variables de bien-être, ce qui est peut-être plus surprenant, mais, d'une part, ce trait a rarement été étudié en relation avec le bien-être, d'autre part, les items qui le mesurent sont surtout des items

[50] : Ce trait fait référence à l'un des cinq facteurs du "big-five". Il n'a été mesuré qu'à partir de la quatrième vague du panel (1987) mais, sur la base de sa stabilité, les scores obtenus à cette date sont considérés comme la meilleure approximation de ceux que les individus auraient obtenus les années précédentes.

[51] : Dans le tableau 4.5 de l'article de Headey & Wearing (1991, p.61), le coefficient est négatif. Il est fort probable que cela soit dû à une coquille typographique. En effet, dans leurs commentaires, les auteurs l'interprètent comme un coefficient positif. De plus, dans l'article de 1989, les auteurs rapportent les coefficients de corrélation entre l'extraversion mesurée en 1981 et le nombre d'événements positifs en 1983, 1985 et 1987. Ces coefficients sont tous positifs (respectivement .19, .15 et .14). La coquille a donc été corrigée dans le tableau ci-dessus.

d'"ouverture aux sentiments" ("*openness to feelings*"). Les corrélations avec les variables socio-démographiques sont un peu plus élevées que celles qui sont mises en évidence habituellement. Enfin, les variables de soutien social sont liées au bien-être de la manière attendue.

En ce qui concerne les événements de vie, on notera tout d'abord une influence importante de l'âge. Les personnes les plus âgées connaissent moins d'événements, à la fois positifs et négatifs. Ce résultat n'est pas surprenant en soi, mais il pose tout de même une question — que les auteurs n'évoquent pas — sur la manière dont les listes d'événements sont constituées. Dans l'état actuel des choses, rien ne permet de dire s'il arrive effectivement moins d'événements aux personnes âgées ou si les listes d'événements de vie sous-représentent les événements qui arrivent à ces âges de la vie. Moins attendues peut-être sont les relations entre les traits de personnalité et les événements de vie. Les données longitudinales utilisées dans cette étude permettent d'être assez confiant quant au sens de la causalité (l'influence d'une troisième variable, rendant la corrélation artificielle, étant peu probable), d'autant que des résultats identiques ont été obtenus, de manière indépendante, par Magnus et al. (1993)^[52]. Ces relations sont tout à fait parallèles à celles qui ont été observées pour le bien-être : l'extraversion prédit les événements de vie positifs mais pas les événements de vie négatifs, alors que le névrosisme prédit les événements négatifs mais pas les événements positifs. L'ouverture est corrélée positivement à la fois au nombre d'événements positifs et au nombre d'événements négatifs.

L'ensemble de ces résultats amène les auteurs à faire l'hypothèse que les événements de vie influencent le bien-être au-delà de l'influence des caractéristiques stables des individus. Il faut donc vérifier si les événements de vie permettent de rendre compte d'une certaine part de la variance du bien-être après que les effets dus aux caractéristiques stables aient été contrôlés. Pour cela, il suffit d'effectuer une analyse de régression dans laquelle le bien-être est la variable dépendante ; les caractéristiques

[52] : Cette étude permet notamment de retrouver le fait que le nombre d'événements de vie positifs est corrélé positivement au nombre d'événements de vie négatifs. Mais il y a un résultat beaucoup plus étonnant. Alors que l'extraversion prédit l'occurrence d'événements positifs objectifs — l'objectivité des événements étant déterminée par un groupe d'experts — et que le névrosisme prédit l'occurrence d'événements négatifs objectifs, ces traits de personnalité ne permettent pas de prédire l'occurrence d'événements de vie plus subjectifs. Ce résultat, que relèvent les auteurs sans en proposer d'explication, illustre sans doute l'absence d'une ou de plusieurs variables importantes dans le modèle. On pense bien sûr au bien-être qui peut avoir une influence sur la manière de percevoir, d'interpréter et de se rappeler des événements de vie, influence qui ne serait prise en compte ni par l'extraversion ni par le névrosisme. Cette hypothèse est d'ailleurs corroborée par l'existence d'une corrélation élevée entre les erreurs de mesure associées aux événements positifs et aux événements négatifs ; des erreurs de mesure corrélées entre deux variables traduisent le plus souvent l'absence d'une variable dans le modèle. Sur un plan théorique, les auteurs notent que les résultats montrant l'influence de la personnalité sur les événements de vie ne doivent pas surprendre, car ils sont tout à fait cohérents avec une conception interactionniste. En effet, si la personnalité influence le choix des situations il n'est pas étonnant qu'elle puisse prédire les événements de vie puisque certains événements ont une probabilité plus grande de se produire dans certaines situations que dans d'autres.

stables des individus sont entrés dans l'équation comme un premier bloc et les événements de vie sont entrés ensuite. Les résultats globaux de ces analyses, résumés dans le tableau 2.21, indiquent effectivement que les événements de vie permettent de prédire le bien-être *au-delà de l'influence des caractéristiques stables*. Mais l'originalité du modèle présenté ne réside pas dans l'influence des événements de vie.

% de variance expliquée par	satisfaction ₁₉₈₅	PA ₁₉₈₅	NA ₁₉₈₅
* les caractéristiques stables (1983)	20%	16%	26%
* les événements de vie (1983-85)	8%	7%	4%
Total	28%	23%	30%

Tableau 2.21 : Pourcentages de variance expliquée par les caractéristiques stables des individus et par les événements de vie (d'après Headey et al., 1991, p. 63)

Le modèle de Headey & Wearing (1989, 1991) est original parce qu'il s'agit d'un modèle d'équilibre dynamique. Cela signifie que chaque personne est caractérisée par un pattern normal d'événements de vie, influencé par des caractéristiques individuelles stables. Ainsi, ce ne sont pas les événements de vie en eux-mêmes qui ont un effet sur le bien-être. Un événement n'a un effet sur le bien-être qu'à la condition qu'il s'écarte de manière significative du niveau normal d'événements de vie que connaît une personne pendant un intervalle de temps donné. Autrement dit, seule une déviation par rapport au niveau normal des événements fait que les événements de vie sont associés au bien-être.

Pour tester ce modèle, les auteurs utilisent une méthode astucieuse et simple qui s'appuie sur une analyse de régression multiple. L'idée est de montrer que le bien-être à un moment donné n'est pas influencé par les événements de vie, sauf si ceux-ci s'écartent de leur niveau normal. La meilleure estimation du niveau normal d'événements, du niveau d'équilibre, est la moyenne des événements survenus pendant plusieurs années consécutives. De la même manière, la meilleure estimation du niveau normal de bien-être d'une personne, de son niveau d'équilibre, est la moyenne des scores de bien-être pour les années 1981, 1983 et 1985. L'équation servant à tester le modèle d'équilibre dynamique est donc la suivante :

$$\text{Bien-être}_{1987} = \beta_1 \text{Bien-être}_{1981-85} + \beta_2 \text{EV}_{1981-85} + \beta_3 \text{EV}_{1985-87} + \epsilon$$

La variable dépendante est le bien-être en 1987. La première variable indépendante est le niveau normal de bien-être (i.e. moyenne des scores en 1981, 1983 et 1985). Son introduction permet de supprimer la variance du bien-être en 1987 due au niveau normal de bien-être. Ainsi, la variable dépendante à expliquer par les autres variables indépendantes devient la *dévi*ation du niveau normal de bien-être. La seconde variable

indépendante représente l'effet du niveau normal des événements sur la déviation du niveau normal de bien-être. On s'attend donc à ce que le coefficient β_2 ne soit pas significativement différent de zéro. L'effet de la troisième variable indépendante représente l'effet de la *dévi*ation du niveau normal d'événements, puisque le niveau d'équilibre est déjà pris en compte par la seconde variable indépendante. Les résultats des modèles testés sont présentés dans le tableau 2.22 (les coefficients sont des coefficients de régression standardisés).

	satisfaction	PA	NA
niveau normal de bien-être	.69	.41	.70
niveau normal des événements	-.05 ^{ns}	-.06 ^{ns}	.04 ^{ns}
dévi	.18	.29	.16

Tableau 2.22 : Résultats des régressions testant le modèle d'équilibre dynamique (cf. texte) (d'après Headey & Wearing, 1991, p. 65)

Les hypothèses dérivant du modèle d'équilibre dynamique sont vérifiées. Les coefficients correspondant au niveau normal de bien-être sont bien évidemment triviaux, mais ils sont indispensables pour tester le modèle. Les coefficients de régression correspondant au niveau normal d'événements ne sont pas significatifs ; en revanche, ceux qui correspondent à la déviation par rapport au niveau normal d'événements le sont. Ainsi, la déviation du niveau normal de bien-être est influencée par la déviation du niveau normal d'événements de vie.

En conclusion, le modèle d'équilibre dynamique de Headey & Wearing est sans doute celui qui rend le mieux compte de la stabilité et des changements constatés empiriquement dans les mesures de bien-être. Mais ce modèle est loin d'être complet, ainsi que le reconnaissent les auteurs eux-mêmes. Il reste notamment à mieux comprendre les mécanismes qui relient événements de vie et bien-être. De plus, ce modèle reste exclusivement *bottom-up*. Il n'intègre nullement la possibilité que le bien-être puisse avoir lui-même un effet sur les événements de vie, au-delà de l'effet de la personnalité et des autres caractéristiques stables de l'individu. Que cette possibilité ne soit pas évoquée par les auteurs ayant mis eux-mêmes en évidence la faiblesse des modèles *bottom-up* constitue notre principal regret vis-à-vis de ce travail.

5.6 Conclusion

Arrivé au terme de ce chapitre, on peut reprendre l'interrogation de Dodge (1930) évoquée au début : les théories sur le bien-être ont-elles réellement progressé depuis la philosophie grecque ? Nous avons laissé entendre une réponse positive en arguant du nombre de modèles théoriques proposés. Mais qu'en est-il sur un plan plus qualitatif ?

Il est indéniable que les choses ont beaucoup évolué depuis la question de Dodge, même si on peut reconnaître que tel ou tel modèle s'origine dans des conceptions philosophiques anciennes. Les théories sont devenues beaucoup plus précises et, surtout, elles donnent lieu à des vérifications empiriques.

Il n'est pas certain que l'on dispose actuellement d'un modèle du bien-être satisfaisant. Mais plutôt qu'en ces termes, peut-être faut-il envisager les progrès de façon différente. On ne dispose pas, certes, d'une théorie globale et cohérente. En revanche, on commence à savoir que certains modèles ne sont pas ou que peu plausibles, ce qui est loin d'être négligeable.

C'est le cas notamment des modèles basés sur les conditions de vie. Si celles-ci ont une influence sur le bien-être, il semble qu'elle soit faible et limitée à des circonstances exceptionnelles de déprivation. C'est le cas aussi de modèles qui s'appuient sur des différences calculées entre des évaluations spécifiques faites par les individus de leur situation, car ces modèles, tels qu'ils sont habituellement formulés, ne sont pas réfutables. C'est le cas enfin des modèles qui utilisent, comme variables intermédiaires, des évaluations subjectives directes qui semblent plutôt être les effets du bien-être que ses causes.

Pour le reste, il n'y a aucune certitude, notamment parce que les modèles proposés ne sont pas forcément incompatibles. Ainsi, le modèle d'équilibre dynamique de Headey & Wearing implique forcément que les individus confrontent les événements de vie vécus au niveau d'équilibre. La manière dont s'effectue cette confrontation n'est pas précisée dans le modèle, mais elle peut très bien suivre les principes énoncés, par exemple, par la théorie étendue-fréquence. On peut aussi penser que certains modèles sont complémentaires parce qu'ils sont plus appropriés pour certaines composantes du bien-être que d'autres. Ainsi, certaines caractéristiques des modèles relativistes semblent cohérents avec des aspects cognitifs du bien-être parce qu'ils postulent des processus de jugement. En revanche, les modèles auto-téléologiques qui situent la source du bien-être dans l'activité qui accompagne la poursuite d'un objectif sont peut-être plus à même de rendre compte des aspects plus affectifs du bien-être.

Il y a un point qui nous paraît fondamental, c'est l'opposition entre approches *bottom-up* et *top-down*. Les résultats que nous avons présentés ne laissent guère de doute quant à la faible vraisemblance des modèles exclusivement *bottom-up*. Cependant, de nombreuses questions restent en suspens.

Du point de vue des théories définitoires, les relations entre satisfaction générale et satisfactions plus spécifiques commencent à s'éclairer. Nous avons relaté plusieurs travaux qui semblent montrer que le sens de la relation dépend des domaines de vie. Mais ces travaux sont limités à ce qu'il est convenu de considérer comme les aspects cognitifs du bien-être. Qu'en est-il des aspects plus affectifs ? Si l'affectivité positive et l'affectivité négative sont toutes deux des indicateurs du bien-être, s'ils sont indépendants, il faut bien envisager, dans ce cas, une causalité *bottom-up*. En effet, l'indépendance entre deux indicateurs d'un même concept ne peut se comprendre que s'ils sont assimilés à des indicateurs-causes, ainsi que l'a rappelé Bollen (1984). On ne sait rien de cette causalité d'autant que l'on peut se demander si, selon les aspects affectifs envisagés, il n'y aurait pas à la fois des effets *top-down* et des effets *bottom-up*, comme c'est le cas pour la satisfaction.

Du point de vue des théories nomologiques, il n'est pas concevable de rejeter l'hypothèse d'une causalité *bottom-up* des événements de vie. Mais si les événements de vie ne sont pas totalement exogènes, étant influencés notamment par les traits de personnalité, ne pourraient-ils pas aussi être influencés, d'une manière ou d'une autre, par le bien-être ? Pour que cette hypothèse soit envisageable, il faut supposer que le bien-être a une influence sur les événements de vie au-delà de celle des traits de personnalité habituellement pris en compte. Ceci n'est pas exclu. Dans des conceptions comme celle de Stones & Kozma (1986 a), le bien-être est conçu comme une propension plus large qu'un trait ; dans celle de Tellegen (1985, cité par Watson et al., 1988) l'émotionnalité positive et négative sont deux facteurs primaires de la personnalité. Le bien-être pourrait favoriser la création de conditions de vie spécifiques ayant une influence en retour sur le bien-être, à la manière de l'hypothèse instrumentale de McCrae & Costa (1991, cités par Larsen & Ketelaar, 1991) sur l'influence des traits de personnalité. Pour se convaincre de la vraisemblance d'une telle hypothèse, il suffit de considérer certains événements de vie tels que "se faire de nouveaux amis" ou "avoir des problèmes avec son conjoint". Les travaux sur l'influence de l'humeur ont montré l'effet bénéfique d'une humeur positive sur les relations inter-personnelles qui sont justement prises en compte dans ces deux exemples. L'examen plus attentif du concept d'événements de vie en lui-même semble décisif pour mieux appréhender cette question, car le fait de se limiter à une simple somme des événements vécus est sans doute une démarche beaucoup trop grossière.

A condition d'accepter d'abandonner une opposition trop souvent idéologique entre approches *top-down* et *bottom-up*, on peut envisager cette différenciation d'une manière

qui suggère de toutes nouvelles pistes de recherche. Wood et al. (1989) opèrent une distinction entre perspectives holistique et particulariste, qui se superpose partiellement à celle entre démarches *top-down* et *bottom-up*. L'originalité de ces auteurs est de l'appréhender comme une variable de différences interindividuelles. La perspective holistique correspond à une approche générale de la vie et elle est compatible avec certains résultats qui montrent que les corrélations entre la satisfaction générale et les satisfactions spécifiques ne diffèrent pas selon les domaines. Dans la perspective plus particulariste, l'approche est plus analytique : les relations entre satisfaction générale et satisfactions spécifiques diffèrent selon les domaines et selon les sujets (e.g. la relation entre satisfaction générale et satisfaction avec le mariage est plus élevée pour les hommes que pour les femmes). Si certains individus adoptent une perspective plus particulariste alors que d'autres adoptent une perspective plus holistique, il n'est donc pas étonnant que certains résultats supportent une approche *top-down*, d'autres une approche *bottom-up*. Les différences dans le sens de la relation selon les domaines pourrait trouver un nouvel éclairage avec cette proposition.

Le bilan est donc mitigé, mais les avancées de ces dernières années sont prometteuses. C'est en adoptant de nouvelles perspectives, sans abandonner totalement les anciennes, que les recherches sur le bien-être peuvent progresser. Au-delà des querelles parfois vaines, il y a une place pour les travaux qui appréhendent le bien-être comme une large disposition psychologique plutôt que comme un état. C'est dans ce domaine quelque peu négligé jusqu'à présent que des efforts doivent être consentis.

