
Université des sciences et technologies de Lille (Lille I)
Faculté des sciences économiques et sociales



Parcours, publications, travaux

Habilitation à diriger les recherches par Renaud Fillieule
Maître de conférences en sociologie, Lille I, CLERSÉ

VOLUME 2 : ANNEXES

Garant de l'habilitation : M. le Professeur Claude Giraud, Lille I
M. le Professeur Raymond Boudon, Paris IV
M. le Professeur Pierre Demeulenaere, Paris IV
M. le Professeur Yves Grafmeyer, Lyon II
M. le Professeur Guido Hülsmann, Angers
M. le Professeur Philippe Steiner, Lille III

Parcours, publications, travaux

Sommaire

Première partie : parcours	5
Deuxième partie : la rationalité.....	16
Troisième partie : histoire et méthodes de la sociologie	82
Quatrième partie : la sociologie de la délinquance	107
Cinquième partie : l'école autrichienne d'économie	115

Table des matières

Première partie : parcours

Débuts	6
La question de la rationalité dans les sciences sociales	9
Histoire et méthodes des sciences sociales	10
La sociologie de la délinquance.....	11
L'école autrichienne d'économie	12
Projets	14

Deuxième partie : la rationalité

1. « Frames, inferences and rationality : some light on the controversies about rationality » (1996) <i>Rationality and Society</i> , vol. 8	17
2. « Les pièges de la rationalité : deux études de cas en sociologie de la connaissance scientifique » (1997) <i>L'Année sociologique</i> , vol. 47	33
3. « Un modèle d'analyse des présupposés en sociologie de la connaissance scientifique » (1998) <i>Revue européenne des sciences sociales</i> , vol. 36.....	59
4. « Types de rationalité » (2005) in <i>Dictionnaire de la pensée sociologique</i> , dirigé par M. Borlandi, R. Boudon, M. Cherkaoui et B. Valade, Paris, PUF	77

Troisième partie : histoire et méthodes de la sociologie

1. « Croissance, crise et renouveau » (1996) in <i>Introduction aux sciences sociales</i> , dirigé par B. Valade, Paris, PUF.....	83
2. « System : Social » (2001) <i>International encyclopedia of the social and behavioral sciences</i> , Oxford, Elsevier	87
3. « Neutrality : Axiological » (2001) <i>International encyclopedia of the social and behavioral sciences</i> , Oxford, Elsevier.....	95

4. Table des matières, <i>Les méthodes en sociologie</i> (2002) dirigé par R. Boudon, Paris, PUF	100
5. « Individualisme méthodologique et holisme » (2005) <i>Dictionnaire de la pensée sociologique</i> , dirigé par M. Borlandi, R. Boudon, M. Cherkaoui et B. Valade, Paris, PUF	102

Quatrième partie : la sociologie de la délinquance

1. Table des matières, <i>Sociologie de la délinquance</i> (2001) Paris, PUF	108
--	-----

Cinquième partie : l'école autrichienne d'économie

1. Table des matières, <i>Introduction aux théories de l'école autrichienne</i> (2000) document pédagogique, Université de Lille 1	116
2. « The "values-riches" model : an alternative to Garrison's model in Austrian macroeconomics of growth and cycle » (2005) <i>The Quarterly Journal of Austrian Economics</i> , vol. 8	121
3. « A formal model of the Hayekian structure of production » (2005) article non publié	139
4. « A graphical illustration of Böhm-Bawerk's productivity theory of interest » (2005) article non publié	160
5. « La théorie du marché des fonds prêtables et le paradigme autrichien » (2005) article non publié	171

Première partie : parcours

DÉBUTS

Mon parcours, ou ma (très modeste) histoire intellectuelle, commence par un jour de l'automne 1984, avec la découverte, au hasard des rayonnages de la bibliothèque municipale, du fameux ouvrage épistémologique de Karl Popper *La logique de la découverte scientifique*. « Voici enfin que nous parvient, traduit en français, ce grand et puissant livre... » Ces mots qui ouvrent la préface de Jacques Monod m'intriguèrent et me donnèrent envie d'emprunter un ouvrage dont l'auteur m'était à l'époque totalement inconnu, et dont la notoriété en France était encore assez limitée. J'ai tout de suite été impressionné par ces pages lumineuses dans lesquelles Popper présente les deux grands problèmes épistémologiques, le problème de l'induction (« problème de Hume ») et le problème de la démarcation entre science et métaphysique (« problème de Kant »), et propose des solutions claires et convaincantes pour les résoudre. J'étais alors étudiant en école d'ingénieur, mais ne m'étais jamais posé la question de savoir comment les théories scientifiques que l'on m'enseignait depuis des années étaient reliées aux faits empiriques. Et c'est à cette question que Karl Popper répondait de façon originale et provocatrice : « La science ne repose pas sur une base rocheuse. La structure audacieuse de ses théories s'édifie en quelque sorte sur un marécage. Elle est comme une construction bâtie sur pilotis » (*La logique de la découverte scientifique*, p. 111). La lecture de cet ouvrage, puis dès sa traduction en français celle de *Conjectures et réfutations*, m'ont donné le goût des réflexions abstraites, théoriques et épistémologiques, dont on trouvera des illustrations dans ces deux volumes d'habilitation. Ces textes de Karl Popper sont aussi pour moi des modèles de clarté et d'argumentation critique, beaucoup de travaux philosophiques me paraissant en comparaison verbeux et obscurs. Une fois mes études d'ingénieur terminées, je décidai d'entamer des études de philosophie en 3^e année de licence à Paris IV, en vue de me former à l'épistémologie.

La deuxième rencontre marquante a été celle avec Raymond Boudon. Cette rencontre, purement livresque d'abord, puis personnelle en suivant ses enseignements à Paris IV, m'a conduit à me diriger vers la sociologie, en DEA puis en thèse de doctorat. De nombreux aspects de ses travaux m'ont marqué, mais je souhaiterais surtout évoquer ici la façon dont il analyse les textes sociologiques, en procédant à ce que j'appellerais un « décorticage » des théories et des illustrations développées par les grands auteurs des sciences sociales, pour en tirer les leçons méthodologiques

et épistémologiques. Sa défense de l'individualisme méthodologique dans *La logique du social* et *La place du désordre* s'appuie sur des théories d'auteurs divers ayant travaillé à des époques différentes sur des thèmes variés, et ne consiste pas du tout à concevoir dans l'abstrait un cadre prédéfini dans lequel il faudrait faire entrer de force toutes les théories passées et futures des sciences sociales. Sa démarche, qui fait émerger les méthodes et les épistémologies de l'analyse approfondie de travaux importants – classiques ou contemporains – en sciences sociales, m'a toujours paru très séduisante. Un second mouvement de la pensée est ensuite possible, qui consiste à enrichir les théories à partir de ces considérations de méthode pour en examiner la structure logique, par exemple en distinguant bien et en détaillant les deux moments de la compréhension et de l'agrégation. *A contrario*, l'exemple de ce qu'il ne faut pas faire me semble se trouver chez Mark Blaug dans sa *Méthodologie économique* : bien que son livre soit intéressant et que l'on y trouve matière à réfléchir, sa tentative de plaquer la méthodologie popperienne sur les raisonnements de l'économie me paraît très artificielle et vouée à l'échec (malgré toute l'admiration que j'éprouve pour l'œuvre de Popper, je ne crois pas que ses considérations épistémologiques puissent s'appliquer telles quelles aux sciences sociales). Raymond Boudon montre aussi qu'il faut distinguer, chez les auteurs, les méthodologies et épistémologies « officielles » d'une part, et d'autres part celles qui sont effectivement utilisées dans la conception de leurs théories scientifiques. Par exemple, dans son analyse de la théorie de la magie développée par Durkheim dans *Les formes élémentaires de la vie religieuse*, il met bien en évidence le fait que ce dernier applique dans certains passages de son œuvre des principes « individualistes » de type weberien, qui contredisent son épistémologie positiviste officielle : « Durkheim refuse la "psychologie" dans ses textes doctrinaires, mais elle est omniprésente dans ses analyses » (*Études sur les sociologues classiques*, p. 119). Je me suis inspiré de ces leçons au chapitre 1 du document principal de l'habilitation, pour étudier les textes de sociologie économique des prix.

Le troisième et dernier auteur que je souhaiterais citer ici est Friedrich Hayek, que j'ai beaucoup « utilisé » dans le premier volume de l'habilitation, tout simplement parce que le sujet s'y prêtait. Je l'ai découvert en lisant son traité *Droit, législation et liberté*, un ouvrage fascinant à bien des égards. La distinction qu'il y opère entre les deux types d'ordres – ordre spontané et ordre fabriqué – et les moyens spécifiques de réguler chacun d'eux – respectivement le droit (règles de juste conduite) et la lé-

gislation – ouvre des perspectives qui me paraissent indispensables à toute réflexion approfondie sur l'ordre social et politique. Ses considérations sur les processus d'évolution culturelle et de sélection des règles de conduite, sur les limites de la raison et du constructivisme sont elles aussi d'une grande portée et brassent avec virtuosité les thèmes de l'analyse scientifique de la « Grande Société » et de la philosophie morale et politique. Cette lecture m'a décidé à entreprendre des études d'économie en parallèle avec celles déjà entamées en sociologie, et je dois reconnaître avoir été quelque peu surpris et déçu de constater qu'à l'issue de trois années de cours à Paris I, de la licence au DEA, les travaux de Hayek n'avaient été abordés en tout et pour tout que pendant une trentaine de minutes. La science économique contemporaine, qui me paraît se perdre quelque peu dans sa course à la formalisation mathématique, n'a pas intégré les importantes leçons de ce grand auteur, et plus généralement de l'école autrichienne. Ce thème qui me tient à cœur est amplement développé au chapitre 2 du volume 1 de l'habilitation.

Références

- Blaug, Mark (1982) *La méthodologie économique*, Paris, ECONOMICA.
- Boudon, Raymond (1979) *La logique du social. Introduction à l'analyse sociologique*, Paris, Hachette.
- Boudon, Raymond (1984) *La place du désordre. Critique des théories du changement social*, Paris, PUF.
- Boudon, Raymond (1998) *Études sur les sociologues classiques*, Paris, PUF.
- Durkheim, Émile (1985[1912]) *Les formes élémentaires de la vie religieuse. Le système totémique en Australie*, Paris, PUF.
- Hayek, Friedrich A. (1980[1973]) *Droit, législation et liberté*, vol. 1 : *Règles et ordres*, Paris, PUF.
- Popper, Karl R. (1973[1934]) *La logique de la découverte scientifique*, Paris, Payot.
- Popper, Karl R. (1985[1963]) *Conjectures et réfutations. La croissance du savoir scientifique*, Paris, Payot.

LA QUESTION DE LA RATIONALITÉ DANS LES SCIENCES SOCIALES

Notre travail de thèse a porté sur *La question de la rationalité dans les sciences sociales : histoire et bilan* (1994). Nous y avons étudié les fractures qui apparaissent sur le thème de la rationalité entre différentes disciplines, sociologie de la connaissance, économie, psychologie cognitive, philosophie des sciences. Nous avons pu constater l'absence de dialogues ou de confrontations entre ces disciplines autour d'une notion qui est leur est pourtant centrale. Il nous a semblé important de faire ressortir une distinction méconnue entre ce que nous avons appelé la « rationalité du cadre » (rationalité du choix entre des cadres conceptuels différents) et la « rationalité de l'inférence » (rationalité du choix entre des règles de résolution de problème dans un cadre conceptuel donné). La rationalité du choix entre les paradigmes scientifiques en philosophie des sciences est un exemple typique de rationalité du cadre, et le choix de la règle d'optimisation en économie un exemple typique de rationalité de l'inférence. Un résumé de la thèse est proposé, avec les exemples qui illustrent la distinction entre ces deux formes de rationalité, dans un article publié dans la revue *Rationality and Society* en 1996. Dans l'article suivant, publié dans *L'Année sociologique* en 1997, nous avons cherché à mettre en évidence les confusions qui pouvaient survenir lorsque cette distinction entre les rationalités du cadre et de l'inférence était méconnue et que des problèmes relevant de la rationalité du cadre étaient traités comme relevant de la rationalité de l'inférence. Nous avons aussi commencé un travail de sociologie de la connaissance en tentant d'expliquer pourquoi certains auteurs pouvaient commettre la confusion entre ces deux formes de rationalité. Nous avons poursuivi dans cette veine de la sociologie de la connaissance dans un article publié en 1998 dans la *Revue européenne des sciences sociales*, qui vise à expliquer l'adhésion à des croyances scientifiques fausses en rendant compte de l'adhésion à des présupposés explicites ou implicites qui conditionnent ces croyances. Cette sélection se termine par un article à vocation pédagogique sur les types de rationalité, rédigé pour le *Dictionnaire de la pensée sociologique* publié en 2005, qui distingue et présente les notions de rationalité instrumentale, de rationalité des croyances positives et de rationalité des croyances normatives (rationalité axiologique).

Documents annexés

- ▶ « Frames, inferences and rationality : some light on the controversies about rationality », *Rationality and Society*, 1996, vol. 8, p. 151-165
- ▶ « Les pièges de la rationalité : deux études de cas en sociologie de la connaissance scientifique », *L'Année sociologique*, 1997, vol. 47, n° 2, p. 177-201
- ▶ « Un modèle d'analyse des présupposés en sociologie de la connaissance scientifique », *Revue européenne des sciences sociales*, 1998, vol. 36, n° 111, p. 199-215
- ▶ « Types de rationalité », *Dictionnaire de la pensée sociologique*, dirigé par M. Borlandi, R. Boudon, M. Cherkaoui et B. Valade, Paris, PUF, 2005, p. 591-594

HISTOIRE ET MÉTHODES DES SCIENCES SOCIALES

Dans cette partie sur l'histoire de la pensée sociologique et sur la méthodologie, nous avons rassemblé cinq publications écrites dans un but pédagogique et dont le fil conducteur est une préoccupation constante pour l'individualisme méthodologique. Le premier texte est une contribution au manuel d'*Introduction aux sciences sociales* (1996) de Bernard Valade, ce dernier m'ayant proposé d'en rédiger la partie sur les sciences sociales au XX^e siècle en suivant un plan en trois chapitres, le premier sur les théories de l'interaction sociale (école de Chicago, interactionnisme), le deuxième sur les grands paradigmes et leurs limites (fonctionnalisme, culturalisme, structuralisme), et le dernier sur trois questions fondamentales (rationalité et irrationalité, individualisme et holisme méthodologique, objectivité et relativisme). Les deux textes suivants sont des articles parus dans une encyclopédie en 26 volumes sur les sciences sociales et comportementales. L'article sur les systèmes sociaux (« System : social ») est orienté vers l'histoire de la pensée puisqu'il passe en revue les conceptions historicistes de Comte et de Marx, les conceptions structuro-fonctionnalistes de Radcliffe-Brown, et les conceptions actionnistes de l'économie et de la sociologie weberienne. L'article sur la neutralité axiologique (« Neutrality : axiological ») a une tonalité plus philosophique et emprunte plusieurs illustrations à l'histoire de la pensée économique. Nous avons ensuite participé à la refonte du « Que sais-je ? » de Raymond Boudon sur les méthodes en sociologie, sur la base d'une division en trois chapitres, le premier sur l'analyse causale, le deuxième sur l'individualisme méthodo-

logique, et le dernier sur la rationalité cognitive (rationalité des croyances positives et normatives). Le dernier texte de cette sélection est une entrée pour le *Dictionnaire de la pensée sociologique* (2005), et il reprend de façon synthétique les termes de la confrontation entre individualisme méthodologique et holisme.

Documents annexés

- ▶ Table des matières, « Croissance, crise et renouveau », in *Introduction aux sciences sociales*, dirigé par B. Valade, Paris, PUF, 1996, p. 445-605
- ▶ « System : Social », in *International Encyclopedia of the Social and Behavioral Sciences*, Oxford, Elsevier, 2001, p. 15 418-15 423
- ▶ « Neutrality : Axiological », in *International Encyclopedia of the Social and Behavioral Sciences*, Oxford, Elsevier, 2001, p. 10 621-10 623
- ▶ Table des matières, *Les méthodes en sociologie*, dirigé par R. Boudon, Paris, PUF, coll. « Que sais-je ? », 2002, 125 p.
- ▶ « Individualisme méthodologique et holisme », in *Dictionnaire de la pensée sociologique*, dirigé par M. Borlandi, R. Boudon, M. Cherkaoui et B. Valade, Paris, PUF, 2005, p. 351-354

SOCIOLOGIE DE LA DÉLINQUANCE

À l'issue d'une année en tant qu'attaché temporaire d'enseignement et de recherche à l'université de Lille III, j'ai pu obtenir pendant deux ans, de 1995 à 1997, un poste de chargé de recherche contractuel à l'IHESI (Institut des hautes études de la sécurité intérieure), un institut dépendant du Ministère de l'Intérieur. J'ai profité de cette occasion, et de l'accès quotidien que j'avais à son centre de documentation fort bien fourni, pour approfondir mes connaissances sur les questions de déviance, et pour préparer et donner des enseignements sur ce thème. J'ai eu, plus tard, la possibilité de rédiger un manuel de sociologie de la délinquance, publié en 2001, dans lequel j'ai essayé de rendre compte des apports anglo-saxons (et québécois) que j'avais trouvés très riches, aussi bien en matière de description des phénomènes de délinquance et de victimation que concernant les théories explicatives et les méthodes de prévention. Un point qui m'a beaucoup intéressé a été de constater que le

paradigme dominant en sociologie de la délinquance jusqu'aux années 1960, à savoir le paradigme culturaliste, a été presque complètement abandonné au moment même où il avait atteint son plus haut point d'achèvement. Ce paradigme tentait d'expliquer les phénomènes délinquants comme des phénomènes liés à la socialisation à des valeurs pro-délinquantes dans des groupes d'adolescents. À partir des années 1950, de nouvelles données issues des enquêtes de terrain, des statistiques policières et des sondages de victimation, sont venues mettre en évidence de nombreux faits typiques dont il était incapable de rendre compte, ce qui a conduit à son remplacement par le paradigme actionniste, beaucoup plus riche et beaucoup plus souple.

Document annexé

► Table des matières, *Sociologie de la délinquance*, Paris, PUF, coll. « Premier cycle », 2001, 285 p.

L'ÉCOLE AUTRICHIENNE D'ÉCONOMIE

Mon intérêt pour l'école autrichienne remonte aux débuts de mon apprentissage des sciences sociales, mais j'étais alors surtout attiré par les aspects méthodologiques et épistémologiques, ainsi que par les questions de philosophie politique. Les études universitaires d'économie, bien qu'elles m'aient beaucoup intéressé, m'ont laissé un profond sentiment d'insatisfaction. Les différents paradigmes enseignés se contredisaient sur des points fondamentaux, sans que la moindre argumentation ne vienne éclaircir ces questions basiques – la microéconomie n'évoquait jamais les coûts de transaction alors que ces derniers jouaient un rôle éminent en économie des organisations, les marchés étaient censés s'équilibrer en micro alors qu'en macroéconomie ils étaient censés subir des contraintes de rationnement qui les empêchaient de parvenir à l'équilibre (équilibres non walrasiens). Il était difficile voire impossible de retirer de tout cela des bases de raisonnement solides et cohérentes, et je me retrouvai au terme de trois années d'études à un haut niveau universitaire dans l'incapacité de tenir un discours sur l'économie, ou d'analyser les propos qui étaient tenus par les médias ou tout simplement par les gens sur la conjoncture éco-

nomique et sur les mesures gouvernementales aptes à améliorer la situation. Dès que j'en ai eu le temps, au moment de ma nomination comme maître de conférences en 1997, je me suis intéressé de très près aux théories économiques de l'école autrichienne. Je me suis peu à peu aperçu que leurs principes étaient profondément différents de ceux de l'analyse économique standard, en théorie des prix de par l'utilisation systématique du principe de la tendance à la disparition des profits/pertes entrepreneuriaux, et en macroéconomie de par le recours à la notion de structure de production. C'est donc sur ces deux points que j'ai cherché à approfondir mes connaissances puis que j'ai orienté mes recherches. En théorie des prix, j'ai rédigé des articles, non publiés, sur « L'économie de la coordination » (2000), puis sur « L'équilibre final » (2003) et sur « Les modèles de coordination par les prix d'Adam Smith à Frank Knight » (2004). Le chapitre 2 du premier volume de l'habilitation est issu lui aussi de cette préoccupation, et vise à mettre en évidence la spécificité des modèles autrichiens (ou classico-autrichiens) et ce qui les sépare des modèles standard. Pour la macroéconomie et la notion de structure de production, mon premier travail a été un article non publié qui présente de façon synthétique les conceptions macroéconomiques de Roger Garrison, un auteur important de l'école autrichienne contemporaine (Garrison, *Time and money. The macroeconomics of capital structure*, Londres, Routledge, 2001). La réflexion sur son modèle macroéconomique m'a conduit à proposer un modèle différent dans un article publié dans le *Quarterly Journal of Austrian Economics* en 2005, modèle qui met l'accent sur certains phénomènes qui n'apparaissent pas dans la présentation graphique de Garrison. J'ai aussi développé un modèle formel d'un cas particulier de structure de production hayekienne, entièrement mathématisé, qui permet de mettre en évidence des formules méconnues comme celle qui relie le taux d'intérêt et la longueur moyenne de la structure. Deux articles plus récents examinent la question de la détermination du taux d'intérêt, l'un sur la théorie productiviste de Böhm-Bawerk, et l'autre sur la théorie du marché de fonds prêtables. Cette question du taux d'intérêt est importante mais délicate puisqu'elle ne fait pas l'objet d'un consensus au sein de l'école autrichienne, Hayek ayant généralisé la conception productiviste (objectiviste) de Böhm-Bawerk, von Mises ayant repris et fondé plus rigoureusement la théorie subjectiviste de Frank Fetter, Rothbard ayant développé la théorie de l'échange de biens présents contre biens futurs issue de Böhm-Bawerk, et Garrison et d'autres macroéconomistes « autrichiens » contemporains utilisant la théorie du marché des fonds prêtables.

Documents annexés

- ▶ Table des matières, *Introduction aux théories économiques de l'école autrichienne*, document pédagogique, Université de Lille I
- ▶ « The “values-riches” model : an alternative to Garrison’s model in Austrian macroeconomics of growth and cycle », *The Quarterly Journal of Austrian Economics*, 2005, vol. 8, n° 2, p. 3-18
- ▶ « A formal model of the Hayekian structure of production » (2005) article non publié
- ▶ « Böhm-Bawerk’s productivity theory of interest » (2006), article non publié
- ▶ « La théorie du marché des fonds prêtables et le paradigme autrichien » (2006) communication au *Austrian Research Seminar*, Université de Paris II

PROJETS

Du côté de la sociologie économique, notre projet est de placer dans la perspective autrichienne, non plus seulement le thème de la sociologie des prix, mais celui, plus général, de la sociologie des marchés. Le travail de Fligstein (*The Architecture of Markets. An Economic Sociology of Twenty-First-Century Capitalist Societies*, Princeton, Princeton University Press, 2001) nous paraît constituer un point de départ intéressant pour cette enquête, pour déterminer dans quelle mesure il y a complémentarité, et des possibilités d'enrichissement mutuel, ou au contraire contradiction entre les points de vue économique et sociologique. La question de la destruction créatrice, par exemple, est envisagée par Fligstein dans une perspective assez différente de celle de Schumpeter, comme on a pu le voir brièvement dans le premier volume (chap. 2, sect. 1.3).

Du côté de l'école autrichienne, nous souhaitons continuer à approfondir les deux axes dans lesquels nous sommes déjà engagés. En théorie des prix, il nous paraît intéressant d'essayer d'énumérer les applications très diverses du principe de la tendance à l'uniformisation du taux d'intérêt originaire, puis d'examiner plus soigneusement les relations possibles – les convergences et les divergences – entre ce principe et les principes de raisonnement standard. En macroéconomie, nous souhaitons approfondir encore le thème de la théorie de l'intérêt, et aussi procéder à une

confrontation détaillée entre les analyses autrichiennes fondées sur la notion de structure de production et les analyses standard.

Deuxième partie : la rationalité

« Frames, inferences and rationality : some light on
the controversies about rationality » (1996)

Rationality and Society

Vol. 8, p. 151-165

FRAMES, INFERENCES AND RATIONALITY

SOME LIGHT ON THE CONTROVERSIES ABOUT
RATIONALITY

Renaud Fillieule

ABSTRACT

This paper adopts a distinction between two uses of the notion of rationality, the 'rationality of frame' and the 'rationality of inference'. Rationality of frame is the rationality of the choices made between conceptual frames, for instance, between paradigms in science. Rationality of inference is the rationality of choices made between inferences, i.e. between rules which, in a given conceptual frame, allow actors to reach a decision or to solve a problem; many examples can be found in microeconomics (the rule of optimization) and in cognitive psychology (models of cognition). This distinction is suitable for the study of the controversies and misunderstandings that arise concerning the notion of rationality and it throws some light on controversies about rationality as optimization in microeconomics (the neo-classical versus evolutionary paradigm), choice anomalies (neo-classical economics versus cognitive psychology), and scientific creativity (cognitive psychology versus philosophy of science).

The notion of rationality is obviously very important for social science. This notion, however, is not easy to specify because it is *polythetic* in the sense of Needham (1975):¹ it has different meanings according to the paradigms and contexts in which it is used, but these meanings retain enough in common that a single word—rationality—can take them all into account. Raymond Boudon shows in his latest book (1995, Ch. 14) that the notion of rationality fits the Wittgensteinian conception of meaning: its many uses show at the same time some irreducible differences and an undeniable *family resemblance* in the sense of Wittgenstein (1958, 67).

Boudon's theory of rationality elaborates on this feature. It distinguishes three major uses of the term, namely, *instrumental rationality* (the rational-choice model), *cognitive rationality* (explanation of positive beliefs), and *axiological rationality* (explanation of normative beliefs). There are some irreducible differences between these uses: instrumental rationality operates within a frame of given beliefs, while cognitive and

axiological rationality are used to explain beliefs; cognitive rationality explains beliefs that can be tested against reality, while axiological rationality explains value judgments that cannot. In spite of their differences, however, these uses have an incontestable 'family resemblance'. We can, for instance, give an interpretation of all three of them via a theory of *good reasons*. Instrumental rationality: people act this way rather than another because they have 'good reasons' for that (for instance, it is in their interest). Cognitive rationality: people adhere to this belief rather than that one because they have 'good reasons' for it (for instance: in their situation, they perceive phenomena that seem to corroborate this belief, and not those that contradict it). Axiological rationality: people hold this value judgment rather than that one because they have 'good reasons' for it (see the examples in Boudon 1996). This conception of rationality can be represented by a picture where the different aspects of the notion of rationality are discs that partially overlap one another:

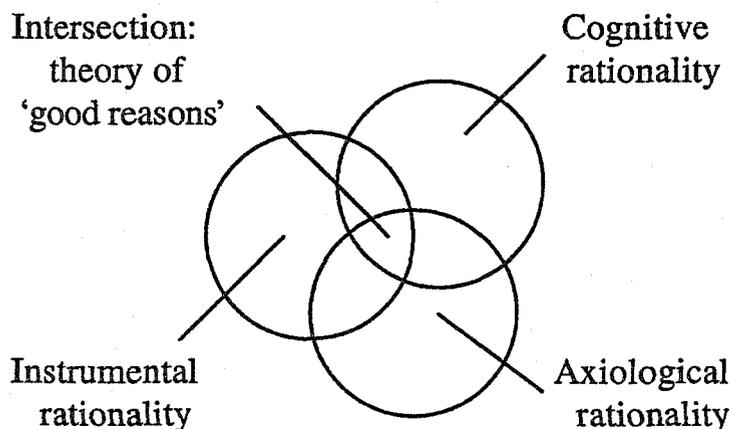


Figure 1. Polythetical character of 'rationality': first example

In this paper, I give another instance of the polythetical character of the notion of rationality, an instance based on the distinction between a *rationality of frame* and a *rationality of inference*. This distinction does not serve as a substitute for the one propounded by Boudon, nor for other possible distinctions (such as those between objective and subjective rationality, or between procedural and substantive rationality, etc.), but only aims at removing some ambiguities in the debates about the question of rationality.

Rationality of Frame and Rationality of Inference

Among the theories of rationality in the different paradigms of the human sciences, two kinds can usefully be distinguished: those that study the rationality of the choices *between conceptual frames*, and those that study the rationality of the choices *between rules of inference* (i.e. rules that are carried out within a given conceptual frame in order to solve a problem or to reach a decision; the word 'inference', here, is not restricted to logical or statistical operations). The former are theories of 'rationality of frame' and the latter theories of 'rationality of inference'. There follow some examples that specify this distinction. The first is borrowed from economics and the other two from psychology.

Rationality of Inference

Neo-classical microeconomics is the best example of a theory of rationality of inference. The theory of the consumer models economic choice as a utility maximization program. If u is the utility function and R the income of the consumer, and if we suppose that there are only two goods of prices p_1 and p_2 , the program is:

$$\begin{aligned} &\max u(c_1, c_2) \\ &\text{budget constraint : } p_1c_1 + p_2c_2 = R \end{aligned}$$

The solution to this program is the maximizing choice, consumption c^*_1 of the first good and c^*_2 of the second. Rationality, in this example, is not so much the result of the maximization as the maximizing rule itself. Other decision rules might have been adopted. For instance: choose at random a consumption bundle (c_1, c_2) meeting the budget constraint. Or: assign identical parts of income to the consumption of each good ($c_1 = R/2p_1$, $c_2 = R/2p_2$). From a mathematical point of view, there is an infinite number of rules of inference. If the goal is to reach the highest level of utility, then the best rule is obviously the rule of maximization. It is thus rational to choose the latter. Rationality is here a rationality of choice between rules of inference: it is a rationality of inference.

The inference of maximization necessarily operates within a conceptual frame. In the consumer's choice, this frame is made up of the scale of preferences (utility function u), the resources (income R), and the situation of the market (prices p_1 and p_2). In this theory of the consumer, the frame is given and not an object of explanation. The theory is thus not a theory of rationality of frame.

The same can be said of the theory of the firm. The conceptual frame

of decision-making in this case is made up of the production function, price of output and prices of inputs. When the firm is in a situation of perfect competition, this frame is given and only a rationality of inference is taken into account, through the rule of maximization.²

Rationality of Frame

We find interesting examples of rationality of frame in two experiments carried out in psychology. The first one is an experiment on perception, related by Kuhn in his famous book on scientific revolutions (1970). A subject is shown playing cards one by one, and he is asked to identify each one of them. Most are normal, but some are anomalous, for instance a red six of spades or a black four of hearts. At first, the presentations are rather short and the subjects do not perceive anything wrong, even when they are shown an anomalous card. The subjects take the red six of spades for a six of hearts or for a six of spades; but then the time of exposure to anomalous cards increases and they begin to realize that something is wrong. The subjects hesitate and even feel uneasy for a while, then suddenly they understand what is wrong: the colors are inverted. They immediately readjust their perception and no longer have any difficulty in identifying the anomalous cards.

In this experiment, two conceptual frames are successively held by the subjects. In the first, the shape of a spade is associated with the color black and the shape of a heart with the color red. When a subject is asked to identify a card, he operates an inference inside of this frame. If, for instance, he is shown a black four of hearts, he may notice the shape of the symbol—in which case he identifies a four of hearts—or the color—in which case he identifies a four of spades. In the second frame, the different shapes are no longer necessarily related to their usual color. In order to identify the cards, the subjects must carefully separate the shape and the color. They carry out the following inference: to identify a card, check the shape *and* the color. The second frame is without doubt better than the first one, since it allows reality to be perceived in a better way. Choosing this second frame is rational in the sense of rationality of frame.

Here is another classic experiment; six matches are lying on a table and a subject is asked to try and join the matches end to end so that they make a picture with four triangles. The subjects move the matches on the table and try to construct triangles with them; but they realize more or less rapidly that it is impossible to solve the problem. Some subjects give up, but others have the idea of trying to join the matches in space

rather than on the plane. The latter quickly reach the solution and build a pyramid (the sides are four triangles). The subjects who solve the problem hold successively two conceptual frames: the first one is bi-dimensional (plane) and the second tri-dimensional (space). In each frame, the subjects carry out an inference: try to build as many triangles as possible with the matches. Here again, the second frame is better than the first one, since only the former makes it possible to solve the problem. This is another example of rationality of frame.

These examples show the differences between rationality of frame and rationality of inference. But they also show that these two categories have something in common: the choice of the best—choice of the best available frame in one case, choice of the best available inference in the other. The polythetical character of rationality can once again be pictured with the help of discs partially overlapping one another:

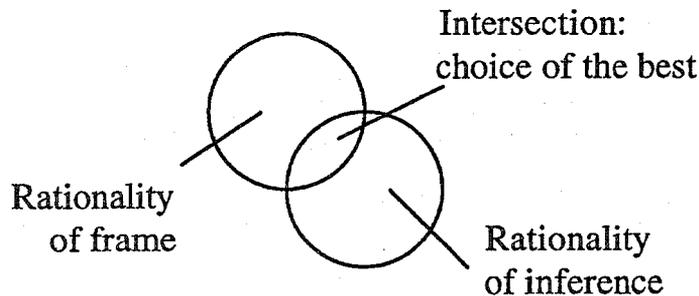


Figure 2. Polythetical character of 'rationality': second example

This distinction can be used in order to study some controversies and misunderstandings concerning rationality: the controversy over rationality as optimization in economics; the controversy over choice anomalies between economics and cognitive psychology; and the controversy over scientific creativity between cognitive psychology and philosophy of science.

Rationality and Optimization

There are two main paradigms in microeconomics, the neo-classical paradigm (founded by Walras [1874], among others) and the 'evolutionary' paradigm (mainly developed by the Austrian school). This major fracture in microeconomics rests upon a distinction between two kinds of models of human action: those that use exclusively a rationality

of inference (neo-classical models), and those that also call upon a rationality of frame (evolutionary models).

Neo-classical economists often tend to identify individual rationality with optimization procedures. Such an identification is obviously very reductionist, since it means the disappearance of the category of rationality of frame. Their tendency is nevertheless easy to explain. In the mathematical models that are their main achievement, there is *no place at all* for a rationality of frame, since these models rest upon the equilibrium hypothesis: the agents are always set in a situation where all they have to do is maximize their utility or their profit (this is also true in the case of game theory). Only a rationality of inference is taken into account.

As we have seen, in the simplest case, the conceptual frame of the consumer is made up of his or her utility function, his or her income and the prices of the goods; the frame of the firm is made up of its production function, the price of its output and the prices of its inputs. The following question arises: how have the economic agents obtained the information composing these frames, information absolutely necessary to carry out the optimizations? It could be supposed that they have been obtained from a prior optimization, through an economics of information *a la* Stigler (1961). But, on the one hand, this would necessitate very heavy hypothesis, and, on the other, it would open the way to an infinite regression (those prior optimizations would necessitate frames that would in their turn have to be explained by optimizations, and so on). It is much simpler to suppose that these agents got their information through a rationality of frame. But as soon as we call upon the latter kind of rationality, we are no longer in the neo-classical paradigm: we are shifting towards the 'evolutionary' paradigm.

Let us briefly examine two theories belonging to this paradigm, those of Kirzner (1973) and Nelson and Winter (1982). Kirzner is a member of the Austrian school of Mises (1966) and Hayek (1937). He has developed a theory of the market process that rests on the notion of unexpected discovery. If a consumer discovers by chance that he can buy a good at a lower price than he used to (for instance, in another shop), then he takes this information into account and changes the conceptual frame of his decision. A consumer can unexpectedly discover new goods, and a producer can discover that better production techniques are available or that profit opportunities can be grasped. In all these cases, the agent adopts a new conceptual frame in which he can subsequently carry out his optimization. This new frame is better than the old one, since it gives a better picture of the economic reality, and it

is thus rational to adopt it; but it is probably not the best of all possible frames: if the agent were omniscient, he could adopt one that would be even better. In this sense, rationality of frame is not an optimization but rather an adaptation. In Kirzner's theory, the requisite that the discoveries by the agents be unexpected is very important because these kinds of discoveries are the very roots of the market process. When information is deliberately looked for (instead of being unexpectedly discovered), the process is already on its way and can be studied with the help of a rationality of inference only.

Nelson and Winter's theory belongs to the tradition of Schumpeter (1934) and studies the nature of technical progress and its consequences. It uses a rationality of frame that is partial because it applies only to the production techniques: prices and wages are calculated at equilibrium and given to the firms without search. Firms look for better production techniques, but they do not know which one is going to bring the highest profit in the future configuration of the market (prices, wages, etc.). They therefore proceed by trial and error: they keep a technique as long as it creates a profit, but as soon as it makes a loss they try out other techniques until they find one that is profit-earning again.

The examination of this fracture between neo-classical and evolutionary paradigms teaches us that the equilibrium hypothesis and rationality of inference go together. As soon as one of them is forgone, the other must also be given up. If the equilibrium hypothesis is abandoned, then the study focuses on processes and a rationality of frame must be called upon. If the rationality of inference is given up, then the equilibrium hypothesis must be abandoned and processes become the objects of study.

Rationality and Framing

The distinction between rationality of frame and rationality of inference helps to specify the nature of a fracture between psychology and economics. An important criticism has been raised against the rational-choice model in economics by the cognitive psychologists Tversky and Kahneman (1981). Here we concentrate on what is perhaps the most basic part of that criticism, that concerning the phenomenon of *framing*. Tversky and Kahneman have established, through experiment, that the preferences of people can be influenced by the way the options of choice are formulated. They consider how this fact questions the ability of any theory of rationality to explain some choices. Here I try and show that

they are wrong: 'framing' is not necessarily an irrational phenomenon, because it is possible to explain it in terms of a rationality of frame. Tversky and Kahneman are right, however, when they say that this phenomenon cannot be explained by the *economic* model of rational choice: as we have seen, this model is not designed for phenomena that belong to a rationality of frame, and that is why we shall need a much more general model, the 'cognitive model' of Boudon.

Let us study one of their experiments. A first group of subjects is asked to choose an alternative in the following problem (Tversky and Kahneman 1981, 453):

Problem 1 [sample: 152 persons]:

Imagine that the US is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

If program *A* is adopted, 200 people will be saved. [72% have chosen this alternative]

If program *B* is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved. [28% have chosen this alternative]

Which of the two programs would you favor?

A second group of subjects was asked to answer the problem formulated as follows:

Problem 2 [sample: 155 persons]:

(same cover story as problem 1)

If program *C* is adopted, 400 people will die. [22% have chosen this alternative]

If program *D* is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die. [78% have chosen this alternative]

Which of the two programs would you favor?

The formulations of problems 1 and 2 are different, but it is easily seen that the problems are in fact identical. There is no difference—in substance—between programs *A* and *C*, and no difference between programs *B* and *D*. Nevertheless, people tend to answer *A* in the first formulation and *D* in the second one: these choices are contradictory. The preferences are influenced by the formulation of the question independently of the content.

Let us study the nature of these choices. In problem 1, the conceptual frame is made up of the two options *A* and *B*, and of the preference order between them. There are two possible frames: (1) *A*, *B*, *A* preferred to *B*,

(2) A, B, B preferred to A . Subjects choose a frame—they tend to choose the first one—then they carry out the inference of maximization: choose the preferred option. The rationality of inference is here perfectly taken into account, since each subject chooses the best option given his preferences. If we admit that the economic model of rational choice is strictly limited to a rationality of inference, then this model is in no way refuted by this experiment. The anomaly originates in the preferences, that is, in the choice of the conceptual frame. If a rationality is at work, it can only be a rationality of frame.

Tversky and Kahneman do not explain why subjects prefer A to B in the first formulation, and D to C in the second one. They just state that when a problem is formulated in terms of gain (lives saved) people tend to be risk averse (choice of A), and when it is formulated in terms of losses (lives lost) people tend to be risk taking (choice of D). Thanks to the cognitive model of Boudon, we can go a step further and try to explain via a rationality of frame why the preferences have this peculiar structure. We start with the hypothesis that the subjects have good reasons to adopt the preference order that they do. These reasons rest on the information that is communicated to them in the terms of the problems. Since this information conveys different aspects of reality according to the terms of the problems, it gives birth to different reasons, and ultimately to different preferences, among the subjects.

Let us consider the first problem. Program A appears to the subjects as a success, because only its favorable consequences are presented: the ability to save 200 persons. Can program B do better? It offers an opportunity to save all the sick, but this advantage over program A is neutralized by the fact that this opportunity is not certain but only probable (probability $1/3$). The unfavorable consequences of program B are by contrast explicitly stated: ' $2/3$ probability that no people will be saved'. In other words, in the case of program A only the positive consequences are specified (200 persons are saved), while in the case of program B the negative consequences are also specified ($2/3$ probability to lose all the sick). Subjects have thus good reasons to prefer program A to program B .

Let us consider the second problem. Program C must at first sight appear to the subjects as a failure, because only its unfavorable consequences are presented: the inability to save 400 persons. Program D may also be a failure (' $2/3$ probability that 600 people will die'), but the fact that this failure may be worse than with program C ($600 > 400$) is compensated by the fact that it is not certain but only probable (probability $2/3$). This time, the favorable consequences of program D

catch the eye of the subjects. They are explicitly stated: '1/3 probability that nobody will die'. In other words, in the case of program C only the negative consequences are specified (400 persons are lost), while in the case of program D the positive consequences are also specified (1/3 probability to save all the sick). Subjects have thus good reasons to prefer program D to program C.

To sum up, this anomaly of choice is here explained by the fact that in one case only the favorable consequences are specified (program A), and in the other only the unfavorable consequences (program C). It is possible that some subjects reach through imagination a neutral representation of the first program: 'If program A/C is adopted, 200 persons will be saved and 400 persons will die'. This requires an effort of abstraction from the text; it seems likely that in general the subjects use the written text that is handed to them without thinking to complete it using their imagination.

The choice between frames is here rational in the sense that the subjects have good reasons—taking into account the way they are given information—to choose in the majority this or that frame. This explanation is, of course, a conjecture. It would be necessary to check empirically that they reason the way I suppose they do (it would also be interesting to know the reasoning of the subjects that have in each case chosen the program against the majority).

The explanation offered here seems in any case better than the one propounded by social psychologist Donald Campbell, who offers an evolutionary explanation:

In the formative ecology of perceptual and brain processes adaptation was achieved by modularly separate focal attention to desired objects on the one hand and to feared objects (aversive 'goals') on the other. . . . A brain thus shaped is ill adapted, in ways that Tversky and Kahneman specify, to two-valued logics or to an exhaustive-set, three-valued logic. (1987, 175)

This conjecture is not convincing because it is ad hoc—it only explains the phenomenon it is supposed to explain, and does not rest upon any independent argument.³

I have propounded the idea that if we want to explain the choice anomalies related to 'framing', we must transcend the model of rational choice and state the problem in terms of a rationality of frame. Tversky and Kahneman do not state the problem in these terms, maybe because they are under the influence of the economic model of rational choice in which one is content with describing the conceptual frames of the choices without trying to explain them.

Rationality and Creativity

Major controversy exists about the nature of scientific creativity between philosophy of science and cognitive psychology. Some philosophers of science consider that this phenomenon does not belong to the category of rationality. Some cognitive psychologists consider on the contrary that scientific creativity is a rational phenomenon in the sense that scientists are rational while they are making a discovery. The distinction between a rationality of frame and a rationality of inference is very useful for studying this controversy, as it allows us to understand that these philosophers and psychologists talk about different and complementary aspects of creativity: the former evoke the discovery of new conceptual frames (of new paradigms, for instance), while the latter evoke the use of inferences within given conceptual frames. In other words, the philosophers think of a rationality of frame and the psychologists think of a rationality of inference.

Some important philosophers of science characterize scientific discovery as a purely empirical phenomenon. According to them, scientific discovery can be described—it is a succession of mental states in the mind of the scientist—but it is outside the category of rationality. It is only once the discovery has been made that rationality comes into operation in order to study the validity of this discovery with the help of conceptual arguments or empirical tests. This idea is upheld by Popper: ‘The question of how it happens that a new idea occurs to a man . . . may be of great interest to empirical psychology; but it is irrelevant to the logical analysis of scientific knowledge’ (1959, 31). It is also upheld by Lakatos: ‘[The] problem of appraisal is completely different from the problem of why and how new theories emerge. . . . I take this Kantian demarcation between the “logic of appraisal” and the “psychology of discovery” for granted. Attempts to blur it have only yielded empty rhetoric’ (1978, 168–9).

Some cognitive psychologists—prominent among them is Herbert A. Simon—offer strenuous opposition to this sharp distinction between a rational logic of appraisal and an empirical psychology of discovery. As early as the 1960s, Simon considered how the process of discovery by scientists could be modeled as a rational process with the help of the human problem-solving theory in cognitive psychology (Newell and Simon 1972). However, his examples were a bit simplistic and unconvincing (see Simon 1977, Ch. 5.2 and 5.4). At the end of the 1970s, Pat Langley used this theory of problem-solving in order to develop a set of computer programs supposed to model the creative process of the

discovery of some of the greatest laws of physics: Kepler's third law, Galileo's law of uniform acceleration, Ohm's law, etc. (see Langley et al. 1987). In order to understand the nature of these models, it is necessary to present briefly the theory of problem-solving of Newell and Simon.

Newell and Simon's theory studies the way human beings solve 'complex' problems, i.e. problems that cannot be solved with the help of memory only, or by very simple calculations, but need a search by trial and error. For instance: finding a good move in chess, or finding a proof of a theorem in logic. This is a theory of rationality of inference, since Newell and Simon compare the efficiency in terms of problem-solving of two kinds of inferences: (1) an *exhaustive search* surveying in a systematic way the set of all alternatives until the solution is found, and (2) a *heuristic search* using selective principles towards the solution or at least towards a good alternative. The best of these two inferences is obviously the heuristic search, because it allows the solution to be reached much faster than the exhaustive search. It is thus rational to choose the former inference rather than the latter (the exhaustive search may in certain cases be impracticable; in chess, for instance, it is generally impossible to compute the best possible move because the number of alternatives is too great to be processed within the required time). These inferences are carried out in a conceptual frame that Newell and Simon call a *problem space*, which is a representation of the problem in view of its solution: it includes the data of the problem, its goal, and the set of operators (the structure of rules) that direct the search towards its goal (in the case of chess, for instance, the operators are the moves). When they face a 'complex problem', people begin by adopting a problem space, but this first step lies outside this theory of problem-solving: the problem space is a datum.

The theory of Newell and Simon can be applied to the study of scientific discovery. In a space problem made up of the pertinent variables, the empirical data, and the kind of law to be discovered between these variables, a heuristic search is able to discover the law that relates the data. Langley et al. (1987, 83-4) give (among others) the example of Galileo's law of falling bodies. The problem-solving program is supplied with idealized data relating the time and distance of free fall:

Time (T)	0.1	0.2	0.3	0.4	0.5	0.6
Distance (D)	0.098	0.392	0.882	1.568	2.450	3.528

The program must discover a law of the kind $T^a D^b = k$, where k is a

constant, and where a and b are whole numbers (positive or negative). It uses the following heuristic:

If the values of a term are constant, then infer that the term always has that value. If the values of two numerical terms increase together, then consider their ratio. If the values of one term increase as those of another decrease, then consider their product. (Langley et al. 1987, 66)

Since D and T increase together, the program considers their ratio and creates a new variable $V_1 = D/T$. There are now three variables: D , T and V_1 . Since V_1 and T increase together, the program considers their ratio and creates a new variable $V_2 = V_1/T = D/T^2$. The program computes the values of V_2 and establishes that they all are equal to 9.8, and infers that V_2 always has this value. The law is $V_2 = 9.8$, or else $D/T^2 = 9.8$, the law of uniform acceleration. This process is a heuristic search. If the program had made an exhaustive search, much more time would have been necessary to reach the solution. It would have had to try many formulas before finding the good one: $D = k$, $T = k$, $D^2 = k$, $T^2 = k$, $D \times T = k$, $D/T = k$, $D^3 = k$, $T^3 = k$, $D^2 \times T = k$, $D \times T^2 = k$, $D^2/T = k$, $D/T^2 = k$ (this series is an abridgment because I have taken some symmetries into account: for instance, $T/D = k$ is the same law as $D/T = k$).

A heuristic search saves computation time relative to an exhaustive search, and is thus a better inference than the latter. Cognitive psychologists are right, then, to state that scientific discovery can be considered as rational. But it is a rationality of inference, one that deals with only one aspect of the process of discovery. These models of the creative process, interesting as they may be, neglect a crucial point: where does the problem space (specifying the variables and the nature of the law) come from? The problem space is the conceptual frame of the problem, and the theory of Langley and Simon, being a theory of rationality of inference, can obviously not answer such a question. Consider the example of Galileo once again. The most important part of his discovery is not the law of uniform acceleration itself, but rather a new way of thinking about movement. Galileo has discovered a new conceptual frame—a new paradigm—characterized by the ‘geometrization of the universe’ (see Koyre 1966). This essential aspect of his discovery is a datum in the models of problem-solving of Langley and Simon.

Can the discovery in science of a new conceptual frame (i.e. of a new paradigm) be considered rational? Philosophers of science are probably right when they answer: ‘No’. A new paradigm means a new organization of the categories of experience: it cannot be induced from empirical data. If the paradigm is really a new one, it cannot be inferred

from a pre-existing paradigm because there is an incommensurability between them (as between the Galilean and the Aristotelian paradigms, or between the Newtonian and the Einsteinian paradigms).

The question of the rationality of discovery is very different depending on whether it concerns the discovery of a new conceptual frame or not. If the question bears on the discovery of a scientific law with the help of an inference within a given conceptual frame (a given problem space), cognitive psychologists have conclusively shown that the discovery process can be considered as rational in the sense of the rationality of inference. If the question bears on the discovery of the conceptual frame itself (in which the discovery of the law will subsequently be made), then the theory of cognitive psychologists has no value whatsoever, and until proof of the contrary is presented we can admit that this part of the creative process does not belong to the category of rationality.⁴

Conclusion

The controversies and misunderstandings analyzed in this paper arise from the fact that rationality is a 'polythetical' concept. In each paradigm of the human sciences, scientists develop their own conception of rationality and tend to neglect or ignore the conceptions worked out in other paradigms. Some economists and some psychologists, for instance, tend to neglect rationality of frame and to reduce rationality to rationality of inference. The distinction we have drawn between rationality of frame and rationality of inference is precisely intended to avoid this kind of category mistake. It shows that it is very important to take the irreducible diversity of the notion of rationality into account.

NOTES

1. See Boudon (1993, Ch. 8).
2. Let f be the function of production, p the price of the output, and w_i the price of the input i (for $i = 1$ to n). The profit is: $\pi = pf(q_1, \dots, q_n) - \sum_1^n w_i q_i$, where q_i is the quantity of input i . The maximizing point (q^*_1, \dots, q^*_n) for which profit is the highest, can in principle be calculated from the system of n equations: $\pi'_i(q^*_1, \dots, q^*_n) = 0$ for $i = 1$ to n , where π'_i is the i th derived function of π .
3. See the comparison made by Boudon (1996), in the second section of his paper, between evolutionary explanations and the cognitive model.
4. This idea is expressed by Popper in the following way: 'There is no such thing as a logical method of having new ideas, or a logical reconstruction of this process. My view may be expressed by saying that every discovery contains "an irrational element", or "a creative intuition", in Bergson's sense' (1959, 32).

REFERENCES

- Boudon, Raymond. 1993. *The Art of Self-persuasion*. Oxford: Blackwell.
- . 1995. *Le juste et le vrai. Etudes sur l'objectivité des valeurs et de la connaissance*. Paris: Fayard.
- . 1996. 'The "Cognitivist Model": A Generalized "Rational-choice Model".' *Rationality and Society* 8(2): 123–50.
- Campbell, Donald T. 1987. 'Rationality and Utility from the Standpoint of Evolutionary Biology.' In *Rational Choice: The Contrast Between Economics and Psychology*, ed. Robin M. Hogarth and Melvin W. Reder, 171–80. Chicago, IL: University of Chicago Press.
- Hayek, Friedrich A. 1937. 'Economics and Knowledge.' *Economica* 4: 33–54.
- Kirzner, Israel M. 1973. *Competition and Entrepreneurship*. Chicago, IL: University of Chicago Press.
- Koyre, Alexandre. 1966. *Etudes galiléennes*. Paris: Hermann.
- Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions*. 2nd ed. Chicago, IL: University of Chicago Press.
- Lakatos, Imre. 1978. *The Methodology of Scientific Research Programmes: Philosophical Papers Volume 1*. Cambridge: Cambridge University Press.
- Langley, Pat, Herbert A. Simon, Gary L. Bradshaw and Jan M. Zitzow. 1987. *Scientific Discovery: Computational Explorations of the Creative Processes*. Cambridge, MA: MIT Press.
- Mises, Ludwig von. 1966. *Human Action: A Treatise on Economics*. 3rd ed. Chicago, IL: Henry Regnery.
- Needham, R. 1975. 'Polythetic Classification: Convergence and Consequences.' *Man* 10: 349–69.
- Nelson, Richard R. and Sidney G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press.
- Newell, Allen and Herbert A. Simon. 1972. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Popper, Karl R. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.
- Schumpeter, Joseph A. 1934. *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.
- Simon, Herbert A. 1977. *Models of Discovery*. Dordrecht: Reidel.
- Stigler, George J. 1961. 'The Economics of Information.' *Journal of Political Economy* 69: 213–25.
- Tversky, Amos and Daniel Kahneman. 1981. 'The Framing of Decisions and the Psychology of Choice.' *Science* 221: 453–8.
- Walras, Léon V. 1874. *Eléments d'économie politique pure*. Lausanne: L. Corbaz & Co.
- Wittgenstein, Ludwig. 1958. *Philosophical Investigations*, 2nd ed. Oxford: Blackwell.

RENAUD FILLIEULE (PhD Université de Paris-Sorbonne, 1994) is a research worker in criminology at the Institut des Hautes Etudes de la Sécurité Intérieure (IHESI, Paris). His current research focuses on rationality, sociology of scientific knowledge, and computer simulated modeling in social science.

ADDRESS: 4, rue Saint Hubert, 75011 Paris, France.

« Les pièges de la rationalité : deux études de cas en
sociologie de la connaissance scientifique » (1997)

L'Année sociologique

Vol. 47, n° 2, p. 177-201

LES PIÈGES DE LA RATIONALITÉ : DEUX ÉTUDES DE CAS DE SOCIOLOGIE DE LA CONNAISSANCE SCIENTIFIQUE

Renaud FILLIEULE

RÉSUMÉ. — Cet article opère une distinction entre deux catégories de la rationalité, la *rationalité du cadre* et la *rationalité de l'inférence*. Les psychologues cognitifs et les économistes néoclassiques méconnaissent cette distinction. Ils tendent à généraliser outre mesure le champ de la rationalité de l'inférence au détriment de celui de la rationalité du cadre. Cela les a conduits à soutenir des thèses erronées dans des controverses scientifiques majeures centrées sur la notion de rationalité : la question de la créativité intellectuelle (en psychologie) et la question de l'efficacité relative d'un système planifié et d'une économie de marché (en économie). Comment ces auteurs ont-ils pu se faire piéger dans leur propre domaine de compétence par le concept de rationalité ? La réponse que nous apportons met en évidence deux effets de distorsion (un *effet de filtrage* et un *effet d'expansion*) produits par l'emploi de leurs paradigmes respectifs, le paradigme du traitement de l'information pour les psychologues cognitifs et le paradigme néoclassique pour les économistes.

ABSTRACT. — This paper makes a distinction between two categories of rationality, *rationality of frame* and *rationality of inference*. Cognitive psychologists and neo-classical economists fail to recognize this distinction. They tend to generalize beyond measure the domain of rationality of inference to the prejudice of the domain of rationality of frame. This has led them to defend wrong thesis in major scientific controversies focused on the notion of rationality: the question of intellectual creativity (in psychology) and the question of the relative efficiency of central planification and market economy (in economics). How have these authors been trapped in their own domain of competence by the concept of rationality? Our answer brings to light two effects of distortion (a *filtering effect* and an *expansion effect*) produced by the use of their respective paradigms, the information processing paradigm for cognitive psychologists and the neo-classical paradigm for economists.

L'importance de la notion de rationalité pour les sciences de l'homme n'est plus à démontrer. Mais c'est une notion complexe qui ne se laisse pas enfermer dans des définitions étroitement circonscrites. Cette nature protéiforme de la rationalité peut donner naissance à des erreurs en cas de confusion entre ses différents aspects : on est induit en erreur si l'on croit par exemple qu'un type de phénomènes appartient à une certaine catégorie de rationalité alors qu'en réalité il appartient à une autre. Quand on fait usage,

L'Année sociologique, 1997, 47 n° 2

dans une discipline ou un paradigme des sciences de l'homme, d'une conception particulière de la rationalité, on peut être tenté de la généraliser à un type de phénomènes pour lequel c'est une autre conception qui convient. Des auteurs de premier plan ont commis des *erreurs scientifiques majeures* en généralisant ainsi de manière abusive un aspect de la notion de rationalité au détriment d'autres aspects.

Nous en donnerons deux exemples dans cet article. Le premier en *psychologie cognitive* : les psychologues cognitifs de l'école de Herbert Simon sont aujourd'hui persuadés que les modèles de cognition développés par le paradigme du traitement de l'information en psychologie sont capables de rendre compte du processus de la créativité scientifique. Nous verrons qu'ils se trompent de catégorie de rationalité et nous essayerons d'expliquer pourquoi ils commettent cette erreur. Le second en *économie* : pendant plusieurs décennies (des années 1930 aux années 1970), des économistes néoclassiques ont développé des modèles de planification censés démontrer la possibilité de planifier une économie nationale, et la supériorité d'un système planifié sur une économie de marché. Ces modèles reposent sur une conception de la rationalité qui ne permet pas de traiter le problème de façon correcte, et il s'agira là aussi d'essayer d'expliquer pourquoi ces économistes se sont trompés de catégorie de rationalité.

La présente étude cherche à expliquer l'adhésion des scientifiques à des théories qui sont élaborées dans leur propre domaine de compétence. Elle se rattache à un paradigme de la sociologie de la connaissance qui a été proposé par Raymond Boudon, et qui vise à expliquer les croyances fausses par de « bonnes raisons »¹. Nous appliquerons ce paradigme à la connaissance scientifique : il s'agira de rechercher les « bonnes raisons » qu'ont eues les scientifiques de se tromper, ou en d'autres termes d'*expliquer l'erreur scientifique par des raisons*. L'erreur scientifique est le signe que la vocation de la connaissance scientifique à atteindre la vérité et l'universalité n'a pas été satisfaite. L'étude de l'erreur conduit donc à mettre en évidence les limites qu'impose à la connaissance humaine le contexte localisé de sa production. Elle correspond bien à l'objectif que se fixe traditionnellement la sociologie de la connaissance, à savoir la

1. Voir R. Boudon, 1986, *L'idéologie*, Paris, Fayard ; 1990, *L'art de se persuader*, Paris, Fayard ; 1995, *Le juste et le vrai*, Paris, Fayard. Voir également la thèse récente de P. Demeulenaere, 1994, *Enquête sur les principes des actions économiques*, thèse de doctorat de sociologie de l'université Paris IV.

mise en relation de la connaissance avec le contexte dans lequel elle est produite ou acquise.

Deux catégories de la rationalité

L'un des principaux mécanismes générateurs de l'erreur scientifique est la généralisation excessive de la portée d'un paradigme : les scientifiques appliquent leur paradigme au-delà de son domaine de validité, c'est-à-dire qu'ils l'appliquent à des phénomènes auxquels il n'est pas adapté et dont il ne peut rendre compte². Ce mécanisme appartient à un type d'erreur très général qui consiste à attribuer une validité universelle à des propositions qui n'ont qu'une validité locale. La notion de rationalité se prête à ce type de dérapage car elle a donné naissance à plusieurs catégories différentes suivant les paradigmes où elle est employée. Pour les besoins de notre étude, nous distinguerons deux catégories que nous appellerons respectivement la *rationalité de l'inférence* et la *rationalité du cadre*. Ces catégories ne prétendent en aucun cas fournir un recouvrement exhaustif de la notion de rationalité. Elles ne servent qu'à distinguer des aspects qui méritent de l'être. Nous verrons en effet qu'il existe chez certains psychologues cognitifs et certains économistes une tendance à généraliser outre mesure la catégorie de la rationalité de l'inférence, et à vouloir faire entrer dans cette catégorie des phénomènes qui relèvent de la rationalité du cadre.

La rationalité de l'inférence

La « rationalité de l'inférence » est la rationalité du choix opéré *entre des règles d'inférence* qui permettent de parvenir à une décision ou de résoudre un problème *dans un cadre conceptuel donné*.

Le meilleur exemple de ce type de rationalité nous est offert par la théorie du choix du consommateur en microéconomie. Dans cette théorie, l'agent économique possède une échelle donnée de fins (représentée par une fonction d'utilité) et dispose de moyens donnés (ses ressources). Le cadre conceptuel de sa décision étant ainsi fixé, l'agent doit décider comment il va allouer ses moyens en fonction des fins qu'il poursuit. Il a le choix entre de nombreuses inférences, c'est-à-dire entre de nombreuses façons de mettre ses moyens au service de ses fins. Parmi toutes ces inférences, il en est

2. Voir R. Boudon, *L'idéologie*, op. cit., p. 213-229.

une qui va lui permettre de réaliser ses fins le mieux possible compte tenu des moyens dont il dispose. C'est la règle d'optimisation qui détermine l'allocation maximisant la fonction d'utilité compte tenu de la contrainte qui pèse sur les ressources. La rationalité consiste à choisir cette règle d'optimisation puisqu'elle est meilleure que les autres. On a bien affaire ici à une rationalité de l'inférence : choix de la meilleure inférence parmi un ensemble d'inférences qui s'appliquent dans un cadre conceptuel donné. Ce modèle économique du choix rationnel est un cas particulier pour lequel la meilleure inférence disponible est aussi la meilleure inférence possible compte tenu des données du problème de décision. Nous évoquerons dans la première étude de cas des modèles plus généraux de rationalité de l'inférence, qui ont été proposés par des psychologues cognitifs (au premier rang desquels Herbert Simon).

La rationalité du cadre

La mise en œuvre d'une rationalité de l'inférence présuppose l'existence d'un cadre conceptuel. La « rationalité du cadre » est la rationalité du choix opéré *entre les cadres conceptuels eux-mêmes*.

Reprenons l'exemple du choix du consommateur. Supposons que notre agent économique découvre par hasard un nouvel objectif plus intéressant que les objectifs qu'il poursuivait jusqu'alors, ou bien qu'il découvre un nouveau moyen plus efficace que ceux dont il disposait auparavant pour atteindre ses fins. Dès le moment où il fait cette découverte inattendue, l'inférence d'optimisation qu'il avait précédemment mise en œuvre ne sert plus à rien puisque le cadre conceptuel dans lequel doit s'effectuer la décision n'est plus le même. La rationalité de l'inférence doit maintenant être appliquée dans le nouveau cadre. Ce nouveau cadre est meilleur que l'ancien puisqu'il offre une image plus fidèle de la réalité économique, et il est donc rationnel de l'adopter. Des théories économiques fondées sur ce type de rationalité du cadre ont été développées par des économistes appartenant à l'école autrichienne³.

La philosophie des sciences nous offre elle aussi des exemples de rationalité du cadre. Larry Laudan a proposé une intéressante théorie de la rationalité des changements de paradigmes⁴. Cette théorie définit une rationalité du cadre dans la mesure où un para-

3. Voir par exemple I. M. Kirzner, 1973, *Competition & Entrepreneurship*, Chicago, The University of Chicago Press.

4. L. Laudan, 1987, *La dynamique de la science*, Bruxelles, Mardaga, chap. 3.

digme constitue, dans un domaine donné de la science, un cadre conceptuel dans lequel vont pouvoir être élaborées des théories explicatives. Le paradigme, en tant que cadre conceptuel, précise quelles sont les entités fondamentales qui existent dans le domaine en question, et de quelle façon ces entités interagissent pour donner naissance aux processus étudiés. Laudan présente une combinaison de critères permettant de choisir rationnellement entre des paradigmes, et donc de choisir rationnellement entre des cadres conceptuels⁵.

Affirmer l'existence d'une rationalité du cadre dans le développement de la connaissance scientifique revient à accepter l'existence du progrès scientifique lors des transitions d'un paradigme à un autre. C'est là une question très controversée depuis la publication de l'ouvrage de Kuhn sur les révolutions scientifiques⁶. Nous nous contenterons ici de signaler que des arguments — qui sont à notre sens décisifs — ont été avancés par Laudan à l'encontre des deux piliers du relativisme kuhnien : tout d'abord contre l'idée que l'incommensurabilité entre des paradigmes successifs supprime la possibilité d'un progrès, et ensuite contre l'idée que les changements de paradigmes auraient un caractère « holiste » en ce sens qu'ils concerneraient simultanément les théories, les méthodologies et les épistémologies⁷.

Les modèles psychologiques du processus de créativité scientifique

Depuis la fin des années 1970, des psychologues cognitifs de l'école de Herbert Simon ont développé une théorie du processus de créativité scientifique. Cette théorie s'appuie sur des modèles de cognition programmés sur ordinateur et censés retracer le processus intellectuel qui a conduit à la découverte de quelques-unes des plus importantes lois de la physique : troisième loi de Kepler,

5. Il faut cependant noter qu'un paradigme ne se réduit pas à un cadre conceptuel. Il possède également une dimension heuristique qui indique dans quelle direction rechercher les théories et avec quels éléments les construire. Cette composante heuristique ou méthodologique des paradigmes relève de la rationalité de l'inférence puisqu'elle indique comment atteindre un objectif (en l'occurrence, comment élaborer une théorie) dans un cadre conceptuel donné.

6. T. S. Kuhn, 1983, *La structure des révolutions scientifiques*, Paris, Flammarion, coll. « Champs ».

7. Voir L. Laudan, 1984, *Science and Values*, Berkeley, University of California Press.

loi de la chute des corps de Galilée, loi d'Ohm, etc.⁸ Nous allons commencer par voir que ces modèles reposent sur une rationalité de l'inférence. Nous montrerons ensuite que les grandes découvertes scientifiques relèvent pour l'essentiel de la rationalité du cadre et ne se réduisent donc pas à une rationalité de l'inférence. Il nous restera à expliquer pourquoi ces psychologues cognitifs ont pu ainsi exagérer démesurément l'importance d'un aspect de la rationalité (la rationalité de l'inférence) tout en négligeant totalement un autre aspect qui s'avère pourtant indispensable (la rationalité du cadre).

Ces modèles de créativité relèvent d'une rationalité de l'inférence

Les modèles de Pat Langley et Herbert Simon en psychologie de la découverte scientifique constituent une application de la théorie de la résolution de problème élaborée principalement par ce dernier avec son collègue Allen Newell à partir du milieu des années 1950⁹. Cette théorie a marqué l'avènement de ce que l'on appelle aujourd'hui la « science cognitive ». Elle généralise la théorie économique du choix rationnel en tenant compte des limites de l'esprit humain du point de vue de ses capacités de traitement de l'information. Il existe des problèmes que l'on peut qualifier de « complexes », dont la solution ne peut être trouvée par l'application d'un algorithme simple et direct comme celui de l'optimisation (auquel font appel les modèles économiques du choix rationnel). La recherche de la solution de ce type de problèmes doit suivre un chemin plus contourné, construit par essais et erreurs dans un cadre conceptuel que Newell et Simon appellent un « espace du problème » (*problem space*). L'espace du problème est une représentation du problème en vue de sa solution : il comprend les données du problème, son objectif, et l'ensemble des règles qui permettent de passer d'un état de connaissance vers un autre. Cet espace du problème est considéré dans cette théorie comme une donnée : son origine n'est pas expliquée.

Réduite à sa plus simple expression, la théorie de la résolution de problème de Newell et Simon est une théorie de la rationalité de

8. P. Langley, H. A. Simon, G. L. Bradshaw et J. M. Zitkow, 1987, *Scientific Discovery. Computational Explorations of the Creative Processes*, Cambridge, Mass., The MIT Press.

9. A. Newell et H. A. Simon, 1972, *Human Problem Solving*, Englewood Cliffs, NJ, Prentice-Hall.

l'inférence : elle montre la supériorité d'un type d'inférence — la *recherche heuristique* — sur un autre — la *recherche exhaustive* — pour résoudre les problèmes « complexes ». La recherche exhaustive passe en revue de façon systématique l'ensemble de toutes les possibilités jusqu'à ce que l'on tombe sur la solution, alors que la recherche heuristique utilise des principes de sélection et de guidage vers la solution ou tout au moins vers une solution acceptable. La meilleure de ces deux inférences est évidemment la recherche heuristique, puisqu'elle permet de parvenir beaucoup plus rapidement à la solution des problèmes complexes que la recherche exhaustive. La rationalité consiste donc en l'occurrence à adopter une recherche heuristique plutôt qu'une recherche exhaustive.

L'exemple paradigmatique de ce type de problèmes est celui du jeu d'échecs. L'espace du problème (c'est-à-dire le cadre conceptuel) est constitué par la configuration courante des pièces sur l'échiquier (données), l'objectif d'aboutir au mat ou du moins à une position avantageuse, et les règles du mouvement des pièces. Dans ce cadre, les joueurs effectuent une séquence d'explorations successives des conséquences du mouvement de leurs différentes pièces. Dans l'immense majorité des cas, il leur est impossible d'explorer l'ensemble de l'arbre du jeu, c'est-à-dire l'ensemble des coups et réponses possibles de l'adversaire jusqu'au mat. Il existe un meilleur coup possible, mais les joueurs ne peuvent le déterminer du fait des limites de leur capacité de mémorisation et de leur vitesse d'exploration des branches. Seule une recherche exhaustive permettrait de spécifier le ou les meilleurs coups possibles (ceux qui permettent de parvenir le plus rapidement possible au mat ou au nul). Comme une telle recherche est hors de portée car elle prendrait trop de temps, les joueurs ont recours à une recherche « heuristique » : ils sélectionnent un petit nombre de coups qui leur paraissent prometteurs, explorent d'une façon de plus en plus approfondie leurs conséquences en essayant de tenir compte de la meilleure réponse possible de l'adversaire (procédure du minimax), puis choisissent le coup qui leur paraît conduire à la situation stratégique la plus favorable pour eux.

Passons maintenant à l'application de cette théorie de la résolution de problème à l'étude du processus de découverte scientifique. Voici le modèle que proposent Langley et Simon pour expliquer la découverte par Galilée de la loi du mouvement uniformément accé-

léré¹⁰. On fournit au programme informatique de résolution de problème des données idéalisées qui indiquent la distance parcourue par un solide en chute libre en fonction du temps mis à la parcourir :

Temps (T)	0,1	0,2	0,3	0,4	0,5	0,6
Distance (D)	0,098	0,392	0,882	1,568	2,450	3,528

Le programme doit chercher à relier les variables T et D par une loi du type $T^a D^b = k$, où k est une constante, et où a et b sont des entiers relatifs qui ne peuvent être nuls simultanément. L'ensemble constitué par les variables pertinentes (ici, le temps T et la distance D), les données empiriques, et le type de relation à découvrir entre ces variables, constitue l'« espace du problème ».

Le programme met en œuvre une *recherche heuristique*, c'est-à-dire une recherche guidée par des principes d'orientation et de sélection. Ces principes sont en l'occurrence les suivants : si les valeurs d'un terme sont constantes, inférer que ce terme a toujours cette valeur ; si les valeurs de deux termes numériques augmentent ensemble, considérer leur rapport ; si les valeurs d'un terme augmentent alors que celles d'un autre diminuent, considérer leur produit. Constatant que D et T augmentent ensemble, le programme considère leur rapport et construit une nouvelle variable $V_1 = D/T$. Il dispose maintenant de trois variables : D , T et V_1 . Constatant que V_1 et T augmentent ensemble, le programme considère leur rapport et construit une nouvelle variable $V_2 = V_1/T = D/T^2$. En calculant les valeurs de la variable V_2 , le programme constate qu'elles sont toutes égales à 9,8. Il infère donc que V_2 a toujours cette valeur. La loi recherchée est : $V_2 = 9,8$ ou encore $D/T^2 = 9,8$. Il s'agit bien de la loi de la chute des corps de Galilée.

La rationalité de ce processus apparaît quand on compare cette recherche heuristique avec une recherche exhaustive qui consisterait à passer en revue de façon systématique l'ensemble de toutes les possibilités de lois jusqu'à ce que l'on tombe sur la bonne. Une recherche exhaustive devrait tester successivement toute une série de formules avant de trouver la bonne : $D = k$, $T = k$, $D^2 = k$, $T^2 = k$, $D \times T = k$, $D/T = k$, $D^3 = k$, $T^3 = k$, $D^2 \times T = k$, $D \times T^2 = k$, $D^2/T = k$, $D/T^2 = k$ (cette série est abrégée car j'ai tenu compte de certaines

10. Voir P. Langley, H. A. Simon *et al.*, *Scientific Discovery*, *op. cit.*, p. 83-84.

symétries dont une recherche exhaustive n'aurait pas tenu compte). Quand le programme conduit une recherche exhaustive, il met beaucoup plus de temps pour parvenir à la formule qu'en conduisant une recherche heuristique. Cette dernière permet d'économiser des calculs, et constitue donc une meilleure inférence qu'une recherche exhaustive. La recherche heuristique qui est censée modéliser le processus de découverte est rationnelle au sens de la rationalité de l'inférence.

Les découvertes théoriques majeures ne se réduisent pas à une rationalité de l'inférence

Les modèles du processus créatif proposés par Langley et Simon, malgré leur intérêt incontestable, négligent une question cruciale : d'où vient l'espace du problème qui définit les variables pertinentes et indique la forme de la loi ? À cette question, une théorie qui se limite à une rationalité de l'inférence ne peut évidemment apporter aucune réponse puisque *l'espace du problème est un cadre conceptuel* et en tant que tel relève d'une rationalité du cadre.

Reprenons l'exemple de la loi de la chute des corps de Galilée. L'aspect le plus important de cette découverte ne réside pas tant dans l'expression même de la loi du mouvement uniformément accéléré que dans une nouvelle façon de penser le mouvement, c'est-à-dire dans l'élaboration d'un nouveau cadre conceptuel. Voici ce qu'Alexandre Koyré a pu écrire à ce sujet :

« J'ai essayé d'analyser [dans les *Études galiléennes* (1940)] la révolution scientifique du XVII^e siècle, à la fois source et résultat d'une profonde transformation spirituelle qui a bouleversé non seulement le contenu, mais les cadres mêmes de notre pensée : la substitution d'un Univers infini et homogène au cosmos fini et hiérarchiquement ordonné de la pensée antique et médiévale, implique et nécessite la refonte des principes premiers de la raison philosophique et scientifique, la refonte aussi de notions fondamentales, celles du mouvement, de l'espace, du savoir et de l'être. C'est pourquoi la découverte de lois très simples, telle la loi de la chute des corps, a coûté à de très grands génies de si longs efforts qui n'ont pas toujours été couronnés de succès. »¹¹

La découverte de Galilée, lorsqu'on la considère dans toute son ampleur, ne se réduit pas à une rationalité de l'inférence. Elle fait apparaître un nouveau cadre conceptuel, c'est-à-dire un nouveau

11. A. Koyré, 1973, *Études d'histoire de la pensée scientifique*, Paris, Gallimard, p. 13.

paradigme, qui se caractérise essentiellement par la géométrisation de l'univers et donc de la physique. La rationalité qui entre en jeu est une rationalité du cadre. Elle ne porte pas sur le processus même de découverte du nouveau paradigme, mais opère *a posteriori* en comparant l'apport des différents paradigmes en concurrence. Le cadre élaboré par Galilée a peu à peu montré sa supériorité sur ses concurrents en donnant naissance à des théories qui permettaient de poser et de résoudre un certain nombre de problèmes cruciaux qui étaient mal posés dans les autres paradigmes (paradigme aristotélicien et paradigme de l'*impetus*).

On voit donc à quel point la modélisation de Langley et Simon donne une vision tronquée de la découverte de Galilée. La difficulté pour ce dernier n'était pas d'établir une loi que l'on peut après coup considérer comme simple, mais bien de construire un cadre conceptuel entièrement nouveau dans lequel cette découverte pourrait être effectuée. Cette difficulté est totalement escamotée dans leur théorie psychologique puisque le cadre conceptuel y est considéré comme une donnée qui ne fait l'objet d'aucune explication en termes de rationalité.

Langley et Simon ont conscience du type d'objection que nous soulevons ici :

« Les processus que nous avons postulés peuvent-ils vraiment expliquer les grandes découvertes qui ont produit des révolutions scientifiques par changements majeurs de paradigmes ? De tels doutes resteront dans quelques esprits, et c'est légitime, jusqu'au jour où un programme [informatique] comprenant des processus normaux de résolution de problème produira un changement majeur de paradigme et initiera une révolution scientifique. »¹²

Cette dernière phrase montre néanmoins qu'ils ne comprennent pas que leurs modèles emploient une rationalité qui ne peut être mise en œuvre que dans un cadre conceptuel préexistant. Or un changement majeur de paradigme consiste généralement en l'avènement d'un cadre conceptuel incommensurable avec ceux qui l'ont précédé¹³. Comment donc un programme informatique pourrait-il transcender le cadre conceptuel dans lequel il a été conçu et qui lui a été inculqué par son programmeur ? Les modèles que nous présentent Langley et Simon sont, par leur nature même, incapables d'un tel accomplissement.

12. P. Langley, H. A. Simon *et al.*, *Scientific Discovery*, *op. cit.*, p. 339.

13. T. S. Kuhn, *La structure des révolutions scientifiques*, *op. cit.*, chap. 8 et 9 ; L. Laudan, *La dynamique de la science*, *op. cit.*, chap. 4.

Pourquoi ces psychologues commettent-ils cette erreur ?

Lorsqu'ils appliquent aux découvertes scientifiques majeures leur paradigme du traitement de l'information, les psychologues cognitifs de l'école de Herbert Simon se trouvent en dehors de ses limites de validité. Or, ils ne s'en rendent pas compte. Pourquoi ? Cela s'explique selon nous par un *effet de filtrage* produit par leur paradigme lui-même. Un paradigme est un instrument de connaissance très élaboré, qui définit les phénomènes pertinents d'un champ d'investigation et les soumet à ses principes d'explication. Cet instrument vise l'objectivité, et il n'est donc pas neutre : il focalise l'attention sur certains types de phénomènes au risque de conduire à en négliger certains autres. Il se peut qu'il opère comme un filtre visuel, en ne rendant « visible » qu'un aspect de la réalité (c'est-à-dire une partie seulement du « spectre des fréquences », si l'on veut filer la métaphore de la vision). C'est bien ce qui se passe dans le cas d'espèce. Le paradigme du traitement de l'information, nous allons le voir, est incapable de prendre en compte la rationalité du cadre. Cette forme de rationalité, du fait qu'elle échappe à ses principes d'explication, ne peut apparaître dans le champ d'analyse de ses partisans.

Ce paradigme est né au milieu des années 1950, en même temps que les premiers ordinateurs. Il cherche à expliquer le fonctionnement de la pensée en utilisant comme modèles des programmes informatiques. Il repose sur l'idée suivante : les accomplissements intellectuels de l'esprit humain — apprentissage, résolution de problème, induction — consistent en manipulations de symboles, c'est-à-dire en traitement de l'information. Comme les programmes informatiques sont eux aussi capables de traiter l'information, on peut les utiliser comme modèles de la pensée humaine. Ce type de modèle présente un avantage en termes de rigueur : le processus de traitement de l'information doit être décrit de façon suffisamment rigoureuse pour que le programme puisse de fait fonctionner sur un ordinateur.

Une tâche « complexe » étant donnée, il existe bien entendu de nombreux programmes différents capables de l'accomplir. Il faut donc savoir quel est le type de programmes effectivement employé par les êtres humains, compte tenu en particulier de leurs capacités

limitées de mémorisation. On recourt pour cela à des expériences basées sur l'enregistrement de « protocoles verbaux » : des personnes disent à haute voix à quoi elles pensent durant l'accomplissement de la tâche, ce qui permet de déterminer quels opérateurs elles utilisent et dans quel ordre elles le font. Il ne reste plus qu'à élaborer un modèle qui passe *grosso modo* par les mêmes étapes¹⁴. Il faut se plonger dans les articles techniques de la psychologie cognitive contemporaine pour mesurer l'apport considérable que constitue ce paradigme. Il a permis de réaliser des progrès remarquables en fournissant toute une batterie de concepts, en permettant d'atteindre une grande rigueur dans la formalisation, et en offrant la possibilité d'opérer des tests empiriques grâce à l'utilisation de protocoles verbaux¹⁵. Tout ceci explique fort bien le succès qu'il a remporté auprès des psychologues cognitifs.

Le paradigme du traitement de l'information, malgré l'apport conséquent qu'il représente, ne permet de prendre en compte que les phénomènes qui relèvent de la rationalité de l'inférence : dès lors que l'on a modélisé un processus de pensée sous la forme d'un programme informatique, on a placé ce processus dans un certain cadre conceptuel qui ne pourra pas être transcendé, c'est-à-dire qui ne pourra pas donner naissance à un cadre nouveau et incommensurable avec lui. Dans ces conditions, il se produit l'« effet de filtrage » dont nous avons parlé. Les phénomènes qui appartiennent à la catégorie de la rationalité du cadre ne peuvent pas être expliqués par ce paradigme, et ils se retrouvent donc « filtrés » : ils se voient attribuer un statut épistémique qui oscille suivant les cas entre l'incompréhension et l'inexistence pure et simple.

Une grande découverte scientifique est un phénomène qui présente un double aspect : un aspect — primordial — de rationalité du cadre et un aspect — subsidiaire — de rationalité de l'inférence. L'effet de filtrage place le premier aspect en dehors des capacités de perception des psychologues cognitifs, et les conduit d'une part à identifier la rationalité à la rationalité de l'inférence, et d'autre part à affirmer que les découvertes scientifiques sont rationnelles. Lorsqu'ils sont mis en présence d'un phénomène rationnel qui échappe totalement à la catégorie de la rationalité de l'inférence, ces psychologues cognitifs s'avèrent incapables d'en saisir la nature

14. K. A. Ericsson et H. A. Simon, 1993, *Protocol Analysis. Verbal Reports as Data*, Cambridge, Mass., The MIT Press, éd. rév.

15. Voir p. ex. les articles rassemblés dans H. A. Simon, 1979, *Models of Thought*, New Haven, Yale University Press.

rationnelle. Le cas s'est produit pour Newell et Simon lors de l'une de leurs expériences. Ils demandaient à un sujet de résoudre l'énigme cryptarithmétique suivante :

$$\begin{array}{rcccccc}
 & & L & E & T & S & \\
 + & & W & A & V & E & \\
 \hline
 & L & A & T & E & R &
 \end{array}$$

Chaque lettre représente un chiffre entre 0 et 9, et le but est de découvrir la valeur de chacune de ces lettres de sorte que l'addition puisse s'effectuer conformément aux lois de l'arithmétique¹⁶. Le sujet a commencé par essayer de découvrir si le sens des mots (*lets*, *wave* et *later*) donnait une indication quant aux nombres qu'ils représentaient. Puis il a essayé d'attribuer à chaque lettre le nombre correspondant à sa place dans l'alphabet ($A = 1$, $E = 5$, etc.). Il s'est rendu compte que cela ne marchait pas. Il a aussi essayé de voir si la forme même des lettres ne donnait pas une indication sur le chiffre se dissimulant derrière (par exemple : $E = 3$, $S = 2$, etc.). Ce phénomène a beaucoup intéressé Newell et Simon, mais il leur est davantage apparu comme une curiosité que comme une anomalie¹⁷. Or c'est bien d'une *anomalie empirique* qu'il s'agit. Le sujet met en œuvre une rationalité du cadre : il essaye successivement un certain nombre de cadres qui ne conviennent pas, jusqu'à ce qu'il tombe sur le bon cadre. Cette catégorie de la rationalité est totalement étrangère à la théorie de Newell et Simon. Ces derniers ont eu « devant les yeux » une manifestation très claire des limites de leur paradigme, mais ils n'ont pas pu l'interpréter de façon correcte.

Il est important de noter que les expériences nombreuses et très détaillées (reposant sur l'analyse de protocoles verbaux enregistrés en laboratoire) auxquelles se sont livrés ces psychologues ne pouvaient pas les conduire à une prise de conscience de l'existence de la catégorie de la rationalité du cadre. En effet, les tâches accomplies par les sujets de ces expériences font appel à des cadres conceptuels (des « espaces du problème ») qui sont presque évidents. Ces cadres reposent par exemple sur les règles du jeu d'échecs, sur les opérations arithmétiques, ou sur les principes de l'équivalence logique. La phase d'adoption du cadre, dans ces circonstances, n'est qu'une phase sans grand intérêt, une phase secondaire en comparaison de

16. La solution est la suivante : $1\ 567 + 9\ 085 = 10\ 652$, soit $L = 1$, $E = 5$, $T = 6$, $S = 7$, $W = 9$, $A = 0$, $V = 8$, $R = 2$.

17. A. Newell et H. A. Simon, *Human Problem Solving*, op. cit., p. 273-282.

l'accomplissement qui va suivre. Dans les découvertes scientifiques majeures du point de vue théorique, c'est exactement l'inverse : la phase d'élaboration du cadre conceptuel (du paradigme) est de loin la plus intéressante et la plus importante, et la découverte de l'expression même des lois est secondaire en comparaison. Le contexte expérimental dans lequel le paradigme du traitement de l'information a été testé est donc biaisé, du point de vue qui nous occupe. Ce biais est aisément compréhensible. Il est relativement simple de réaliser des protocoles verbaux pour des tâches intellectuelles courantes, bien délimitées dans le temps (pas trop longues). Il paraît en revanche tout à fait impossible, même en principe, d'appliquer cette procédure expérimentale à des découvertes scientifiques majeures. Si cela avait été possible, il ne fait pas de doute que ces psychologues auraient perçu les limitations de leur paradigme. La mise à l'épreuve expérimentale de ces théories psychologiques se prête beaucoup mieux à la confirmation qu'à la réfutation de type poppérien¹⁸.

Les psychologues auraient également pu prendre conscience des limitations de leur paradigme s'ils avaient mené une confrontation de fond avec les philosophes des sciences. Ces derniers ont abandonné depuis environ un siècle la quête d'une rationalité de la découverte scientifique. Ils se sont en effet aperçus que les découvertes théoriques majeures modifiaient les généralisations empiriques dont on croyait disposer¹⁹, et étaient incompatibles avec les théories ou paradigmes antérieurs : il ne pouvait donc pas exister de chemin rationnel entre la connaissance factuelle ou théorique antécédente et la nouvelle théorie ou le nouveau paradigme²⁰. Si nous traduisons cette idée dans notre vocabulaire, cela revient à dire qu'une découverte majeure ne peut se réduire à une rationalité de l'inférence car elle constitue nécessairement un changement de cadre conceptuel. La théorie de Langley et Simon a pris le contre-pied de cette conception bien implantée en philosophie des

18. Voir les remarques classiques formulées sur ce point par T. S. Kuhn dans *La structure des révolutions scientifiques*, *op. cit.*, p. 202-203.

19. La découverte de la gravitation universelle de Newton a par exemple montré que contrairement à ce que l'on pouvait croire depuis Kepler, les trajectoires des planètes n'étaient pas exactement des ellipses puisque l'on devait en toute rigueur tenir compte de l'attraction entre les planètes.

20. K. R. Popper, 1991, « Le but de la science » (1957), repris dans *La connaissance objective*, Paris, Aubier, chap. 5 ; I. Lakatos, 1978, *Mathematics, Science and Epistemology. Philosophical Papers*, Cambridge, Cambridge University Press, vol. 2, p. 130-131 ; L. Laudan, 1980, « Why was the Logic of Discovery Abandoned ? », repris dans L. Laudan, 1981, *Science and Hypothesis*, Dordrecht, Reidel, p. 181-191.

sciences. Il ne s'est pourtant produit aucune discussion critique approfondie entre psychologues et philosophes. Les premiers se sont contentés de constater leur désaccord avec les seconds. À aucun moment ils n'ont examiné les arguments de fond que les philosophes des sciences avaient présentés à l'encontre de la rationalité de la découverte scientifique. À aucun moment ils n'ont expliqué en quoi ces arguments pourraient être erronés²¹.

Cette absence d'argumentation critique sur les questions de fond provient de ce que l'on peut appeler un *effet d'expansion*. Quand un paradigme progresse, c'est-à-dire quand il connaît un élargissement ou un approfondissement de ses domaines d'application, l'attention de ses partisans tend à être monopolisée par cette expansion. Cela est légitime : c'est le travail normal et même prioritaire des scientifiques que d'explorer les capacités explicatives du paradigme qu'ils utilisent. Cette exploration risque néanmoins de les détourner des questions de fond, et les en détournera d'autant plus sûrement que les progrès accomplis leur sembleront plus significatifs. C'est dans ce type de situation que s'est trouvé le paradigme du traitement de l'information au cours des années 1980. Pat Langley a appliqué la pièce maîtresse de ce paradigme (la théorie de la résolution de problème de Newell et Simon) aux processus de découverte scientifique, et il est parvenu à développer des modèles formalisés de cognition applicables à cette classe très importante de phénomènes. Cette généralisation de la théorie ne pouvait évidemment apparaître que comme très stimulante. Elle a donné naissance à de nombreux problèmes techniques (des « énigmes » au sens de Kuhn) auxquels les psychologues cognitifs ont en priorité consacré leurs efforts, laissant de côté la réflexion à mener sur les principes fondateurs de leur paradigme. Cet effet d'expansion explique au moins en partie le caractère *anomique* que présente la division du travail entre psychologues cognitifs et philosophes des sciences autour de la question de la rationalité²².

21. Pour se convaincre de l'inexistence de cette confrontation, il suffit de consulter les textes de référence : H. A. Simon, 1977, *Models of Discovery*, Dordrecht, Reidel, chap. 5.2 et 5.4, ainsi que P. Langley, H. A. Simon *et al.*, *Scientific Discovery*, *op. cit.*

22. On peut également noter que l'absence de discussion critique avec des spécialistes de l'histoire des sciences se révèle préjudiciable pour ce qui est du cas de Galilée. Langley et Simon se font une idée tout à fait fautive de la façon dont il a réalisé sa découverte. Ils croient qu'il a élaboré sa loi de la chute des corps à partir d'expériences qu'il aurait réalisées avec des billes en mouvement sur des plans inclinés. En d'autres termes, ils croient que Galilée a opéré une induction des données vers la théorie, comme le fait leur programme d'intelligence artificielle. C'est inexact. Galilée a procédé à partir d'expériences de pensée. Citons à nouveau Koyré : « L'assertion galiléenne de la chute simultanée des corps graves [telle qu'elle est présentée dans les *Discours et*

Dans sa phase de développement et de progrès, le paradigme du traitement de l'information prend pour ses partisans un caractère hégémonique (effet d'expansion). Placés devant une réalité hétérogène, mélange de rationalité du cadre et de rationalité de l'inférence, les psychologues cognitifs ne sont attentifs et réceptifs qu'aux phénomènes empiriques et aux arguments conceptuels qui vont dans le sens de leur propre paradigme, c'est-à-dire de la rationalité de l'inférence (effet de filtrage). Nous comprenons donc pourquoi, dans le contexte théorique et expérimental qui était le leur, ces psychologues cognitifs ont pu commettre l'erreur de croire que leur paradigme pouvait rendre compte de la créativité scientifique.

Les modèles économiques de planification décentralisée

À partir des années 1930, des économistes néoclassiques partisans du collectivisme ont élaboré une théorie dite de la « planification décentralisée ». Cette théorie répondait à l'objection qui avait été soulevée à l'encontre de la planification centralisée par Ludwig von Mises dans un article célèbre de 1920 sur le calcul économique en régime collectiviste²³. Jusqu'à la publication de cet article, nombreux étaient ceux qui croyaient que la planification intégrale d'une économie nationale ne requerrait pas l'existence d'un système de prix. Il suffirait de mettre fin à l'« anarchie du marché » en confiant le pouvoir exécutif de commandement à un Bureau administratif central composé d'économistes compétents et dévoués au bien commun. Mises a montré que si l'on voulait organiser un ordre économique rationnel du point de vue de la satisfaction de la demande des consommateurs, on ne pouvait pas se passer d'un système de prix pour les biens de production.

Son argument initial est, en résumé, le suivant. Une activité de production donnée peut être réalisée de multiples façons diffé-

*démonstrations mathématiques concernant deux sciences nouvelles (1638)] ne repose, on s'en est bien rendu compte, que sur des raisonnements a priori, et des expériences imaginaires. Jamais, jusqu'ici, nous n'avons été mis en présence d'une expérience réelle ; et aucune des données numériques, que Galilée avait invoquées, n'exprimait des mesures effectivement exécutées. » (A. Koyré, *Études d'histoire de la pensée scientifique*, p. 257.) Il va même jusqu'à dire que « Galilée se trompe à chaque fois qu'il s'en tient à l'expérience » (p. 83). Langley et Simon font montre ici d'un *préjugé inductiviste* qui correspond bien à leur conception particulière de la rationalité puisque l'induction peut être conçue comme une rationalité de l'inférence.*

23. L. von Mises, 1920, « Die Wirtschaftsrechnung im sozialistischen Gemeinwesen », *Archiv für Sozialwissenschaften*, vol. 47 ; en anglais, « Economic Calculation in the Socialist Commonwealth », dans F. A. von Hayek (ed.), 1935, *Collectivist Economic Planning*, Londres, George Routledge & Sons, chap. 3.

rentes, en particulier en ce qui concerne les procédés techniques de production. Cette activité doit conduire à une création de valeur, sans quoi elle ne fera que dégrader la situation de l'économie. En régime de marché, il existe un critère très simple qui permet de déterminer si une activité de production a contribué ou non à créer de la valeur : la comparaison, en termes monétaires, entre ce qu'elle a rapporté, et ce qu'elle a coûté. En régime collectiviste, il n'existe pas de marché des biens de production. Il n'existe donc pas de prix pour ces biens, et il est de ce fait impossible de réaliser un calcul monétaire permettant d'indiquer s'il y a création ou perte de valeur. Le Bureau de planification se retrouve dans l'incapacité d'évaluer la rationalité économique des activités de production. Il ne peut orienter le système productif vers une meilleure satisfaction des demandes des consommateurs. Ce Bureau est un peu dans la situation d'un automobiliste qui serait censé trouver son chemin sans carte, sans boussole, et alors que tous les panneaux indicateurs auraient été enlevés.

Les économistes néoclassiques ont reconnu la pertinence de cet argument. Ils se sont attachés, à partir des travaux précurseurs de Oskar Lange²⁴, à élaborer des modèles permettant de concilier planification et système de prix concurrentiels. On montrera dans un premier temps que ces modèles de planification décentralisée relèvent d'une rationalité de l'inférence, et dans un deuxième temps que Mises et les partisans de l'école autrichienne, dans leur critique de la planification, se placent du point de vue d'une rationalité du cadre. La partie de sociologie de la connaissance scientifique proprement dite consistera, dans un troisième temps, à expliquer pourquoi ces économistes néoclassiques ont cru qu'ils avaient réfuté l'objection de Mises alors que ce n'est pas le cas : leurs modèles ne traitent pas les difficultés posées par ce qui est en l'occurrence la plus importante des deux catégories de la rationalité, à savoir la rationalité du cadre.

La planification décentralisée repose sur une rationalité de l'inférence

Il existe plusieurs types de procédures de planification décentralisée. Il n'est pas question pour nous d'entrer dans des débats techniques sur les qualités respectives de ces différentes procédures.

24. O. Lange, 1936-1937, « On the Economic Theory of Socialism », *Review of Economic Studies*, 4, n° 1 et 2.

Nous nous contenterons de présenter succinctement un exemple simple, celui des procédures inspirées du « tâtonnement walrasien »²⁵. Nous reprenons les notations employées par Pierre Picard dans son utile ouvrage de synthèse²⁶.

Soit une économie comprenant au total n biens de consommation et de production ($i = 1$ à n). La consommation totale pendant la période considérée est un vecteur noté $c = (c_1, \dots, c_i, \dots, c_n)$, où c_i est la consommation finale du bien i . Cette économie se compose de m entreprises ($j = 1$ à m). L'activité de chaque entreprise j peut être représentée par un vecteur $y_j = (y_{1j}, \dots, y_{ij}, \dots, y_{nj})$ qui indique quelle est la part y_{ij} de chaque bien i qui est utilisée ou produite par l'entreprise j . Par convention, une coordonnée y_{ij} est positive si le bien i est produit par l'entreprise j , négative si ce bien est une ressource utilisée par cette entreprise dans son processus de production. Les vecteurs de production ne sont pas tous possibles : il faut tenir compte des contraintes techniques. Pour chaque entreprise j , il existe donc un ensemble de production Y_j qui définit quels sont les vecteurs techniquement possibles.

L'argument de Mises consistait à dire qu'en l'absence de marchés pour les biens de production, le Bureau de planification serait incapable de prendre des décisions rationnelles en vue de la satisfaction des demandes des consommateurs. D'un point de vue néo-classique, on pouvait lui répondre que si ce Bureau disposait d'une information parfaite sur les techniques de production et sur les ressources disponibles, il lui suffirait de résoudre un simple problème d'optimisation de l'utilité sociale U (en appliquant par exemple le théorème de Kühn et Tücker) :

$$\left\{ \begin{array}{l} \max U(c_1, \dots, c_i, \dots, c_n) \\ \text{sous les contraintes :} \\ c_i - \sum_{j=1}^m y_{ij} < \omega_i, \forall i = 1, \dots, n \\ y_j \in Y_j, \forall j = 1, \dots, m \\ c \in X \end{array} \right.$$

L'utilité sociale U est fonction des consommations finales de chacun des n biens. L'économie fait face à une contrainte sur la quantité des ressources disponibles : la quantité du bien i utilisée

25. Voir l'article de O. Lange cité dans la note précédente, ainsi que celui de K. J. Arrow et L. Hurwicz, 1960, « Decentralization and Computation in Resource Allocation », dans *Essays in Economics and Econometrics*, Pfouts.

26. P. Picard, 1979, *Procédures et modèles de planification décentralisée*, Paris, Economica.

pour la consommation (c_i) et pour la production ($-\sum_{j=1}^m y_{ij}$, pour les $y_{ij} < 0$) doit être inférieure à la quantité initiale de ce bien (ω_i) plus la quantité produite par les entreprises ($\sum_{j=1}^m y_{ij}$, pour les $y_{ij} > 0$).

Les économistes néoclassiques ont interprété l'argument de Mises comme un rejet de l'hypothèse d'information parfaite. Il est incontestable qu'en pratique, le Bureau ne dispose jamais de toute l'information qui lui serait nécessaire pour résoudre son programme de planification. Lange a le premier découvert un moyen de contourner cette difficulté : le Bureau de planification peut acquérir progressivement l'information qui lui fait défaut en échangeant des messages avec les entreprises. Il annonce un système de prix fictifs pour les différents biens, $\hat{p} = (\hat{p}_1, \dots, \hat{p}_i, \dots, \hat{p}_n)$ où \hat{p}_i est le prix fictif du bien i . Compte tenu de ce système de prix, chaque entreprise détermine quelle quantité de chaque bien elle devrait produire ou utiliser si elle souhaitait obtenir un profit maximal (sous la contrainte de possibilité technique). Ces informations sont communiquées au Bureau, qui peut alors calculer le vecteur de consommation \hat{c} (fictif) qui maximise le surplus d'utilité sociale $U(c) - \hat{p} \cdot c$ ($\hat{p} \cdot c$ est le produit scalaire des vecteurs \hat{p} et c). La consommation et la production ainsi calculées peuvent ne pas être compatibles. Le Bureau doit donc modifier le vecteur prix en augmentant le prix des biens pour lesquels il y a excès de demande et en réduisant le prix des biens pour lesquels il y a excès d'offre. On peut démontrer que sous certaines conditions mathématiques, la procédure permet de converger vers l'optimum, c'est-à-dire d'aboutir au système de prix concurrentiel p qui correspond à la solution du problème d'optimisation initial²⁷. Cette planification est *décentralisée* en ce sens que certaines des informations sur lesquelles elle s'appuie ne sont pas détenues par le Bureau central, mais seulement par les entreprises (par exemple, le Bureau n'a pas besoin de connaître les techniques de production mises en œuvre par les firmes).

Il était nécessaire de présenter avec un minimum de détails le principe de la planification décentralisée. Mais ce qui nous inté-

27. Il faut remarquer que toute cette procédure repose sur l'hypothèse que les directeurs d'entreprise jouent le jeu et fournissent honnêtement au Bureau central les informations qui leur sont demandées. Il serait pourtant dans leur intérêt de mentir et d'annoncer des performances sous-évaluées par rapport aux capacités réelles de leur entreprise, de façon à pouvoir atteindre plus facilement les objectifs qui leur seraient fixés. Les économistes autrichiens ont utilisé cet argument des incitations pour critiquer la planification décentralisée, mais leur principal argument est plus profond et reste valable même si l'on suppose que tous les directeurs d'entreprise sont parfaitement honnêtes (voir la sous-section suivante).

resse, en définitive, c'est de savoir à quelle catégorie de rationalité appartient ce type de procédure. Le Bureau ne dispose pas des informations qui lui permettraient de réaliser directement une optimisation. Il opère donc une recherche heuristique dans un cadre conceptuel donné comprenant les différents types de biens et les différentes entreprises qui existent. Il s'agit bien d'une rationalité de l'inférence, mais qui présente la particularité d'être « composée » en ce sens qu'elle intègre des inférences rationnelles de niveau inférieur. Pour déterminer le système de prix optimal, le Bureau effectue une recherche programmée en plusieurs étapes, chacune d'entre elles étant à son tour une inférence rationnelle (maximisation du surplus d'utilité sociale $U(c) - \hat{p} \cdot c$).

La rationalité du progrès économique ne se réduit pas à une rationalité de l'inférence

Hayek offre un bon résumé de la position de l'école autrichienne dans l'article qu'il a consacré au « socialisme de marché »²⁸. Il reconnaît que la méthode par essais et erreurs proposée par Lange pourrait en principe permettre au Bureau de planification de faire converger le système économique vers une situation d'équilibre concurrentiel (au sens de la concurrence pure et parfaite). Mais il précise que cette méthode serait inadéquate pour promouvoir un progrès économique comparable à celui qui existe dans un système capitaliste. Le problème, en l'occurrence, ne consisterait pas à résoudre un programme de planification dans un cadre économique donné, mais à opérer des ajustements aussi rapides que possible aux modifications incessantes des conditions économiques²⁹. C'est le cadre même de l'économie — goûts des consommateurs, techniques, ressources — qui se modifie en permanence sous l'effet des découvertes imprévues effectuées par les agents économiques (découvertes de nouveaux biens, de nouvelles techniques, de nouvelles ressources). Traduite dans notre vocabulaire, cette critique revient à dire que le Bureau de planification aurait à mettre en œuvre une rationalité du cadre, et non pas seulement une rationalité de l'inférence.

Schumpeter est l'auteur qui a exprimé avec le plus de clarté et de vigueur l'idée, que nous tenons pour vraie, selon laquelle la ratio-

28. F. A. von Hayek, 1940, « The Competitive Solution », repris en 1948, dans *Individualism and Economic Order*, Chicago, The University of Chicago Press, p. 181-208.

29. F. A. von Hayek, *Individualism and Economic Order*, op. cit., p. 187-188.

nalité économique du capitalisme ne peut se réduire à une rationalité de l'inférence. Il mérite en l'occurrence d'être cité assez longuement :

« Le problème généralement pris en considération est celui d'établir comment le capitalisme gère les structures existantes, alors que le problème qui importe est celui de découvrir comment il crée, puis détruit ces structures [...] [L']attention du théoricien continue à rester exclusivement fixée sur les modalités d'une concurrence enserrée dans un système de conditions (notamment de méthodes de production et de types d'organisation industrielle) immuables. Mais, dans la réalité capitaliste (par opposition avec l'image qu'en donnent les manuels), ce n'est pas cette modalité de concurrence qui compte, mais bien celle inhérente à l'apparition d'un produit, d'une technique, d'une source de ravitaillement, d'un nouveau type d'organisation (par exemple l'unité de contrôle à très grande échelle) [...] L'action de cette modalité de concurrence dépasse celle de la concurrence des prix tout autant que les effets d'un bombardement dépassent ceux d'une pesée sur une porte. »³⁰

Si un Bureau de planification se donnait pour objectif d'égaliser voire de surpasser la rationalité économique de l'économie de marché, il faudrait qu'il soit capable d'employer une rationalité du cadre. Car il ne suffit pas de vouloir optimiser l'« utilité sociale » (pour autant que ce concept ait un sens). Encore faut-il déterminer dans quel cadre on va effectuer cette optimisation. Il faut pour cela comparer différentes structures de production du point de vue de l'objectif que l'on se fixe. L'économie de marché connaît des « révolutions » de la structure de production, qui correspondent à ce que Schumpeter appelle la « destruction créatrice », et qui représentent une force essentielle pour le progrès économique. En économie de marché, ces révolutions ne sont évidemment décidées par personne. Elles résultent d'un vaste processus de sélection *a posteriori*. Mais en économie planifiée, elles devraient faire l'objet de la décision *a priori* d'un agent constitué (le Bureau), décision qui porterait sur un changement de cadre conceptuel et serait donc censée obéir à une rationalité du cadre. Les modèles de planification décentralisée, qui reposent exclusivement sur une rationalité de l'inférence, sont par nature incapables de prendre en compte cette réalité.

30. J. A. Schumpeter, 1951, *Capitalisme, socialisme et démocratie*, Paris, Payot, p. 118.

Comment ces économistes ont-ils pu commettre cette erreur ?

Les économistes néoclassiques n'ont pas correctement posé le problème de l'efficacité relative du collectivisme et de l'économie de marché, puisqu'ils ont totalement négligé la rationalité du cadre. Cette négligence peut s'expliquer, comme dans l'étude de cas précédente, par un *effet de filtrage* : la catégorie de la rationalité du cadre leur est d'une certaine façon invisible, car elle ne joue absolument aucun rôle dans leurs modèles et leurs théories. Le paradigme néoclassique, par sa conception même, ne peut avoir affaire qu'à une rationalité de l'inférence. Il repose en effet sur l'hypothèse d'équilibre. Or, à l'équilibre, les agents économiques sont dans une situation où ils n'ont plus qu'à maximiser leur utilité ou leur profit : le cadre de leur décision (types de biens, prix) est fixé de telle sorte que les actions optimales soient toutes compatibles. Il n'est donc pas étonnant qu'en cherchant à modéliser les décisions rationnelles d'un éventuel Bureau de planification, ces économistes s'en soient tenus à la seule forme de rationalité qu'ils avaient l'habitude d'utiliser, à savoir la rationalité de l'inférence.

Toujours comme dans l'étude de cas précédente, il faut expliquer pourquoi les tenants de la planification décentralisée n'ont pas pris en compte la critique de fond des « Autrichiens ». Nous voyons là encore apparaître un *effet d'expansion*. Dans la période pendant laquelle des économistes néoclassiques ont développé les modèles de planification décentralisée (principalement des années 1930 aux années 1960), leur paradigme a connu une expansion tout à fait remarquable : première démonstration rigoureuse de l'existence d'un équilibre économique général³¹, développement de la théorie des jeux³², formalisation mathématique systématique de la théorie économique³³, application de la topologie mathématique à la théorie de l'équilibre général³⁴. En dehors des innovations formelles que nous venons de citer, il faut également mettre au crédit du paradigme néoclassique sa capacité à s'implanter en dehors des limites

31. A. Wald, 1936, « Über einige Gleichungssysteme der mathematischen Ökonomie », *Zeitschrift für Nationalökonomie*, 7, p. 637-670.

32. J. von Neumann et O. Morgenstern, 1944, *Theory of Games and Economic Behavior*, Princeton, NJ, Princeton University Press ; J. F. Nash, 1950, « Equilibrium Points in n -Person Games », *Proceedings of the National Academy of Sciences, USA*, 36, p. 48-49.

33. P. A. Samuelson, 1947, *Foundations of Economic Analysis*, Cambridge, Mass., Harvard University Press.

34. G. Debreu, 1959, *Theory of Value*, New York, Wiley.

de l'économie *stricto sensu*, par exemple en science politique³⁵ et en sociologie³⁶. Ce dynamisme exceptionnel ne pouvait que détourner l'attention des économistes néoclassiques de la remise en question des fondements conceptuels de leurs théories. La critique de Mises et de Hayek n'a pas eu d'impact car elle était placée du mauvais côté de la trajectoire du paradigme néoclassique, du côté opposé à celui où se produisaient les avancées théoriques.

Aborder la question des fondements s'avérait d'autant plus difficile pour les économistes néoclassiques que la polarisation entre leur paradigme et le paradigme autrichien n'a fait que s'accroître au cours du temps. Les économistes « autrichiens » de la génération postérieure à celle de Hayek ont de plus en plus renoncé au concept d'équilibre au profit de celui de *processus*³⁷. Chez les économistes néoclassiques, la tendance a été exactement inverse : ils ont considérablement affiné le concept d'équilibre, en particulier grâce à l'apport de la théorie des jeux. Or selon que l'on raisonne dans le cadre d'un équilibre ou dans le cadre d'un processus, les concepts les plus fondamentaux de la théorie économique (concurrence, monopole, profit, monnaie, etc.), prennent des significations radicalement différentes³⁸. L'incommensurabilité entre les deux paradigmes, qui était encore assez peu visible dans les années 1920, est devenue de plus en plus flagrante. Les économistes néoclassiques se sont donc retrouvés dans une situation où leur paradigme gagnait en rigueur et en étendue, alors même que la distance qui le séparait de son principal concurrent dans le domaine de la microéconomie ne cessait de s'accroître. De ce fait, les critiques en provenance du paradigme autrichien, bien qu'elles fussent de plus en plus profondes, sont restées sans le moindre effet.

Dans le débat sur la possibilité du calcul économique en régime collectiviste, chacun est donc resté sur ses positions : les économistes néoclassiques sont persuadés d'avoir réfuté l'argument de

35. A. Downs, 1957, *An Economic Theory of Democracy*, New York, Harper ; J. Buchanan et G. Tullock, 1962, *The Calculus of Consent*, Ann Arbor, The University of Michigan Press.

36. G. S. Becker, 1957, *The Economics of Discrimination*, Chicago, The University of Chicago Press ; 1968, « Crime and Punishment : an Economic Approach », *Journal of Political Economy*, 76, p. 169-217.

37. I. M. Kirzner, 1992, *The Meaning of Market Process. Essays in the Development of Modern Austrian Economics*, Londres, Routledge ; L. M. Lachmann, 1986, *The Market as an Economic Process*, Oxford, Basil Blackwell.

38. I. M. Kirzner, *Competition & Entrepreneurship*, op. cit., chap. 1.

von Mises³⁹, et les économistes de l'école autrichienne, de leur côté, maintiennent que leur argument reste valable et que le collectivisme est impraticable⁴⁰. Cette incompréhension réciproque vient de ce que l'argument de von Mises sur la « rationalité » d'un système économique est ambigu. On peut l'interpréter de deux façons différentes selon que l'on se place du point de vue d'une rationalité du cadre ou d'une rationalité de l'inférence. Les économistes néo-classiques pensent (à tort) la rationalité d'un système économique comme une rationalité de l'inférence (convergence vers un optimum d'un système dont les conditions économiques — préférences des consommateurs, techniques de production, ressources initiales — sont données), alors que les économistes de l'école autrichienne pensent la rationalité d'un système économique avant tout comme une rationalité du cadre (capacité d'un système à s'adapter à des conditions nouvelles et toujours changeantes).

Conclusion

Les psychologues cognitifs et les économistes néoclassiques généralisent de façon abusive la rationalité de l'inférence à des phénomènes qui relèvent de la rationalité du cadre. Nous avons expliqué cette erreur par deux types d'effets liés à la nature et à la situation de leurs paradigmes respectifs (paradigme du traitement de l'information pour les psychologues, paradigme néoclassique pour les économistes) : un effet de filtrage tout d'abord, qui dissimule les aspects de la réalité qui ne peuvent être pris en compte dans le champ explicatif du paradigme, et un effet d'expansion ensuite, qui focalise l'attention sur les progrès en cours du paradigme et la détourne de la discussion critique des principes fondateurs.

Cette étude appartient à la sociologie de la connaissance scientifique plutôt qu'à l'épistémologie dans la mesure où elle a cherché à expliquer l'adhésion à des croyances, et non à évaluer la validité de ces croyances. L'épistémologie joue néanmoins un double rôle dans cette forme de sociologie de la connaissance. Il est tout d'abord indispensable de tenir compte des évaluations normatives des scientifiques dont on étudie les théories. Lorsque par exemple

39. J. Bénard, 1987, « L. von Mises : précurseur ou attardé ? », *Revue d'économie politique*, 97, p. 907-917.

40. D. Lavoie, 1985, *Rivalry and Central Planning. The Socialist Calculation Debate Reconsidered*, Cambridge, Cambridge University Press.

nous avons dit qu'il était légitime que les partisans d'un paradigme donnent la priorité à l'exploration de ses capacités explicatives sur la remise en cause de ses principes constitutifs, nous nous sommes appuyé sur un jugement de valeur accepté explicitement ou implicitement par ces scientifiques. Il est ensuite nécessaire d'aborder le paradigme auquel appartiennent les théories scientifiques étudiées à l'aide d'un paradigme de validité supérieure. On perçoit ainsi les failles de ces théories, c'est-à-dire les aspects de la connaissance que l'on va pouvoir expliquer par le contexte dans lequel ils ont été produits plutôt que par leur capacité à rendre mieux compte du réel que des théories antérieures.

Renaud FILLIEULE
Université de Lille I

« Un modèle d'analyse des présupposés en sociologie
de la connaissance scientifique » (1998)

Revue européenne des sciences sociales

Vol. 36, n° 111, p. 199-215

Renaud FILLIEULE

UN MODÈLE D'ANALYSE DES PRÉSUPPOSÉS EN SOCIOLOGIE DE LA CONNAISSANCE SCIENTIFIQUE

La sociologie de la connaissance scientifique est-elle possible ? On pourrait aujourd'hui en douter. Les tentatives les plus notoires qui ont été effectuées pour la fonder depuis le début des années 1970 – le *programme fort* et le *constructivisme* – n'ont pas tenu leurs promesses. Face à cet échec, une première solution est celle du repli stratégique : abandonner l'étude de la connaissance scientifique à la philosophie des sciences et revenir à une sociologie de la science à la Merton¹, c'est-à-dire à une sociologie étudiant l'environnement social de la science sans plus se mêler d'expliquer le contenu même des croyances scientifiques. Mais il existe une autre solution, qui permet d'éviter de retomber dans cette conception très étroite du rôle de la sociologie et qui repose sur une idée féconde de l'épistémologie contemporaine : *toute connaissance est conditionnée par des présupposés explicites ou implicites, conscients ou méta-conscients*². Cette idée permet de concevoir un modèle d'analyse des croyances scientifiques qui intègre les apports de la théorie de la connaissance tout en se distinguant des modèles développés par la philosophie des sciences³. Je vais dans un premier temps en offrir une présentation synthétique, et dans un second temps l'illustrer par une étude de cas portant sur la sociologie de la connaissance elle-même.

LE MODÈLE D'ANALYSE DES PRÉSUPPOSÉS ERRONÉS

En sociologie de la connaissance scientifique, un modèle devrait d'une part *être capable d'expliquer des croyances scientifiques*, ou en d'autres termes d'expliquer pourquoi des scientifiques croient à des propositions développées dans leur propre domaine de compétence, et d'autre part *se distinguer des modèles qui ont cours en philosophie des sciences* de façon à représenter un champ de recherche spécifique.

Le modèle que je vais présenter, et que j'appellerai le modèle d'*analyse des présupposés erronés*, poursuit ce double objectif.

¹ R. K. Merton, *The sociology of science*, Chicago, The University of Chicago Press, 1973.

² Voir F. A. Hayek, *Studies in philosophy, politics, and economics*, London, Routledge & Kegan Paul, 1967, pp. 60-63 ; K. R. Popper, *Objective knowledge*, Oxford, Clarendon Press, éd. rév., 1979, p. 71 ; R. Boudon, *L'art de se persuader*, Paris, Fayard, 1990, pp. 105-106.

³ Ce modèle a déjà été exploré par R. Boudon, *op. cit.*, et appliqué entre autres par P. Demeulenaere, *Homo oeconomicus*, Paris, PUF, 1996, et R. Fillieule, *Les pièges de la rationalité*, in « L'Année sociologique », XXXVII, 1997.

Le modèle

Il part de l'idée que tout domaine de connaissance scientifique repose sur un ensemble de présupposés, c'est-à-dire sur des propositions qui sont explicitement ou implicitement tenues pour acquises par les scientifiques qui y travaillent. Une analyse de ce domaine de connaissance peut alors être opérée en deux temps : dans un premier temps, on met en évidence certains de ses présupposés (nous verrons lesquels juste après), et dans un second temps, on essaie de déterminer pour quelles raisons les agents les ont adoptés, consciemment ou non.

Quels sont, pour le domaine de connaissance auquel on s'intéresse, les présupposés que l'on va chercher à mettre en évidence ? Toute connaissance, toute croyance, repose sur une infinité de présupposés⁴. Comme il est impossible de tous les prendre en compte, il est nécessaire d'opérer un choix parmi eux. Intervient ici un critère de sélection tout à fait crucial : *on choisit un ou des présupposés qui se sont avérés faux lors des progrès ultérieurs de la connaissance*. Pourquoi s'intéresser à des croyances fausses et dépassées du point de vue scientifique ? Parce que l'erreur nous offre un point de vue privilégié sur la connaissance : elle nous indique les présupposés qui, à la lumière de l'évolution ultérieure de la connaissance, se sont avérés erronés alors qu'ils étaient auparavant tenus pour vrais – et même pour évidents – par les scientifiques concernés. L'erreur nous donne ainsi accès à des croyances clés des systèmes conceptuels auxquels on s'intéresse (paradigmes, théories), elle nous donne accès à ce que Mannheim appelait des « idéologies totales »⁵, à savoir les éléments de la pensée qui apparaissent aux agents comme les plus évidents alors qu'ils sont en fait conditionnés par un « habitat social » particulier. Cependant, Mannheim croyait à tort que les sciences de la nature étaient débarrassées de ces « idéologies totales », et il limitait arbitrairement ce concept aux croyances politiques ou économiques.

Ce principe de sélection – choix d'un présupposé erroné – présente une caractéristique assez singulière : il est indéterministe, dans la mesure où le progrès de la connaissance est lui-même indéterminé. Les analyses que fournit le modèle sont liées à la façon dont la connaissance a effectivement progressé. Si elle avait progressé suivant une autre voie, d'autres présupposés auraient été remis en cause, et le point de vue explicatif offert par modèle sur un système conceptuel donné aurait donc été différent. On peut noter au passage que tout progrès majeur de la pensée scientifique ouvre la perspective de nouvelles théories pour la sociologie de la connaissance scientifique telle qu'elle est conçue ici.

Une fois que l'on a détecté un présupposé pertinent, au sens défini ci-dessus, il s'agit d'expliquer pourquoi les scientifiques concernés y ont adhéré. Cette explication va consister à mettre en évidence les « bonnes raisons » pour lesquelles les personnes auxquelles on s'intéresse l'ont adopté. Je me réfère ici aux travaux de Boudon⁶ qui montrent que l'on peut *avoir de bonnes raisons d'adhérer à des croyances fausses*. Ces bonnes raisons peuvent être très diverses – constats empi-

⁴ Si on suppose que toute connaissance possède au moins un présupposé, alors ce présupposé – qui est une connaissance – possède lui aussi au moins un présupposé, et ainsi de suite. On en conclut que toute connaissance possède une infinité de présupposés.

⁵ K. Mannheim, *Ideology and utopia*, London, Routledge & Kegan Paul, 1936.

⁶ R. Boudon, *L'idéologie*, Paris, Fayard, 1986, *L'art de se persuader*, op. cit.

riques, argumentations conceptuelles⁷ –, mais elles ont pour caractéristique commune d'être *situées* : dans la situation qui est (ou était) la leur, les agents ont (ou avaient) de bonnes raisons d'adopter tel présupposé plutôt que tel autre. La « situation » dont il est question ne se réduit évidemment pas à des variables sociologiques élémentaires comme la place dans une hiérarchie sociale ou l'appartenance à un groupe particulier, mais intègre des représentations, des théories auxquelles adhèrent les agents. La rationalité, ici, consiste donc en des raisonnements explicites ou implicites que conduisent les agents en s'appuyant sur les différents éléments qui composent leur situation : ce qu'ils « voient » (constats empiriques) et ce qu'ils savent ou croient savoir (argumentations conceptuelles).

En résumé, le modèle d'analyse des présupposés erronés se décompose en deux phases successives :

1. Une phase *descriptive* dans laquelle, d'une part on indique le présupposé erroné auquel on s'intéresse, et d'autre part on montre en quoi il est erroné ;
2. Une phase *explicative* dans laquelle on explique l'adhésion à ce présupposé par de « bonnes raisons » au sens de la rationalité située.

Bien qu'il offre un point de vue spécifique sur le processus de progrès de la connaissance, ce modèle se trouve dans un rapport de subordination vis-à-vis de ce processus. Il suppose en effet qu'un travail scientifique a préalablement été effectué, qui permet de savoir si une théorie, un concept, ou un paradigme est dépassé, et si oui pourquoi.

Analyse des présupposés erronés et philosophie des sciences

Dès lors que l'on se place dans le cadre de l'épistémologie contemporaine et que l'on a recours à la notion de rationalité, on supprime la barrière infranchissable que de nombreux sociologues ont établie entre sociologie de la connaissance scientifique et philosophie des sciences. Ces deux disciplines ne sont plus en contradiction radicale l'une avec l'autre, mais elles ne se confondent pas pour autant. La philosophie des sciences est une discipline qui a essentiellement pour rôle de déterminer les critères permettant de décider de la *supériorité objective* d'une théorie scientifique sur une autre ou d'un paradigme sur un autre⁸. Elle est donc tournée vers le vrai, c'est-à-dire vers ce que l'on peut considérer comme l'aspect *universel* de la connaissance scientifique. La sociologie de la connaissance scientifique, à travers le modèle que nous présentons, s'intéresse au revers de la médaille, à savoir à l'échec d'une théorie par rapport à une autre ou d'un paradigme par rapport à un autre. Ce type d'échec prend la forme d'une remise en cause de propositions que l'on avait crues vraies et qui s'avèrent fausses, ou de propositions que l'on avait crues universelles et qui s'avèrent particulières. Dans tous les cas, les erreurs scientifiques mettent en évidence les limites qu'impose à

⁷ J'emprunte la distinction entre problèmes empiriques et problèmes conceptuels à L. Laudan, *Progress and its problems*, Berkeley, University of California Press, 1977, chap. I et II.

⁸ K. R. Popper, *La logique de la découverte scientifique*, Paris, Payot, 1973 [1934]; I. Lakatos, *The methodology of scientific research programmes*, in I. Lakatos et A. Musgrave (eds.), *Criticism and the growth of knowledge*, Cambridge, Cambridge University Press, 1970, pp. 91-196; L. Laudan, *Progress and its problems*, *op. cit.*

la connaissance humaine le contexte localisé de sa production. Elles sont le signe que la vocation de la connaissance scientifique à atteindre le vrai et l'universel n'a pas été satisfaite. La sociologie de la connaissance scientifique est en ce sens tournée vers l'aspect *contextuel* de la connaissance scientifique.

On pourrait penser que la sociologie de la connaissance scientifique se distingue de la philosophie des sciences en ce qu'elle est une discipline explicative ou positive, alors que cette dernière est une discipline normative. Cela ne me paraît pas exact. La philosophie des sciences possède une dimension explicative : dans la mesure où les scientifiques obéissent aux normes qu'elle définit, elle permet d'expliquer pourquoi ils adhèrent à telle théorie ou tel paradigme plutôt qu'à tel autre. Quant à la sociologie de la connaissance scientifique (au sens où elle est définie ici), elle possède une dimension normative : en faisant prendre conscience aux scientifiques des sources possibles d'erreurs, elle peut éventuellement les aider à les éviter. Si par exemple un scientifique s'aperçoit que des erreurs peuvent être dues à des présupposés implicites présents dans le paradigme auquel on a recours, alors il pourra décider d'être plus attentif aux discussions critiques avec les partisans des paradigmes concurrents, ce qui lui permettra éventuellement d'améliorer la valeur de ses travaux.

La différence entre sociologie de la connaissance scientifique et philosophie des sciences est en définitive une *différence de point de vue sur le processus de connaissance*. Quand un philosophe des sciences étudie la controverse entre le paradigme newtonien (gravitation universelle) et le paradigme einsteinien (relativité), il se demande toujours pourquoi Einstein a eu raison et jamais pourquoi Newton s'est trompé. Or l'erreur de Newton mérite bien une explication, qui ne peut se limiter à dire que Newton s'est trompé « parce que c'est Einstein qui avait raison ». Il y a donc bien place pour deux types d'études du processus de connaissance scientifique : l'étude du succès, orientée vers l'universel, et qui correspond à ce que propose traditionnellement la philosophie des sciences, et l'étude de l'échec ou de l'erreur, orientée vers le contextuel, et qui correspond à ce que propose traditionnellement la sociologie de la connaissance (à savoir l'étude de la relation entre la connaissance et le contexte dans lequel elle est produite ou acquise).

Analyse des présupposés erronés et rationalité

La notion de rationalité indispose beaucoup de sociologues de la connaissance car elle leur apparaît comme un concept philosophique plutôt que scientifique. Or, c'est un fait : la rationalité est à l'œuvre dans de nombreux processus sociaux, et en particulier dans le processus de connaissance scientifique. Si on nie cette réalité, on se prive de toute possibilité de pouvoir expliquer le contenu des théories scientifiques. Le sociologue de la connaissance qui essaie d'expliquer des croyances scientifiques sans recourir à la rationalité fait penser à un artisan qui refuserait d'utiliser un marteau pour enfoncer un clou.

Ce rejet de la rationalité remonte à l'avènement de la sociologie de la connaissance « classique », avec des auteurs comme Mannheim et Merton⁹. Ces derniers

⁹ K. Mannheim, *Ideology and utopia*, op. cit.; R. K. Merton, *The sociology of knowledge*, in R. K. Merton, *Social theory and social structure*, New York, The Free Press, éd. augm., 1968, pp. 510-542.

ont pensé que seule la philosophie des sciences était capable d'expliquer les croyances scientifiques. Pour eux, l'expression même de «sociologie de la connaissance scientifique» aurait été une contradiction dans les termes. Dans un deuxième temps, les tenants du programme fort¹⁰ et du constructivisme¹¹ ont au contraire estimé qu'il était possible d'expliquer les croyances scientifiques d'une façon radicalement distincte de la philosophie des sciences. Comment? En renonçant complètement à la notion de rationalité, c'est-à-dire en renonçant à l'idée que les scientifiques adhèrent à leurs théories pour des raisons. Ainsi, la sociologie de la connaissance scientifique devenait pour eux, non seulement possible, mais même nécessaire pour étudier scientifiquement (ou en tout cas sérieusement) les croyances scientifiques. Pourtant, ces tentatives récentes se sont heurtées à une difficulté insurmontable: en rejetant la notion de rationalité, elles se sont du même coup privées de la capacité d'expliquer de façon satisfaisante les croyances scientifiques. Ce dernier point fait encore l'objet de controverses, mais je le tiendrai pour un acquis de la discussion critique qui s'est déroulée depuis une quinzaine d'années en provenance de la philosophie des sciences¹², de la psychologie cognitive¹³ et de la sociologie¹⁴.

Il s'est donc produit depuis Mannheim un mouvement de balancier. Dans la première phase, les sociologues de la connaissance (Mannheim, Merton, etc.) ont renoncé à expliquer les croyances scientifiques car il leur paraissait impossible de proposer une explication concurrente de celle proposée par les philosophes des sciences. Dans la seconde phase, les sociologues de la connaissance (Bloor, Latour, etc.) ont fondé une sociologie de la connaissance scientifique parfaitement distincte de la philosophie des sciences, mais dont le pouvoir explicatif s'est révélé très faible (voire nul). Dans les deux cas, l'impossibilité de fonder une sociologie de la connaissance scientifique est venue du fait que l'on a réservé l'emploi de la notion de rationalité à la philosophie des sciences. Or il faut admettre que la sociologie de la connaissance scientifique ne peut se passer de cette notion, et simultanément remarquer qu'elle fait appel à une modalité de la rationalité qui est différente de celle employée par la philosophie des sciences. Cette dernière s'intéresse à la supériorité d'une théorie sur une autre ou d'un paradigme sur un autre, et juge cette supériorité sur des *résultats*: la meilleure théorie est par exemple capable de résister à des tests empiriques plus sévères que la moins bonne, le meilleur paradigme donne naissance à des théories qui permettent de résoudre une série de problèmes empiriques ou conceptuels qui n'ont pas pu être résolus par les théories du moins bon paradigme, et ainsi de suite. Laudan¹⁵ propose une bonne énumération des types de raisons que l'on peut avoir de préférer

¹⁰ D. Bloor, *Knowledge and social imagery*, Chicago, The University of Chicago Press, 2^e éd., 1991 [1976]; B. Barnes, *Scientific knowledge and sociological theory*, Londres, Routledge & Kegan Paul, 1974.

¹¹ B. Latour et S. Woolgar, *Laboratory life*, London, Sage, 2^e éd., 1986 [1979].

¹² L. Laudan, *The pseudo-science of science?*, in «Philosophy of the social sciences», XI, 1981, pp. 173-198.

¹³ P. Slezak, *Scientific discovery by computer as empirical refutation of the strong programme*, in «Social studies of science», XIX, 1989, pp. 563-600.

¹⁴ R. Boudon et M. Clavelin (sous la dir. de), *Le relativisme est-il résistant?*, Paris, PUF, 1994.

¹⁵ L. Laudan, *Progress and its problems*, op. cit., chap. III et IV.

rer une théorie à une autre ou un paradigme à un autre¹⁶. Le modèle d'analyse des présupposés erronés, en se tournant vers les *sources* du processus de connaissance (les présupposés) et en s'intéressant à l'erreur scientifique, fait appel à des raisons d'un ordre différent – les raisons qui font que les agents, dans la situation où ils se trouvent, sont incapables de percevoir le caractère erroné d'une proposition qu'ils considèrent comme une évidence.

Analyse des présupposés erronés et réflexivité

La question de la réflexivité constitue un thème majeur des discussions contemporaines dans le domaine de la sociologie de la connaissance. L'idée que la sociologie de la connaissance doit être «réflexive», c'est-à-dire que ses schémas d'explication doivent en principe pouvoir s'appliquer à elle-même, a été proposée par Bloor¹⁷ dans son programme fort, et reprise par les auteurs constructivistes. Chez Bloor, la réflexivité ne pose aucune difficulté de principe. Pour lui, la science a une vocation explicative, et la sociologie de la connaissance scientifique doit être une science qui explique de façon causale et a-rationnelle les paradigmes et les théories auxquels adhèrent les scientifiques; en tant que paradigme scientifique, la sociologie de la connaissance scientifique peut donc constituer son propre objet: elle peut, sans se contredire, faire l'objet d'une explication causale a-rationnelle. Ces considérations de Bloor sur la réflexivité se limitent malheureusement aux questions de principe, et ne trouvent aucune application dans ses travaux. Il se contente en d'autres termes d'affirmer la possibilité de principe d'une réflexivité du programme fort. Chez les auteurs constructivistes, l'application du principe de réflexivité pose des problèmes un peu plus délicats. Le constructivisme conduit en effet à une remise en cause radicale du statut de la connaissance scientifique, puisqu'il affirme que les scientifiques ne découvrent pas le réel grâce à leur raison, mais le construisent ou le déconstruisent à travers des interactions sociales localisées et arbitraires (selon le constructivisme, il n'existe pas de réel objectif, ni de rationalité au sens épistémologique du terme). Un sociologue constructiviste qui applique le principe de réflexivité remet en cause, *ipso facto*, le statut de la connaissance qu'il produit lui-même. Il est impossible d'échapper à cette conséquence destructrice: dès lors que l'on nie la raison, le succès scientifique ne peut plus s'expliquer que par la mise en œuvre d'une persuasion forcée – et ceci, par réflexivité, vaut également pour le succès du sociologue constructiviste qui étudie la façon dont s'impose la connaissance scientifique. Les auteurs constructivistes, non seulement sont conscients de cette situation, mais ils l'admettent et même l'assument¹⁸.

¹⁶ Selon Laudan, une théorie constitue un *progrès* sur une autre si son efficacité à résoudre les problèmes est plus grande: «The overall problem-solving effectiveness of a theory is determined by assessing the number and importance of the empirical problems which the theory solves and deducting therefrom the number and importance of the anomalies and conceptual problems which the theory generates» (*Progress and its problems*, p. 68). Quant aux paradigmes, ils peuvent être évalués, soit de façon «synchronique» (par l'efficacité à résoudre les problèmes des théories les plus récentes auxquelles ils ont donné naissance), soit de façon «diachronique» (en comparant par exemple l'efficacité de leurs théories les plus récentes à celle de leurs théories les plus anciennes).

¹⁷ D. Bloor, *Knowledge and social imagery*, op. cit., p. 7.

¹⁸ Latour et Woolgar reconnaissent sans ambages dans la dernière phrase de leur livre que le succès ou l'échec scientifique de leur paradigme est une question de persuasion autoritaire, et ne peut

Suivant le modèle de sociologie de la connaissance adopté dans le présent article, la question de la réflexivité apparaît sous un jour très différent de celui que l'on trouve dans la littérature citée au paragraphe précédent. Dès lors que l'on considère qu'une connaissance, pour être analysée d'un point de vue sociologique, doit être placée dans la perspective d'une connaissance de validité supérieure, il devient impossible d'adopter une démarche réflexive. *La réflexivité constitue tout simplement une impossibilité logique*: je ne peux pas porter un regard réflexif sur le modèle de sociologie de la connaissance que je défends ici, puisque cela supposerait que je puisse me placer du point de vue d'un paradigme ultérieur de la sociologie de la connaissance scientifique, paradigme dont je ne dispose pas à l'heure actuelle (et dont je ne pourrai éventuellement disposer que dans l'avenir).

Attention, il est tout à fait possible d'étudier à l'aide du modèle d'analyse des présupposés erronés certains paradigmes de la sociologie de la connaissance – c'est même précisément ce que je vais faire dans l'étude de cas qui suit. Mais il est un paradigme qui, par définition, est hors de portée d'une telle étude: celui, justement, que je présente. Il nous est radicalement impossible – à moins de pouvoir inverser la flèche du temps – d'étudier dans le cadre de notre paradigme (donc, de façon réflexive) la conception de la sociologie de la connaissance scientifique qu'il constitue, car cela impliquerait de connaître un paradigme qui lui soit supérieur et qui n'advientra éventuellement que dans le futur. Cette considération nous ramène à un élément central du modèle: la contextualisation de la connaissance suppose l'existence d'un point de vue plus englobant, c'est-à-dire suppose qu'un progrès scientifique a été accompli qui place en perspective la connaissance que l'on cherche à expliquer.

Il faut souligner que cette impossibilité logique de la réflexivité n'est absolument pas gênante. Ce qui importe, pour un paradigme scientifique quel qu'il soit, c'est qu'il permette d'élaborer des théories explicatives convaincantes des phénomènes qu'il est censé étudier, et non pas qu'il obéisse à des critères définis dans l'abstrait. La procédure de Bloor, qui consiste à assigner *a priori* la caractéristique de réflexivité (et quelques autres) à la sociologie de la connaissance, ne me paraît pas scientifique.

ÉTUDE DE CAS: LE POSTULAT DE MERTON EN SOCIOLOGIE DE LA CONNAISSANCE

Les principes d'un modèle d'analyse en sociologie de la connaissance scientifique étant posés, il faut maintenant montrer que l'on peut les mettre en œuvre. J'ai choisi de les appliquer à la sociologie de la connaissance elle-même, et plus précisément à un aspect du paradigme mertonien tel qu'il est présenté dans un article célèbre de Merton publié à l'origine en 1945 et intitulé «The sociology of knowledge»¹⁹.

en aucun cas être une question de supériorité intellectuelle: «How much further research, investment, redefinition of the field, and transformation of what counts as an acceptable argument are necessary to make this [constructivist] account more plausible than its alternatives?» (*Laboratory life, op. cit.*, p. 285).

¹⁹ Cet article est repris in R. K. Merton, *Social theory and social structure, op. cit.*, chap. XIV, et in R. K. Merton, *The sociology of science, op. cit.*, chap. I. Dans ce dernier ouvrage, il porte le titre *Paradigm for the sociology of knowledge*.

Le postulat de Merton

Le présupposé erroné auquel je vais m'intéresser a été formulé de façon tout à fait explicite par Merton. Je l'appellerai le *postulat de Merton*. Voici comment cet auteur le présente²⁰:

Un point central d'accord entre toutes les approches de la sociologie de la connaissance est la thèse que la pensée a une base existentielle dans la mesure où elle n'est pas déterminée de façon immanente, et dans la mesure où l'un ou l'autre de ses aspects peut être dérivé de facteurs extra-cognitifs.

Merton définit ici une division du travail qui se veut claire et nette entre l'épistémologie et la sociologie de la connaissance. L'épistémologie permettrait d'expliquer les croyances rationnelles (croyances déterminées «de façon immanente»), alors que la sociologie de la connaissance permettrait d'expliquer les croyances a-rationnelles (celles auxquelles on adhère, non pour des raisons, mais pour des causes extra-cognitives). Or cette distinction, qui pourrait sembler aller de soi, et que l'on retrouve à l'origine de plusieurs développements ultérieurs en sociologie de la connaissance scientifique²¹, est en réalité *erronée*.

L'erreur de Merton

L'erreur fondamentale de Merton, à mon sens, consiste à opposer «base existentielle» et «facteurs immanents» («La pensée a une base existentielle dans la mesure où elle n'est pas déterminée de façon immanente»). Cette erreur l'induit à penser qu'il existe deux sortes d'explications des croyances, exclusives l'une de l'autre: des explications sociologiques (par des facteurs extra-cognitifs) et des explications rationnelle (par des raisons). Pour se convaincre que ce sont bien là des erreurs, il suffit d'analyser quelques exemples classiques.

Tocqueville²² explique les penchants révolutionnaires des intellectuels français du XVIII^e siècle par leur éloignement du gouvernement, alors que les intellectuels anglais, associés de plus près aux décisions publiques, sont davantage conscients des difficultés de l'action publique, et donc beaucoup plus raisonnables. Il relie ainsi un facteur extra-cognitif (le degré d'association à la prise de décision publique) à des croyances (théories politiques respectivement révolutionnaires et raisonnables). Mais on s'aperçoit, dès que l'on explicite cette relation, qu'elle met en jeu des facteurs cognitifs: du fait de leur éloignement des décideurs publics, les intellectuels français dont parle Tocqueville ne peuvent pas percevoir certaines difficultés d'exécution ou certains effets pervers de leurs projets de réforme. Ils manquent tout simplement d'information par rapport à leurs homologues anglais, et il s'agit bien là d'un facteur cognitif lié à l'évaluation rationnelle des théories politiques en jeu.

²⁰ «A central point of agreement in all approaches to the sociology of knowledge is the thesis that thought has an existential basis in so far as it is not immanently determined and in so far as one or another of its aspects can be derived from extra-cognitive factors», cf. *Social theory and social structure*, op. cit., p. 516.

²¹ Voir ci-dessous la sous-section sur le basculement anti-rationaliste de la sociologie de la connaissance scientifique.

²² A. de Tocqueville, *L'Ancien Régime et la Révolution*, Paris, Robert Laffont, Coll. «Bouquins», 1986 [1856], p. 1039.

Autre exemple. Les sociétés archaïques sont des sociétés animistes: leurs membres rendent compte des phénomènes naturels dont ils sont témoins par l'action d'esprits (esprits des ancêtres, divinités commandant aux éléments de la nature). On constate donc à nouveau ici une relation entre un facteur extra-cognitif – la nature « archaïque » d'une société – et une croyance – la croyance animiste. Robin Horton²³ propose une explication séduisante et convaincante de cette relation. Il part de l'idée qu'une pensée qui vise à expliquer, prédire ou contrôler un type de phénomènes, peut s'appuyer sur des analogies, sur des modèles empruntés à des domaines familiers. Ainsi, le physicien anglais Ernest Rutherford a proposé dans les débuts de la science quantique son célèbre modèle de l'atome, qui reposait sur une analogie avec le système planétaire. Ce modèle a progressivement été modifié de sorte que les « planètes » puissent effectuer des changements brusques d'orbite (les sauts quantiques des électrons), le « soleil » a été remplacé par un assemblage de particules différentes (protons et neutrons), etc. Le résultat de cette élaboration a donné un « système planétaire » à l'allure bâtarde, car « sa structure est un hybride entre celle de son prototype céleste et celle perçue dans l'ensemble d'observations qu'il est sommé d'expliquer »²⁴. En analysant de façon approfondie les croyances religieuses des Kalabaris (un peuple africain du delta du Niger), Horton s'est aperçu que les traits *a priori* surprenants de ces croyances (existence de plusieurs types de divinités, chacun ayant des attributs et des pouvoirs particuliers ainsi que des relations bien précises avec les autres divinités) devenaient compréhensibles dès lors qu'on les interprétait comme le résultat d'un mécanisme d'élaboration et d'hybridation tout à fait semblable à celui constaté dans le cas du modèle de Rutherford: les Kalabaris avaient fondé leur modèle explicatif des phénomènes naturels sur une analogie avec leur propre organisation sociale (les esprits peuplant le modèle jouent le rôle des êtres humains dans la société kalabari), et ce modèle avait été progressivement modifié de façon à pouvoir rendre compte des problèmes qui se posaient à eux dans leur environnement particulier. En explicitant la relation entre une société archaïque et ses croyances animistes, Horton a ainsi pu mettre en évidence un facteur cognitif que l'on retrouve jusques et y compris dans le développement de la science moderne: l'utilisation de modèles conçus à partir d'une analogie avec des phénomènes familiers (système planétaire pour Rutherford, organisation sociale pour les Kalabaris et plus généralement pour les peuples primitifs), modèles qui sont ensuite amendés au fur et à mesure qu'on leur demande de rendre compte des observations récalcitrantes.

Troisième et dernier exemple. Ludwig von Mises²⁵ a formulé une critique célèbre de la planification économique centralisée. Il a montré que cette forme de planification allait s'avérer beaucoup moins efficace, du point de vue de la création des richesses et de l'allocation des ressources, qu'un système fondé sur la propriété privée des moyens de production. Considérant ce résultat, il en est venu à se demander pourquoi de nombreux ouvriers avaient pu croire – à tort – depuis le XIX^e siècle qu'un système centralement planifié serait plus efficace qu'un sys-

²³ R. Horton, *The Kalabari world-view: an outline and interpretation*, in «Africa», XXXII, 1962, pp. 197-220.

²⁴ R. Horton, *op. cit.*, p. 211.

²⁵ L. von Mises, *Socialism*, Indianapolis, LibertyClassics, 1981 [1922].

tème capitaliste. Son explication consiste à établir une relation entre un facteur social (la condition d'ouvrier en usine) et une croyance (l'efficacité de la planification économique centralisée):

Le travailleur d'une entreprise capitaliste grande ou moyenne ne voit et ne sait rien des relations unissant les parties individuelles du travail au système économique pris comme un tout. Son horizon en tant que travailleur et producteur ne s'étend pas au-delà du processus qui constitue sa tâche. Il croit que lui seul est un membre productif de la société, et pense que tous ceux, ingénieurs et contremaîtres aussi bien qu'entrepreneurs, qui ne s'occupent pas des machines ou ne portent pas des poids comme il le fait sont des parasites (...) De là où il se trouve, le travailleur ne peut pas voir comment les choses tiennent ensemble. Il pourrait le découvrir en réfléchissant beaucoup et avec l'aide des livres, mais jamais à partir des faits de son propre environnement de travail. Tout comme l'homme moyen peut seulement conclure des faits de son expérience quotidienne que la Terre est au repos et que le soleil se déplace d'est en ouest, l'ouvrier, jugeant par sa propre expérience, ne peut jamais arriver à une connaissance vraie de la nature et du fonctionnement de la vie économique²⁶.

Cette relation entre facteur social et croyance prend la forme d'un « effet de position » au sens de Boudon²⁷: le raisonnement typique (facteur immanent) de l'ouvrier typique, s'appuyant sur les données que lui apporte sa « position » (en l'occurrence, son environnement de travail), peut tendre à l'induire en erreur sur la question de la faisabilité et de l'efficacité d'un système centralement planifié.

Ces différents exemples nous montrent bien que les facteurs extra-cognitifs et les facteurs rationnels, loin de s'exclure mutuellement, sont souvent des compléments indispensables dans l'explication des croyances. Contrairement à ce que laisse entendre le postulat de Merton, il est en général impossible de distinguer les idées qui s'expliquent par des facteurs rationnels de celles qui s'expliquent par des facteurs extra-cognitifs. Ces deux types de facteurs sont simultanément présents, et il s'ensuit que la division du travail entre épistémologie et sociologie de la connaissance, telle que la conçoit Merton, est inadéquate.

Il faut remarquer que l'on trouve chez Mannheim²⁸ une erreur tout à fait similaire. Ce dernier fonde sa théorie sur une distinction fondamentale entre deux types de modes de pensée: tout d'abord, les modes de pensée logique, mathématique, philosophique et scientifique, et ensuite, les modes de pensée qui guident nos décisions et nos actions à visée pratique dans le domaine de la vie politique, économique et sociale. Selon lui, les croyances doivent être expliquées de façon très différente suivant qu'elles appartiennent au premier ou au second type de mode de pensée. Si elles appartiennent au premier type (mode de pensée logico-scientifique), c'est leur validité propre qui explique leur apparition, leur développement et l'adhésion dont elles font l'objet: l'explication des croyances ou des idées se réduit dans ce cas à une question de validité intrinsèque, c'est-à-dire à une question purement épistémologique (qui ne met en jeu que des facteurs immanents). Si les croyances appartiennent au second type de mode de pensée, en revanche, on peut négliger dans une très large mesure, voire complètement, la question épistémologique de leur validité intrinsèque. L'épistémologie laisse alors

²⁶ L. von Mises, *op. cit.*, pp. 319-320.

²⁷ R. Boudon, *L'idéologie, op. cit.*, chap. VI.

²⁸ K. Mannheim, *Ideology and utopia, op. cit.*, chap. I.

la place à la sociologie de la connaissance : pour expliquer l'apparition et le développement d'une idée, il s'avère nécessaire de mener à bien une investigation purement empirique de ses déterminants sociaux (appartenance à un groupe, à une classe, ou à une génération, existence de groupes antagonistes, participation à une action collective, etc.). Cette distinction entre explication purement épistémologique (par validité intrinsèque) et explication purement empirique (par détermination sociale) des croyances est tout aussi erronée que l'opposition formulée par Merton entre facteurs immanents et base existentielle, puisque comme nous l'avons vu dans les exemples ci-dessus, raisons épistémologiques et causes sociales non seulement ne s'excluent pas mais en général se combinent.

Explication de l'erreur

Passons à la phase de sociologie de la connaissance scientifique proprement dite. Pourquoi Mannheim et Merton – et après eux, comme nous le verrons, les partisans du programme fort et du constructivisme – commettent-ils l'erreur d'opposer facteurs immanents et base existentielle ?

Voici l'explication que je propose. Ces sociologues sont victimes d'une combinaison d'*effets de visibilité et d'invisibilité* qui concernent respectivement les facteurs rationnels et les facteurs sociaux dans l'explication des croyances :

- dans le cas des croyances scientifiques, la base existentielle est dissimulée en ce sens qu'elle reste implicite (effet d'invisibilité), alors que les facteurs immanents sont au contraire bien en évidence (effet de visibilité) ; en se fiant à ce qu'ils « voient » dans les travaux épistémologiques, ces auteurs sont conduits à penser qu'il peut exister une détermination « purement » immanente ou rationnelle des croyances, sans interférence d'une base existentielle quelle qu'elle soit,
- dans le cas des croyances de sens commun, les effets jouent en sens inverse : ce sont les facteurs immanents qui ne se laissent pas facilement percevoir (effet d'invisibilité), alors que la base existentielle saute aux yeux (effet de visibilité) ; nos auteurs sont donc amenés à penser que la rationalité ne joue aucun rôle dans l'explication de ce second type de croyance, et qu'il peut exister une détermination « purement » sociologique des croyances.

Explicitons le premier point. La base existentielle de la connaissance scientifique subit un effet d'invisibilité car elle ne joue pas de rôle direct dans les explications des croyances scientifiques que nous offrent les philosophes des sciences. Lorsque ces derniers reconstruisent les argumentations qui ont conduit les scientifiques à adopter telle théorie plutôt que telle autre, ils passent sous silence la combinaison souvent très complexe de facteurs sociaux qui a rendu possible la discussion critique. Par exemple, la réalisation d'une expérience coûteuse et délicate, qui a pu mobiliser des dizaines de personnes pendant plusieurs mois sur du matériel de haute technologie, se résume en un seul et unique argument : dans quelle mesure le fait mis en évidence a permis de corroborer telle théorie. Plus profondément, le rôle des institutions de la rationalité scientifique – c'est-à-dire des institutions qui permettent la critique intersubjective des données et des théories (ouvrages, revues, bibliothèques, colloques, etc.) – est implicitement tenu

pour acquis. En utilisant le langage de Merton, on peut dire que la « fonction » de la base existentielle de la connaissance scientifique est de se rendre aussi transparente que possible, de façon à laisser place exclusivement aux facteurs immanents. Mais bien qu'on ne la « voie » pas, cette base existentielle de la rationalité est absolument indispensable à la production de la connaissance scientifique. Il est tout à fait erroné de l'opposer aux facteurs immanents, ou de considérer qu'elle interfère avec eux, puisqu'en réalité elle les conditionne.

Cet effet d'invisibilité, qui induit une méconnaissance du rôle des facteurs sociaux dans la constitution de la rationalité scientifique, est attesté par une autre conséquence : il conduit Mannheim et Merton à adhérer à une conception erronée de l'objectivité scientifique. Au lieu de considérer que l'objectivité est le produit d'une base existentielle – le produit des institutions de communication qui rendent possible la critique publique et intersubjective (à commencer par le langage et l'écriture) –, ils la considèrent à tort comme le produit de normes et de valeurs internalisées par chaque scientifique. Pour Mannheim, par exemple, le sociologue de la connaissance atteint l'objectivité en prenant conscience des déterminations sociales qui pèsent sur ses propres croyances politico-économiques, et en s'affranchissant du même coup de ces déterminations. Le sociologue est, en d'autres termes, censé faire la synthèse des différents points de vue existants en essayant consciemment d'être impartial²⁹. Pour Merton³⁰, le scientifique – pas seulement le sociologue de la connaissance cette fois – obéit à un ensemble de normes spécifiques qui le guident dans son travail (universalisme, scepticisme contrôlé, désintéressement, communalisme). Ces représentations du scientifique comme un être impartial et désintéressé sont à l'évidence beaucoup trop simplistes. Elles ont constitué des cibles faciles pour Barry Barnes lorsque celui-ci a posé les prémisses du programme fort³¹.

Je passe maintenant à l'explicitation du second point, à savoir l'effet de visibilité qui porte sur la base existentielle et l'effet d'invisibilité qui porte sur les facteurs immanents dans le cas des croyances de sens commun. La base existentielle est maintenant très visible, car c'est elle qui définit la situation : rapports entre les élites intellectuelles et le pouvoir politique dans l'exemple emprunté à Tocqueville, nature archaïque d'une société dans l'exemple emprunté à Horton, environnement de travail de l'ouvrier dans l'exemple emprunté à von Mises. Pour ces différents exemples (et pour beaucoup d'autres que l'on pourrait citer), la relation entre base existentielle et croyances tient à des constats très élémentaires : les

²⁹ « If the investigator, instead of at once taking a definite position, will incorporate into his vision each contradictory and conflicting current, his thought will be flexible and dialectical, rather than rigid and dogmatic (...) [I]t is our aim to bring all that is ambiguous and questionable in our contemporary intellectual life within the scope of overt consciousness and control by constantly pointing out the often concealed and carefully disguised elements in our thinking », cf. K. Mannheim, *Ideology and utopia*, op. cit., p. 88. Cette conception est critiquée de façon fort appropriée par K.R. Popper, *The open society and its enemies*, vol. II, London, Routledge & Kegan Paul, 5^e éd., 1966[1945], pp. 215-217.

³⁰ R.K. Merton, *The normative structure of science*, in *The sociology of science*, op. cit., pp. 267-278.

³¹ Voir sa critique de la théorie de l'« ethos » scientifique de Merton dans B. Barnes et R.G.A. Dolby, *The scientific ethos : a deviant viewpoint*, in « Archives européennes de sociologie », XI, 1970, pp. 3-25.

intellectuels français du XVIII^e siècle ne sont pas associés aux décisions publiques, les peuples archaïques sont animistes, les ouvriers ont souvent été favorables à la planification centrale. Les facteurs immanents sont en revanche beaucoup plus difficiles à percevoir. On ne les découvre que si l'on franchit un pas supplémentaire dans l'analyse en s'intéressant aux raisonnements typiques que peuvent tenir les agents dans la situation qui est la leur. Or ces raisonnements ne sont pas directement observables et doivent être reconstruits par le sociologue. Ils sont d'autant moins observables que les sociologues – et c'est bien compréhensible – ont tendance à focaliser leur attention sur les facteurs sociaux, les facteurs rationnels n'étant pris en compte que de façon implicite. Lorsque par exemple Mannheim se propose d'expliquer l'apparition de la pensée sophiste en Grèce, il met en évidence un facteur social qui lui paraît déterminant, à savoir la concurrence politico-économique qui s'instaure entre deux groupes sociaux, un artisanat urbain qui monte en puissance, et une noblesse en déclin. Ce facteur explique selon lui le scepticisme caractéristique du sophisme, car il conduit deux visions du monde à entrer en concurrence, une pensée mythologique traditionnelle défendue par la noblesse et un mode de pensée plus analytique défendu par de l'artisanat urbain : « Etant donné que pour toute décision morale on disposait d'au moins deux étalons, et pour tout événement cosmique ou social d'au moins deux explications, il n'est pas étonnant qu'ils aient eu une conception sceptique de la valeur de la pensée humaine. »³² Cette explication comporte une composante rationnelle incontestable : un penseur qui se trouve en présence de deux systèmes de pensée antithétiques (faisant chacun valoir de bons arguments) a de fort bonnes raisons de devenir sceptique puisqu'il perd tout point de référence. Mais cette composante rationnelle n'est pas prise en compte comme telle dans l'analyse de Mannheim. Elle ne fait l'objet d'aucun approfondissement. L'intérêt théorique, en se concentrant sur les facteurs sociaux, s'est de fait détourné des facteurs rationnels.

Cet effet d'invisibilité qui opère sur les facteurs immanents permet d'expliquer pourquoi Mannheim et Merton ne sont pas parvenus à spécifier correctement le lien de détermination qui pouvait relier base existentielle et croyances. Mannheim est resté très flou sur la nature de ce lien. Dans une note de bas de page, il écrit : « Nous n'entendons pas par 'détermination' une séquence de cause à effet mécanique : nous laissons ouverte la signification du terme 'détermination', et seule l'investigation empirique nous montrera à quel point est stricte la corrélation entre situation de vie et processus de pensée, ou quelle latitude existe pour des variations dans la corrélation. »³³ Quant à Merton, il note que « bien que ce problème [de la relation entre connaissance et base existentielle] forme à l'évidence le cœur de toute théorie en sociologie de la connaissance, il a souvent été traité par implication plutôt que directement »³⁴. Or, ce lien, nous l'avons vu, est constitué par le postulat de rationalité des agents : une situation sociale a des répercussions directes sur la façon dont les agents vont pouvoir raisonner, sur les données dont ils vont disposer, sur les théories plus ou moins élaborées qu'ils vont appliquer, sur les arguments qui leur paraîtront solides ou au contraire sans valeur. Ainsi, les croyances pourront être expliquées par une rationalité en situation. Mais dès lors

³² K. Mannheim, *Ideology and utopia*, op. cit., p. 8.

³³ K. Mannheim, op. cit., p. 239.

³⁴ R.K. Merton, *Social theory and social structure*, op. cit., p. 530.

que l'on ne perçoit pas la présence de ces facteurs rationnels, la relation entre base existentielle et croyances ne peut plus être explicitée de façon convaincante, et l'on est obligé d'avoir recours à des métaphores peu scientifiques: on parle par exemple de «congruence» entre contexte social et croyances, d'«harmonie», etc.³⁵

La combinaison des différents effets de visibilité et d'invisibilité qui viennent d'être mis en évidence rend bien compte de l'adhésion au postulat de Merton, puisqu'elle conduit à distinguer deux types d'explications reposant sur des principes inconciliables: une explication épistémologique (par facteurs immanents), et une explication sociologique (par base existentielle).

Importance du postulat de Merton dans le basculement anti-rationaliste de la sociologie de la connaissance scientifique

L'erreur de Mannheim et Merton est d'autant plus importante à signaler et à expliquer que leur postulat a joué un rôle essentiel dans les développements qui ont marqué la sociologie de la connaissance scientifique à partir du début des années 1970.

Avec l'avènement du programme fort de David Bloor, puis du constructivisme de Latour et Woolgar, la sociologie de la connaissance *bascule* dans l'option du rejet systématique de la rationalité des agents pour expliquer l'adhésion aux croyances, et en particulier l'adhésion aux croyances scientifiques. Alors que Mannheim et Merton avaient maintenu une sorte d'équilibre entre épistémologie et sociologie de la connaissance en leur attribuant des domaines respectifs bien définis, ces nouveaux paradigmes se veulent des concurrents directs de la philosophie des sciences. Ainsi, selon Barry Barnes, il est possible de proposer des explications causales a-rationnelles de l'évolution des croyances scientifiques. David Bloor reprend la même idée, à savoir que ce sont des facteurs psychologiques et sociaux a-rationnels qui causent – et donc expliquent – de façon déterministe *toutes* les croyances, qu'elles soient considérées comme vraies ou comme fausses, comme rationnelles ou comme irrationnelles, comme scientifiques ou comme magiques. Quant à Latour et Woolgar, ils estiment que les faits scientifiques sont construits dans les laboratoires par le biais de microprocessus sociaux a-rationnels (négociations locales, pratiques institutionnalisées, etc.), microprocessus que l'on peut mettre en évidence grâce à une méthode pseudo-anthropologique (ethnologie du laboratoire).

Comment s'explique ce basculement anti-rationaliste? Pourquoi, en d'autres termes, ces sociologues de la connaissance en sont-ils venus à croire que l'on pouvait expliquer *toutes* les croyances – y compris les croyances scientifiques – sans du tout faire appel à la rationalité des agents? Comme nous allons le voir, ce basculement s'explique sans trop de difficulté dès lors que l'on tient pour acquise l'adhésion de ces sociologues au postulat de Merton.

Commençons donc par remarquer que les partisans de ces nouveaux paradigmes considèrent le postulat de Merton comme une évidence, et qu'ils ont adopté ce présupposé fondateur de la sociologie de la connaissance «classique»

³⁵ Pour une liste de ces métaphores, voir R. K. Merton, *Social theory and social structure*, op. cit., p. 515.

comme point de départ de leur propre entreprise sociologique. David Bloor, par exemple, considère qu'il existe deux types inconciliables de modèles explicatifs des croyances, le *modèle téléologique* (explication par des raisons) et le *modèle causal* (explication par des facteurs psychologiques et sociaux a-rationnels): « Les modèles téléologique et causal constituent des programmes exclusifs l'un de l'autre. En effet, ce sont deux points de vue métaphysiques opposés »³⁶. Pour ce qui est du constructivisme, Bruno Latour nous apprend qu'il est né lors d'une étude anthropologique menée en Côte d'Ivoire, étude qui aurait montré « que de nombreux traits étaient attribués un peu trop vite à 'l'esprit' africain, alors qu'ils pouvaient être expliqués de façon plus simple par des facteurs sociaux »³⁷. On retrouve donc bien là aussi une opposition tranchée entre explication cognitive et explication sociologique des croyances.

Le basculement anti-rationaliste peut alors être expliqué en termes de *rationalité située*: l'adhésion des sociologues du programme fort et du constructivisme au postulat de Merton les place dans une situation où il est rationnel pour eux de choisir un modèle a-rationnel d'explication des croyances. De fait, ce basculement correspond trait pour trait à l'un des critères de rationalité du choix entre paradigmes proposés par Larry Laudan: le *contexte de poursuite* dans une *évaluation diachronique*³⁸. Selon lui, un programme de recherche qui n'en est encore qu'à ses débuts, doit être évalué non pas de façon synchronique mais de façon diachronique. En d'autres termes, il ne faut pas comparer les capacités explicatives globales d'un programme naissant (par exemple, le programme fort) aux capacités du programme de recherche en vigueur (en l'occurrence, la philosophie des sciences): dans une telle évaluation synchronique le challenger se retrouve systématiquement battu car il n'a pas le temps de faire ses preuves. Ce qu'il faut prendre en compte, pour savoir si un nouveau programme de recherche mérite d'être poursuivi (contexte de poursuite), c'est son taux de progrès (évaluation diachronique), c'est-à-dire sa capacité à générer assez rapidement un bon nombre de théories explicatives intéressantes. Il est tout à fait rationnel de poursuivre un programme de recherche nouveau qui présente un taux de progrès plus élevé que ses rivaux, même si globalement il explique moins de phénomènes, puisqu'on peut penser que ce nouveau programme finira par surpasser les plus anciens, ou que certains de ses éléments pourront leur être incorporés, etc. A titre d'illustration, Laudan cite le cas de Galilée: « Si Galilée a été pris au sérieux par les scientifiques ultérieurs du XVII^e siècle, ce n'est pas parce que son système pouvait globalement expliquer davantage que ne faisaient ceux de ses prédécesseurs médiévaux et de la Renaissance (car il ne le pouvait manifestement *pas*), mais plutôt parce qu'il s'est montré prometteur en étant capable, en un temps assez court, d'offrir des solutions aux problèmes qui constituaient des anomalies pour les autres traditions de recherche dans le domaine »³⁹.

C'est précisément dans cette situation que se trouvent le programme fort et le constructivisme à l'époque où ils sont élaborés. À partir du début des années 1970, les études de cas de sociologie de la connaissance scientifique se multiplient, qui mettent en évidence le rôle et l'importance des facteurs sociaux sur le

³⁶ D. Bloor, *Knowledge and social imagery*, op. cit., p. 12.

³⁷ B. Latour and S. Woolgar, *Laboratory life*, op. cit., p. 273.

³⁸ L. Laudan, *Progress and its problems*, op. cit., p. 110.

³⁹ L. Laudan, op. cit., pp. 112-113.

contenu de la pensée scientifique : l'influence des croyances religieuses, l'impact des intérêts politiques en présence, les décisions partisans des institutions académiques de diffuser ou au contraire d'étouffer dans l'œuf des travaux de recherche, etc. On trouvera une ample moisson de faits de cet ordre dans les études de cas rassemblées par Callon et Latour⁴⁰. Dans ce contexte, il est parfaitement rationnel, du point de vue d'une évaluation diachronique, de poursuivre le développement du modèle a-rationnel. Bloor lui-même présente ce type d'argument rationnel en faveur de son programme :

Les modèles téléologique et causal constituent des programmes exclusifs l'un de l'autre (...) [La fausseté de la conception téléologique] ne doit-elle pas être établie avant que le programme fort puisse prétendre continuer son chemin ? La réponse est 'non' (...) Il est improbable que des arguments (« grounds ») décisifs et indépendants puissent être invoqués a priori pour prouver la vérité ou la fausseté des branches de cette alternative métaphysique majeure. Si des objections et des arguments sont proposés contre l'une des deux théories, on découvrira qu'ils dépendent de l'autre et la présupposent, ce qui représente une pétition de principe. Tout ce que l'on peut faire, c'est vérifier la cohérence interne des différentes théories, puis voir ce qui se produit en pratique lorsque des recherches et des théorisations se fondent sur elles⁴¹.

En résumé : les partisans des paradigmes anti-rationalistes en sociologie de la connaissance scientifique, partant de l'idée (fausse) selon laquelle ils se trouvent face à deux modèles concurrents d'explication des croyances (postulat de Merton), sont rationnellement conduits à conclure à la supériorité du modèle a-rationnel en se fondant sur l'accumulation des données empiriques en sa faveur⁴². Comme les facteurs extra-cognitifs sont omniprésents dans l'activité scientifique (que ce soit dans les études de cas d'histoire des sciences ou dans la pseudo-anthropologie de la science d'un Latour), ces auteurs vont trouver autant d'éléments de preuve empirique qu'ils le souhaitent à l'appui de leurs conceptions a-rationnelles. Cela n'empêche évidemment pas que sur le fond, ils se trompent, puisque leur interprétation de la réalité est conditionnée par l'adhésion au présupposé erroné que constitue le postulat de Merton.

CONCLUSION

Cet article s'est attaché à montrer la possibilité, aussi bien du point de vue du principe que du point de vue de l'application, d'une sociologie de la connaissance scientifique.

Le modèle qui a été présenté – le modèle d'analyse des présupposés erronés – repose sur un principe de *contextualisation* qui le distingue des modèles dévelop-

⁴⁰ M. Callon et B. Latour (sous la dir. de), *La science telle qu'elle se fait*, Paris, La Découverte, 2^e éd., 1991[1982].

⁴¹ D. Bloor, *Knowledge and social imagery*, op. cit., p. 12.

⁴² Cette accumulation apparaissait comme d'autant plus probante que le postulat de rationalité subissait dans le même temps une attaque d'une qualité sans précédent dans les travaux de T. S. Kuhn, *La structure des révolutions scientifiques*, Paris, Flammarion, 1983[1962]. Je signale à ce propos que l'on trouve chez L. Laudan, *Science and values*, Berkeley, University of California Press, 1984, une réfutation tout à fait convaincante du relativisme kuhnien.

pés en philosophie des sciences. On explique l'adhésion aux théories scientifiques, non pas en montrant leur supériorité empirique ou conceptuelle sur des théories antérieures (comme on le fait en philosophie des sciences), mais au contraire en mettant en évidence leur *infériorité* vis-à-vis de théories plus solides. On les place en d'autres termes dans un contexte élargi, dans une perspective plus profonde, ce qui permet de faire apparaître les présupposés erronés sur lesquels sont fondées ces théories à expliquer. On peut dès lors se poser une question précise : quelles sont les raisons qui ont conduit les scientifiques, dans tel domaine et à telle époque, à tenir pour acquis tel présupposé, dont nous savons aujourd'hui qu'il est faux ? La contextualisation nous permet donc de définir l'objet à expliquer – l'adhésion à tel présupposé erroné –, et le postulat de rationalité cognitive⁴³ nous offre un principe d'explication de cet objet grâce à la notion de « bonnes raisons ».

Ce modèle spécifique ne prétend évidemment pas à l'exclusivité, et laisse ouverte la possibilité de développer en sociologie de la connaissance scientifique d'autres types de modèles, reposant sur des principes différents.

*Institut de Sociologie
Université de Lille I*

⁴³ R. Boudon, *Le juste et le vrai*, Paris, Fayard, 1995.

« Types de rationalité » (2005)

Dictionnaire de la pensée sociologique

Dirigé par M. Borlandi, R. Boudon, M. Cherkaoui et B. Valade

Paris, PUF, p. 591-594

Types de rationalité

La réflexion sur les types de rationalité présente une importance fondamentale pour le paradigme actionniste en sciences sociales. Dans ce paradigme, les phénomènes sociaux sont conçus comme le résultat agrégé d'actions individuelles, et ces actions sont à leur tour conçues comme ayant un sens pour les acteurs qui les effectuent. Dire qu'une action a un sens pour l'acteur signifie essentiellement qu'elle s'appuie sur des raisons que l'acteur lui-même juge valables : elle ne se réduit pas à un comportement déterminé par des instincts, des pulsions ou des dispositions. Donner un sens à une action, comprendre une action, c'est donc avant tout en retrouver la rationalité du point de vue même de l'acteur qui l'effectue (Boudon, 1995). Cependant, les actions ne reposent pas toutes sur le même type de rationalité. En effet, elles se composent de divers éléments : des règles d'inférence (méthodes de raisonnement), des croyances factuelles ou causales (croyances « positives »), des jugements de valeur (croyances normatives). Selon l'action étudiée, c'est tel ou tel de ces éléments qui réclame une explication, et à chacun d'eux correspond un type de rationalité : respectivement une rationalité instrumentale, une rationalité des croyances et une rationalité axiologique.

La typologie würrmienne

Les discussions sur les types de rationalité trouvent leur source chez Weber. Ce dernier distingue deux types d'« actions sociales » : les actions rationnelles, d'une part, et les actions déterminées par des sentiments ou des habitudes, d'autre part. Parmi les actions rationnelles, il distingue entre les actions obéissant à une *rationalité instrumentale* et celles obéissant à une *rationalité axiologique*. Une action « rationnelle du point de vue instrumental » est celle que l'acteur effectue après avoir soigneusement analysé et pesé les différentes fins qu'il pourrait poursuivre, les moyens dont il dispose pour les atteindre, et toutes les conséquences directes et indirectes des actions envisageables. Une action est « rationnelle du point de vue axiologique », c'est-à-dire du point de vue des valeurs, si elle est accomplie en application stricte de certains jugements de valeurs (moraux, esthétiques, religieux, ou autres), et quoi qu'il en coûte à l'acteur d'adhérer à ces valeurs.

Les distinctions sur lesquelles repose cette typologie webérienne ne sont pas pleinement satisfaisantes. Mises (1933, 82-85) fait remarquer que les actions que Weber exclut de la sphère de la rationalité, à savoir celles qui trouvent leur source dans des sentiments ou des habitudes, peuvent parfaitement être considérées comme rationnelles. L'influence d'un sentiment intense, comme par exemple le désir de vengeance, peut modifier subitement et grandement l'échelle des valeurs qu'un acteur établit entre ses fins. Mais l'existence et l'influence de ce sentiment ne remettent pas en cause le caractère rationnel (intentionnel et calculé) des actions qui seront effectuées pour accomplir cette vengeance. Dans cette perspective, même la frontière entre les deux sous-catégories de rationalité – instrumentale et axiologique – tend à s'effacer. Lorsqu'un acteur est prêt à renoncer à des avantages matériels au nom de valeurs morales ou religieuses (« rationalité axiologique »), il peut le faire après avoir analysé les autres systèmes de valeurs et après avoir soigneusement déterminé les conséquences directes et indirectes que cela aura pour lui et pour autrui (« rationalité instrumentale »). Mises illustre ce point avec l'exemple d'un aristocrate propriétaire terrien qui refuse que son nom ou son blason soit utilisé comme marque pour vendre les produits de sa terre : il renoncera au revenu supplémentaire que lui apporterait cette pratique parce qu'il lui paraît préférable de se conformer à la tradition aristocratique. Ainsi, les sentiments, les habitudes, les besoins matériels et les règles morales influencent les échelles de préférence des acteurs entre leurs fins, mais ne modifient pas le caractère fondamentalement rationnel – au sens de la mise en relation des fins préférées et des moyens disponibles – de toute action consciente et intentionnelle. En définitive, Mises reproche à Weber d'opérer ses distinctions à partir de jugements de valeur arbitraires : dans une analyse scientifique de l'action, il n'y a pas de raison de placer les sentiments et les habitudes « au-dessous » du calcul économique ou des règles morales.

Malgré les limites indiquées par Mises, cette typologie webérienne reste très suggestive. Si l'on interprète ses concepts de base d'une façon un peu différente, elle s'avère utile aujourd'hui encore pour délimiter de grands domaines de discussion autour de la notion de rationalité.

La rationalité instrumentale

Éloignons-nous quelque peu de la définition webérienne (trop large), et définissons ici la « rationalité instrumentale » comme la rationalité des raisonnements ou des *règles d'inférence* que l'acteur utilise pour résoudre les problèmes auxquels il est confronté (on pourrait aussi la nommer la « rationalité référentielle »).

L'illustration la plus classique de la rationalité instrumentale ainsi définie est le « modèle du choix rationnel ». Ce modèle constitue l'un des développements majeurs de la réflexion des économistes depuis la révolution subjectiviste et marginaliste des années 1870. Sa principale application en économie est la formalisation de l'inférence qui permet à un acteur de déterminer l'allocation optimale de ses ressources (que ce soit des facteurs de production ou une quantité de monnaie), c'est-à-dire l'allocation qui lui apportera une satisfaction maximale (Jevons, 1871, 58-61). Cette rationalité peut être exprimée sous la forme d'un calcul mathématique : la maximisation d'une fonction d'utilité sous contrainte de ressources ou de budget. Des théorèmes mathématiques, comme celui de Kuhn et Tucker, sont appliqués à la résolution rigoureuse de ces problèmes d'optimisation. Ce modèle a été appliqué à partir des années 1950 à des questions qui n'étaient habituellement pas conçues comme relevant de la science économique : discrimination, concurrence politique en régime démocratique, criminalité, mariage et fertilité, altruisme, etc. (les travaux de Becker).

Cependant, la rationalité optimisation du modèle du choix rationnel n'est pas applicable à tous les types de problèmes. C'est dans le domaine de la théorie des organisations que la nécessité de disposer d'une conception élargie de la rationalité instrumentale est d'abord apparue. Dans son étude sur les processus de décision dans les organisations administratives, Simon (1947) a mis en évidence l'existence de toute une série de limites qui pèsent sur la rationalité du choix : les données sur lesquelles s'appuient les décisions sont souvent fragmentaires et parfois erronées ; les séquences causales qui relient les actions à leurs conséquences ne sont pas toujours connues de façon parfaite ; enfin, par manque de temps ou par ignorance, seul un petit nombre des actions possibles est examiné. Simon a donc été conduit à distinguer d'une part une rationalité « substantive » ou « parfaite » qui est celle des modèles du choix économique optimal, et d'autre part une rationalité « limitée » (*bounded rationality*). Dans sa théorie, l'organisation a précisément pour but d'aider à surmonter ces limites de la rationalité, d'améliorer la rationalité instrumentale des actions (ce thème sera repris dans les travaux de la nouvelle économie institutionnelle [Williamson, 1975]).

Dans ses recherches ultérieures du côté de la psychologie cognitive, Simon (1981) met l'accent sur les limites des capacités de l'esprit humain à traiter l'information. Face à des problèmes complexes (l'exemple typique est celui du jeu d'échec), l'acteur doit d'appuyer sur des « heuristiques », c'est-à-dire sur des règles d'inférences simplifiées qui lui permettent de parvenir à une solution approximativement correcte, même si ce n'est pas la meilleure possible. La règle optimale est inapplicable parce

que sa mise en œuvre dépasse les capacités de traitement de l'information dont dispose l'acteur. Dès lors, la question qui se pose du point de vue de la rationalité devient celle de choisir une règle d'inférence qui, à défaut d'être optimale, s'avère « satisfaisante ».

À partir du début des années 1970, les psychologues cognitifs s'aperçoivent dans des expériences de laboratoire que les gens emploient intuitivement des règles d'inférence qui sont souvent invalides, même pour résoudre des problèmes relativement simples (Kahneman, Slovic et Tversky, 1982). Les erreurs d'inférence qui ont été ainsi mises en évidence sont tellement systématiques que l'on pourrait estimer que le raisonnement statistique et probabiliste de sens commun est tout bonnement et simplement irrationnel. Boudon (1990) a cependant montré de façon convaincante que l'on pouvait retrouver les « bonnes raisons » qu'avaient les acteurs d'appliquer ces règles d'inférence intuitives, même si ces dernières pouvaient conduire dans certains cas à des résultats objectivement erronés.

La rationalité des croyances

Ce type de rationalité s'applique aux croyances positives au sens large (« positives » s'oppose ici à « normatives ») : croyances de sens commun, croyances magiques, croyances idéologiques, et bien sûr aussi théories et paradigmes scientifiques. La typologie webérienne porte sur les actions et non sur les croyances, et c'est pourquoi ce type de rationalité n'y est pas explicitement pris en compte. On peut cependant considérer qu'il relève de la catégorie de la « rationalité instrumentale » telle que Weber la définit.

Le débat sur la rationalité des croyances magiques court tout au long de l'histoire de la pensée sociologique et anthropologique. Les partisans de la rationalité opèrent un rapprochement entre les croyances magiques primitives et les théories scientifiques modernes. Ils ne confondent certes pas la magie ou la religion primitive avec la science, mais selon eux les croyances magiques jouent dans les sociétés primitives un rôle similaire à celui que jouent les théories scientifiques dans notre propre société : ces croyances ont dans les deux cas pour but d'expliquer et de prévoir les phénomènes naturels ou sociaux, et de permettre d'agir sur ces derniers. Dans les deux cas aussi, ce sont des entités invisibles qui forment la base des explications : les esprits des ancêtres ou de la nature, et, pour nous, les forces, les atomes, les fonctions d'utilité, etc. Cette tradition dite « intellectualiste » en anthropologie est issue de Frazer et Tylor ; elle s'oppose à la tradition « symboliste » selon laquelle les rituels magiques primitifs doivent être plutôt compris comme des formes d'expression artistique : le point de vue symboliste est défendu entre autres par Beattie (1966). Durkheim propose une autre facette

de la tradition intellectualiste lorsqu'il explique comment les catégories de la pensée scientifique et même de la pensée logique ont pu naître de la pensée religieuse et magique primitive. Comme Horton (1973) l'a montré dans une exégèse détaillée, Durkheim s'oppose ici totalement aux thèses a-rationalistes de Lévy-Bruhl sur la pensée primitive comme « mentalité pré-logique ». Weber adopte lui aussi un point de vue rationaliste à la base de sa sociologie de la religion primitive. Horton (1967) propose une version très élaborée de cette tradition intellectualiste. Il montre en particulier comment les systèmes de pensée magique s'articulent à l'expérience du monde social et des phénomènes naturels : selon lui, ces systèmes trouvent leur origine dans une des analogies avec l'organisation sociale existante, puis ils sont peu à peu élaborés et complétés de façon à améliorer leur pouvoir explicatif et prédictif. À ces points de vue, la ressemblance avec l'élaboration des modèles scientifiques lui apparaît tout à fait frappante. Il cite l'exemple du modèle de l'atome de Rutherford, qui a trouvé naissance dans une analogie avec le système solaire, puis s'est peu à peu éloigné de cette analogie initiale pour prendre des caractéristiques spécifiques : ce mécanisme d'« hybridation », comme il le nomme, s'applique aussi bien aux modèles scientifiques qu'à la pensée magique primitive.

Dans le domaine de l'explication des croyances idéologiques, Boudon met en évidence une série d'effets rationnels (liés à l'exercice par les acteurs de leur raison) qui peuvent tendre à faire passer pour vraies des idées qui, en réalité, sont fausses. Ainsi, par exemple, la croyance très répandue selon laquelle « le progrès technique est la source du chômage » illustre les « effets de situation » : le public étant informé des destructions d'emploi dues au progrès technique, mais pas des créations d'emploi liées à la mise en œuvre de ce même progrès, il peut avoir de « bonnes raisons » de croire à l'idée, fautive, selon laquelle la machine produit le chômage. Un mot seulement sur l'explication des croyances scientifiques, pour signaler les travaux importants de Laudan (1977 et 1984), qui tentent de réhabiliter la rationalité de la science en répondant point par point aux arguments anti-rationalistes de Kuhn (1962).

La rationalité axiologique

La définition abstraite de la rationalité axiologique – rationalité des valeurs – proposée par Weber est assez peu satisfaisante. Si on la prend au pied de la lettre, elle se contente de décrire une conformité à des valeurs morales, religieuses ou esthétiques. Or, si l'on souhaite analyser de façon approfondie les actions obéissant à des valeurs, on ne peut pas se limiter à une description pure et simple d'une adhésion inconditionnelle des principes mo-

raux. La définition de Weber est trop étroite, et c'est peut-être pourquoi sa notion de « rationalité axiologique » n'a obtenu par la suite qu'un succès mitigé.

Le problème est en effet de comprendre pourquoi l'acteur adhère aux valeurs qui sont les siennes, d'expliquer cette adhésion aux valeurs (entendues ici au sens de principes moraux, religieux, ou esthétiques). La notion de rationalité peut-elle contribuer à le résoudre ? La réponse est affirmative, et l'on trouve dans l'œuvre même de Weber, dans sa sociologie de la religion en particulier, des analyses qui peuvent nous en convaincre. Il explique par exemple la séduction exercée par le culte de Mithra sur les officiers subalternes (centurions) et les fonctionnaires civils de la Rome antique, par le fait qu'il était organisé, tout comme une administration, sur la base d'une hiérarchie de grades religieux.

Le concept de « rationalité axiologique » a connu récemment un regain d'intérêt à travers l'interprétation proposée par Boudon (1995). Ce dernier a élaboré un modèle qu'il qualifie de « cognitivisme », selon lequel lorsqu'un acteur adhère à une valeur, c'est en général pour une série de raisons. Ce modèle rationaliste s'oppose radicalement aux modèles d'après lesquels l'adhésion aux valeurs s'opérerait selon un mécanisme déterministe a-rationnel. Dans la formulation de base des théories positivistes ou culturalistes, par exemple, les valeurs de groupe déteignent (si l'on peut dire) sur l'individu qui en est membre, sans que la raison de celui-ci ne soit exercée de quelque manière que ce soit. Dans le modèle cognitiviste de la rationalité axiologique, en revanche, l'explication des valeurs – principes moraux, sentiments de justice, goûts esthétiques – passe par la mise en évidence de raisons qui justifient, du point de vue de l'acteur, l'adhésion à ces valeurs. Selon Boudon, ce modèle englobe les théories « conséquentialistes » des valeurs, comme les théories utilitaristes ou les théories fonctionnalistes. Ces théories sont dites « conséquentialistes » parce qu'elles expliquent les jugements favorables ou défavorables portés sur les phénomènes ou les institutions par l'évaluation de leurs conséquences, bonnes ou mauvaises, pour les membres de la société. Suivant les théories fonctionnalistes, par exemple, un phénomène ou une institution est jugé « bon » par un acteur s'il facilite le fonctionnement du système social. Ainsi, les inégalités sociales sont bien acceptées lorsqu'elles sont perçues comme nécessaires au bon fonctionnement du système social. Mais le modèle cognitivisme permet d'aller au-delà de ces théories conséquentialistes et d'expliquer la réprobation morale générale qui s'exprime à l'encontre de phénomènes comme les vols bénins ou le plagiat littéraire. Boudon estime que cette réprobation ne peut pas s'expliquer par les conséquences de ces actes, puisqu'ils n'entraînent pas de désagréments sérieux pour leurs

victimes (dans le cas où un plagiat est découvert, l'auteur plagié peut même bénéficier d'un regain de notoriété). Si ces actes sont réprouvés, c'est pour une « bonne raison » : parce qu'ils violent l'une des règles majeures de la vie en société, à savoir qu'à toute rétribution doit correspondre une contribution. C'est là, selon Boudon (1999), l'explication en termes de rationalité axiologique de la réprobation qu'ils suscitent.

► ABELL P., « Sociological theory and rational choice theory », in B. S. TRUNER (éd.), *Blackwell Companion to Social Theory*, Oxford, Blackwell, 2000. – BEATTIE J., « Ritual and social change », *Man*, 1966, vol. 1, 60-74. – BOUDON R., *L'Art de se persuader*, Paris, Fayard, 1990 ; *Le Juste et le Vrai*, Paris, Fayard, 1995 ; « The "cognitivist model" : A generalized "rational-choice model" », *Rationality and Society*, 1996, 8, 123-150 ; *Le Sens des valeurs*, Paris, PUF, 1999. – COLEMAN J. S., *Foundations of Social Theory*, Cambridge (MA), Harvard Univ. Press, 1990. – DAMASIO A. R., *Descartes' Error. Emotion, Reason and the Human Brain*, New York, Putnam's Sons, 1994. – DENZAU A. T., NORTH D. C., « Shared mental models », *Kyklos* 1994, 47, 3-29. – EARL P., *Lifestyle Economics*, New York, St. Martin's Press, 1986 ; éd., *Behavioural economics*, vol. 1, 2, Aldershot, E. Elgar, 1988. – ESSER H., *Soziologie. Spezielle Grundlagen*, Band 1, Francfort-sur-le-Main, Campus, 2000. – FAVEREAU O., « L'incomplétude n'est pas le problème, c'est la solution », in B. REYNAUD (éd.), *Les Limites de la rationalité*, t. 2, Paris, La Découverte, 1997, 219-234. – FRANK R. H., *Passions within Reason*, New York, Norton, 1988 ; (1992), « Frames of reference and the intertemporal wage profile », in G. F. LOEWENSTEIN, J. ELSTER (éd.), *Choice over Time*, New York, Russell Sage, 1992, 371-382. – FREY B. S., *Not Just for the Money*, Brookfield, E. Elgar, 1997. – FRIEDMAN J. (éd.), *The Rational Choice Controversy*, New Haven, Yale Univ. Press, 1996. – GIGERENZER G., TODD P. M., *Simple Heuristics that Make us Smart*, Oxford, Oxford Univ. Press, 1999. – HARSANYI J. C., *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge, Cambridge Univ. Press, 1977. – HECHTER M., *Principles of Group Solidarity*, Berkeley, Univ. of California Press, 1987. – HECKATHORN D., « Sociological rational choice theory », in G. RITZER, B. SMART (éd.), *Handbook of Social Theory*, Londres, Sage, 2001. – HEDSTROM P., « Rational choice and social structure », in B. WITTRICH (éd.), *Social Theory and Human Agency*, Londres, Sage, 1996. – HORTON R., « African traditional thought and Western science », *Africa*, 1967, vol. 38, 50-71 et 155-187 ; « Lévy-Bruhl, Durkheim and the scientific evolution », in R. FINNEGAN, R. HORTON (éd.), *Modes of Thought*, Londres, Faber and Faber, 1973, 249-305. – KAHNEMAN D., « Objective happiness », in D. KAHNEMAN, E. DIENER, N. SCHWARZ (éd.), *Well-Being : the Foundations of Hedonic Psychology*, New York, Russell Sage, 1999, 3-25. – KAHNEMAN D., SLOVIC P., TVERSKY A., *Judgment under Uncertainty. Heuristics and Biases*, Cambridge, Cambridge Univ. Press, 1982. – LAUDAN L., *Progress and its Problems. Towards a Theory of Scientific Growth*, Berkeley, Univ. of California Press, 1984 ; *Science and Values. The Aims of Science and their Role in Scientific Debate*, Berkeley, Univ. of California Press, 1984. – LINDENBERG S., « Social rationality versus the theory of rational egoists », in J. TURNER (éd.), *Handbook of Sociological Theory*, New York, Plenum, 2001 ; LINDENBERG S., FREY B., « Alternatives, frames, and relative prices », *Acta Sociolo-*

gica, 1993, 36, 191-205. – LOEWENSTEIN G. F., ELSTER J. (éd.), *Choice over Time*, New York, Russell Sage, 1992. – MACY M., « Identity, interest and emergent rationality », *Rationality and Society*, 1997, 9, 427-448. – MISES L. von (1933), *Epistemological Problems of Economics*, New York, New York Univ. Press, 1981. – MONTGOMERY J., OPP K. D., « Contending conceptions of the theory of rational action », *Journal of Theoretical Politics*, 1999, 11, 171-201. – RAUB W., WEESIE J., « The management of matches », *Netherlands' Journal of Social Sciences*, 2000, 36, 71-88. – SCHEFF T. J., « Towards a sociological model of consensus », *ASR*, 1967, 32, 32-46. – SCITOVSKY T., *The Joyless Economy*, Oxford, Oxford Univ. Press, 1976. – SIMON H. A. (1947), *Administrative Behaviour*, New York, The Free Press, 1976 ; *The Sciences of the Artificial*, Cambridge, The MIT Press, 1981. – STINCHCOMBE A., « Reason and rationality », *Sociological Theory*, 1986, 4, 2, 151-166. – THALER R. H., *Quasi Rational Economics*, New York, Russell Sage, 1991. – WILLIAMSON O. E., *Markets and Hierarchies*, New York, Free Press, 1975 ; « Calculativeness, trust, and economic organization », *Journal of Law and Economics*, 1993, 36, 453-486.

Siegwart LINDENBERG (La théorie du choix rationnel),
Renaud FILLIEULE (Types de rationalité)

→ Action ; Coleman ; Économie et sociologie ; Individualisme méthodologique/holisme ; Utilitarisme.

Troisième partie : histoire et méthodes de la sociologie

« Croissance, crise et renouveau » (1996)

Table des matières

Introduction aux sciences sociales

Dirigé par B. Valade

Paris, PUF, p. 445-605



Collection
Premier
Cycle

Introduction aux sciences sociales

BERNARD VALADE

Professeur à la Sorbonne-Paris V

AVEC LA COLLABORATION DE

RENAUD FILLIEULE

Chargé de recherche à l'IHESI

*Presses
Universitaires
de France*



- D'un autre paysage des sciences sociales, 390
 - Mineurs et majeurs, 390
 - La logique sociale de Tarde, 397
 - Simmel et la philosophie des formes sociales, 407
 - La théorie politique de Mosca, 421
- Conclusion, 427

CINQUIÈME PARTIE

CROISSANCE, CRISE ET RENOUVEAU

Introduction, 435

I – Le développement des théories de l'interaction sociale, 445

L'école de Chicago, 446

- La microsociologie de Georg Simmel, 447
- La sociologie urbaine et rurale de Thomas et Znaniecki, 449
- L'écologie urbaine, 453
- Les relations ethniques et la délinquance, 456

Les apports de la psychologie sociale, 459

- L'esprit et la société, 459
- La théorie des opinions, 464
- La dynamique des groupes, 472
- La mémoire collective, 477

L'interactionnisme, 479

- L'interactionnisme symbolique, 480
- La phénoménologie sociale, 484

II – Intérêt et limites des systèmes théoriques, 491

Le fonctionnalisme, 491

- La notion de fonction chez Durkheim, 492
- Le fonctionnalisme en anthropologie, 494
- L'analyse fonctionnelle en sociologie, 497
- Les limites du fonctionnalisme, 502

Le culturalisme, 505

- La diversité des cultures, 506
- Culture et personnalité, 513
- La relativité des cultures, 516
- Les limites du culturalisme, 519

Le structuralisme, 521

- Le structuralisme en linguistique et en anthropologie, 522
- La notion de structure sociale en anthropologie, 529

La notion de structure sociale en sociologie, 533
Les limites du structuralisme, 538

III – Trois questions fondamentales pour les sciences sociales, 541

Actions et croyances des acteurs sociaux : rationnelles ou irrationnelles ?, 542

La notion de rationalité, 542

La notion d'irrationalité, 551

Fécondité de l'hypothèse de rationalité, 556

La méthode des sciences sociales : individualiste ou holiste ?, 567

L'individualisme méthodologique, 567

Le holisme méthodologique et ses limites, 580

De quelques malentendus concernant l'individualisme méthodologique, 588

La connaissance du social : objective ou relative ?, 591

L'objectivité des sciences sociales, 591

Les critiques contre la thèse de l'objectivité, 599

Conclusion, 605

Bibliographie, 613

Index, 625

« System : Social » (2001)

International encyclopedia of the social and behavioral sciences

Oxford, Elsevier, p. 15 418-15 423

System: Social

A system is a set of differentiated and interdependent components. With such an abstract definition, the notion of system explicitly or implicitly underlies all or most of the theories of the social sciences. It is therefore useful to distinguish three ways of conceptualizing social systems, each one corresponding to a major paradigm. Firstly, the *historicist* paradigm, according to which Society constitutes a global system subject to a necessary law of historical evolution; this paradigm was mainly represented by Comte and Marx in the nineteenth century. Secondly, the *structural-functional* paradigm, according to which a society is a system whose components are the social institutions (the family, religion, etc.); this approach was developed in anthropology by Radcliffe-Brown and Malinowski in the 1930's, and then imported into sociology by Merton and Parsons. And thirdly, the *actionist* paradigm, in which the social systems are constituted by actors and their interactions; one of the main sources of this paradigm is the subjectivist and marginalist revolution that happened in the 1870's in economics.

1. The Historicist Conception of Social Systems

According to the historicist conception, Society is destined to go through a necessary succession of great historical stages, each of which is a social system. The aim of sociology, then, is to determine the nature of these systems, and especially the historical "law" governing their succession. Thanks to this historical "law", it becomes possible to know the nature of the social system that will appear in the future, the nature of the next great historical stage. What is the use of knowing in advance an inescapable future? Comte and Marx have the same answer: it gives us the power to facilitate and accelerate the advent of the ideal future society, since according to them the future social system will be the ultimate stage in the progress of humanity.

1.1 Comte and Marx

In his *positivist* theory, Comte (1822) analyses the past and future evolution of society as the succession of three great social systems. The *theological* system, first, was represented by the European absolute monarchies: social phenomena were then interpreted in the frame of the Christian religion, the spiritual power was held by the priests and the temporal power by the aristocracy (the sword). Then came the *metaphysical* system ushered in by the French Revolution: social phenomena were interpreted in the frame of the philosophy of the Enlightenment, the spiritual power gave way to freedom of conscience and the temporal power came into the hands of the people (democracy). According to Comte, this metaphysical system is only transitional. It has a destructive nature and is thus unable to constitute a lasting social order. It will be replaced in the future by the *positive* system, in which sociology will become a "positive" science (founded, just like modern physics and chemistry, upon observation and invariable natural laws), and in which the spiritual power will be held by the scientists and the temporal power by the producers (the manufacturers). Each social system is thus defined by its "civilization" (the state of our understanding of social phenomena) and by its "social organization" (as defined by who holds the spiritual and of the temporal powers). Social organization is here supposed to be determined by the state of civilization, and the latter is determined by scientific progress. Comte believed that this progress obeyed the *law of the three stages*: he thought that sociology, like the other positive sciences, would appear at the end of a three-stage process ("theological", "metaphysical", and "positive") through which knowledge is progressively liberated from its animistic and anthropomorphic bias. In this way, the components of each

social system and social evolution itself are determined by the progress of knowledge, and this progress is in turn determined by the law of the three stages.

Marx's historicism (1859), or *dialectical materialism*, describes a succession of four "modes of production" (or social systems), antique, feudal, capitalistic, and socialistic. Each one of these systems is made up of two major components: the "material productive forces" (techniques, tools, machinery) and the "production relations" (property laws). The productive forces are supposed to determine the production relations, and also all the other ideological features of society (politics, religion, esthetics, philosophy). Marx illustrated this idea with a famous sentence: the hand mill gives you feudal society, the steam mill industrial capitalism (*Misère de la philosophie*, 1847). In a watered-down version of his historicism, Marx (1859, Preface) explains that the succession of these systems is a historical necessity arising from the development of the productive forces (i.e., of the technical capacities of production). Technical and economic progress end up being incompatible with existing production relations. When this contradiction reaches its climax, the production relations are unshackled, and a great social revolution takes place, marking the transition between two modes of production. In *Das Kapital* (vol. 1, 1867), Marx attempts to study in a much more detailed way, with an actionist methodology, the necessary future succession between capitalism and socialism (about actionism, see Sect. 3 below).

The historicist view was widely held in the nineteenth century. It is to be found in Tocqueville's analysis of the secular transition between aristocratic and democratic society (*De la Démocratie en Amérique*, vol. 1, 1835). Durkheim, greatly influenced by Comte, attempts to explain in *De la Division du Travail Social* (1893) the necessary transition between the primitive societies, where social cohesion is "mechanical" (cohesion comes from the similarity between individuals), and modern society, where social cohesion is "organic" (cohesion arises from the differences between and the complementarity of individuals). In Spencer's historicism (*The Principles of Sociology*, 1891), the historical law of succession between the primitive and modern societies is a law of evolution from a simple to a complex organization, based in part on the biological principles of Darwinism.

1.2 Limits and Errors of Historicism

It is easy today to perceive the limits of historicist theories. From a static point of view, social systems are represented in an oversimplified way, and their internal causal relations are debatable (the alleged primacy of "civilization" over "social organization" in Comte, and of "productive forces" over "production relations" in Marx). But these theories are especially questionable from a dynamic point of view: regardless of the obvious errors of prediction of a Comte or a Marx, the historicist project may be conclusively refuted with the help of an epistemological argument. Historicist "laws" are supposed to be the equivalent in the field of social phenomena of physical laws in the field of natural phenomena. But physical laws (for instance, the law of falling bodies) explain *reproducible* phenomena, i.e., phenomena that may be repeatedly observed because they recur a great number of times. Historicist "laws", on the other hand, try to explain a *unique* phenomenon, human history. This phenomenon, which happens only once and is definitely not reproducible, cannot be explained by a scientific law in the sense of the natural sciences (Popper 1957). In other words, the claim of historicism to possess a scientific character is completely unfounded. Historicist authors are unable to establish, in a convincing way, the causal relations that are supposed to determine the necessary succession between the great social systems or historical stages (for an analysis of other deep contradictions in the positivism of Comte see Hayek 1952, and in the dialectical materialism of Marx see Mises 1957).

2. The Structural-Functional Conception of Social Systems

In the structural-functional paradigm, social institutions are the components of society considered as a system. On the one hand, these institutions play a role in the functioning of society (*functionalism*), and on the other hand are held together by relationships of complementarity that constitute a structure (*structuralism*). The historicist sociologists of the nineteenth century were immersed in a vast process of sociological, economic, and political change. This process seemed to be explainable in terms of a necessary succession of great social systems. The structural-functional paradigm was born in a totally different intellectual context. It originated from the anthropological study of primitive societies, whose long-term history was unknown. As a consequence, this paradigm was completely severed from the most questionable aspect of historicism, namely the search for a global historical “law” of succession between great social systems.

2.1 Society as a Functional System

As formulated by anthropologist Radcliffe-Brown (1935), the basic idea of functionalism is that in a given society, every institution, every social process, every cultural form possesses a “function”, i.e., it plays a definite and even indispensable role in the survival and perpetuation of this system. This reasoning is borrowed from biology, where each organ plays a necessary role in the survival of the organism. Faced with a social institution, whatever it may be, the social scientist is supposed to establish its function from the standpoint of the global system constituted by society. For instance, Radcliffe-Brown explains the existence of ancestor worship in traditional societies by the function of strengthening solidarity: thanks to these rites, individuals learn that they can count on the magical help of their forebears, and also that they must submit to the traditional rules. A narrower conception of functionalism was offered by Malinowski (1944). According to him, the economic, educational and political institutions of a society have the function of satisfying—directly or indirectly—the “primary” needs related to the biological survival of individuals (food, medical care, reproduction). These primary needs give birth to “derivative” needs. Satisfying the need for food, for instance, implies a cooperation between the members of society. Since this cooperation requires rules of conduct, these rules themselves become a need, but an “indirect” one in the sense that it is derived from a primary need.

Functionalism was then imported into the sociology of complex societies by Merton (1949). His *functional analysis* is more qualified than the basic or “absolute” functionalism developed by the anthropologists of traditional societies. Merton begins by distinguishing the functions that are “manifest” (i.e., deliberately sought for or designed by the actors taking part in the system) and the functions that are “latent” (i.e., those that play a role unknown to the participants in the system). The function of the cult of the ancestors according to Radcliffe-Brown is an example of a latent function. Merton then observes that in modern societies some institutions can be functional for a given group, and at the same time dysfunctional for another group (a technical advance, for instance, can increase the average productivity of labor, but also create a redundancy in some branches of production and compel some workers to search for a new job). A modern society, considered as a system, shares none of the “functional unity” of biological organisms. Furthermore, some institutions or customs may not have any function. Finally, Merton criticizes the idea that there exist a certain number of functions that are necessary to the survival of any society, and that the same specific institutions would be indispensable to perform these functions across all societies.

2.2 *The Structure of Social Systems*

The conception of society as a functional system of institutions leads quite naturally to an analysis of *social structure*, that is, to an analysis of the complementarity of these institutions. A famous illustration is given by Parsons (1951). He observes that a highly industrialized society requires actors to be socially and geographically mobile. But this mobility is possible only if the actors are not too closely connected to their parents: there must be a clear-cut separation between an actor's "orientation" family (where he or she was born) and the "procreation" family (the one created by his or her marriage). If we follow this reasoning to its logical conclusion, industrialization is seen to be incompatible with powerful and extensive traditional family ties. Those two components—industrialization on the one hand and the separation between orientation and procreation families on the other—constitute a consistent social structure typical of modern societies. Conversely, the absence of industrialization and extensive and powerful family ties are the components of another stable social structure, typical of traditional societies. This example shows how the existence of one institution can imply the existence of another institution, their combination representing a stable and coherent social structure.

In anthropology, with Lévi-Strauss (1958), the study of social structure became a *structuralism* largely based upon linguistics. Lévi-Strauss realized that in traditional societies kinship relations (brother/sister), (husband/wife), (uncle/nephew), (father/son) constitute a system characterized by a law of opposition. If, in a given society, the relations between a brother and his sister are generally "good" (friendly), then (husband/wife) relations are "bad" (reserved or hostile); in the same way, if (uncle/nephew) relations are "good", then (father/son) relations are "bad", or vice versa. The system of kinship (brother/sister)(husband/wife)/(uncle/nephew)(father/son) can thus exist in only four structures in the different societies: "+ -/+ -", "+ -/- +", "- +/+ -", or "- +/- +." The first of these structures, "+ -/+ -", means that (brother/sister) relations are good, as symbolized by the "+" sign, (husband/wife) relations are bad, "-" sign, (uncle/nephew) relations are good, and (father/son) relations are bad. This law of opposition, in which relations take precedence over the components they relate, is perfectly similar to the law that had been previously discovered by the linguists N. S. Troubetzkoy and R. Jakobson in the area of structural phonology (these authors showed that language as a communication system rests on a law of opposition between the phonemes).

2.3 *The Explanatory Limits of Structural-Functionalism*

This paradigm has mainly been criticized for its lack of explanatory power. In principle, the "absolute" functionalism of a Radcliffe-Brown or a Malinowski permits us to explain the existence of an institution by the role it plays in the working and thus in the existence of a given society. But the more qualified conception of functionalism to be found in Merton does not enable us to explain the existence of an institution by its latent function, since this institution or function is not considered as a functional requisite of the society anymore. The existence of an institution, then, can only be explained with the help of a "genetic" analysis consisting in the study of the historical origin of this institution. Two important criticisms have been aimed at structural-functionalism from an actionist viewpoint by Homans (1987): firstly, this paradigm takes for granted the actors' obedience to social rules, and does not explain why the actors choose to follow these rules; secondly, the existence of social structures is also taken for granted, and no explanation is offered as to how these structures emerge from social interaction.

3. The Actionist Conception of Social Systems

The actionist paradigm aims at discovering the causes of social phenomena in the *actions* of individuals, and more precisely in the combination of these actions (see *Action, Theories of Social, and Methodological Individualism in Sociology*). Social systems are defined as systems of actions or of interactions, from the smallest (two actors) to the largest ones (for instance, the international division of labor). In this paradigm, individual actions are generally considered as rational. This rationality may be utilitarian (appropriateness of means to ends), cognitive (rationality of positive beliefs, i.e., of factual propositions), or axiological (rationality of normative beliefs). The focus of attention, here, is neither on the global evolution of society (historicism), nor on social institutions seen from the standpoint of their functional role and of their mutual relationships (structural-functionalism), but on human action: social phenomena are considered as emergent and possibly unintentional effects whose causes are actions constituting a system. The actionist paradigm is widely used in economics as well as in sociology. It is quite old since it implicitly underlies, for instance, the works of a classical economist like A. Smith (*An Inquiry into the Nature and Causes of the Wealth of Nations*, 1776). Its importance was reinforced in the 1870's by the marginalist and subjectivist revolution that took place in economics. This revolution consisted in a much more rigorous analysis of human choices and actions in an environment characterized by the scarcity of goods, and made possible the development of really satisfying theories of prices and interest rates.

3.1 The Systems of Economic Exchange

The most simple system of exchange is composed of two actors, either of which possesses a good and decides or refuses to exchange it against the other good, according to his or her subjective preferences. For an exchange to take place, the good supplied must have less subjective value than the good demanded (and this is true of each of the two actors). This system can be generalized to a market comprising many sellers and buyers trading one good for another. It may then be generalized to a whole economic system made up of a great numbers of markets. In the tradition originating in L. Walras (*Éléments d'Économie Politique Pure*, 1874), the representation of a global economic system rests on the concept of *general equilibrium*. In a situation of general equilibrium, all the actions of all the actors on all the markets are compatible with one another: they can all be carried out without any actor feeling any regret or disappointment. Thanks to a mathematical formalization of the subjective preferences of the consumers and of the possibilities of production of the producers, the existence of a general equilibrium can be logically demonstrated: on some conditions, there exists a set of prices of consumers goods and of factors of production, such that all the actions of buying and selling are compatible (Arrow and Hahn 1971). This theory is rather complex from a mathematical point of view, and it suffers from two main defects: it does not demonstrate that the system converges towards a general equilibrium, and it does not take into account the radical uncertainty of the future.

In another school of thought, the Austrian school (Mises 1949), the focus is on *market processes*. Economic analysis, here, focuses on the consequences in the economic system of a change in one of its elements. If, for instance, the consumers change their preferences and decide to spend more than before on good *A* and less on good *B*, the demand for *A* rises and its price tends to rise. The rate of profit also tends to rise in the branch of production of good *A* (since the price of this product rises when its costs of production remain the same). If this rate of profit becomes higher than the current rate of profit, more resources will be invested in the production of good *A*. Two consequences follow: firstly, the demand for the factors of production of *A* will rise and thus the prices of these factors will tend to rise, and secondly,

the supply of *A* will increase and thus its price will tend to diminish. In this way, the rate of profit in the branch of production of *A* is brought down to the current level (since the price diminishes and the costs rise). Conversely, as the demand for *B* and its price are lowered, the rate of profit falls below the current level, and a disinvestment occurs. Demand for the factors of production of *B* decreases, resulting in lower prices. As a consequence, the supply of *B* tends to drop and its price to rise. As a consequence, the rate of profit in the production of *B* is brought back to the current level (the price rises and the costs diminish). A first adaptation of the economic system to the change of demand is thus achieved: the branch of production of *A* expands and the branch of *B* contracts. But the changes are not restricted to these two branches of production. To the extent that the demand for the factors of production of *A* or *B* (and for the factors of production of these factors) is also changed, the changes diffuse throughout the system and a new balance between the various branches of industry is established. It is possible to study in the same way the diffusion of the effects of a technical advance, or of a variation in the supply of a natural resource, in the quantity of money, in savings, and so on.

3.2 *The Actionist Systems in Sociology*

The actionist tradition in sociology was born in Germany in the end of the nineteenth century. Weber (1956), for instance, developed a “comprehensive” method, which consists in trying to understand actions from the point of view of the actors themselves. This method constitutes a generalization of the analysis of human choices derived from the subjectivist revolution of the 1870’s in economics. The German sociological tradition had a major influence on sociology in the twentieth century, for instance through the “interactionist” approach (see *Interactionism, Symbolic*). The use of actionist systems enables us to explain in a convincing way a number of important and puzzling phenomena. Boudon (1973) has thus explained “Anderson’s paradox” in the field of sociology of education. From the 1960’s on, an important democratization of schooling took place, as the proportion of children of low socio-economic status increased in the colleges and universities. Since future social status is strongly related to academic achievement, vertical social mobility was expected to increase in developed countries. But, and this is the paradox, it didn’t. In spite of a reduction of this inequality, social mobility was not greater than before. Boudon showed that this paradox could be explained thanks to an actionist system based on the motivations and choices of individuals. Many other examples could be taken from the various fields of sociology (see Boudon 1979 for a survey).

4. The General System Theory

In the 1940’s, Bertalanffy (1968) began to develop a “general system theory”, aiming at transcending the frontiers between a wide range of disciplines—physics, chemistry, biology, psychology, and the social sciences. His basic idea was that the systems analyzed in these different branches of learning share a number of features that can and should be the subject of a science of systems as such. But if it is admitted that the aim of science is to find satisfactory explanations of important phenomena, then the analogies that can be discovered between the systems originating from different branches of science do not matter very much. These analogies can constitute genuine scientific contributions only if they suggest new ways of explaining phenomena, or if they improve the existing explanations. Structural anthropology (Lévi-Strauss 1958) is a good example of a set of theories that originated from an abstract theory of the structure of systems (see Sect. 2 above), but it is an exception. Most of the time, the study of social systems can dispense with a general science of systems (obviously, the opposite is not true).

Bibliography

- Arrow K J, Hahn F H 1971 *General Competitive Analysis*. North-Holland, Amsterdam
- Bertalanffy L von 1968 *General System Theory: Foundation, Development, Applications*. George Braziller, New York, NY
- Boudon R 1973 *L'Inégalité des Chances. La Mobilité Sociale dans les Sociétés Industrielles*. Armand Colin, Paris [1974 *Education, Opportunity, and Social Inequality*. Wiley, New York, NY]
- Boudon R 1979 *La Logique du Social. Introduction à l'Analyse Sociologique*. Hachette, Paris [1981 *The Logic of Social Action: An Introduction to Sociological Analysis*. Routledge and Kegan Paul, Boston]
- Comte A 1822 *Plan des Travaux Nécessaires pour Réorganiser la Société Européenne*
- Hayek F A von 1952 *The Counter-Revolution of Science: Studies on the Abuse of Reason*. The Free Press, Glencoe, IL
- Homans G C 1987 *Certainties and Doubts: Collected Papers, 1962-1985*. Transaction Books, New Brunswick, NJ
- Lévi-Strauss C 1958 *Anthropologie Structurale*. Plon, Paris [1963 *Structural Anthropology*. Basic Books, New York, NY]
- Malinowski B 1944 *A Scientific Theory of Culture and Other Essays*. The University of North Carolina Press, Chapel Hill
- Merton R K 1949 Manifest and Latent Functions. In: Merton R K 1968 *Social Theory and Social Structure*, enlarged edn. The Free Press, New York, NY
- Marx K 1859 *Zur Kritik der politischen Oekonomie*. Duncker, Berlin [1904 *A Contribution to the Critique of Political Economy*. Kerr, Chicago, IL]
- Mises L von 1949 *Human Action: A Treatise on Economics*. Yale University Press, New Haven, CT
- Mises L von 1957 *Theory and History: An Interpretation of Social and Economic Evolution*. Yale University Press, New Haven, CT
- Parsons T 1951 *The Social System*. The Free Press, New York, NY
- Popper K R 1957 *The Poverty of Historicism*. Routledge and Kegan Paul, London
- Radcliffe-Brown A R 1935 On the Concept of Function in Social Science. In: Radcliffe-Brown A R 1952 *Structure and Function in Primitive Society*. The Free Press, Glencoe, IL
- Weber M 1956 *Wirtschaft und Gesellschaft. Grundriss der verstehenden Soziologie*, 4th edn. Mohr, Tübingen [1978 *Economy and Society: An Outline of Interpretive Sociology*. University of California Press, Berkeley and Los Angeles, CA]

R. Fillieule

« Neutrality : Axiological » (2001)

International encyclopedia of the social and behavioral sciences

Oxford, Elsevier, p. 10 621-10 623

Neutrality, Axiological

“Axiological neutrality”—or “ethical neutrality”—is an imperative imposed on social scientists. It requires of them, when they make policy prescriptions or recommendations, to distinguish clearly between: on the one hand the purely scientific content of their presentation, i.e., the facts they are seeking to explain and the theoretical explanations themselves, and on the other hand the ethical value judgments necessary to reach these prescriptions. This distinction cannot be avoided, since social science theories as such are totally unable to express ethical value judgments—judgments about what social states are good or bad, just or unjust, desirable or undesirable. The axiological neutrality imperative would be transgressed by a social scientist who pretended to “prove scientifically” the desirability or undesirability of a state of affairs. Such a “proof” would necessarily rest upon the concealment—involuntary or not—of implicit ethical value judgments.

1. The Logical Foundation of Axiological Neutrality

The source of axiological neutrality is to be found in a double logical impossibility:

- (a) the impossibility to deduce a value judgment from factual propositions,
- (b) the impossibility to deduce a factual proposition from value judgments.

A factual proposition is true if it corresponds—and false if it does not correspond—to the empirical phenomenon that it is supposed to describe. A value judgment (“Social state *A* is better than social state *B*”), on the other hand, is a *preference* between social facts, and it does not “correspond” to any empirical phenomenon: it cannot be said “true” or “false.” Since a proposition logically derived from factual propositions is also a factual proposition (true or false), it cannot be a value judgment. Conversely, a proposition logically derived from a value judgment is also a value judgment (neither true nor false), and thus cannot be a factual proposition (Ayer 1936).

The consequence of this dichotomy between facts and values is that social science theories *alone* are unable to express value judgments. These theories are supposed to provide causal (deductive) explanations of social facts. They exclusively consist of factual and logical propositions: their premises cannot contain any value judgment, because it would prevent them from explaining empirical phenomena (impossibility (b) above). And it is therefore impossible to deduce from these theories any value judgment (impossibility (a)). From this point of view, social science is not different from any other empirical science. In and of itself, a theory seeking to explain the origin of an inequality cannot yield a value judgment approving of or condemning morally this inequality. For the same reason, a physical theory explaining the solidity of a building cannot yield a value judgment in favor of or against this solidity.

In order to prevent a possible misunderstanding, it must be observed that the factual propositions making up the theories of the social sciences may be propositions *describing* the moral preferences, the esthetic judgments, or the tastes of the agents of the system under study. But when the social scientist describes or explains these preferences, they are considered as mere facts (the fact that an agent has a given preference). It may of course happen that he or she approves of or condemns these preferences, but this value judgment plays no role whatsoever in the explanatory (deductive) theories themselves.

As a consequence of the double impossibility above, social scientists cannot prescribe or recommend a policy—unless they *combine* the scientific facts and theories with an ethical value judgment specifying the end(s) to be achieved. Given an end *E*, scientific theories can in principle determine the means *M* that can be used to reach it (or they can tell us whether this end is impossible to reach). Actually, the proposition “If *M*, then *E*”—“If the means *M* are used, then the end *E* will be reached”—is a factual proposition: it can in principle be said to be

“true” or “false.” On the other hand, propositions such as “It is morally desirable to attain end *E*” and “It is morally desirable to use means *M*” are value judgments: they are not factual propositions, and are neither true nor false. In other words, a social scientist can use scientific knowledge of cause and effect in order to determine the means to an end, but not in order to advocate this end or these means from a moral point of view (Simon 1976). It goes without saying that if the theories are wrong or the relevant facts not fully known, the means in question will (generally) not actually permit one to reach the end.

In the 1940’s, the founders of the “New Welfare Economics” (economists such as John Hicks and Paul Samuelson) believed that they could overcome the dichotomy between facts and values by resorting to the unanimity principle, that is, by using only value judgments accepted by all the members of the society (“Pareto-optimality” seemed to be such a unanimous value judgment). Against this view, a number of objections have been raised. First, even if a value judgment is unanimous, it remains a value judgment and can never be ethically neutral (Sen 1970). Second, really unanimous (and thus non-conflicting) value judgments are probably less numerous than economists would like to believe. Last, such value judgments are morally questionable insofar as they tend to promote the status quo (Rothbard 1976).

2. Axiological Neutrality as a Moral Imperative

This dichotomy between facts and values implies that social scientists have a responsibility towards their audiences, be it their scientific colleagues, the general public (which they may reach through the media), or more especially the “captive” audience of their students. The imperative of axiological neutrality commands social scientists not to make their audience believe that science alone—as a system of factual and logical propositions—can demonstrate the value of their own moral or political convictions. The social scientist, in other words, must not indulge in moral teaching while pretending to be engaged in pure empirical science. In his classic text about axiological neutrality, Weber (1917) emphasized that the professors of social sciences had to respect this imperative, first by being aware of it themselves, second by making their audience and especially their students aware of the ethical value judgments founding their prescriptions. When social science theories are put into practice, value judgments are necessarily required. Such judgments are not scientific (empirical) propositions, and their existence and content should be explicitly stated—especially when these propositions may elicit interpersonal disagreement.

Weber went very far in the defense of neutrality since he recommended that professors confine themselves to purely scientific presentations, and if possible totally refrain from expressing their ethical preferences in front of their students. But on this point, his conception of axiological neutrality was too demanding. For if empirical theories are devoid of value judgments, the latter cannot be excluded from scientific *activity*. A social scientist cannot do science—try and discover the true explanations of phenomena considered as important—without first laying down value judgments. To decide to do science is to express a preference and thus to follow a value judgment. The selection of phenomena considered as important and deserving of a scientific explanation also rests on a value judgment. Social scientists may thus be guided by their ethical or political convictions in the choice of their subject matter: poverty, unemployment, violence, ethnic relations, etc. This is in part what makes the imperative of axiological neutrality necessary, since a clear distinction must therefore be made between value judgments and scientific, factual propositions. Axiological neutrality does not mean that scientific activity can get rid of all value judgments. It is obvious that if scientists had to remain perfectly neutral towards value judgments, they would not be able to choose an interesting subject matter, or to prefer truth to error. Science would simply become impossible (Gouldner 1963, Homans 1978).

Axiological neutrality is a moral duty inherent in intellectual honesty. Transgressions of this imperative may of course originate in dishonesty, but also in honest error. Here is a famous example of an involuntary confusion between facts and values chosen from the field of economics. After the law of marginal utility was discovered in the 1870's, some of the greatest economists at the time (such as Edgeworth and von Wieser) thought they could scientifically justify an egalitarian policy. The law of marginal utility, from which contemporary price theory originates, shows that when the stock of a good possessed by an agent grows, the subjective utility (satisfaction) for this agent of each additional unit of the good diminishes. The first unit is used to satisfy the most urgent need, the second one a less urgent need, the third one a still less urgent need, and so on: each additional unit is allocated to the satisfaction of less and less urgent needs. From this law is deduced the fact that, all other things being equal, the greater the quantity of money (monetary income) possessed by an agent, the smaller the subjective utility for this agent of an additional unit of money. Conversely, the smaller the monetary income of an agent, the greater the subjective utility of an additional unit of money. It would then seem as if the following reasoning could be made on purely empirical grounds: If the State decides to reduce the inequalities of income by a monetary transfer from the rich to the poor, and if this transfer does not reduce production (i. e. does not have any adverse effect on work and investment), then the total utility of society will grow, since the utility lost by the rich will be less than that gained by the poor. Robbins (1935) has shown that this deduction is not scientific: the law of marginal utility can only be applied to *one* agent, and does not enable us to make comparisons of utility *between* persons. It is thus impossible to deduce from this law the proposition that the utility of the income taken from the rich is smaller than the utility of the income given to the poor. This proposition is an ethical value judgment—a preference given to what the poor gain over what the rich lose—that has been confused with a factual proposition. In order to deduce the egalitarian policy, both this ethical value judgment and the (factual) law of marginal utility are needed.

3. An Objection Against Axiological Neutrality

Strauss (1953) has raised an important objection against the weberian imperative of axiological neutrality: if value judgments are not scientific (see impossibility (a) in Sect. 1 above), then they are not based on reason and are thus *arbitrary*; but if value judgments are arbitrary, why should they be obeyed? Why should social scientists respect the moral imperative of axiological neutrality? Furthermore, if values are arbitrary then scientific activity is also arbitrary since it requires a series of value judgments (see Sect. 2). In other words, if the logical foundation of axiological neutrality is accepted, then the rational basis for the formulation of value judgments is undermined, and scientific activity becomes utterly arbitrary. The ideal of neutrality is supposed to protect science from the unwarranted interference of value judgments, but it results in the destruction of science since it implies that the selection of scientific problems and the practical utilization of science are purely arbitrary. According to Strauss, in order to avoid these nihilistic consequences the foundation of axiological neutrality must be questioned, and it must be admitted that a science of morals, a science of value judgments is possible.

Very briefly, a reply to this argument of Strauss is that the impossibility scientifically (empirically) to derive any value judgment does not imply that these judgments are necessarily irrational. The notion of rationality is not limited to the empirical domain. There also exists a rationality of norms, of values, an *axiological rationality* in the sense of Boudon (1995): a value can be justified by reasons, by arguments whose soundness and validity can be rationally discussed. Just because value judgments cannot be established on a purely empirical basis does not mean that they are arbitrary. It is true that among the advocates of axiological neu-

trality are authors believing in the pure irrationality and subjectivity of values (like Weber himself), but there are also among them authors defending the rationality of values (Boudon 1995, Rothbard 1976).

Bibliography

- Ayer A J 1936 *Language, Truth, and Logic*. Victor Gollancz, London
- Boudon R 1995 *Le Juste et le Vrai: Études sur l'Objectivité des Valeurs et de la Connaissance*. Fayard, Paris
- Gouldner A W 1963 Anti-Minotaur: The Myth of a Value-Free Sociology. *Social Problems* 9: 199-231
- Homans G C 1978 What Kind of a Myth is the Myth of a Value-Free Social Science? In: Homans G C 1987 *Certainties and Doubts: Collected Papers, 1962-1985*. Transaction Books, New Brunswick, NJ
- Robbins L 1935 *An Essay on the Nature and Significance of Economic Science*, 2nd edn. Macmillan, London
- Rothbard M N 1976 Praxeology, Value Judgments, and Public Policy. In: Rothbard M N 1997 *The Logic of Action, I: Method, Money, and the Austrian School*. Edward Elgar, Cheltenham
- Sen A K 1970 *Collective Choice and Social Welfare*. North-Holland, Amsterdam
- Simon H A 1976 *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations*, 3rd edn. The Free Press, New York, NY
- Strauss L 1953 *Natural Right and History*. The University of Chicago Press, Chicago, IL
- Weber M 1917 The Meaning of "Ethical Neutrality" in Sociology and Economics. In: Schils E A and Finch H A (trans. and eds.) 1949 *Max Weber on the Methodology of the Social Sciences*. The Free Press, Glencoe, IL

R. Fillieule

Les méthodes en sociologie (2002)

Table des matières

Dirigé par R. Boudon

Paris, PUF, coll. « Que sais-je ? »

QUE SAIS-JE ?

Les méthodes en sociologie

RAYMOND BOUDON

Professeur à l'Université de Paris-Sorbonne
Membre de l'Institut

RENAUD FILLIEULE

Maître de conférences à l'Université de Lille I

Douzième édition mise à jour

103^e mille

TABLE DES MATIÈRES

Introduction	3
Chapitre I – L'analyse causale	5
I. La formulation des hypothèses, 6 – II. La construction du plan d'observation, 9 – III. La construction des variables, 19 – IV. L'analyse des relations entre variables, 28.	
Chapitre II – L'individualisme méthodologique	41
I. Le dépassement de l'analyse causale, 41 – II. Les principes de l'individualisme méthodologique, 53 – III. Diversité des applications, 76.	
Chapitre III – La rationalité cognitive	91
I. L'explication des croyances positives, 92 – II. L'explication des croyances normatives, 111.	
Conclusion	121
Bibliographie	123



« Individualisme méthodologique et holisme » (2005)

Dictionnaire de la pensée sociologique

Dirigé par M. Borlandi, R. Boudon, M. Cherkaoui et B. Valade

Paris, PUF, p. 351-354

INDIVIDUALISME MÉTHODOLOGIQUE/ HOLISME

Les sciences sociales doivent-elles adopter une méthode « individualiste » ou « holiste » (*holon* signifie « le tout » en grec) ? Dans le premier cas, les phénomènes sociaux sont censés s'expliquer à partir des actions individuelles qui les composent. Dans le second, les phénomènes sociaux sont considérés comme des « touts » irréductibles aux individus. Historiquement, la méthode « individualiste » a été formulée à la fin du XIX^e s. par l'École sociologique allemande représentée par M. Weber et G. Simmel, même si son application est beaucoup

plus ancienne et remonte au moins au XVIII^e s. La méthode « holiste » se rattache à l'école française de sociologie fondée par A. Comte puis É. Durkheim. Elle est d'application plus récente et fait partie du mouvement « positiviste » qui, à partir du XIX^e s., a touché l'ensemble des sciences et pas seulement la sociologie.

Les principes de l'individualisme méthodologique

La méthode « individualiste » consiste à expliquer les phénomènes sociaux en montrant qu'ils sont les résultats d'une *agrégation* d'actions « compréhensibles ». Soit la très forte augmentation du nombre des cambriolages en France ou aux États-Unis dans les décennies 1960 et 1970 : elle s'explique par la multiplication des passages à l'acte individuels (Cohen et Felson, 1979) ; si ces derniers se sont multipliés, c'est parce que leurs cibles potentielles sont devenues de plus en plus attractives et de moins en moins bien gardées avec le temps : l'équipement des ménages en matériels divers s'est nettement accru au cours de cette période, mais il s'est trouvé de plus en plus souvent sans surveillance du fait notamment de l'extension du travail féminin. On « comprend » donc que des individus de plus en plus nombreux aient décidé de commettre des cambriolages à des fréquences de plus en plus élevées.

Tous les phénomènes sociaux peuvent en principe être expliqués de cette manière (Boudon, 1984, 1986) : les événements singuliers, les régularités (corrélations statistiques, tendances), les croyances collectives (la diffusion des idéologies, etc.). La méthode s'applique tout autant à l'étude du fonctionnement des systèmes sociaux qu'à celle des changements sociaux.

Les deux concepts fondamentaux de l'individualisme méthodologique sont l'action et l'agrégation.

L'action. – Les systèmes sociaux étudiés se composent d'individus qui *agissent*. Leurs actions sont supposées « compréhensibles » au sens webérien. En essayant de se mettre « à la place » de l'acteur, le sociologue tente de comprendre les raisons que l'acteur se donne pour agir de telle façon ou pour adhérer à telle croyance. Pour éviter que son interprétation des actions ne tombe dans l'arbitraire, il doit s'informer sur le contexte de décision des acteurs. De telles actions sont *rationnelles* au sens où elles s'appuient du point de vue de l'acteur sur des raisons. Deux significations du terme « rationalité » peuvent être distinguées. La rationalité *instrumentale* désigne l'adéquation des fins visées aux moyens disponibles. Dans sa version la plus étroite, cette rationalité est celle de l'*homo œconomicus*. Mais elle peut être élargie à des fins altruistes et à des moyens non pécuniaires. La rationalité *cognitive* vise à expliquer par de bonnes raisons – « bon-

nes » du point de vue de l'acteur mais pas nécessairement de l'observateur – l'adhésion aux croyances de toutes sortes (croyances de sens commun, croyances scientifiques, croyances morales, idéologies, etc.). On a objecté que les gens agissent sous l'effet de forces irrationnelles. C'est une possibilité, mais qu'il ne faut se résoudre à adopter qu'en dernier recours, lorsque tous les efforts d'interprétation par de « bonnes raisons » ont échoué.

L'agrégation. – La combinaison des actions individuelles peut donner naissance à plusieurs types d'effets d'agrégation (Boudon, 1979) : des effets additifs simples, des réactions en chaîne, des effets cumulatifs (avec causalité circulaire), des effets de reproduction, des effets de contradiction, cette énumération n'étant pas exhaustive. L'exemple de l'augmentation des cambriolages illustre l'effet d'agrégation élémentaire qui consiste en une simple *addition* des actions individuelles. Le progrès technique peut déclencher des effets de *réaction en chaîne*. L. White (1962) a montré comment des innovations telles que la charrue à soc métallique ou le mors du cheval ont entraîné, par le biais de multiples successions de causes et d'effets, des mutations économiques et sociales très profondes dans les sociétés du Moyen Âge. Un effet *cumulatif* est un effet qui se renforce par l'intermédiaire de ses propres conséquences. Lorsque, dans un pays, la quantité de monnaie augmente, il s'ensuit une inflation ; les gens accélèrent leurs achats, pour parer à l'augmentation des prix, mais ils contribuent ainsi à les faire monter (causalité circulaire). Le « paradoxe d'Anderson » est un effet classique de *reproduction* : la démocratisation de l'éducation n'a eu, dans les dernières décennies, que peu ou pas d'effet sur la mobilité sociale : bien que le niveau d'instruction influence fortement le statut social, et bien que les enfants d'ouvriers aient une probabilité de plus en plus grande d'atteindre des niveaux scolaires élevés, leur probabilité d'occuper dans l'échelle sociale une place relativement plus élevée que celle de leurs parents a peu varié. R. Boudon (1973) a montré que cette reproduction de la structure sociale était une conséquence agrégée d'une interaction complexe entre les acteurs concernés. Dans un effet de *contradiction*, les conséquences macrosociologiques des actions individuelles sont contraires à celles recherchées par les acteurs. Dans les pays pauvres, le taux de natalité est souvent élevé, les parents comptant sur les enfants pour les aider dans leurs vieux jours. Or, l'accroissement de la population provoqué par cette forte natalité peut accroître la pauvreté de tous.

La méthodologie individualiste peut être difficile ou impossible à mettre en œuvre en pratique, soit parce que les contextes de décision des acteurs sont mal connus, soit parce que les effets d'agrégation sont trop complexes. Dans l'expression « individualisme méthodologique », le mot « individua-

lisme » est, par ailleurs, susceptible de prêter à confusion. Il est faux de croire que cette méthode ne s'applique qu'aux sociétés modernes « individualistes » (selon l'expression d'A. de Tocqueville) : elle permet aussi d'expliquer les croyances et le fonctionnement des sociétés traditionnelles (Horton, 1967 ; Popkin, 1979). Il est tout aussi faux de considérer l'individualisme méthodologique comme une méthode « libérale » au sens politique. Elle est utilisée par des économistes britanniques classiques ou les économistes libéraux de l'École autrichienne (von Mises, 1949 ; Hayek, 1952), mais aussi par des économistes socialistes : par K. Marx dans sa théorie sur le caractère autodestructeur du système capitaliste (chaque capitaliste veut maximiser son profit, et contribue ce faisant à conduire le capitalisme à sa perte) et par O. R. Lange (1936) dans sa théorie sur la possibilité et la rationalité d'une planification étatique collectiviste.

Les sources de l'individualisme méthodologique

L'individualisme méthodologique est employé depuis longtemps dans les sciences sociales. Il sert de fondement à la théorie de Hume (1739-1740) sur l'émergence des normes telles que le droit de propriété ou le respect des contrats (Demeulenaere, 1996, 39-51), ainsi qu'à la théorie du contrat social de Rousseau (Boudon, 1979, 220-222). En économie, on le retrouve dans tous les modèles formulés par les économistes classiques, et en particulier par D. Ricardo (1817) : loi des rendements décroissants, loi d'airain des salaires, loi de la rente, tendance générale à l'égalité des taux de profit entre les différentes branches de production, loi de l'offre et de la demande, loi des coûts, théorie quantitative de la monnaie, loi des avantages comparatifs, etc. En sociologie, Tocqueville fait usage de l'individualisme méthodologique dans sa sociologie historique de la Révolution française (1856) comme dans sa sociologie de la société démocratique (1840).

Au cours des années 1870, la *révolution marginaliste* (Jevons, 1871) a donné naissance à l'économie dite « néoclassique ». Cette révolution a consisté à formaliser les choix individuels, et a donc permis de mieux comprendre les actions « rationnelles » au sens instrumental du terme. L'individualisme méthodologique en est sorti renforcé.

Une étape décisive est franchie à la fin du XIX^e et au début du XX^e s. par l'École sociologique allemande. Selon Weber, la sociologie « ne peut procéder que des actions d'un, de quelques, ou de nombreux individus séparés. C'est pourquoi elle se doit d'adopter une méthode strictement "individualiste" ». L'individualisme méthodologique apparaît ainsi comme une méthode à part entière, et reconnue comme telle, qui permet d'étendre à la sociologie la méthodologie des économistes. Cette généralisation impliquait de pouvoir analyser une très large gamme d'actions et de croyances : notam-

ment des actions et des croyances non instrumentales. Weber a donc été conduit à adopter un critère large de rationalité fondé sur le postulat de la « compréhension ».

L'économie n'est pas la seule discipline à avoir inspiré la conceptualisation de l'individualisme méthodologique : la réflexion des sociologues allemands sur la philosophie de l'histoire à la fin du XIX^e siècle a joué elle aussi un rôle important (Simmel, 1892).

Le holisme méthodologique

Le holisme méthodologique, dans sa version forte, exclut l'action comme source causale des phénomènes sociaux. Le holisme dérive du *positivisme* de Comte, selon lequel la sociologie ne peut être une science que si elle conserve vis-à-vis des phénomènes la même extériorité que le physicien ou le chimiste.

Le holisme a été développé par l'École française de sociologie issue de Comte. Durkheim (1895) considère que ce sont les phénomènes sociaux qui exercent une contrainte causale sur les individus, et non l'inverse. Il en conclut que les faits sociaux s'expliquent par d'autres faits sociaux et non par des « états de la conscience individuelle ». C. Lévi-Strauss (1958, 71-73) estime que le langage est l'institution humaine qui se prête le mieux à une étude scientifique, les règles linguistiques étant appliquées inconsciemment par les individus, et l'influence de l'observateur sur l'objet observé étant négligeable. Les règles « holistes » issues du positivisme (exclusion de la conscience, extériorité radicale de l'observateur) sont incompatibles avec la phase de « compréhension » de l'individualisme méthodologique.

Les deux grands paradigmes qui se rattachent au holisme méthodologique sont l'*analyse causale* et le *structuro-fonctionnalisme*.

L'analyse causale est illustrée par l'étude de Durkheim (1897) sur le suicide. Les corrélations qu'il a relevées l'ont conduit à formuler l'hypothèse selon laquelle le taux de suicide augmente, soit lorsque l'intégration aux groupes sociaux est trop faible, soit au contraire lorsqu'elle est trop forte.

Dans le structuro-fonctionnalisme, les institutions sociales sont considérées comme les éléments d'un système : elles remplissent des fonctions du point de vue de la société dans son ensemble (aspect fonctionnel) et elles entretiennent des relations de complémentarité (aspect structurel). Ce paradigme est né chez les anthropologues : étudiant des sociétés traditionnelles dont ils ignoraient l'histoire longue, ils pouvaient trouver avantageux d'utiliser une méthode synchronique. Ainsi, A. R. Radcliffe-Brown (1935) explique l'institution du culte des ancêtres dans les sociétés traditionnelles par sa fonction : renforcer la solidarité ; car ce culte permet à l'individu de comprendre qu'il peut compter

sur l'action magique et bénéfique de sa lignée, à condition de se soumettre à l'autorité des règles traditionnelles.

Lorsque la société est conçue comme un système fonctionnel d'institutions, on passe facilement à l'analyse de la structure sociale qui manifeste la complémentarité entre ces institutions. Ainsi, Parsons (1951) explique que, dans une société industrialisée, le marché du travail incite les agents à la mobilité. Or, cela implique que les agents ne soient pas liés de façon trop étroite à leurs parents. Selon ce raisonnement, l'industrialisation est incompatible avec des liens familiaux étendus et forts de type traditionnel. L'industrialisation et la coupure entre famille de procréation et famille d'orientation seraient ainsi deux éléments complémentaires, constituant une structure sociale cohérente.

En anthropologie, l'étude des structures sociales a pris avec Lévi-Strauss (1958) la forme d'un structuralisme qui s'inspire des travaux des linguistes. Lévi-Strauss s'est aperçu que les relations de parenté ou d'alliance (frère/sœur) (mari/femme)/(oncle/neveu) (père/fils) forment, au sein de diverses sociétés, des systèmes caractérisés par des lois d'opposition. Ces systèmes de relations sont représentés par quatre structures (au lieu des seize possibles), quelles que soient les sociétés considérées : + -/+ -, + -/- +, - +/+ -, ou - +/- + (la structure + -/+ -, par exemple, signifie que les relations frère/sœur sont bonnes, les relations mari/femme mauvaises, les relations oncle/neveu bonnes, et les relations père/fils mauvaises). Ces lois d'opposition sont similaires à celles que N. Troubetzkoy et R. Jakobson ont établies dans le domaine de la phonologie structurale (la langue repose, en tant que système de communication, sur des lois d'opposition entre les phonèmes).

Individualisme méthodologique et holisme : opposition ou complémentarité ?

Dans la forme que lui ont donnée Durkheim ou Lévi-Strauss, le holisme s'oppose à l'individualisme méthodologique. Si on suivait à la lettre leurs directives, des pans entiers de la sociologie et la plus grande partie de la science économique disparaîtraient.

Le holisme méthodologique a lui aussi son efficacité. Il a permis de produire des théories importantes dans divers champs des sciences sociales. L. Dumont (1983), notamment, a rapporté l'unité de son investigation, commencée en 1957 avec une thèse sur une sous-caste de l'Inde du Sud, à la mise en œuvre des principes du holisme méthodologique dont on pourra toujours discuter la validité (Valade, 2001). Mais l'individualisme méthodologique est souvent en mesure de « récupérer » les théories holistes. Les résultats d'une analyse causale (méthode holiste) sont souvent complétés par un raisonnement en termes d'individualisme méthodolo-

gique, comme dans l'exemple de l'explosion du nombre des cambriolages au cours des décennies 1960 et 1970 : on comprend que les délinquants potentiels décident de plus en plus fréquemment de cambrioler des habitations dans la mesure où celles-ci sont de moins en moins bien surveillées. La théorie de la famille de Parsons suppose que les individus profitent des occasions de promotion sociale que leur ouvrent leurs études et leurs possibilités de mobilité géographique : ce sont ces choix individuels qui provoquent, au niveau macrosociologique, la coupure entre « famille d'orientation » et « famille de procréation ».

► BOUDON R., *L'Inégalité des chances. La mobilité sociale dans les sociétés industrielles*, Paris, Colin, 1973 ; *La Logique du social. Introduction à l'analyse sociologique*, Paris, Hachette, 1979 ; *La Place du désordre. Critique des théories du changement social*, Paris, PUF, 1984 ; *L'Idéologie, ou l'origine des idées reçues*, Paris, Fayard, 1986 ; « Individualisme et holisme dans les sciences sociales », in P. BIRNBAUM, J. LECA (éd.), *Sur l'individualisme*, Paris, FNSP, 1986 ; « Individualisme et holisme : un débat méthodologique fondamental », in H. MENDRAS, M. VERRET (éd.), *Les Champs de la sociologie française*, Paris, A. Colin, 1988. – COHEN L. E., FELSON M., « Social change and crime rates trends : a routine activity approach », *ASR*, 1979, 44, 588-608. – DEMEULENAERE P., *Homo æconomicus. Enquête sur la constitution d'un paradigme*, Paris, PUF, 1996. – DUMONT L., *Essais sur l'individualisme*, Paris, Le Seuil, 1983. – DURKHEIM É. (1895), *Les Règles de la méthode sociologique*, Paris, PUF, 1986 ; (1897) *Le Suicide. Étude de sociologie*, Paris, PUF, 1986. – HAYEK F. A. (1952), *Scientisme et sciences sociales. Essai sur le mauvais usage de la raison*, Paris, Plon, 1953. – HORTON R., « African traditional thought and Western science », *Africa*, 1967, vol. 38, 50-71 et 155-187. – HUME D. (1739-40), *Traité de la nature humaine*, Paris, Aubier, 1973. – JEVONS W. S. (1871), *The Theory of Political Economy*, New York, Augustus M. Kelley, 1965. – LANGE O. R. (1936), « On the economic theory of socialism », in B. E. LIPPINCOTT (éd.), *On the Economic Theory of Socialism*, New York, McGraw-Hill, 1964, 55-143. – LÉVI-STRAUSS C. (1958), *Anthropologie structurale*, Paris, Plon, 1985. – MISES L. von (1949) *L'Action humaine. Traité d'économie*, Paris, PUF, 1985. – PARSONS T., *The Social System*, New York, The Free Press, 1951. – POPKIN S., *The Rational Peasant*, Berkeley, Univ. of California Press, 1979. – RADCLIFFE-BROWN A. R. (1935), « On the concept of function in social science », in A. R. RADCLIFFE-BROWN, *Structure and Function in Primitive Society*, Glencoe, The Free Press, 1952. – RICARDO D. (1817), *On the Principles of Political Economy and Taxation*, Cambridge, Cambridge Univ. Press, 1951. – SIMMEL G. (1892), *Les Problèmes de la philosophie de l'histoire. Une étude d'épistémologie*, Paris, PUF, 1984. – TOCQUEVILLE A. de (1840), *De la démocratie en Amérique*, vol. 2, Paris, Gallimard, 1961 ; (1856), *L'Ancien Régime et la Révolution*, Paris, Gallimard, 1952. – VALADE B., « De l'explication dans les sciences sociales : holisme et individualisme », in J.-M. BERTHELOT (éd.), *Épistémologie des sciences sociales*, Paris, PUF, 2001. – WEBER M. (1922), *Économie et société*, Paris, Plon, 1971. – WHITE L., *Medieval Technology and Social Change*, Oxford, Clarendon Press, 1962.

Renaud FILLIEULE

→ Action ; Économie et sociologie ; Mises ; Weber M.

Quatrième partie : la sociologie de la délinquance

Sociologie de la délinquance (2001)

Table des matières

Paris, PUF



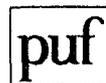
Collection
Premier
Cycle

Sociologie de la délinquance

RENAUD FILLIEULE

Maître de conférences en sociologie
à l'Université Lille I

*Presses
Universitaires
de France*



Sommaire

Introduction, 1

- La notion de délinquance, 5
- Plan de l'ouvrage, 6

PREMIÈRE PARTIE

FAITS

I – Les statistiques officielles de la réaction judiciaire, 11

Les statistiques policières, 12

- Les catégories d'infractions, 13
- Les comptages, 14
- Les taux d'élucidation, 17
- Les personnes mises en cause, 18

Les tendances d'évolution de la délinquance enregistrée, 19

Les tendances générales, 19

L'évolution globale, 19 ; L'évolution des grandes catégories d'infractions, 21 ; L'évolution des taux d'élucidation, 22 ; Les évolutions concernant les personnes mises en cause, 23.

Les tendances pour certaines infractions, 24

Les vols, 24 ; Les violences physiques, 25.

Comparaison avec les États-Unis, 25

Le « chiffre noir » de la délinquance, 28

La notion de « chiffre noir », 28

L'évaluation du « chiffre noir », 29

Les sondages de victimation, 29 ; Les évolutions du « chiffre noir », 32 ; Les limites des sondages de victimation, 34.

La décision de ne pas porter plainte, 37

La construction des statistiques policières, 38

La sélection des plaintes, 38

L'enregistrement des plaintes, 40

Les statistiques judiciaires, 41

L'orientation des affaires, 42

Les condamnations, 44

Les incarcérations, 45

II – **Délinquants et délinquance**, 49

Les sources de données, 49

Les données officielles, 50

Les données autorévélees, 51

Le profil sociodémographique des délinquants, 52

Deux concepts descriptifs : participation et fréquence, 52

L'âge, 54

La courbe des âges, 54 ; *La thèse de l'invariance*, 58.

Le genre, 59

La classe sociale, 62

Délinquants « occasionnels » et délinquants « chroniques », 66

La délinquance en groupe, 69

Le déroulement des actes de délinquance typiques, 72

Les vols, 72

Les vols de voitures, 72 ; *Les cambriolages*, 73.

Les agressions, 75

Les attaques à main armée, 75 ; *Les viols*, 76 ; *Les homicides*, 77.

III – **Les victimes**, 79

Les risques de victimation, 80

L'évaluation des risques par types d'infractions, 80

Une comparaison internationale, 84

Le profil sociodémographique des victimes, 87

L'explication de la victimation : la théorie des styles de vie, 90

L'impact de la délinquance, 98

Le coût monétaire de la délinquance, 98

Les émotions suscitées chez les victimes, 101

Le sentiment d'insécurité, 105

L'évaluation du sentiment d'insécurité, 105

Sentiment d'insécurité et risque de victimation, 107

La rationalité du sentiment d'insécurité, 108

Sentiment d'insécurité et style de vie, 111

DEUXIÈME PARTIE

THÉORIES

IV – **Théories actionnistes et contrôle social**, 117

La théorie du contrôle social, 118

Le « lien » social, 118

Les institutions du contrôle social, 121

La famille, 122 ; *L'école*, 123 ; *Les copains*, 124.

Théorie du contrôle social et profil sociodémographique des délinquants, 126

Fins et moyens de l'action délinquante, 127

La typologie de Cusson, 127

Le choix de la délinquance comme moyen : une théorie des opportunités, 132

L'explication du penchant délinquant : le « présentisme », 134

La nature de la délinquance typique, 134

Des caractéristiques de la délinquance à celles du délinquant, 136

Les sources du présentisme, 138

Les conséquences du présentisme, 140

Présentisme et courbe des âges, 140 ; *La différence entre les genres*, 143 ;
Présentisme et classe sociale, 143 ; *La versatilité des délinquants*, 144.

La théorie des activités routinières et les occasions de délinquance, 145

La théorie, 146

Une explication des statistiques de victimation, 148

Une explication de la croissance de la délinquance, 149

V – **Théories multifactorielles et carrières délinquantes**, 153

La « carrière » délinquante, 154

La notion de « carrière », 154

Les variables descriptives de la « carrière », 156

La recherche des causes de la délinquance, 158

La méthode, 158

L'observation par enquête longitudinale, 158 ; *L'analyse causale*, 159.

Les facteurs de risque individuels, 160

Les facteurs démographiques, 160 ; *Les facteurs psychologiques*, 161 ; *Les facteurs biologiques*, 163.

Les facteurs de risque familiaux, 164

L'éducation parentale, 164 ; *La taille de la famille*, 165 ; *Les conflits parentaux et les familles monoparentales*, 166 ; *La délinquance des parents*, 167.

Les facteurs de risque collectifs, 168

Le chômage, 168 ; *Médias et violence*, 170.

La critique actionniste des théories multifactorielles, 173

VI – Les théories classiques : culturalisme et étiquetage, 177

Les théories culturalistes, 177

Les théories de la transmission culturelle, 178

L'association différentielle (Sutherland), 178 ; *Les aires de délinquance (Shaw et McKay)*, 180.

Les théories de la « tension », 184

L'origine sociale (Merton), 185 ; *La masculinité (Parsons)*, 187 ; *L'adolescence (Bloch et Niederhoffer)*, 188.

Les théories mixtes, 188

L'apparition des sous-cultures délinquantes (Cohen), 189 ; *Sous-cultures délinquantes et « opportunités » (Cloward et Ohlin)*, 192.

Critique des théories culturalistes, 194

La théorie de l'étiquetage, 198

La fabrication du criminel par l'étiquetage (Tannenbaum), 199

La théorie des « outsiders » (Becker), 201

La déviance est « créée par la société », 201 ; *Les carrières déviantes*, 203 ; *Création et imposition des normes*, 206.

Critique de la théorie de l'étiquetage, 207

TROISIÈME PARTIE

PRÉVENTION

VII – La dissuasion, 215

La dissuasion pénale, 216

La doctrine de la dissuasion, 216

La dissuasion générale, 217

Principe, 217 ; *Évaluation*, 218.

La dissuasion spécifique, 221

Principe, 221 ; *L'évaluation de la délinquance « avant-après » une sanction (Murray et Cox)*, 223 ; *La dissuasion de la violence domestique et ses dilemmes (Sherman)*, 224.

Les critiques contre la dissuasion pénale, 228

La prévention des incivilités, 230

La théorie de la vitre cassée, 231

La police de proximité, 233

La prévention situationnelle, 236

Principes, 237

La classification des techniques, 238

Accroître l'effort, 239 ; Augmenter les risques, 240 ; Réduire les gains, 242.

Évaluation, 244

Les critiques contre la prévention situationnelle, 245

*Ne vaut-il pas mieux traiter les « causes profondes » de la délinquance ?, 245 ; Le déplacement de la délinquance, 247 ; Big Brother nous surveille..., 251.***VIII – La réhabilitation, 253**

L'inefficacité des mesures classiques de réhabilitation, 254

Rien ne marche ? (Martinson), 254

L'instruction et la formation professionnelle, 256 ; Le conseil individualisé ou en groupe, 257 ; La transformation de l'environnement institutionnel, 258 ; La réhabilitation hors les murs, 259.

Les raisons de l'échec, 261

La prévention développementale : un renouveau de la réhabilitation ?, 263

IX – La neutralisation, 269

Neutralisation et « carrières » délinquantes, 270

La neutralisation collective, 271

La neutralisation sélective, 274

Bibliographie, 279

Cinquième partie : l'école autrichienne d'économie

Introduction aux théories de l'école autrichienne (2000)

Table des matières

Document pédagogique non publié

TABLE DES MATIÈRES

Première partie : action et valeur

Chapitre I : les biens

1. L'action humaine
 - 1.1. *Le postulat de l'action*
 - 1.2. *La science sociale comme praxéologie*
2. Biens et conditions générales de l'action
3. Les types de biens
 - 3.1. *Biens de consommation et facteurs de production*
 - 3.2. *La notion de structure de production : premier aperçu*
 - 3.3. *Facteurs originaires et facteurs produits*
4. Le temps
5. Les techniques

Chapitre II : la valeur

1. La valeur
 - 1.1. *La théorie de la valeur subjective*
 - 1.2. *La représentation des valeurs*
 - 1.3. *L'ordinalité des valeurs*
2. Les lois de la valeur
 - 2.1. *La loi d'imputation de la valeur (loi de Menger)*
 - 2.2. *Loi d'imputation contre théorie de la valeur travail*
 - 2.3. *La loi de l'utilité marginale*
 - 2.4. *Résolution du « paradoxe de la valeur »*
 - 2.5. *La loi de l'utilité totale*
 - 2.6. *La valeur des paniers de biens*
3. Le coût

Deuxième partie : échange et prix

Chapitre III : l'échange

1. La nature de l'échange
 - 1.1. *Les formes élémentaires d'interaction : échanger, prendre, donner*
 - 1.2. *Possession et propriété*
 - 1.3. *Échange et marché*
 - 1.4. *Les droits de propriété dans une économie de marché libre*
2. L'explication de l'échange
 - 2.1. *L'action d'échanger : valeur d'usage et valeur d'échange*
 - 2.2. *La double inégalité des valeurs*
 - 2.3. *La nature des biens échangés*
 - 2.4. *L'échange entre des stocks*
 - 2.5. *Le mythe de l'égalité des valeurs dans l'échange*
3. Les gains de l'échange
 - 3.1. *Les gains mutuels ex ante*
 - 3.2. *L'échange considéré ex post*

Chapitre IV : la division du travail

1. Division du travail et accroissement de la production
 - 1.1. *La loi des avantages absolus*
 - 1.2. *La loi des avantages comparatifs (ou loi d'association de Ricardo)*
 - 1.3. *Généralisation à l'économie monétaire*
 - 1.4. *Les différents aspects productifs de la division du travail*
 - 1.5. *Les sources de la division du travail*

2. Division du travail et société
 - 2.1. *Société et coopération*
 - 2.2. *L'origine des sociétés*
 - 2.3. *L'évolution des sociétés*
 - 2.4. *Division du travail et progrès*
- Annexe : démonstration des lois des avantages absolus et comparatifs

Chapitre V : les prix en économie de troc

1. Le prix
2. La formation du prix en concurrence unilatérale
 - 2.1. *Concurrence entre les acheteurs*
 - 2.2. *Concurrence entre les vendeurs*
3. La formation du prix en concurrence bilatérale
 - 3.1. *Les échelles de préférences des acteurs*
 - 3.2. *Le processus d'équilibrage entre les quantités offertes et demandées*
 - 3.3. *L'offre et la demande*
 - 3.4. *Les prix et quantité d'équilibre*
 - 3.5. *Acteurs marginaux, supramarginaux, et inframarginaux*
4. Les variations d'offre et de demande
 - 4.1. *Représentation simplifiée de l'offre et de la demande*
 - 4.2. *Les variations de demande*
 - 4.3. *Les variations d'offre*
 - 4.4. *Influence de ces variations sur les prix et quantités d'équilibre*

Chapitre VI : les prix en économie monétaire

1. Échange direct et échange indirect
 - 1.1. *Les limites de l'échange direct*
 - 1.2. *Le principe de l'échange indirect*
2. L'émergence de la monnaie
 - 2.1. *La recherche de biens liquides dans l'échange indirect*
 - 2.2. *Facteurs qui accroissent la liquidité des biens*
 - 2.3. *Le processus cumulatif conduisant à l'apparition de la monnaie*
 - 2.4. *Quelques exemples historiques de monnaies*
3. La formation des prix en monnaie
 - 3.1. *Le prix des biens en monnaie et les prix de la monnaie en biens*
 - 3.2. *L'utilité marginale de la monnaie*
 - 3.3. *Le prix monétaire d'équilibre en concurrence bilatérale*
4. La régression monétaire
5. Les relations entre les prix monétaires des biens de consommation
 - 5.1. *L'élasticité de la demande*
 - 5.2. *Les relations entre les prix des biens substituables*
 - 5.3. *Les relations entre les prix des biens complémentaires*
 - 5.4. *Tableaux récapitulatifs des relations entre les prix*

Troisième partie : production et taux d'intérêt

Chapitre VII : la valeur des facteurs de production

1. Les processus de production : aspects quantitatifs
 - 1.1. *La notion de fonction de production*
 - 1.2. *La loi des rendements d'échelle constants*
 - 1.3. *Les notions de produit physique total, moyen, et marginal*
 - 1.4. *La loi des rendements décroissants*
 - 1.5. *La relation entre produit physique marginal et produit physique moyen*
2. L'économie de l'acteur isolé ou « économie de Robinson Crusoe »
3. La valeur des facteurs de production pour l'acteur isolé
 - 3.1. *La distinction entre productivité et produit physique*
 - 3.2. *Le principe fondamental de valorisation des facteurs*
 - 3.3. *Première application*

- 3.4. *Facteurs spécifiques et facteurs convertibles*
- 3.5. *La valeur d'un facteur spécifique*
- 3.6. *Les facteurs convertibles : valeur, allocation, coût d'opportunité*
- 3.7. *Le travail comme facteur de production*
- 3.8. *La valeur d'un facteur d'ordre supérieur*
- 4. Les changements de valeur du produit : conséquences sur la valeur des facteurs

Chapitre VIII : la préférence pour le présent

- 1. Valeur et temps
 - 1.1. *Fins futures et biens futurs*
 - 1.2. *La loi de la préférence pour le présent*
 - 1.3. *Importance de la condition « toutes choses égales par ailleurs »*
 - 1.4. *La valeur des biens durables*
- 2. La production comme choix intertemporel
 - 2.1. *L'action dans le temps*
 - 2.2. *Les notions d'épargne et d'investissement*
 - 2.3. *Un cas élémentaire : la consommation différée*
 - 2.4. *Une décision d'investissement*
 - 2.5. *Le choix de la période de production*

Chapitre IX : le taux d'intérêt

- 1. La notion de taux d'intérêt
 - 1.1. *Le taux d'intérêt dans l'échange intertemporel*
 - 1.2. *Le prêt/emprunt de monnaie comme échange intertemporel*
 - 1.3. *La production marchande comme échange intertemporel*
- 2. L'explication du taux d'intérêt par la préférence pour le présent
- 3. La détermination du taux d'intérêt pur d'équilibre
 - 3.1. *Le marché du temps*
 - 3.2. *Un cas simple à deux acteurs*
 - 3.3. *Le cas général*
- 4. Analyse critique de trois théories concurrentes
 - 4.1. *La théorie de la productivité du capital*
 - 4.2. *La théorie de l'usage*
 - 4.3. *La théorie de l'exploitation*

Chapitre X : les prix des facteurs de production

- 1. Le prix d'une unité d'un facteur de production convertible
 - 1.1. *L'offre d'un facteur de production*
 - 1.2. *La demande d'un facteur de production*
 - 1.3. *Le prix d'équilibre*
 - 1.4. *La loi des coûts*
- 2. La prise en compte du temps
 - 2.1. *L'actualisation*
 - 2.2. *La capitalisation*
- 3. Les facteurs « terre »
- 4. Les facteurs travail
- 5. Les biens de capitaux

Chapitre X : la structure de production

- 1. La structure de production dans l'économie de Robinson Crusoe
 - 1.1. *Représentation d'une structure élémentaire*
 - 1.2. *Le maintien ou « équilibre » de la structure*
 - 1.3. *La structure de production globale*
- 2. La structure de production d'une économie développée
 - 2.1. *Première esquisse*
 - 2.2. *Une représentation typique*
- 3. L'équilibre final ou « économie en rotation uniforme » (ÉRU)
 - 3.1. *La notion d'équilibre final*
 - 3.2. *La distinction entre équilibre final (ÉRU) et équilibre de marché*

- 3.3. *L'uniformité du taux d'intérêt en ÉRU*
- 4. Une structure de production simple à l'équilibre final
 - 4.1. *Hypothèses simplificatrices*
 - 4.2. *Types d'acteurs et de revenus*
 - 4.3. *Représentation d'une étape*
 - 4.4. *Représentation de la structure*
 - 4.5. *Les marchés du temps d'une étape*
 - 4.6. *Les flux monétaires : dépenses et revenus*
 - 4.7. *Investissement brut et maintien de la structure de production*
 - 4.8. *À quoi sert la notion de structure de production ?*

Quatrième partie : l'entrepreneuriat

Chapitre XI : entrepreneuriat et profit/perte

- 1. Incertitude et action
 - 1.1. *L'omniprésence de l'incertitude*
 - 1.2. *Incertaineté et vigilance entrepreneuriale*
 - 1.3. *Les deux phases de l'action humaine : entrepreneuriat et maximisation*
 - 1.4. *Les gains ex ante de l'action : « profit » et « récompense »*
 - 1.5. *Les résultats ex post de l'action : « profit/perte » et « récompense/perte »*
- 2. L'entrepreneuriat en économie monétaire
 - 2.1. *La fonction entrepreneuriale et son revenu net : le profit/perte monétaire*
 - 2.2. *Le calcul du profit/perte monétaire : un exemple*
 - 2.3. *La signification du profit/perte monétaire*
 - 2.4. *L'origine du profit/perte monétaire*
 - 2.5. *Les fonctions de l'entrepreneur et du profit/perte monétaire*
 - 2.6. *La tendance à la disparition du profit/perte monétaire*
 - 2.7. *L'imputation du profit/perte aux revenus des facteurs originaires*
 - 2.8. *Ce que le profit/perte monétaire n'est pas*
- 3. La composante entrepreneuriale du taux d'intérêt

« The “values-riches” model : an alternative to Garrison’s model
in Austrian macroeconomics of growth and cycle » (2005)

The Quarterly Journal of Austrian Economics

Vol. 8, n° 2, p. 3-18

THE “VALUES-RICHES” MODEL: AN ALTERNATIVE TO GARRISON’S MODEL IN AUSTRIAN MACROECONOMICS OF GROWTH AND CYCLE

RENAUD FILLIEULE

The “values-riches” model is a pedagogical and graphical representation of Austrian macroeconomics, combining four elements: the concept of “invariable money” (Reisman 1996), the capitalization theory of interest of Fetter (1904, 1914) and Mises (1998), the Hayekian structure of production (Hayek 1931), and the growth theory of Böhm-Bawerk (1959). This model offers a simple visual illustration of equilibrium, growth, and business cycle as they may be conceived in an Austrian framework. The graphical approach of this model is heavily influenced by Garrison (2001). It will be argued below that Garrison’s model represents and relates real macroeconomic magnitudes. The “values-riches” model,¹ on the other hand, seeks to display the relations between the great macroeconomic nominal variables (“values”) and the flows of quantities of consumer goods (“riches”). The two models are therefore to some extent complementary, offering two different viewpoints on the production process. But there are also a number of theoretical disagreements between them, specially about the theory of interest, that must not be overlooked and will be pointed out in this paper.

MACROECONOMIC EQUILIBRIUM

The Three Panels of the Model

The values-riches model is composed of three interdependent panels.

(1) The first panel (figure 1) represents the *line of aggregate expenditure*. The annual aggregate expenditure E in the economic system is the sum of the

RENAUD FILLIEULE is *maître de conférences* at the University of Lille 1 (USTL), Villeneuve d’Ascq, France, and member of the CLERSÉ research unit (Centre lillois d’études et de recherches sociologiques et économiques). The author wishes to thank the two anonymous referees for remarks that have helped to improve the final version of this paper.

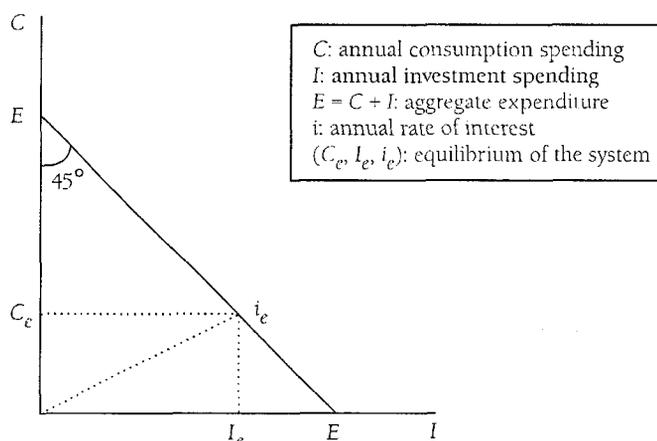
¹This name refers to the title of chapter 20 in Ricardo’s *Principles* (“Value and Riches, their Distinctive Properties”).

annual consumption spending C and the annual investment spending² I : $E = C + I$. This panel illustrates the concept of an “invariable money” of Reisman (1996, pp. 536-40). According to him, macroeconomic reasoning should begin by supposing that aggregate expenditure is a constant, i.e., that money is “invariable.” This assumption makes it possible to understand that an increase in investment spending implies, everything else equal, a corresponding diminution of consumption spending. It is therefore essential in order to study the economic growth originating in a lowering of time preference (see the section below “Lowering of the Preference for the Present”). And then, in a second moment, monetary interferences should be taken into account through variations of aggregate expenditure E (see “Money and Business Cycle” below).

This first panel shows the equilibrium point of the economic system (C_e , I_e , i_e). Everything else equal, i.e., if preferences do not change (particularly intertemporal preferences), if there is no technical progress, and if offers of originary factors remain the same, then the economic system converges towards a final equilibrium with equilibrium values C_e for consumption spending, I_e for investment spending and i_e for the annual rate of interest. Saying that money is “invariable” means that the sum ($C + I$) is a constant and that the point representing the system moves along the line EE in figure 1 (this line cuts the two axes for the value E).

How are saving/investment and the rate of interest to be explained? Skousen (1990, p. 204) and Garrison (2001, pp. 36-40) resort to the theory of

Figure 1
The Line of Aggregate Expenditure EE
(Equation $E = C + I$)



²This investment spending I encompasses the *totality* of the annual expenditures on factors of production, whatever their nature, including the expenditures on the intermediary goods (that are excluded from the calculation of the narrow aggregate called “investment” in the Gross Domestic Product).

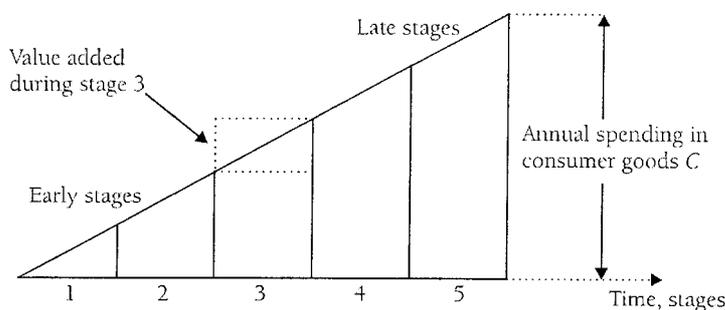
the market for loanable funds. But this theory does not belong to the Austrian tradition. It is found, neither in Böhm-Bawerk (1959), neither in Mises nor Rothbard (both of them explicitly reject it),³ neither in Hayek (who upholds a productivity theory in Hayek 1941). The theory used in the values-riches model is the capitalization theory developed by Fetter (1904) and adopted by Mises (1998, chap. 19): the rate of interest is explained by the preference for the present, which implies that factors of production are bought with a discount, i.e., at a price lower than that of the good they will produce in the future. This price differential is the (originary) interest. Entrepreneurial competition tends to equalize these rates between firms and sectors of production, and a unique rate would prevail if a final equilibrium was reached (but recurrent and unanticipated dynamic changes prevent this outcome to happen). If the preference for the present diminishes among the agents, then the demand for (and the value of) present goods is reduced and the demand for (and the value of) factors of production increases: the originary rate of interest (relative price differential between the factors and their produce) and the ratio C_e/I_e are both lowered. Under these circumstances, with an “invariable money,” the point of equilibrium in panel 1 moves downward along the line EE : C_e diminishes, I_e increases, the sum $C_e + I_e$ remains constant ($= E$), and i_e diminishes (see figure 6).

(2) The second panel (figure 2) represents the *Hayekian structure of production* (Hayek 1931; Rothbard 1962; Skousen 1990). The simplest illustration of this structure is found in Garrison (2001, p. 47): a triangle whose horizontal side represents time and the successive stages of production, and whose vertical side represents prices (values). The structure in figure 2 has five stages and its horizontal side measures the total length of the period of production. Supposing that each stage lasts one year, the total length of the structure is here five years. The vertical side on the right measures the annual aggregate value of consumer goods. The third side (hypotenuse) represents the added value at each stage, which is the sum of the prices of originary factors used at this stage (non produced factors of production: labor, standing room, natural resources) and of the interest on capital invested at this stage. The sum of all these added values is the aggregate value of the annual output of consumer goods. Extractive industries like mining are to be found in the early stages (on the left of the structure, far away from final consumption), manufacturing industries are located towards the center of the structure, and retailing industries in the last stage (on the right of the structure).

Given the equilibrium values for consumption, investment and the rate of interest (C_e , I_e , i_e), and with two supplementary hypotheses, one on the length

³“Originary interest is not a price determined on the market by the interplay of the demand for and the supply of capital or capital goods” (Mises 1998, pp. 523-24). “[The rate of interest] is determined by factors that have nothing to do with the usual discussion by economists of the producers’ loan market” (Rothbard 1962, p. 364).

Figure 2
The Structure of Production



Source: Adapted from Garrison (2001, p. 47)

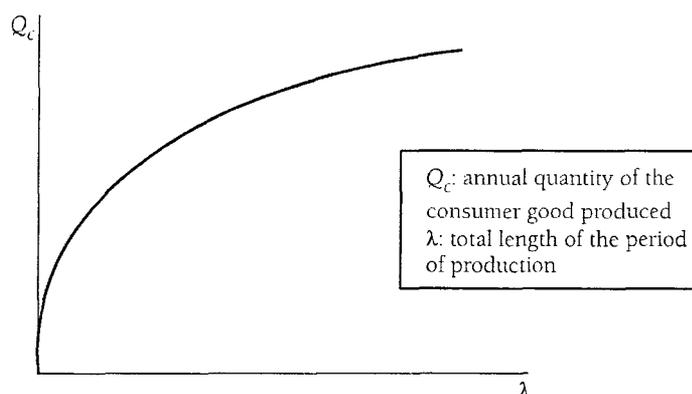
of each stage of production and another on the ratio of the value of originary factors to investment at each stage, the structure of figure 1 is *univocally* determined.⁴ There is thus a relation of determination from the point representing the economic system on the line of aggregate expenditure (figure 1) to the shape and length of the structure of production (figure 2).

(3) The third panel (figure 3) represents the *global production function* of the economic system. According to Böhm-Bawerk's theory of roundaboutness in production (1959, pp. 81-88), the quantity (or quality) of goods produced increases, everything else equal, when the period of production of the structure increases. The *ceteris paribus* assumption implies that the total quantity of labor used in the structure of production is a constant. In order to make things as simple as possible, only the total period of the structure will be considered here.⁵ This period is equal to the length of the horizontal side of the triangle in figure 2. Böhm-Bawerk's theory of growth must not be misunderstood. It does not mean that just any lengthening of the structure is productive, but only that among the countless possibilities of lengthening, some are productive (particularly those that would be chosen by economic agents in the case of an increase of investment relative to consumption). This physical productivity of the lengthening of the period of production is explained by the fact that it enables labor to use additional natural forces and resources (Hayek 1941, pp. 60-64).

⁴See the calculation in Fillieule (2004). Supposing that each stage lasts one year and that the ratio of the value of originary factors to investment at each stage is a constant a (for instance, at each stage, the aggregate value of originary factors divided by investment at this stage is equal to 20%, $a = 0.2$), Fillieule (2004) shows that the average length of the structure is $\lambda = (1 + i_e)/(i_e + a)$, where i_e is the annual equilibrium rate of interest.

⁵The more rigorous concept of *average* length should be used here, but it requires a detailed description of the structure of production that goes beyond the scope of this paper (see the previous footnote and the reference therein).

Figure 3
The Production Function of the Economic System



Let us suppose that only one kind of consumer good is produced by the economic system. The annual quantity produced is Q_c . With a given quantity of labor and a given level of technology, the quantity Q_c increases with the total length of the structure. But the returns of the lengthening of the period of production are decreasing since the quantity of labor is given and constitutes a fixed factor. If λ is the length of the structure, then the production function may for instance be written $Q_c = A\lambda^\alpha$, with $0 < \alpha < 1$. Coefficient A is a parameter that augments if the quantity of labor increases and if technologies progress. This function is monotonic (theory of roundaboutness) and concave (decreasing returns of lengthening), as shown in figure 3.⁶

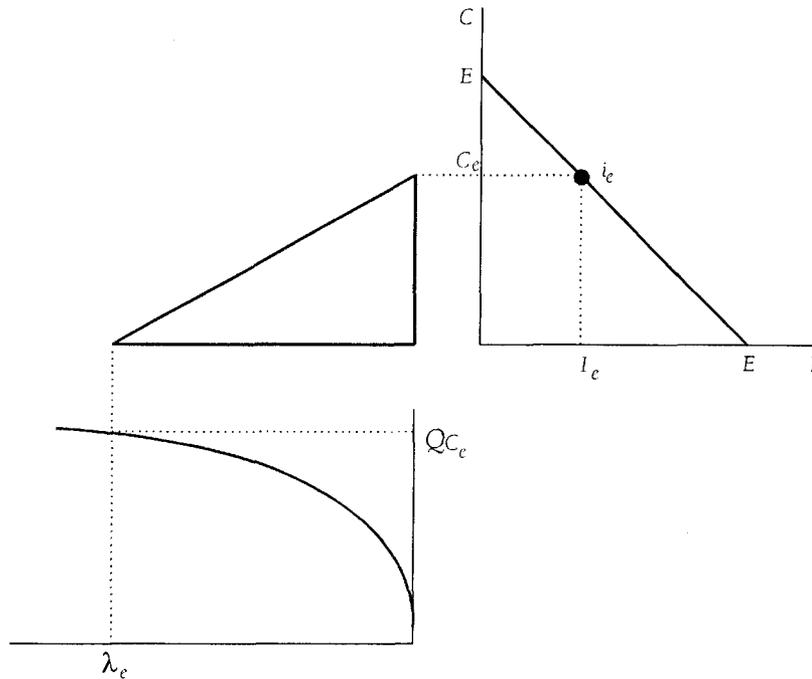
The Complete Model

The three interdependent panels are shown in figure 4. They allow us to visualize the relations between macroeconomic variables in equilibrium: consumption C_e , investment I_e , rate of interest i_e , length of the structure λ_e , and quantity of consumer goods produced Q_{c_e} . In panel 1, the point on the line of aggregate expenditure shows macroeconomic monetary spending at equilibrium ("values"). In panel 2, the structure of production relates these monetary variables (measured along the vertical dimension) with the length of the structure of production, i.e., with a period of time (measured along the horizontal dimension). In panel 3, the length of the structure is related by the production function to the quantity of consumer goods annually produced ("riches").⁷

⁶The lengthening of the structure is obviously a very complex phenomenon, because of the heterogeneity and various degrees of specificity of capital goods. This underlying complexity will not be explored any further in the present paper (see Garrison 1986 for a classical statement on this subject).

⁷This scheme neglects the necessary decreasing returns in the extractive industries. Everything else equal, it becomes more and more difficult to collect ore, natural gas and oil from the ground. With time, the physical productivity of the economic system diminishes and the function of production rotates counterclockwise.

Figure 4
Macroeconomic Equilibrium in the Values-Riches Model

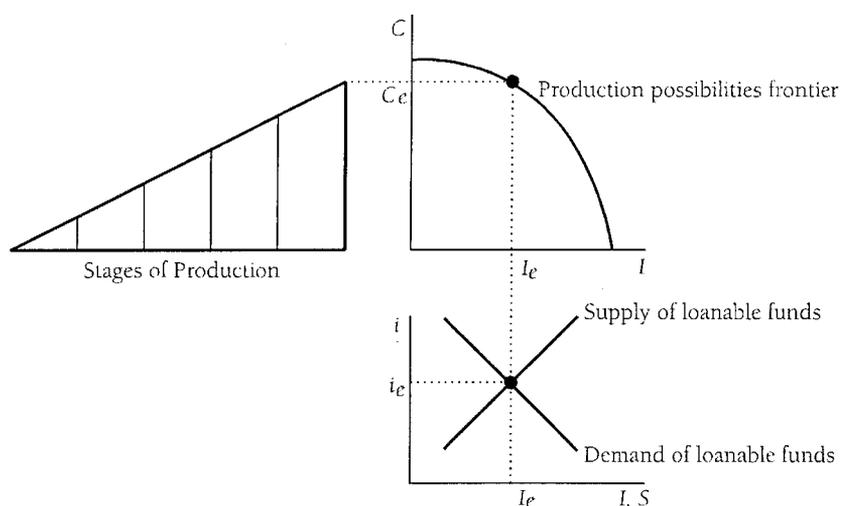


Comparison with Garrison's Model

Garrison (2001, p. 50) has propounded a model represented in figure 5 and composed of three panels: (1) the market for loanable funds, (2) the production possibilities frontier, and (3) the structure of production.

In spite of its undeniable pedagogical and theoretical qualities, Garrison's model appears unsatisfactory for two reasons. The first one has already been suggested and has to do with the theory of interest. The panel 1 on the lower left illustrates the theory of the market for loanable funds, which may be considered as defective and in any case does not belong to the Austrian tradition (see above and also Fetter 1915, p. 308; Salerno 2001, pp. 52-54). This panel has therefore not been integrated to the values-riches model, when it would have been very easy to do so since there is an unoccupied place under the line of aggregate expenditure. Our second disagreement concerns the production possibilities frontier. This concept is well known in macroeconomic analysis (see for instance Samuelson and Nordhaus [1989, p. 28], and the famous example of guns and butter). Garrison uses it in order to explain "the fundamental trade-off between consumer goods and capital goods" (Garrison 2001, p. 41). It is doubtful whether the production possibilities frontier is in fact appropriate to illustrate this kind of trade-off. Does it not rather pertain to the trade-off between two consumption goods? But the important point here is that this frontier relates variables C and I that are *quantities* of goods, when the structure of production and the market for

Figure 5
Garrison's Model



Source: Adapted from Garrison (2001, p. 50)

loanable funds relate variables C and I that are *monetary* expenses respectively on consumer and on producer goods (ordinary factors and capital goods). In Garrison's own words: "The vertical leg [of the structure of production] measures the *value* of the consumable output of the production process" (Garrison 2001, p. 46; emphasis added). To the extent that the production possibilities frontier serves as an intermediary between the market for loanable funds and the structure of production, a confusion seems to occur in Garrison's model between real and nominal variables (see also Hülsmann 2001, p. 40). This confusion is all the more worrying in that these values and quantities may in some cases move in opposite directions, for instance the values may diminish while the quantities of goods increase (see below figure 6 where the nominal value of consumption diminishes while the quantity of consumer goods increases). This is precisely the point that was underlined by Ricardo (1951) in his chapter on "Value and Riches" (Reisman 1996, p. 496 and pp. 712-15, offers an in-depth analysis on this subject). The only way to avoid this difficulty is, in our opinion, to interpret all the variables of Garrison's model in real terms. But then a complementary model is needed in order to visualize the relations between the main real and nominal variables—and this is precisely the function of the "values-riches" model.

GROWTH

The main sources of growth are (1) a lowering of the preference for the present (which favors investment and thus growth), (2) technical progress

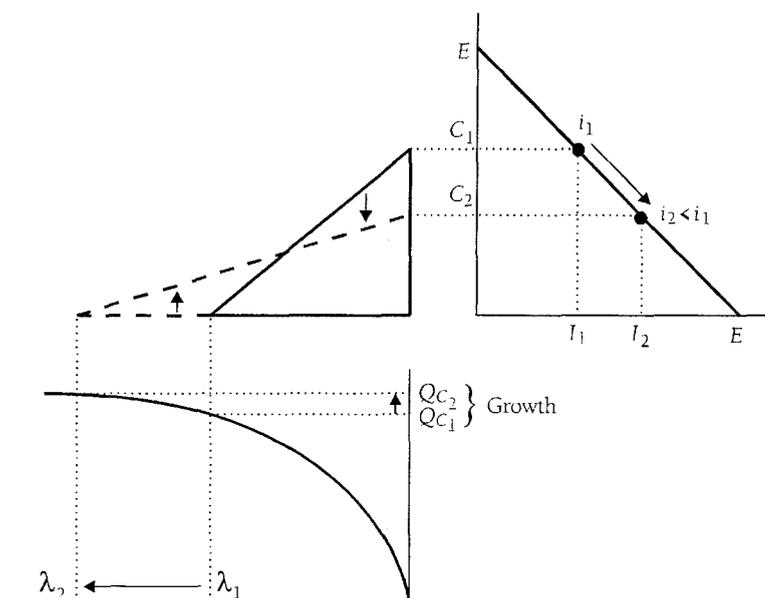
(including organizational progress or an intensification of the division of labor), and (3) an increase of the quantity (or quality) of labor. Since in the models propounded here the quantity of labor is a constant, only the first two sources of growth are taken into account (for present purposes, the trade-off between labor and leisure is not taken into account either). It should be pointed out that growth is conceived here, not as a continuous increase in the flow of consumption goods, but rather as a one-shot movement from a stationary equilibrium to another with a higher level of production.

Lowering of the preference for the present

When the preference for the present diminishes, agents reduce their consumption spending and add to their investment spending. The prices of consumer goods tend to diminish and those of factors of production tend to rise. As a consequence, the equilibrium rate of interest (relative price differential between stages) becomes smaller. Under the “invariable money” assumption, the point of equilibrium of the system moves along the line of aggregate spending EE from point (C_1, I_1, i_1) to (C_2, I_2, i_2) in figure 6. The shape of the structure changes: its height diminishes (since C becomes smaller) and its length augments (since I increases), $C + I = E$ remaining a constant. This lengthening from λ_1 to λ_2 leads to an increase of the annual quantity of consumer goods produced from Q_{C1} to Q_{C2} , this increase constituting of course the phenomenon of growth.

Growth does not happen immediately because, at first, economic agents give up a part of their (real) consumption in order to finance the additional

Figure 6
Growth Induced by a Lowering of the Preference for the Present



investment and the intertemporal restructuring of capital. After a while, when the new structure is finally functional, the quantity of consumer goods produced begins to rise and ends up above the former level of production because the longer production process is more productive. The transition period (when production Q_C goes down) is not represented in figure 6. Another remark: this kind of growth may to a certain extent be endogenous. Once people have become richer after a first episode of growth, they may decide to save and invest more. In this case, their time preference is again reduced and gives birth to a new process of growth. It may therefore happen that a first lowering of the preference for the present triggers a second one, then a third one, and so on.

Technical progress

Technical progress is very easily represented by a rotation clockwise of the production function. Figure 7 shows that this rotation implies an increased annual production Q_C for a structure of a given length.

The question is: will this growth through technical progress change the shape of the structure of production? This shape will change only if there is a modification of time preference (monetary interferences are excluded here and will be treated below in "Money and the Business Cycle"). Technical progress has made people richer (on average). Will this induce a change in their time preference? Three cases may occur: (1) "neutral" technical progress that does not change time preference, (2) technical progress favoring investment (lower time preference), and (3) technical progress favoring consumption (higher time preference). The basic case is the "neutral" technical progress represented in

Figure 7
Growth Induced by "Neutral" Technical Progress

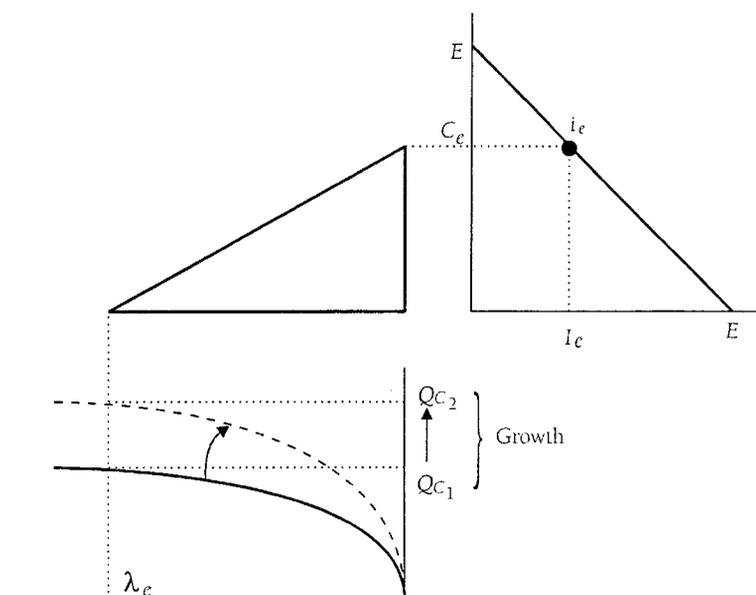
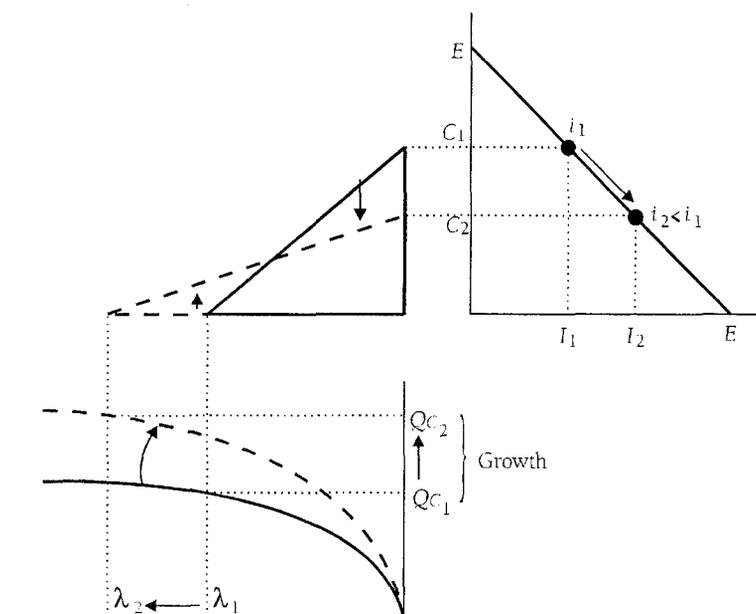


figure 7: thanks to technical progress, agents are richer in the present but also in the future, and they choose not to change their relative evaluation of present and future goods; the rate of interest and the ratio C/I do not vary. The case of technical progress favoring investment is represented in figure 8: agents have become richer and decide to save more; their time preference has thus diminished, which implies a lowering of the interest rate and of the ratio C/I ($C + I$ remaining constant according to the “invariable money” assumption). Growth then comes from two sources adding up: first from technical progress, and second from a lowering of time preference.⁸

The models of figures 7 and 8 may form the basis of a theory of endogenous growth, i.e., of a growth that persists, at least temporarily, because of its own consequences: a lowering of time preference may bring about another lowering (see above) and technical progress may provoke a lowering of time preference. In these ways, an initial shock to time preference or to technology may lead the economic system along a path of growth. But in all likelihood this growth would not be long-lasting. A new exogenous lowering of time preference or (above all) new technical progress would be required in order to bring additional increases of the annual production of consumer goods.

Figure 8
Growth Induced by Technical Progress Favoring Investment



⁸A change of time preference has not the same meaning as in the previous Subsection. In the present context, it does *not* mean that the maps of intertemporal preferences (the intertemporal utility functions) of the agents have changed, but rather that in their new economic situation—after the growth—the agents decide to alter their ratio of consumption to investment (a lower time preference implying a lower ratio C/I , and conversely).

Figure 9
Growth Induced by Technical Progress Favoring Consumption

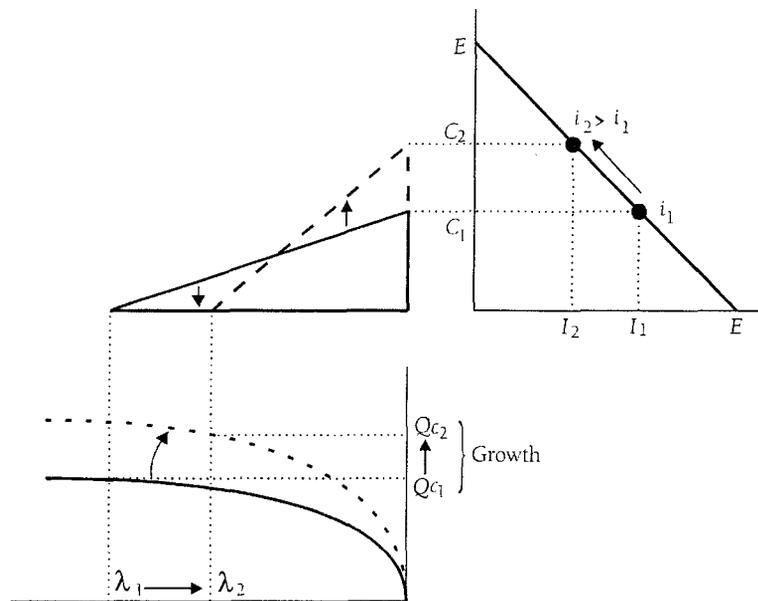


Figure 9 shows the case when technical progress favors consumption instead of investment: while people become richer thanks to technical progress, they decide to spend relatively more on consumption. The structure of production shortens and the rate of interest rises. As a result, growth is not as important as it could have been since the effect of the higher time preference is subtracted from the effect of technical progress.

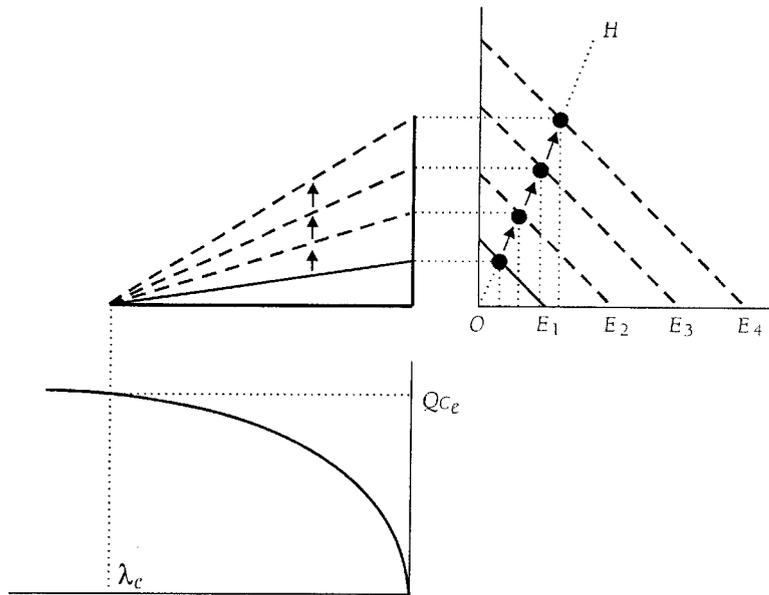
MONEY AND THE BUSINESS CYCLE

In a system with paper money, the central bank expands the quantity of money mainly by means of open market operations and of cutbacks of its discount rate.⁹ If this expansion could be “neutral,” then all prices would rise at the same rate.¹⁰ Even though the increase in the quantity of money is never

⁹In an open market operation, the central bank buys Treasury bills with a newly created standard money that is thus injected in the economic system (and gives birth to a pyramid of fiduciary media through the channel of the banks where it is deposited). When the Central Bank lowers its discount rate, it becomes easier for commercial banks to borrow money from it and then to create new fiduciary money.

¹⁰The concept of “neutrality” is used here in the sense of an equiproportional effect on all the prices of the increase of the quantity of money (see the discussion in Horwitz 2000, pp. 96-97). From the macroeconomic point of view, a much weaker form of “neutrality” is required in which the aggregate prices of factors at each stage are affected equiproportionally.

Figure 10
Effects of Successive and “Neutral” Increases of the Quantity of Money



“neutral” in this sense, it is instructive to study this simple case before approaching the more complex case where the expansion of money is not “neutral” and disturbs the intertemporal equilibrium of the system, giving birth to boom and bust.

A “Neutral” Monetary Expansion

If the quantity of money increases at rate $m\%$ from one year to the next, and if this increase is neutral, then all prices increase at rate $m\%$. As a result, consumption spending C , investment spending I and aggregate expenditure E also rise at rate $m\%$. Figure 10 illustrates the effects of three successive and neutral expansions of the quantity of money: the line of aggregate expenditure EE moves in the upper-right direction; the ratio C/I does not change and the point of equilibrium of the system remains therefore on the same line OH ; the structure expands in the vertical direction at the same rates as the quantity of money; the quantity of consumer goods annually produced Q_C does not vary since only nominal values are modified;¹¹ the nominal rate of interest includes a price premium equal to the increase in the quantity of money and in price.

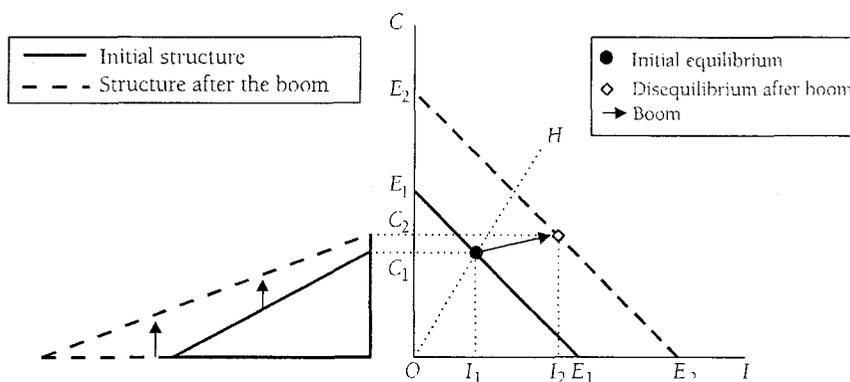
¹¹Some of the costs of inflation might be taken into account (“shoe-leather” costs and “menu” costs), through which the production function would slightly rotate counterclockwise (see Horwitz 2000, pp. 106-07, and for a more detailed treatment Horwitz 2003).

Boom and Bust

Let us suppose that a monetary expansion is not “neutral” and that the new money injected in the system is lent by the banks to businesses.¹² This flow of loanable funds tends to lower the ratio C/I and the rate of interest below their long-term equilibrium values (the nominal rate of interest becomes smaller than the sum of its long-term equilibrium value and the price premium). If this monetary expansion lasts for a sufficiently long time, then entrepreneurs begin to carry out the reallocation of convertible factors of production toward the early stages of production where the profitability has raised more than in the late stages. Figure 11 illustrates this boom. The point representing the state of the system in the right panel follows a path that deviates from the equilibrium line OH : the system follows a trajectory that is beneath this equilibrium line and leads to a point where the rate of return on investment and the ratio C/I are below their equilibrium value ($C_2/I_2 < C_1/I_1$). The point reached by the system (represented by the symbol \diamond in figure 11) is a disequilibrium point. The structure of production expands more in the horizontal than in the vertical direction. The function of production (panel 3) is not represented in figure 11 because it has no meaning outside of equilibrium.¹³

This boom can last for a few months or years, but it cannot last forever because there finally comes a time when the credit expansion slows down or stops altogether (if this slowing down or braking was not decided by the monetary authorities, the economic system would run the risk of suffering from

Figure 11
The Boom

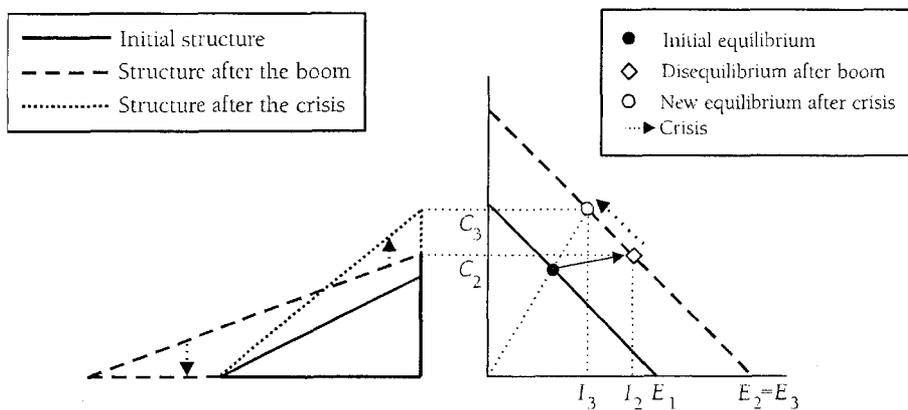


¹²Standard accounts of the Austrian theory of business cycle are found in Hayek (1931), Mises (1998, chap. 20) and Rothbard (1962, chap. 12).

¹³The Böhm-Bawerkian function of production is constructed in a frame of comparative statics. The structure of production, on the other hand, represents flows of spending that exist in equilibrium as well as in disequilibrium. The structure after the boom in figure 11, for instance, is in disequilibrium.

higher and higher levels of inflation, and even from hyperinflation). As soon as this is the case, the agents restore their preferred ratio C/I , which is higher than it was during the boom ($C_3/I_3 = C_1/I_1 > C_2/I_2$). The factors of production (especially labor and land) that had been allocated to early stages suddenly lose a large part of their value. Their profitableness can only be reestablished through a new allocation, this time toward the late stages (i.e., toward the right of the structure). The crisis brings the system back to equilibrium (figure 12) and is characterized by a waste of capital goods and a temporary unemployment.

Figure 12
The Crisis
(With Constant Aggregate Expenditure $E_2 = E_3$)



Deflationist and Stagflationist Crises

In the case of figure 12 above, aggregate expenditure E remains constant during the crisis. This is the simplest case, but of course not the most likely. In figure 13, the cycle ends up with a *deflationist* crisis: aggregate spending diminishes during the crisis, for instance because of a massive destruction of fiduciary money due to a series of bank failures in a system of gold money with fractional reserve banking.

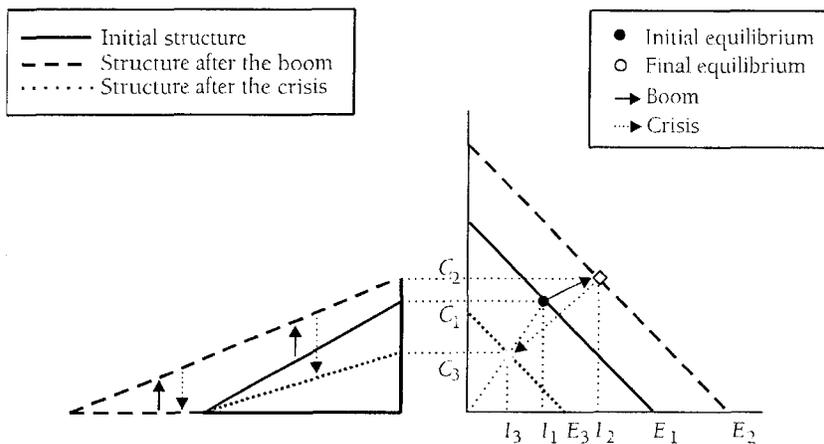
In figure 14, the cycle ends up with an inflationist or a *stagflationist* crisis (unemployment and prices go up together): the quantity of money and aggregate expenditure keep increasing during the crisis, but not enough to prevent the crisis, i.e., to prevent the reallocation of factors from the early to the late stages.

“Non-neutral” Cycle

All the cycles that have been represented until now were “neutral”¹⁴ in the sense that the equilibrium ratio C/I and the length of the structure were the

¹⁴Three very different notions of neutrality have been used in this paper: the neutrality of technical progress, the neutrality of money, and the neutrality of cycles. They obviously must not be confused.

Figure 13
A Cycle With a Deflationist Crisis



same before the boom and after the crisis (the quantity produced Q_C therefore remained the same before and after the cycle). This simplifying assumption may of course be suppressed and more complicated cases can be illustrated. Figure 15 presents the case of a cycle during which time preference becomes higher, whether because of a distribution effect (more riches accrues to agents with a high time preference) or because of an increased distrust of the future. The final structure is shorter than the initial one, the equilibrium ratio C/I is higher after the crisis than before the boom, and the quantity of consumer goods annually produced is consequently reduced (from Q_{C1} to Q_{C3}).

FINAL REMARKS

The “values-riches” model permits an easy and gradual presentation of some of the main theories of Austrian macroeconomics. It may serve as an introduction to the much more detailed works of Skousen (1990), Reisman (1996),

Figure 14
A Cycle With a Stagflationist Crisis

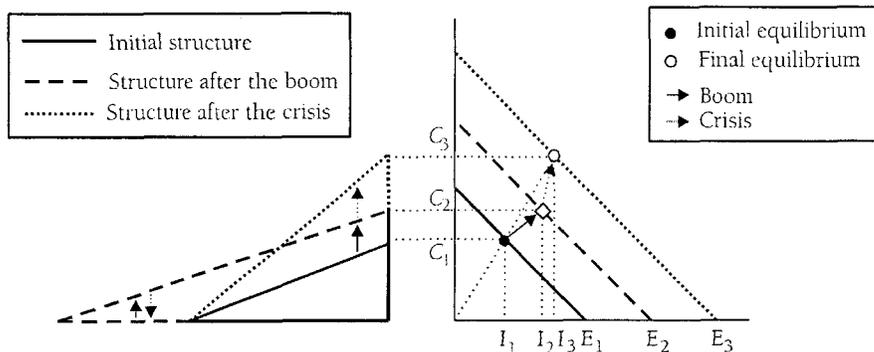
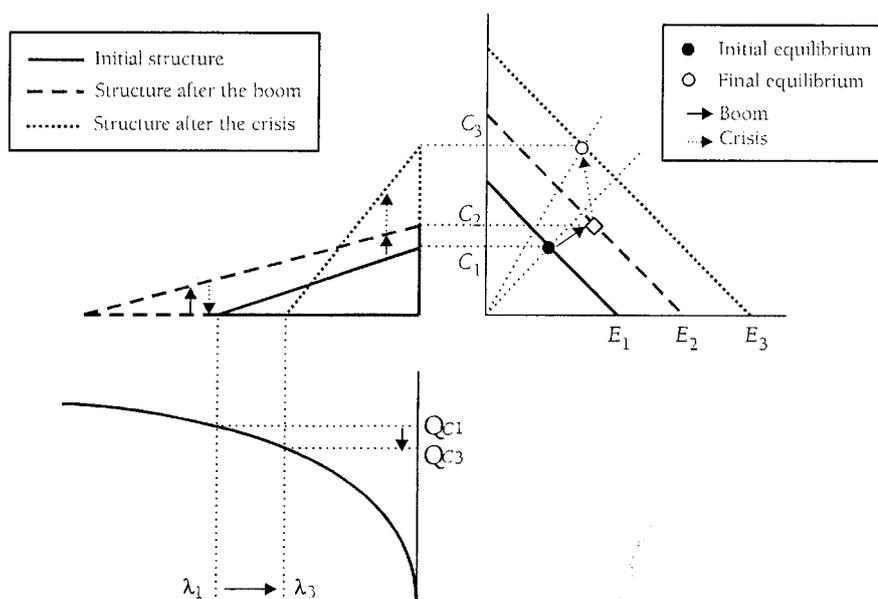


Figure 15
A "Non-neutral" Cycle



Horwitz (2000) and Garrison (2001). But because of its simplicity, this model suffers from a number of shortcomings. Here are three of them. First, the illustration of the structure of production as a triangle, even if it may suffice in a first approach, should be supplemented by a more rigorous representation, in order to be able to draw accurately the changes in the shape of the structure and to calculate the average length of the structure (Fillieule 2004). Second, the movements from one equilibrium to another or from equilibrium to disequilibrium and back again should be analyzed more thoroughly by displaying the underlying microeconomic processes. And third, the values-riches model, contrary to Garrison's model, does not allow to know how (real) consumption varies during the business cycle. It thus cannot answer the question as to whether real consumption increases during the boom, as Garrison asserts (2001, p. 69), or as Hülsmann maintains (2001, p. 39), this is not the case.

REFERENCES

- Böhm-Bawerk, Eugen von. 1959. *Positive Theory of Capital*. 4th ed. South Holland, Ill.: Libertarian Press.
- Fetter, Frank A. 1915. *Economic Principles*. New York: Century.
- . 1914. "Interest Theories, Old and New." *American Economic Review* 4. Reprinted in Frank A. Fetter. *Capital, Interest, and Rent: Essays in the Theory of Distribution*. Kansas City: Sheed Andrews and McMeel, 1977.
- . 1904. *The Principles of Economics*. New York: Century.

- Fillieule, Renaud. 2004a. "A Formal Model of the Hayekian Structure of Production." Unpublished paper.
- Garrison, Roger W. 2001. *Time and Money: The Macroeconomics of the Capital Structure*. London: Routledge.
- . 1986. "A Subjectivist Theory of a Capital-using Economy." In *The Economics of Time and Ignorance*. Gerald P. O'Driscoll, Jr., and Mario J. Rizzo. Oxford: Basil Blackwell.
- Hayek, Friedrich A. 1941. *The Pure Theory of Capital*. Chicago: University of Chicago Press.
- . 1931. *Prices and Production*. London: Routledge.
- Horwitz, Steven. 2003. "The Costs of Inflation Revisited." *Review of Austrian Economics* 16 (1): 77-95.
- . 2000. *Microfoundations and Macroeconomics: An Austrian Perspective*. London: Routledge.
- Hülsmann, Jörg Guido. 2001. "Garrisonian Macroeconomics." *Quarterly Journal of Austrian Economics* 4 (3): 33-41.
- Mises, Ludwig von. 1998. *Human Action: A Treatise on Economics*. 1st ed. reprint. Auburn, Ala.: Ludwig von Mises Institute.
- Ricardo, David. 1951. *On the Principles of Political Economy and Taxation*. 3rd ed. Cambridge: Cambridge University Press.
- Reisman, George. 1996. *Capitalism: A Treatise on Economics*. Ottawa, Ill.: Jameson Books.
- Rothbard, Murray N. 1962. *Man, Economy, and State: A Treatise on Economic Principles*. Princeton, N.J.: D. Van Nostrand.
- Salerno, Joseph. 2001. "Does the Concept of Secular Growth Have a Place in Capital-Based Macroeconomics." *Quarterly Journal of Austrian Economics* 4 (3): 43-61.
- Samuelson, Paul A., and William D. Nordhaus. 1989. *Economics*. 13th ed. New York: McGraw-Hill.
- Skousen, Mark. 1990. *The Structure of Production*. New York: New York University Press.

« A formal model of the Hayekian
structure of production » (2005)

Article non publié

*A Formal Model of the Hayekian
Structure of Production*

Renaud Fillieule

Université de Lille 1 (USTL), CLERSÉ

July 28, 2005

Abstract. With the exception of Reisman (1996), the economists of the Austrian school use rudimentary representations of the structure of production. And yet, in order for Austrian macroeconomics to become a convincing alternative to standard macroeconomics, this core concept should be carefully formalized. The model developed in this paper permits (under appropriate simplifying assumptions): (1) to calculate the values characterizing a structure in static equilibrium, (2) to illustrate a structure and to visualize the distribution of the added value, (3) to demonstrate a series of mathematical formulas describing a structure (for instance the formula giving the average length of a structure), and (4) to analyze the complex phenomena that occur in a dynamic process during which the shape of the structure changes.

The core concept of Austrian macroeconomics is the structure of production, which makes it possible to analyze the intertemporal dimension of an economic system and to understand aspects of equilibrium, growth and trade cycle that are neglected by usual–Keynesian and Neoclassical–macroeconomic models. This concept and a very convenient graphical illustration have been introduced by Hayek (1935, 1941), and then developed mainly by Rothbard (1962), Skousen (1990), Reisman (1996) and Garrison (2001). Reisman has offered the most detailed exposition and most thoroughgoing utilization of the concept of structure of production so far. The present paper seeks to formalize and to generalize some aspects of his approach by presenting, first a systematic method for the calculation and illustration of the structures (under suitable simplifying assumptions), second a number of mathematical formulas previously unnoticed that relate the main macroeconomic variables

(consumption and investment spending, aggregate wages, interest, average length of the structure), and third an application of the model to the issue of macroeconomic dynamics, i.e. to the study of the deformation of the structure.

1. THE STRUCTURE OF PRODUCTION IN STATIC EQUILIBRIUM

1.1 Hypotheses

The institutional setting is a market economy, in which production is carried out through successive stages. Consumer goods are produced by combining the two kinds of factors of production, ordinary factors and capital goods. Ordinary factors are factors that have not been produced, i.e. labor and land (ground and natural resources). Capital goods are produced means of production, resulting from a combination, carried out at the previous stage of production, of ordinary factors and of capital goods. The capital goods of this previous stage are themselves produced with the help of ordinary factors and capital goods used at the preceding stage, and so on *ad infinitum*¹ (see figure 1).

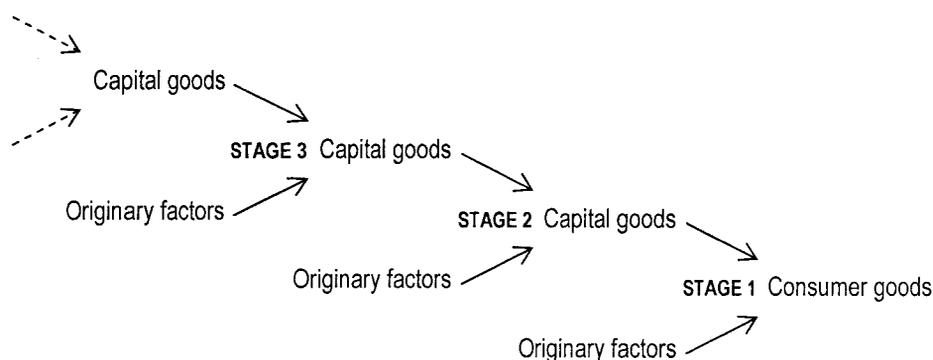


Figure 1. The stages of a structure of production

¹ The structures that will be described below have an infinite number of stages. This hypothesis is obviously unrealistic since after a number of stages the values become negligible and have no meaning anymore in the realm of human planning and acting. But this hypothesis permits to resort to a mathematical apparatus whose usefulness should be clear at the end of the paper.

In static equilibrium, the same structure recurs year after year. There is no entrepreneurship and only two kinds of economic agents or functions exist, owners of ordinary factors of production and capitalists. Capitalists advance wages and rents to owners of ordinary factors, buy capital goods from the capitalists of the previous stage, combine these inputs in a production process, and wait until the end of this process to get their net income which is an interest on invested capital. The following simplifying hypotheses will be adopted all along.²

(1) Each stage of production lasts one year.

(2) All the annual exchanges take place on the first of January. The services of the factors of production are entirely paid at the beginning of each stage (i.e. of each year) and the totality of the product is sold exactly one year later. Hypotheses (1) and (2) imply that the equilibrium rate of interest is unique and that the price of a produced good is $p_G = (1 + i)p_F$, where p_F is the price of its factors and i the annual equilibrium rate of interest.

(3) There are no durable capital goods: they are entirely used up during a stage of production. Durable ordinary factors (land) are rented each year by their owners; they are not sold at their capitalized value. In other words, this paper endorses the “circulating capital” (or “goods in process”) approach, typical of the Austrian perspective, rather than the “fixed capital” approach (Hayek 1941, p. 47).

(4) The proportion of the value of ordinary factors to investment is a constant across stages. Let n be a given stage of production. At this stage n , the total investment spending is I_n , and it is divided among an expenditure I_{OFn} in ordinary factors and an expenditure I_{KGN} in capital goods: $I_n = I_{OFn} + I_{KGN}$. Hypothesis (4) means that the ratio $a_n = I_{OFn}/I_n$ is a constant: whatever the stage of production n , $a_n = a$. This is obviously the simplest hypothesis that can be made.³

² Reisman (1996) uses all of these hypotheses in his macroeconomic theory, including hypothesis 4 (the latter in a slightly different form; see Reisman 1996, p. 844-846).

³ As Hayek (1941, p. 124) wrote: “It is perhaps reasonable to assume that the amount of input which is applied to the stock of intermediate products in each stage will bear a constant proportion to the amount of those intermediate products (or, in more popular but more inexact terminology, that the proportion between capital and labour will be roughly the same in all stages).”

1.2 Determination of the structure

In static equilibrium and under the hypotheses above (particularly hypothesis 4), a structure is entirely determined by the three parameters C , i and a , where C is the annual aggregate spending on consumer goods, i the annual ordinary rate of interest, and a the ratio of ordinary factors to investment at each stage. The calculation begins at stage 1 and then proceeds backwards to stages 2, 3, 4, and so on (see figure 1).

Stage 1. The factors of production of stage 1 produce the consumer goods. Consumer goods have an annual aggregate value C . They are produced by the factors of production of stage 1, whose aggregate value is I_1 (investment at stage 1). In equilibrium, the sale of consumer goods just suffices to cover the expenses of production I_1 and the interest income iI_1 of the capitalists of stage 1: $C = I_1(1 + i)$. The value of I_1 is thus easily deduced from those of C and i .

Stage 2. The factors of stage 2 produce the capital goods of stage 1. Among the factors of production of stage 1, by definition only the capital goods have been produced. Their aggregate value is $I_{KG1} = (1 - a)I_1$ (since it follows from hypothesis (4) that the investment I_1 of stage 1 is the sum of an expense aI_1 in ordinary factors and an expense $(1 - a)I_1$ in capital goods). The capital goods of stage 1 are produced in the course of stage 2 by the factors of production of stage 2. In equilibrium, the sale of the capital goods of stage 1 I_{KG1} just suffices to cover the productive expenses I_2 of stage 2 and the interest income iI_2 on investment: $I_{KG1} = I_2(1 + i)$, or equivalently $(1 - a)I_1 = I_2(1 + i)$, and the value of I_2 is deduced from those of I_1 , i and a .

Stage 3. The factors of stage 3 produce the capital goods of stage 2. The capital goods of stage 2 have an aggregate value $I_{KG2} = (1 - a)I_2$. They are produced by the factors of production used in stage 3. In equilibrium, the sale of the capital goods of stage 2 I_{KG2} covers the expenses I_3 of stage 3 and the interest income iI_3 : $I_{KG2} = I_3(1 + i)$, or $(1 - a)I_2 = I_3(1 + i)$, and the value of I_3 is deduced from those of I_2 , i and a .

Stage n ($n \geq 2$). The factors of stage n produce the capital goods of stage $(n - 1)$: $(1 - a)I_{n-1} = I_n(1 + i)$.

It is possible in this way to calculate step by step all the values that characterize a structure of pro-

duction.

1.3 Calculation

Table 1 shows the data of the first ten stages of a structure characterized by a total annual consumption spending $C = 100$ monetary units (for instance: 1 unit = 10 billions \$ or €), an equilibrium annual rate of interest $i = 10\%$, and a ratio $a = 20\%$ of ordinary factors at each stage.

Table 1. The first ten stages of the structure ($C = 100$, $i = 10\%$, $a = 20\%$)

Stage	Investment	Investment in ordinary factors	Investment in capital goods	Interest
n	I_n	$I_{OFn} = aI_n$	$I_{KGN} = (1 - a)I_n$	$Y_{in} = iI_n$
1	90.91	18.18	72.73	9.09
2	66.12	13.22	52.89	6.61
3	48.08	9.61	38.47	4.81
4	34.97	6.99	27.98	3.50
5	25.43	5.09	20.35	2.54
6	18.50	3.70	14.80	1.85
7	13.45	2.69	10.76	1.34
8	9.78	1.96	7.83	0.98
9	7.11	1.42	5.69	0.71
10	5.17	1.03	4.14	0.52

At each stage n , aggregate investment I_n is calculated with the help of the algorithm of the previous Subsection (C and i determine I_1 ; I_1 , i and a determine I_2 ; I_2 , i and a determine I_3 , etc.). The investments in ordinary factors and in capital goods are respectively deduced from the formulas $I_{OFn} = aI_n$ and $I_{KGN} = (1 - a)I_n$. The interest income Y_{in} on investment I_n is $Y_{in} = iI_n$. The values of table 1 have

been calculated with a computer program. Numbers are rounded at the second decimal, which explains why some equalities are not verified.

Three other important aggregate variables can be calculated:

– the total annual investment spending⁴ $I = I_1 + I_2 + I_3 + \dots + I_n + \dots = \sum I_n$; for the structure ($C = 100$, $i = 10\%$, $a = 20\%$), this infinite sum converges⁵ towards $I = 333.33$,

– the total annual income Y_{OF} of the owners of ordinary factors (aggregate wages and rents): $Y_{OF} = I_{OF1} + I_{OF2} + I_{OF3} + \dots + I_{OFn} + \dots = aI_1 + aI_2 + aI_3 + \dots + aI_n + \dots = a\sum I_n = aI = 0.2 \times 333.33 = 66.67$,

– and the total annual net income Y_i of the capitalists (aggregate interest): $Y_i = iI_1 + iI_2 + iI_3 + \dots + iI_n + \dots = i\sum I_n = iI = 0.1 \times 333.33 = 33.33$.

1.4 Illustration

A typical stage of production is represented in figure 2. Time runs from left to right. Nominal values are measured along the vertical axis. The vertical side on the left of the quadrangle is the investment spending I_n of stage n , split into an investment in ordinary factors I_{OFn} and an investment in capital goods I_{KGn} . The vertical side on the right is the gross income Y_n of the capitalists of stage n (it is equal to the value $I_{KG(n-1)}$ of the capital goods that are bought by the capitalists of stage $n-1$). The net income of the capitalists of stage n is $Y_m = iI_n$ (interest on investment).

⁴ This investment comprises the expenditures made in order to buy all of the factors of production, including intermediate goods. It is of course much larger than the narrow “investment” used in the calculation of the GNP (see Rothbard 1962, p. 340-341, Skousen 1991, Chapter 4, Reisman 1996, Chapter 15).

⁵ When more than 40 stages are taken into account, the calculation with the computer program of the sum $I_1 + I_2 + \dots + I_n$ ($n \geq 41$) gives exactly the same result up to the third decimal (= 333.333).

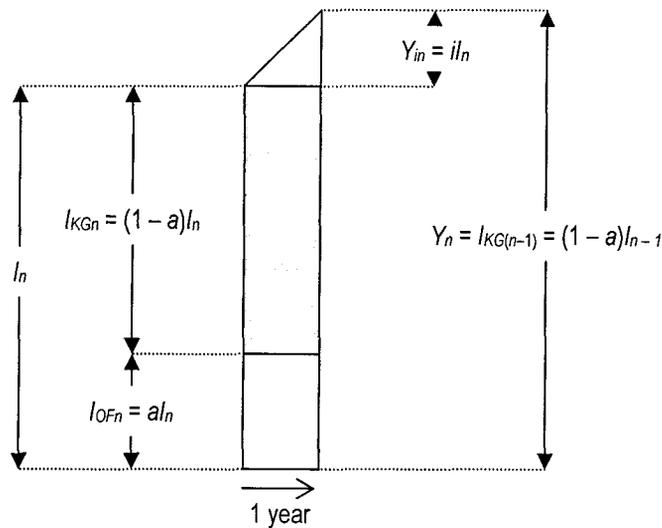


Figure 2. A typical stage of production n

The whole structure ($C = 100$, $i = 10\%$, $a = 20\%$) is illustrated in figure 3. The vertical sides of the white rectangles show the value of the ordinary factors added at each stage, and the horizontal sides show the duration of the process in which these factors are used (one year). The grey rectangles show the value and period of production of the capital goods used at each stage. The vertical side of the white triangles show the interest accruing to capitalists at the end of each stage. The sum of the interest incomes through all stages is Y_i , the aggregate interest received by the capitalists in a year. The sum of the values of ordinary factors through all stages is the aggregate value Y_{OF} received by their owners in a year. The annual spending on consumption C is the sum of the two aggregate incomes Y_i and Y_{OF} . The distribution of the added value between the capitalists and the owners of ordinary factors appears clearly on the figure.

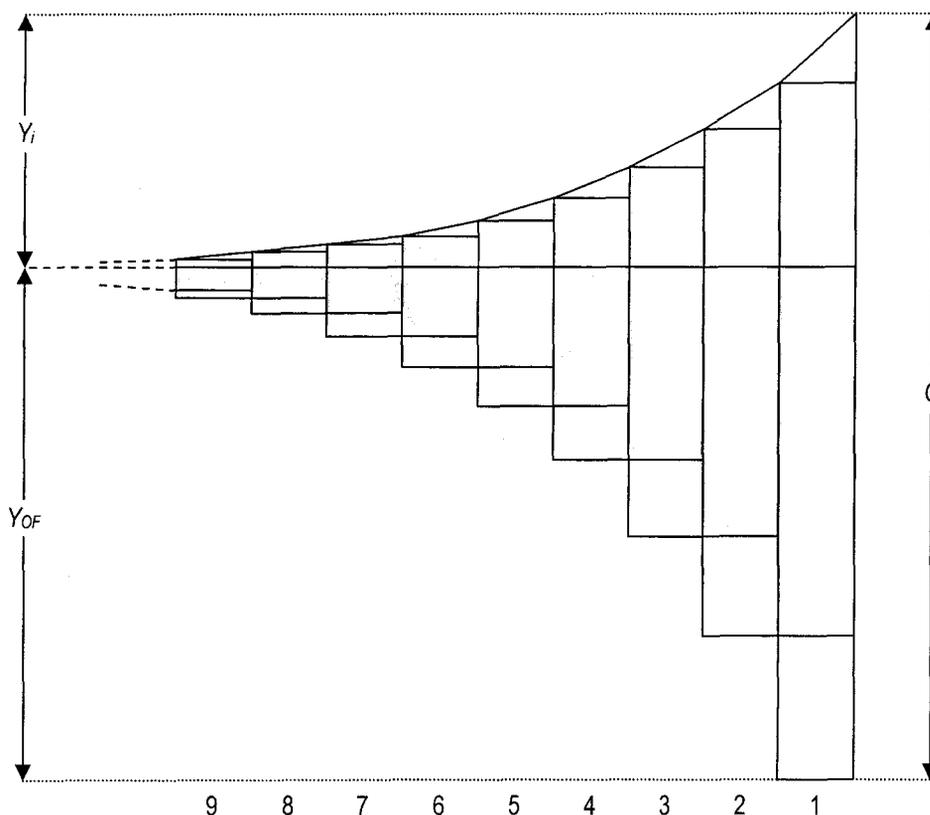


Figure 3. An equilibrium structure de production ($C = 100, i = 10\%, a = 20\%$)

This structure may be interpreted in two ways (Hayek 1941, p. 115-116, Skousen 1990, p. 196-197). First, as the temporal unfolding of the production process of a stationary system during an infinity of successive periods⁶ (successive stages belonging to successive years). Second, and more importantly, as a depiction of the simultaneous or “synchronized” processes of production of all the stages of production during a given year.

2. A MATHEMATICAL DESCRIPTION OF THE STRUCTURE

Austrian economists are usually a bit skeptical about the use of mathematics in economics, and rightly so. It may for instance be argued, from an Austrian point of view, that mathematical formaliza-

⁶ Concerning the infinite number of stages, see the note 1 above.

tion has done a lot of harm in microeconomics. The works of Mises (1998) and Hayek (1948) on price theory, competition, and collectivist planning do not use any mathematical formula, and provide nevertheless a much deeper understanding of the economic process than most—if not all—of the mathematical models piled up in contemporary standard microeconomics textbooks. But when formalization can be used within a valid and pertinent theory, such as the macroeconomics of the structure of production, it must not be rejected because, as will be seen below, it can disclose important relations that would not be available through a purely qualitative analysis.

2.1 The values at a stage n

Annual investment I_n at any stage n may be expressed as a function of the three parameters of the structure (C, i, a) and of the number n . Since:

$$I_1 = \left(\frac{C}{1+i} \right)$$

$$I_2 = I_1 \left(\frac{1-a}{1+i} \right) = \left(\frac{C}{1+i} \right) \left(\frac{1-a}{1+i} \right)$$

$$I_3 = I_2 \left(\frac{1-a}{1+i} \right) = \left(\frac{C}{1+i} \right) \left(\frac{1-a}{1+i} \right)^2$$

it is obvious that:

$$I_n = \left(\frac{C}{1+i} \right) \left(\frac{1-a}{1+i} \right)^{n-1}$$

The values of annual investment in originary factors and in capital goods at stage n are then easily

calculated:

$$I_{OFn} = aI_n = \left(\frac{aC}{1+i}\right)\left(\frac{1-a}{1+i}\right)^{n-1}$$

$$I_{KGn} = (1-a)I_n = C\left(\frac{1-a}{1+i}\right)^n$$

The gross income Y_n and net income Y_{in} of the capitalists of stage n are:

$$Y_n = I_n(1+i) = C\left(\frac{1-a}{1+i}\right)^{n-1}$$

$$Y_{in} = iI_n = \left(\frac{iC}{1+i}\right)\left(\frac{1-a}{1+i}\right)^{n-1}$$

2.2 The aggregate values

Total or aggregate investment I is the sum of the investments made throughout all the stages of production:

$$I = I_1 + I_2 + I_3 + \cdots + I_n + \cdots = \sum_{n=1}^{\infty} I_n$$

In order to calculate I , let us use the formula, given above, of I_n as a function of C , i , a and n :

$$I = \sum_{n=1}^{\infty} I_n = \frac{C}{1+i} \sum_{n=1}^{\infty} \left(\frac{1-a}{1+i}\right)^{n-1}$$

The infinite sum $(1 + q + q^2 + q^3 + \dots + q^n + \dots)$ converges⁷ towards $1/(1 - q)$, if $0 \leq q < 1$. It suffices to define $q = (1 - a)/(1 + i)$ in order to reach the formula of total investment:

$$I = \frac{C}{1+i} \frac{1}{\left(1 - \frac{1-a}{1+i}\right)} = \frac{C}{i+a}$$

The final formula is surprisingly simple:

$$I = \frac{C}{i+a}$$

Investment spending I is directly proportional to consumption spending C and inversely proportional to the sum of the annual rate of interest i and the ratio a of ordinary factors at each stage.

The three other aggregate variables are the total annual interest Y_i , the aggregate income Y_{OF} of the owners of ordinary factors, and the aggregate income Y_{KG} of the owners of capital goods:

$$Y_i = \sum_{n=1}^{\infty} iI_n = i \sum_{n=1}^{\infty} I_n = iI = \frac{iC}{i+a}$$

$$Y_{OF} = \sum_{n=1}^{\infty} I_{OFn} = \sum_{n=1}^{\infty} aI_n = a \sum_{n=1}^{\infty} I_n = aI = \frac{aC}{i+a}$$

$$Y_{KG} = \sum_{n=1}^{\infty} I_{KGn} = \sum_{n=1}^{\infty} (1-a)I_n = (1-a) \sum_{n=1}^{\infty} I_n = (1-a)I = \frac{(1-a)C}{i+a}$$

⁷ Proof : define $s_n(q) = 1 + q + q^2 + \dots + q^n$. Then $qs_n(q) = q + q^2 + q^3 + \dots + q^{n+1}$, and $s_n(q) - qs_n(q) = 1 - q^{n+1}$. Therefore : $s_n(q) = (1 - q^{n+1})/(1 - q)$. If $0 \leq q < 1$ and if n becomes higher and higher, then q^{n+1} converges towards 0, and $s_n(q)$ thus converges towards $1/(1 - q)$.

Taking into account the relation $I = C/(i + a)$, it is perfectly equivalent to define a structure with parameters (C, i, a) , or with (C, I, i) , or with (I, i, a) , or else with (C, I, a) . Since $a = (C/I) - i$, aggregate incomes Y_i , Y_{OF} and Y_{KG} can for instance easily be expressed as functions of C, I and i :

$$Y_i = iI$$

$$Y_{OF} = aI = \left(\frac{C}{I} - i \right) I = C - iI$$

$$Y_{KG} = (1 - a)I = \left[1 - \left(\frac{C}{I} - i \right) \right] I = I(1 + i) - C$$

Aggregate interest Y_i can also be written as a function of C, I and a :

$$Y_i = iI = \left(\frac{C}{I} - a \right) I = C - aI$$

Since aI is the final consumption of the owners of ordinary factors ($aI = Y_{OF}$), and since in equilibrium the total final consumption is divided between these owners and the capitalists, then $(C - aI)$ is the final consumption of the capitalists, or what Reisman calls “net-consumption.”⁸ In equilibrium, aggregate interest is equal to net-consumption (Reisman 1996, Chapter 16).

2.3 The average period of production

Since the number of stages is infinite, the *total* period of production is the sum of an infinite number of years and is thus also infinite. Böhm-Bawerk (1959, p. 86) defines the average period of pro-

⁸ “Considered substantively, and essentially, net consumption is *the consumption expenditure of businessmen and capitalists*” (Reisman 1996, p. 725).

duction⁹ as the “*average* time interval occurring between each expenditure of ordinary productive forces and the final completion of the ultimate consumption good.” If each stage lasts one year, and if labor is homogeneous and is the only ordinary factor of production, then the average time interval between the expenditure of the quantity of labor L_1 used at the first stage and the production of the final consumption good is 1 year $\times (L_1/L)$, L being the total quantity of labor. The average time interval between the expenditure of the quantity of labor L_2 used at the second stage and the final consumption good is 2 years $\times (L_2/L)$, etc., the average time interval between L_n and the final consumption good is n years $\times (L_n/L)$, and so on. The sum of all these products is the Böhm-Bawerkian average period of production λ :

$$\lambda = \sum_{n=1}^{\infty} n \left(\frac{L_n}{L} \right)$$

Under the assumptions of the formal model, the ordinary factors used at stage n are measured by their aggregate value I_{OFn} , and the total quantity of ordinary factors is measured by the total investment in ordinary factors $I_{OF} = Y_{OF}$:

$$\lambda = \sum_{n=1}^{\infty} n \left(\frac{I_{OFn}}{Y_{OF}} \right)$$

I_{OFn} and Y_{OF} have previously been expressed as functions of C , i , a and n :

$$\lambda = \frac{1}{Y_{OF}} \sum_{n=1}^{\infty} n I_{OFn} = \frac{i+a}{aC} \sum_{n=1}^{\infty} n \left(\frac{aC}{1+i} \right) \left(\frac{1-a}{1+i} \right)^{n-1} = \frac{i+a}{1+i} \sum_{n=1}^{\infty} n \left(\frac{1-a}{1+i} \right)^{n-1}$$

Since, when $0 \leq q < 1$, the mathematical limit of the infinite sum $(1 + 2q + 3q^2 + \dots + nq^{n-1} + \dots)$

⁹ For recent applications of the concept of period of production to production theory, see Reisman (1996) and Fillieule (2005).

is¹⁰ $1/(1 - q)^2$, it suffices to replace q by $(1 - a)/(1 + i)$ in order to calculate the average period of production:

$$\lambda = \frac{i + a}{1 + i} \frac{1}{\left(1 - \frac{1 - a}{1 + i}\right)^2} = \frac{1 + i}{i + a}$$

Since $i + a = C/I$, an equivalent formula, again surprisingly simple, is:

$$\lambda = \left(\frac{I}{C}\right)(1 + i)$$

This formula is interesting in that it shows that a diminution of the rate of interest *by itself*—i.e. without any change in the ratio (I/C) —would lead to a *shortening* of the structure. Under the assumptions of this model, it would not be strictly correct to write that “If [the pure rate of interest] is lowered, the stock of invested capital will increase; if it is raised, the stock of invested capital will fall” (Rothbard 1962, p. 496). In fact, in order for the structure of production to become more roundabout, C must diminish or I must rise. And if the interest rate diminishes, the magnitude of the lowering of C and of the rise in I must be important enough to more than offset the effect of this diminution. It may be added that a “feeble” relative variation of (I/C) suffices to counterbalance an “important” relative variation of i . Consider a ratio $(I/C) = 2$, which fits with the data calculated by Skousen (1991, p. 45) for the economy of the United States. If the initial interest rate is 5%, then a 1% rise in (I/C) more than offsets a 20% decrease in the interest rate and leads to a lengthening of the structure of production.¹¹

¹⁰ I wish to thank Alain Fillieule for this result: if $s(q) = 1 + q + q^2 + q^3 + \dots + q^n + \dots = 1/(1 - q)$, then its derivative is $s'(q) = 1 + 2q + 3q^2 + \dots + nq^{n-1} + \dots = 1/(1 - q)^2$.

¹¹ $\lambda = (I/C)(1 + i) = 2(1 + 0.05) = 2.1000$ and $\lambda' = 2(1 + 0.01)[1 + 0.05(1 - 0.20)] = 2.1008 > \lambda$.

3. FORMAL MACROECONOMIC DYNAMICS

A rigorous formalization of the structure is necessary in order to analyze accurately what happens when the shape of the structure changes, for instance in the standard Austrian case when a lowering of time preference leads to a rise in I , and a decrease in C and in i (Hayek 1935, p. 44-54, Rothbard 1962, p. 470-473, Skousen 1990, p. 234-240, Garrison 2001, p. 61-67). Our inquiry will here be restricted to the formal aspects of the problem, or in other words to the pure logic of structural change.

Let us consider the following process:

- until year 2000, the structure of production (C, I, i) is in static equilibrium and remains the same year after year,
- during year 2001 a deformation of the structure occurs,
- this deformation gives birth to a new structure (C', I', i'),
- this new structure appears in 2001 and then remains the same during the following years.

It is possible first to calculate the (*ex post*) aggregate interest in 2001, equal to the gross income of the capitalists in 2001 less the costs of production that they have incurred in 2000. Their gross income in 2001 is the income received in exchange for the sale of consumer goods C' and for the sale of capital goods $Y'_{KG} = I'_{KG}$. The costs of production are the productive expenditures made in 2000, and are equal to the investment spending I in year 2000. Aggregate interest in 2001 is thus:¹²

$$\tilde{Y}_i = (C' + I'_{KG}) - I$$

Since $I'_{KG} = (1 - a')I'$ (see the Subsection “The aggregate values”):

$$\tilde{Y}_i = (C' - a'I') + (I' - I)$$

Aggregate interest is equal to the sum of the “net-consumption” of the *new* structure ($C' - a'I'$) plus

¹² The symbol “ \sim ” means that the value is calculated *between* two successive structures.

the net investment ($I' - I$). As Reisman has explained, when a net investment exists, “the amount of profit [called “aggregate interest” in the present paper] in the economic system turns out to equal the sum of net consumption plus net investment” (1996, p. 744).

With the help of the formal model, a second step can be taken in order to analyze the complex effects triggered on the rates of ordinary interest. The ordinary rate of interest of a stage n for year y is defined here as the *ex post* rate of return on investment. It is, in other words, the relative difference between the aggregate price of the products of this stage in year y and the aggregate price of the factors that have been used at the same stage n but bought during the *preceding* year ($y - 1$). In static equilibrium, under the assumptions of the model, there is one and only one ordinary rate of interest:

$$\forall n \geq 2, i_n = \frac{I_{KG(n-1)} - I_n}{I_n} = \frac{(1-a)I_{n-1} - I_n}{I_n} = i$$

In year 2000 and before, the ordinary rate of interest is unique ($= i$). In year 2002 and after, the ordinary rate is unique again ($= i'$). But during the year 2001, the year when structure (C', I', i') replaces structure (C, I, i) , *there are as many ordinary rates of interest as there are stages of production*. The ordinary rate of interest \tilde{i}_n of a stage n ($n \geq 2$) is the relative difference between the income from the production of stage n in 2001 ($= I'_{KG(n-1)} = (1-a')I'_{n-1}$) and the costs of production, incurred in 2000, of the factors of stage n ($= I_n$):

$$\tilde{i}_n = \frac{(1-a')I'_{n-1} - I_n}{I_n}$$

where I_n is the investment at stage n of the initial structure (C, I, i) , I'_{n-1} the investment at stage $(n - 1)$ of the new structure (C', I', i') , and a' the ratio of ordinary factors at each stage of the new structure. And for stage 1:

$$\tilde{i}_1 = \frac{C' - I_1}{I_1}$$

The rates \tilde{i}_n are the sums of two components, the initial rate of interest i on the one hand, and a transitional rate of pure profit (or pure loss) on the other. They may be written accordingly:

$$\tilde{i}_n = i + \frac{(1-a')I'_{n-1} - (1-a)I_{n-1}}{I_n} = i + \tilde{\pi}_n$$

In order to illustrate these formulas, let us consider the case of a lowering of the preference for the present that triggers an increase in I and a decrease in C and in i . The initial structure is for instance ($C = 100, I = 300, i = 12\%$) and the final structure ($C' = 90, I' = 310, i' = 10\%$).¹³ The aggregate spending ($C + I$) remains constant¹⁴ and the structure lengthens from $\lambda = (300/100)(1.12) = 3.36$ years to $\lambda' = (310/90)(1.10) = 3.79$ years. Table 2 presents the *ex post* originary rates of interest in year 2001, i.e. the rates between the two structures, for the first ten stages.¹⁵

Table 2. Originary rates of interest between the structures ($C = 100, I = 300, i = 12\%$)
and ($C' = 90, I' = 310, i' = 10\%$)

Stage n	1	2	3	4	5	6	7	8	9	10
Rate \tilde{i}_n	0.80	5.63	10.70	16.01	21.58	27.41	33.52	39.92	46.63	53.67
(%)										

¹³ These data, specially the values of the rates of interest, have been chosen in order to facilitate the graphical representation.

¹⁴ Reisman (1996) calls this assumption the “invariable money” hypothesis.

¹⁵ In order to reduce the risk of error, two distinct computer programs have been used, one based on the step by step calculation of the Subsection “Determination of the structure,” and another based on the mathematical formula of I_n in the Subsection “The values at stage n .” These two programs give the same results.

The originary rate of interest is the higher, the further the stage is from final consumption. It is thus easy to explain why the convertible originary factors of production will be reallocated towards the early—far from final consumption—stages: these factors have become relatively more profitable in the early than in the late stages. The two structures are displayed together in figure 4 (compare with the figures in Hayek 1935, p. 44 and 52, Rothbard 1962, p. 472, Garrison 2001, p. 62).

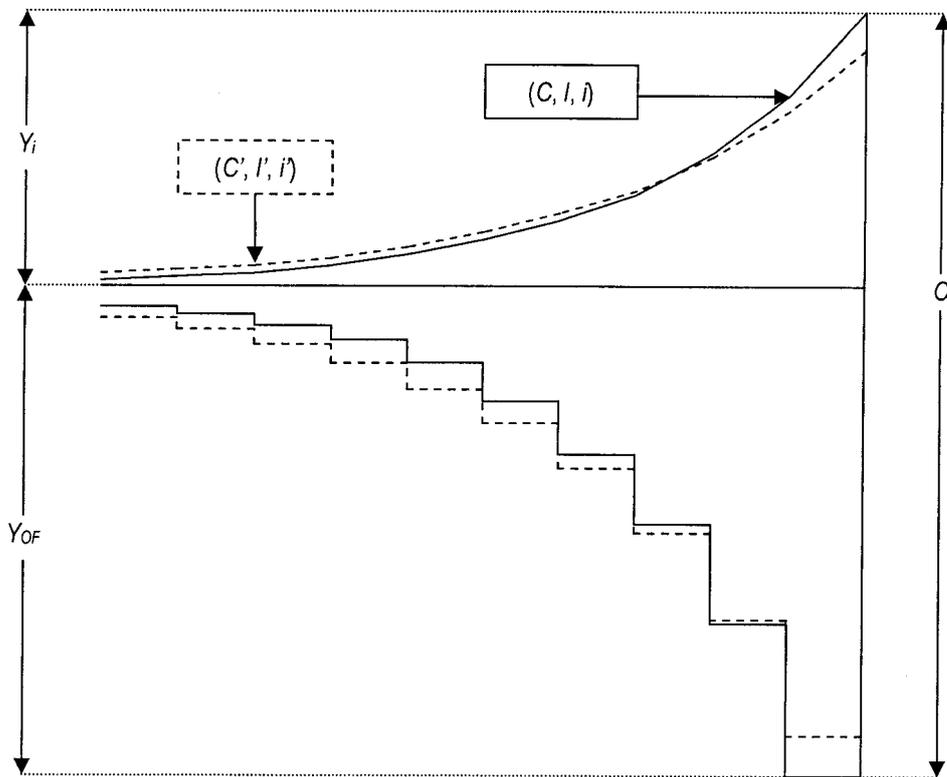


Figure 4. A lengthening of the structure of production from $(C = 100, l = 300, i = 12\%)$ to $(C' = 90, l' = 310, i' = 10\%)$

CONCLUSION

This paper has expounded a formal model of the Hayekian structure of production and its two main applications, first the description of static equilibrium, and second the analysis of the dynamics

or deformation of the structure. This model rests on a series of simplifying hypotheses, the two most important being that all the stages have the same duration, and that the ratio (in monetary values) of originary factors to investment is a constant ($= a$) in all stages of production. The contributions of the paper can be summed up in four points.

(1) *Quantification.* Reisman (1996) has propounded a detailed method for calculating the values of a structure of production. Here, his method is generalized to an algorithm that permits to calculate the spending/incomes of the stages of a structure defined by three parameters: the annual consumption spending C , the annual interest rate i , and the ratio a of originary factors to investment.

(2) *Illustration.* Most of the graphical representations of structures found in books and papers are very rudimentary. Only Rothbard (1962, p. 314) has attempted to apply the calculation of the stages to the illustration a structure. But his figure does not show the interest added at each stage. In other words, he has represented the “input curve” and not the “output curve” (in the sense of Hayek 1941). Yet, the output curve is the most interesting since it permits to visualize the accumulation of compound interest all along the production process. Surprisingly, Reisman does not draw any graphical representation of the structure in his treatise. In the present paper, the graphical representations (see figures 3 and 4) are based on accurate calculations. Furthermore, they show the distribution of added value between the owners of originary factors and the capitalists.

(3) *Mathematical formalization.* Among the mathematical formulas demonstrated above, two stand out as specially noteworthy. First, the annual investment spending I is: $I = C/(i + a)$. The four parameters C , I , i , and a are thus mathematically related. As a consequence, as long as the ratio a of originary factors at each stage is a constant, an equilibrium structure is entirely determined by three parameters only chosen among C , I , i , a . Second: the average length of a structure is $\lambda = (I/C)(1 + i)$. The latter formula has not been noticed before, so far as the author can tell. It shows that a lowering of the rate of interest i by itself would not give birth to a more roundabout structure of production but to a *less* roundabout structure, since the period of production is here directly proportional to the interest rate.¹⁶

(4) *Macroeconomic dynamics.* As far as the dynamic aspects of macroeconomics are concerned,

¹⁶ A variation of i alone is an admittedly unrealistic hypothesis from an economic point of view, since i and (I/C) are interdependent. In the case of a change of time preference, for instance, they vary in the opposite direction.

the notion of interest rate becomes difficult to handle. This is because when the shape of the structure changes, the originary rates of interest differ from stage to stage. The formal model enables us to distinguish between the “structural” rate of interest (relative price difference between the stages of a given equilibrium structure) and the rates of return of investment between two successive structures (relative price differences between the stages of two different structures), and to calculate the originary rates between two given structures.

References

- Böhm-Bawerk, Eugen von. 1959. *Positive Theory of Capital*. 4th ed. South Holland, Ill.: Libertarian Press.
- Fillieule, Renaud. 2005. “The ‘Values-Riches’ Model: An Alternative to Garrison’s Model in Austrian Macroeconomics of Growth and Cycle.” *Quarterly Journal of Austrian Economics* 8 (2): 3-20.
- Garrison, Roger. 2001. *Time and Money: The Macroeconomics of the Capital Structure*. London: Routledge.
- Hayek, Friedrich A. 1935. *Prices and Production*. 2nd ed. London: Routledge and Kegan Paul.
- Hayek, Friedrich A. 1941. *The Pure Theory of Capital*. Chicago: The University of Chicago Press.
- Hayek, Friedrich A. 1948. *Individualism and Economic Order*. Chicago: The University of Chicago Press.
- Mises, Ludwig von. 1998. *Human Action. A Treatise on Economics*. Reprint of the 1st ed. Auburn, Ala.: Ludwig von Mises Institute.
- Reisman, George. 1996. *Capitalism. A Treatise on Economics*. Ottawa, Ill.: Jameson Books.
- Rothbard, Murray N. 1962. *Man, Economy, and State. A Treatise on Economic Principles*. Princeton, N.J.: D. Van Nostrand.
- Skousen, Mark. 1990. *The Structure of Production*. New York: New York University Press.
- Skousen, Mark. 1991. *Economics on Trial: Lies, Myths, and Realities*. Burr Ridge: Irwin.

« A graphical illustration of Böhm-Bawerk's
productivity theory of interest » (2005)

Article non publié

A Graphical Illustration of Böhm-Bawerk's Productivity Theory of Interest

Renaud Fillieule

Université de Lille 1 (USTL), CLERSÉ

2005

Abstract. This paper offers a graphical exposition of Böhm-Bawerk's productivity theory of interest, based on a synthesis of the expositions that have been offered by Wicksell and by Dorfman.

Böhm-Bawerk has formulated a famous and devastating criticism of the “naïve” productivist theories of interest (1959a, 73–121). But he has also developed a productivist theory, according to which the equilibrium rate of interest is determined by three factors: the productivity of capital (productivity of the roundabout methods of production), the total value of capital, and the number of workers (1959b, 347–381). Two main criticisms have been expressed against his theory. Firstly by Clark (1896) and Knight (1933), who also upheld a productivist theory but rejected the Böhm-Bawerkian concept of period of production and thus rejected his conception of the productivity of capital. And secondly by Fetter (1902) and Mises (1949), who defended a subjectivist theory of interest based on intertemporal preferences, and claimed that no productivist theory, whether “naïve” or “sophisticated,” could ever explain interest. The following exposition of Böhm-Bawerk's theory is a synthesis of two ingenious graphical illustrations conceived by Wicksell (1970 [1893]) and by Dorfman (1959a). Böhm-Bawerk and these two followers analyze the market process whereby the rate of interest converges towards its final equilibrium value. Their reasoning rests upon three hypotheses: (1) labor is homogeneous and is the only remunerated factor (firms are vertically integrated: they sell consumer goods only and produce themselves all the capital goods that they use), (2) total capital has a given value K , and (3) there are N workers in the economic system.

1. The Böhm-Bawerkian function of production

According to Böhm-Bawerk (1959b, 82), a more roundabout method of production leads to an increase in the quantity or quality of products, everything else equal, i.e. with no change

in the quantity of labor and no technical progress (it is of course not claimed here that roundaboutness by itself increases the quantity/quality produced: the more roundabout method must be wisely chosen in order to bring about this result). The lengthening of the period of production is physically productive because it permits to use additional natural forces and resources, that would not be available with a shorter process of production (everything else equal). This theory has been endorsed by subsequent economists of the Austrian School (for instance by Mises 1949, 492, Hayek 1941, 59–64, Rothbard 1962, 41–46, Skousen 1990, 238–242). Wickseil illustrates it with the function of production $p(t)$ of figure 1: $p(t)$ is the annual average value produced by a worker in a process that lasts t years. Function $p(t)$ is *monotonic* (theory of roundaboutness: the longer the period of production t of the economic system, the greater the average value $p(t)$ of the annual product of a worker) and *concave* (decreasing returns of capital accumulation).

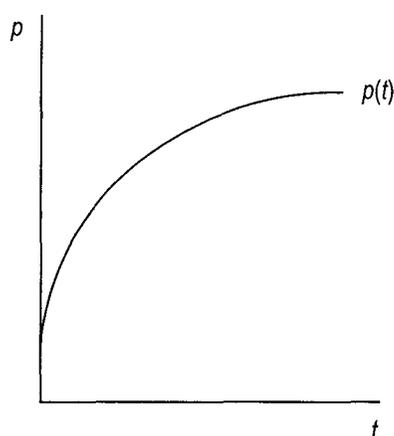


Figure 1. The Böhm-Bawerkian function of production $p(t)$

2. The optimal period of production, with a given production function and a given wage

The annual wage per worker is w , and i is the annual interest rate. If compound interest is neglected, then these variables are related by the following equation:

$$p(t)t = wt + \left(\frac{wt}{2}\right)it$$

The value produced by each worker at the end of the period of production is $p(t)t$ (value

annually produced \times number of years of production). This value is also equal to the costs of production $w t$ (annual wage \times number of years of production) plus the interest on the capital invested. The capital k invested in a worker during a t -years process is, according to Böhm-Bawerk, (approximately) equal to $w t/2$ (and exactly equal to $w(t + 1)/2$). Böhm-Bawerk assimilates the capital to a subsistence fund and calculates the subsistence fund k that would be necessary to maintain a worker during a whole production process. If there were only one process in the economic system, beginning from scratch every t years and producing the consumer goods after t years, then, with an annual wage w , the subsistence fund would be $k = w t$ for each worker. But there is not one process only. Böhm-Bawerk rightly points out that there are t similar processes unfolding in parallel, each one reaching each year a different level of maturation (see figure 2). The economic system is in this case “synchronized” thanks to the overlapping of—and the one year gap between—the different processes. It is thus not necessary for the subsistence fund to maintain the workers during the whole length of a process (i.e. during t years): a part of the subsistence is produced in the course of this process by the *other* processes, when they reach their end in turn. The period during which the subsistence must be provided to a worker is thus measured, so the argument goes, not by the total length of a given process (t years), but rather by the *average* length of a process (this point will be criticized below). Each year, a new process begins that will last t years, another process has begun one year before and will be finished in $(t - 1)$ years, still another has begun two years before and will be finished in $(t - 2)$ years, and so on. The average length of the structure of production is thus:

$$\left(\frac{1}{t}\right)[t + (t - 1) + (t - 2) + \dots + 2 + 1] = \left(\frac{1}{t}\right)\sum_{n=1}^{n=t} n = \left(\frac{1}{t}\right)\frac{t(t + 1)}{2} = \frac{t + 1}{2}$$

- The economic system is composed of five processes
- Each process lasts five years
- Each year, each process reaches a different level of completion
- The average length of the struct. is $(5 + 1)/2$

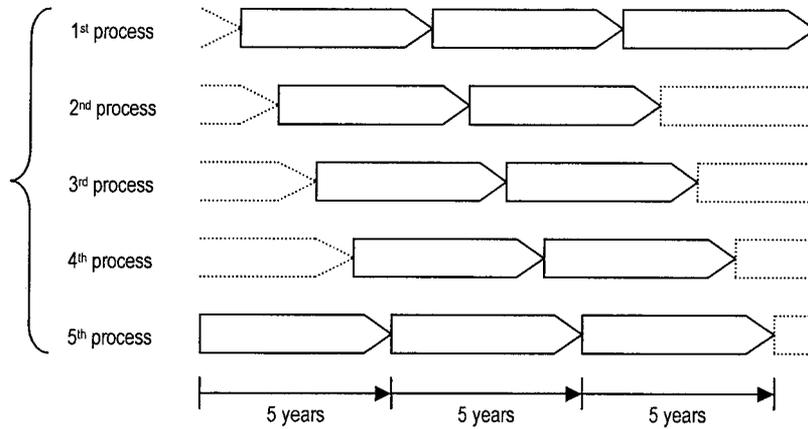


Figure 2. A synchronized economic system with processes of five years period of production

The subsistence fund—and capital invested—for each worker is therefore $w(t + 1)/2$, which is approximately equal to $wt/2$ (if t is high enough). If the annual interest rate is i and if compound interest is neglected, then the interest on capital invested in a worker during a t -years process is thus $(wt/2)it$, from which is deduced the equation $p(t)t = wt + (wt/2)it$. Let us now divide the left and right parts of this equation by wt :

$$\frac{p(t)}{w} = 1 + \frac{it}{2}$$

Let us multiply the numerator and denominator of the right part by $2/i$:

$$\frac{p(t)}{w} = \frac{\frac{2}{i} + t}{\frac{2}{i}}$$

This relation enables us to calculate the interest rate as soon as the annual wage w , the period of production t , and the function of production $p(t)$ are known. Wicksell illustrates this relation by the astute diagram of figure 3. The function $p(t)$ is represented in the right quadrant. If the wage is w_0 and the period of production t_0 , then the mathematical relation above implies that the line connecting the points $(0, w_0)$ and $(t_0, p(t_0))$ cuts the horizontal axis at point $(2/i_0, 0)$. This is an application of the Thales Theorem that is indicated in the upper left part of figure 2: just as the ratio CC'/BB' is equal to the ratio AC/AB , the ratio $p(t_0)/w_0$ is equal

to the ratio $(2/i_0 + t_0)/(2/i_0)$.

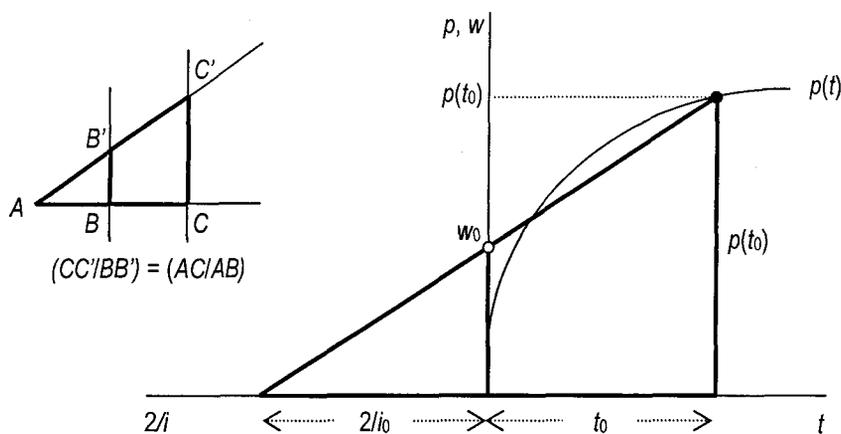


Figure 3. Visualization of the interest rate i_0 (or rather $2/i_0$) corresponding to a given wage w_0 , a given period of production t_0 , and a given function of production $p(t)$

It is then easy to visualize the maximum interest rate that the capitalists can get by choosing the optimal period of production. With a given function $p(t)$ and a given wage w_0 , the line going through the point $(0, w_0)$ and tangent to the curve $p(t)$ *minimizes* the value of $2/i$, and thus *maximizes* the reciprocal $i/2$. It thus maximizes the interest rate i . Figure 4 shows the maximum interest rate i_0^* and the optimal period of production t_0^* corresponding to the wage w_0 .

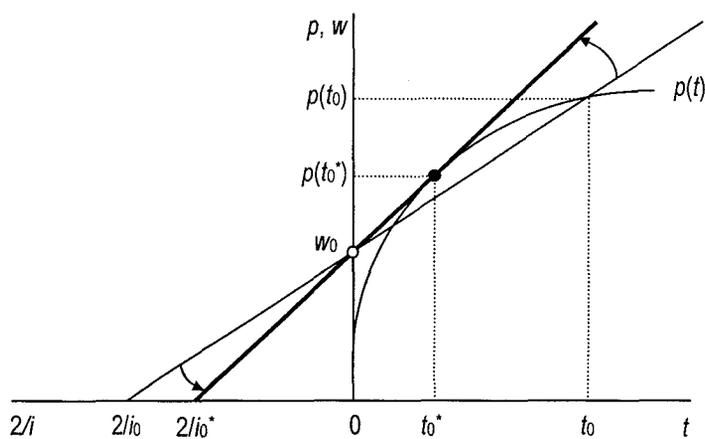


Figure 4. Determination of the maximum interest rate i_0^* , with a given wage w_0 and a given production function $p(t)$

3. The convergence towards final equilibrium through the variation of wage

As we have just seen, with a given wage w , the highest interest rate is reached for an optimal period of production $t^*(w)$. And this optimal period of production is a monotonic function of wage: figure 5 shows that when wage rises, the optimal period of production necessarily rises too.

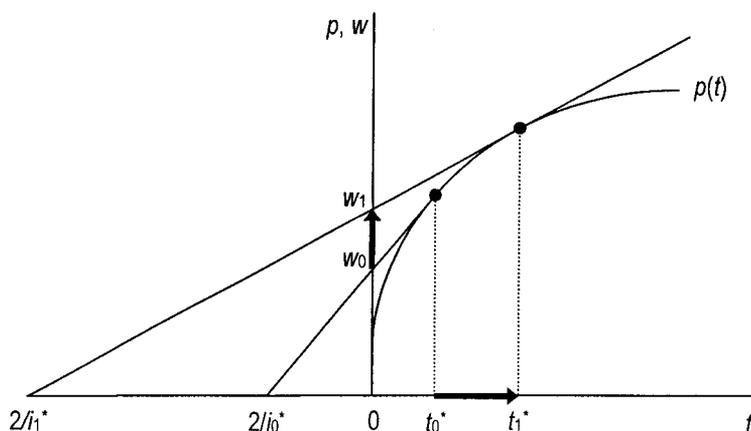


Figure 5. When the wage rises, the corresponding optimal period of production rises and the interest rate diminishes

One last element must now be taken into account, namely the fact that the total capital K and the number of workers N are given. Since the capital invested by worker is $wt/2$, the total capital is:

$$K = N \frac{wt}{2}$$

In equilibrium, the wage w and the corresponding optimal period of production $t^*(w)$ are thus related by the following equation:

$$wt^*(w) = \frac{2K}{N}$$

Figure 6 represents the hyperbola $wt = 2K/N$ to which the point of final equilibrium necessarily belongs. At the left of the hyperbola, wt is lower than $2K/N$, and at the right it is higher

than $2K/N$.

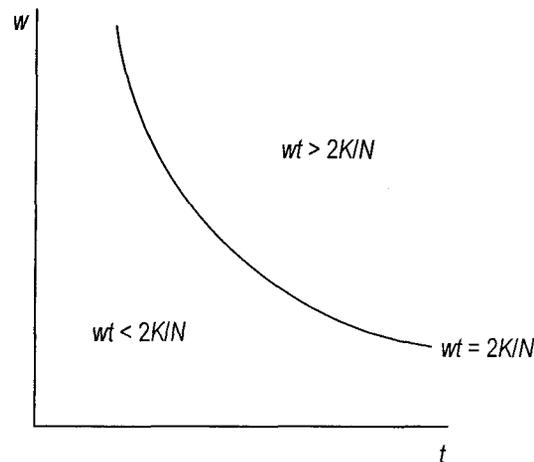


Figure 6. Hyperbola $wt = 2K/N$

The existence and uniqueness of equilibrium is easily demonstrated. Since t^* increases (continuously) with w , the product $wt^*(w)$ also increases continuously with w , and there is one and only value w_e for which the product $w_e t^*(w_e)$ is equal to $2K/N$. The convergence process towards equilibrium takes place in the following way. If the wage w is lower than the equilibrium wage w_e , then $wt^*(w)$ is lower than $w_e t^*(w_e) = 2K/N$. This means that the capital needed to pay the wage w to the N workers is lower than the total capital K . There is, in other words, an excess of available capital to hire the N workers. The capitalists thus compete to hire them, the wage rises and gets closer to the equilibrium wage. Conversely, if w is higher than w_e , then the available capital is insufficient to pay the wage w to all the workers, the latter compete to get hired, the wage tends to decrease and gets closer to the equilibrium wage. Through this tâtonnement process, the economic system converges towards its final equilibrium.

Figure 7 shows a convergence process in which the wage increases towards its equilibrium value. Equilibrium is represented by the two points E and E' , point E showing the optimal period of production t_e , and point E' showing the corresponding equilibrium wage w_e for which $w_e t_e$ just equals $2K/N$.

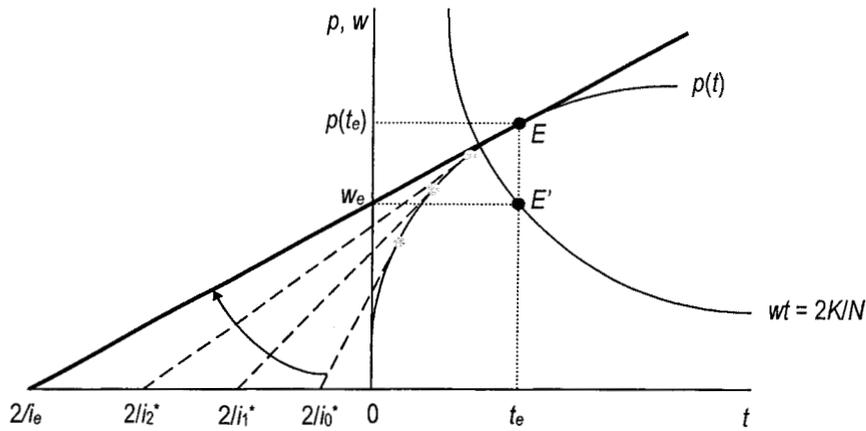


Figure 7. Convergence towards a final equilibrium characterized by the value K of total capital, the number N of workers, and the production function $p(t)$

4 Effects of a variation of the capital, of the number of workers, and of the production function

The Wicksell-Dorfman graphical model can easily be applied—even though none of them has done it—to the representation of the effects of changes in K , N or $p(t)$. An increase in K obviously leads to a fall in the interest rate (the hyperbola $wt = 2K/N$ moves to the right since K increases), and an increase in the number of workers to a rise in the interest rate (the hyperbola moves to the left since N increases). These symmetrical cases are illustrated in figures 8a and 8b.

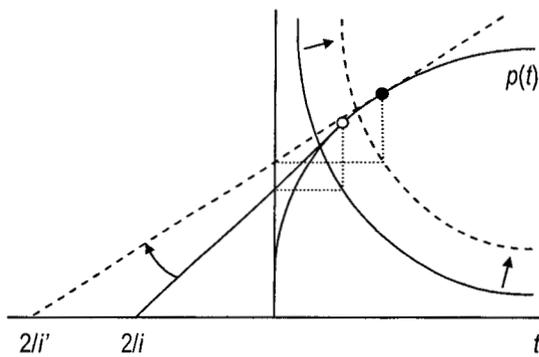


Figure 8a. The total capital K increases and the interest rate falls from i to i'

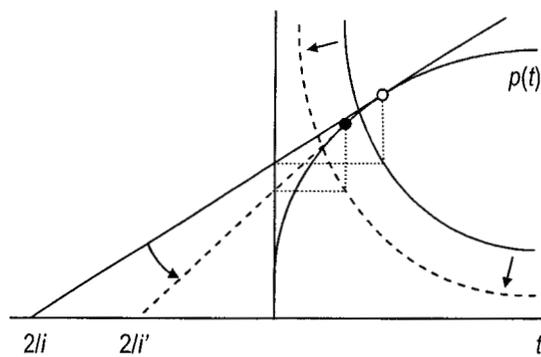


Figure 8b. The number of workers N increases and the interest rate rises from i to i'

The effect of a technical progress, however, is indeterminate in that it may cause, either a rise in the interest rate (figure 9a), or a decrease (figure 9b). Illustration 9b contradicts Böhm-Bawerk's assertion according to which "the rate of interest tends to rise with the discovery of new advantageous methods of production" (1959b, 365).

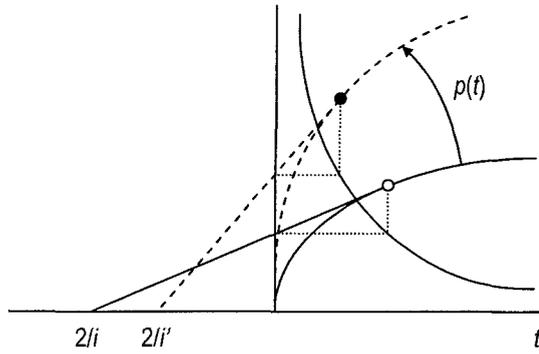


Figure 9a. An increase of the production function may rise the interest rate

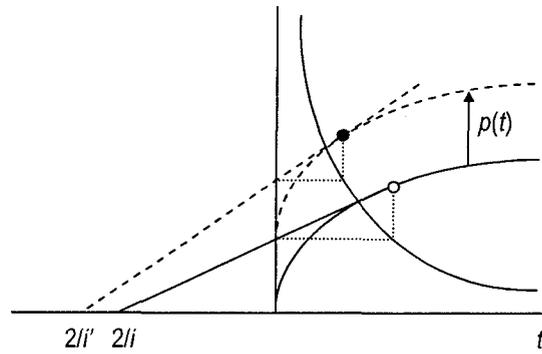


Figure 9b. An increase of the production function may also lessen the interest rate

References

- Böhm-Bawerk, Eugen von. 1959a. *History and Critique of Interest Theories*. 4th ed. South Holland: Libertarian Press.
- . 1959b. *Positive Theory of Capital*. 4th ed. South Holland: Libertarian Press.
- Clark, John B. 1896. The Origin of Interest. *Quarterly Journal of Economics* 10: 257–278.
- . 1899. *The Distribution of Wealth. A Theory of Wages, Interest and Profits*. New York: Macmillan.
- Dorfman, Robert. 1959a. A Graphical Exposition of Böhm-Bawerk's Interest Theory. *The Review of Economic Studies* February: 153–158.
- . 1959b. Waiting and the Period of Production. *Quarterly Journal of Economics* 63.3: 351–372.
- Fetter, Frank A. 1902. The 'Roundabout Process' in the Interest Theory. *Quarterly Journal of Economics* 17, reprinted in Fetter (1977, 172–191).
- . A. 1977. *Capital, Interest and Rent. Essays in the Theory of Interest*. Kansas City: Sheed Andrews and McMeel.
- Hayek, Friedrich A. 1941. *The Pure Theory of Capital*. Chicago: The University of Chicago Press.
- Knight, Frank H. 1933. Capitalistic Production, Time, and the Rate of Return. In *Economic*

- Essays in Honour of Gustav Cassel*, 327–42. London: George Allen and Unwin.
- Mises, Ludwig von. 1949. *Human Action. A Treatise on Economics*. Yale University Press.
- Rothbard, Murray N. 1962. *Man, Economy, and State. A Treatise on Economic Principles*. 2 vols. Princeton: D. Van Nostrand.
- Skousen, Mark. 1990. *The Structure of Production*. New York: New York University Press.
- Wicksell, Knut. 1970 [1893]. *Value, Capital and Rent*. New York: Augustus M. Kelley.

« La théorie du marché des fonds prêtables
et le paradigme autrichien » (2005)

Communication au *Austrian Research Seminar*, Paris II

Article non publié

La théorie du marché des fonds prêtables et le paradigme autrichien

Renaud Fillieule
Université de Lille 1, Laboratoire CLERSÉ
16 novembre 2005

Résumé. Cet article s'interroge sur la place et le statut de la théorie du marché des fonds prêtables comme explication de l'intérêt au sein de l'école autrichienne d'économie. Bien qu'elle soit aujourd'hui acceptée par un certain nombre de macroéconomistes « autrichiens », cette théorie a fait l'objet de critiques de la part d'auteurs de référence plus anciens tels que Mises, Hayek et Rothbard. Cet article vise donc d'une part à indiquer les difficultés posées par cette théorie dans le cadre du paradigme autrichien, difficultés dues en particulier à son incapacité à rendre compte du phénomène de l'intérêt originaire. Mais d'autre part, et de façon plus constructive, l'article présente les bases de la théorie de l'intérêt de Reisman (1996), qui illustre ce que peut être une théorie macroéconomique de l'intérêt à la fois bien ancrée dans la tradition autrichienne et ne s'appuyant pas sur l'existence d'un marché des fonds prêtables.

La théorie du marché des fonds prêtables est la théorie de l'intérêt la plus répandue aujourd'hui chez les macroéconomistes de l'école autrichienne. Elle est utilisée comme *seule et unique* théorie de l'intérêt dans trois des principaux ouvrages publiés dans ce domaine depuis une quinzaine d'années (Skousen 1990, Horwitz 2000, Garrison 2001), ainsi que dans la thèse de Murphy (2003)¹⁷. Cette théorie est aussi celle qui est mobilisée par la macroéconomie néo-classique « standard » pour expliquer la détermination du taux d'intérêt réel de long terme (Mankiw 1998). Elle peut donc apparaître comme parfaitement acceptable, jusqu'à ce que l'on découvre les commentaires qui ont été faits à son sujet par les auteurs de référence de l'école autrichienne. Mises (1949, p. 523-524) a par exemple écrit que « L'intérêt originaire n'est pas un prix déterminé sur le marché par le jeu de la demande et de l'offre de capital ». On lit sous la plume de Hayek (1941, p. 266) que « Le taux de rendement sur l'investissement tel que déterminé par les relations entre les prix des biens de capitaux et des biens de consommation est donc antérieur au, et en principe indépendant du, taux d'intérêt sur les prêts en monnaie ». Dans la même veine, Rothbard (1962, p. 364) explique que « [Le taux d'intérêt] est déterminé par des facteurs qui n'ont rien à voir avec les discussions habituelles des économistes sur le marché des prêts aux producteurs ».

À la lecture de ces commentaires, il est clair qu'une difficulté se présente : quel est au juste le statut de la théorie du marché des fonds prêtables au sein du paradigme autrichien ? Dans quelle mesure est-elle conciliable avec lui ? La réponse que nous apporterons sera nuancée. Nous soulignerons l'utilité de cette théorie pour analyser les phénomènes de déséquilibre conduisant aux crises économiques. Mais nous essayerons surtout de montrer qu'elle peut difficilement être considérée comme *la* théorie autrichienne de l'intérêt. L'article est divisé en trois sections. La première présente brièvement la théorie du marché des fonds prêtables et ses applications. La deuxième section précise les limites et défauts de cette théorie d'un point de vue autrichien. La troisième section, enfin, présente la théorie de l'intérêt de Reisman (1996). Cette dernière est fondée sur la conception misesienne de la préférence pour le présent et sur

¹⁷ O'Driscoll (1977, p. 53-54) pourrait être le premier auteur de l'école autrichienne à avoir accepté la théorie du marché des fonds prêtables comme théorie de base de l'intérêt. Voir aussi O'Driscoll et Rizzo (1985, p. 209-210).

la notion hayekienne de structure de production. Elle illustre fort bien ce que peut être une théorie autrichienne de l'intérêt qui se passe de toute référence au marché des fonds prêtables.

1. LA THÉORIE DU MARCHÉ DES FONDS PRÊTABLES

1.1 Principes

Cette théorie bien connue constitue une application directe de la loi de l'offre et de la demande au marché des fonds prêtables : une offre de fonds prêtables (épargne) croissante en fonction du taux d'intérêt se trouve confrontée à une demande de fonds prêtables (investissement) décroissante avec le taux d'intérêt. Ces courbes sont représentées à la figure 1¹⁸. Le taux d'équilibre i_e est celui pour lequel les fonds prêtables fournis sont égaux aux fonds prêtables demandés. Plus précisément, les épargnants sont disposés à placer des fonds d'autant plus importants que le taux d'intérêt – le revenu de l'épargne – est plus élevé (offre de fonds croissante), et les entrepreneurs sont disposés à emprunter des fonds d'autant plus importants que le taux d'intérêt – le coût du remboursement – est moins élevé (demande de fonds décroissante). Si le taux est supérieur au taux d'équilibre, la quantité de fonds offerts est excédentaire, les épargnants sont en concurrence pour placer leurs fonds et le taux a tendance à baisser vers l'équilibre, et inversement si le taux est inférieur au taux d'équilibre. Quoique l'on pense par ailleurs de cette théorie, il faut bien reconnaître qu'elle a au moins le mérite de la simplicité.

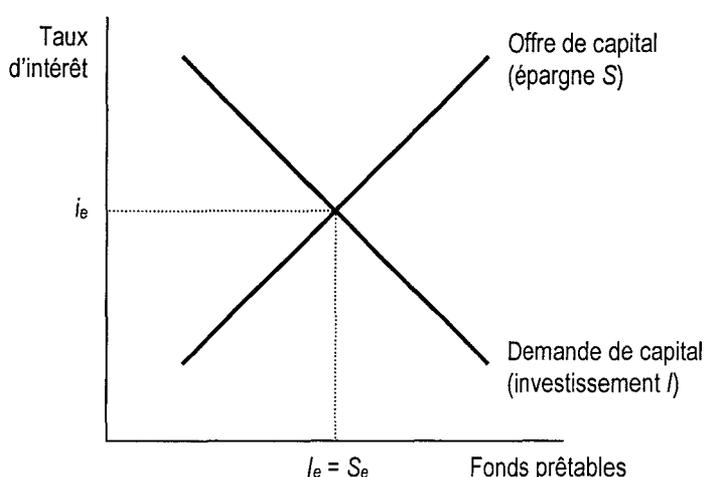


Figure 1. La théorie du marché des fonds prêtables

Garrison précise que l'offre de fonds prêtables consiste en la totalité des fonds destinés à l'investissement, y compris les revenus non distribués (*retained earnings*) et les actions (*equi-*

¹⁸ Skousen présente ce diagramme de la façon suivante : « In this diagram, we have *S*, the supply of capital, which comes from business and personal savings. The *S* schedule has a positive slope, because as the interest rate climbs, savers earn a higher yield and are encouraged to invest more. We also have *D*, the demand for capital by business firms. The *D* schedule has a negative slope because as the interest rate increases, business is less likely to borrow the more expensive money » (1990, p. 204). Les présentations de Horwitz (2000, p. 72) et de Garrison (2001, p. 36) sont tout à fait similaires à celle de Skousen, et tout aussi succinctes.

ty shares)¹⁹. Il précise aussi que la demande de fonds prêtables représente l'intention par les emprunteurs d'investir en moyens de production quels qu'ils soient (biens durables, biens intermédiaires ou capital humain). Quant à Horwitz, il insiste sur le fait que les courbes d'offre et de demande de fonds de la figure 1 sont des courbes *ex ante* : si le taux diffère du taux d'équilibre, il apparaît une inégalité *ex ante* entre l'épargne disponible et l'investissement souhaité à ce taux ; mais en revanche, *ex post*, lorsque les échanges de fonds sont effectués, l'égalité entre épargne et investissement est nécessairement respectée du côté court du marché quel que soit le taux d'intérêt (que ce soit le taux d'équilibre ou non).

La présentation de Garrison (ainsi que celle de Murphy) comporte une particularité, qui est qu'elle intègre les phénomènes liés à l'incertitude et à la confiance. L'incertitude a un effet sur la demande de monnaie, et indirectement sur l'offre de fonds. Si par exemple les épargnants estiment que le degré d'incertitude de l'investissement a augmenté, ils peuvent souhaiter conserver une part plus grande de leurs actifs sous la forme monétaire (moins risquée), d'où une baisse de l'offre de fonds, c'est-à-dire un déplacement de la courbe vers la gauche, entraînant une hausse du taux d'intérêt et une baisse de l'investissement d'équilibre²⁰. D'un autre côté, si les entrepreneurs perdent leur confiance en l'avenir alors la demande de fonds va avoir tendance à baisser, réduisant du même coup le taux et l'épargne-investissement d'équilibre²¹. Cette présentation diffère de celle que l'on trouve chez un auteur néo-classique « standard » comme Mankiw (1998, p. 690-691). Chez ce dernier, la théorie du marché des fonds prêtables est censée expliquer le taux d'intérêt *réel* de long terme, en l'absence d'interférence de la sphère monétaire, et c'est une autre théorie – la théorie keynésienne de la préférence pour la liquidité – qui est mobilisée pour expliquer l'influence à court terme de la demande de monnaie sur le taux d'intérêt. Mankiw (1998, p. 895-897) considère que le niveau général des prix est fixe à court terme, d'où il s'ensuit que c'est le taux d'intérêt qui régule à court terme l'équilibrage entre offre et demande de monnaie (théorie keynésienne de la préférence pour la liquidité) ; à long terme, en revanche, le niveau des prix a selon lui le temps de varier et peut donc réguler l'équilibrage entre offre et demande de monnaie, le taux d'intérêt équilibrant alors l'offre et la demande de fonds prêtables (théorie du marché de fonds prêtables).

Un point est à noter. Dans leurs textes respectifs, nos trois auteurs – Skousen (1990), Horwitz (2000) et Garrison (2001) – n'adoptent pas la moindre distance critique vis-à-vis de la théorie du marché des fonds prêtables. Ils ne se donnent pas la peine de la confronter à d'autres théories pour montrer sa supériorité, et ne citent d'ailleurs *aucune autre théorie*. Ils ne s'appuient pas non plus sur un argument d'autorité qui consisterait à renvoyer le lecteur à une série de textes classiques de l'école autrichienne (ou d'une autre) présentant de façon plus approfondie cette théorie et montrant sa validité. Tout se passe comme si, pour eux, l'acceptation de cette théorie relevait de l'évidence et ne nécessitait aucune justification.

1.2 Applications

Ils utilisent ensuite la théorie du marché des fonds prêtables pour traiter trois thèmes ma-

¹⁹ Inclure les revenus non distribués et les actions au marché des fonds prêtables nous semble pourtant discutable puisque ces fonds *ne sont pas prêtés* mais directement investis par leurs propriétaires.

²⁰ « When individuals perceive greater uncertainty in the impending future, their desire for liquidity (on the margin) will increase. Whereas previously an individual may have been willing to exchange 1,000 present dollars for the (nearly certain) prospect of 1,050 future dollars, due to the changed conditions (or rather, his perception of them) the individual may now only be willing to exchange 1,000 present dollars for the (nearly certain) prospect of 1,100 future dollars. In short, the supply curve of loanable funds shifts to the left, driving up the (nominal) rate of interest » (Murphy 2003, p. 165-166).

²¹ « What manifests itself on the demand side of the loan market as a loss of business confidence manifests itself on the supply side as an increase in liquidity preference » (Garrison 2001, p. 38).

jeurs : la coordination macroéconomique, la croissance par intensification capitaliste, et le cycle.

La coordination macroéconomique. Skousen (1990) consacre une section au rôle du taux d'intérêt monétaire, où d'emblée il présente ce taux comme le *régulateur* de l'économie, qui assure la tendance à l'égalité entre l'offre agrégée et la demande agrégée²² :

« Sans les taux d'intérêt de marché [*market interest rates*] opérant à travers l'ensemble de l'économie, les producteurs seraient incapables de déterminer quelles quantités de ressources devraient être allouées aux divers secteurs intertemporels du marché et les consommateurs ne sauraient pas quelle proportion des fonds dont ils disposent devrait être consacrée à l'épargne et l'investissement. Sans des taux d'intérêt fluctuant librement, le marché serait déstabilisateur » (1990, p. 203).

Les effets déstabilisateurs proviendraient d'interventions gouvernementales visant à fixer autoritairement le taux d'intérêt au-dessus ou au-dessous de sa valeur d'équilibre. Si par exemple le gouvernement imposait un taux d'intérêt supérieur au taux d'équilibre, il en résulterait un déclin de l'activité à cause de la baisse de la consommation (elle-même due à l'excès d'épargne) et de la baisse de l'investissement (due au coût élevé du capital). Garrison, de son côté, explique que « le marché des fonds prêtables facilite la coordination des plans de production avec les préférences des consommateurs » (2001, p. 38), et il précise un peu plus loin (p. 39) que « le taux de prêts qui équilibre [le marché des fonds prêtables] est vital pour traduire les préférences intertemporelles de consommation des ménages sous la forme des plans de production intertemporels du monde des affaires »²³. Le marché des fonds prêtables joue donc bien, dans cette perspective, un rôle essentiel pour la coordination entre la production et la consommation.

La croissance. Cette deuxième application de la théorie du marché des fonds prêtables concerne le cas où la croissance provient, non pas d'un progrès technique, mais d'une augmentation de l'investissement relativement à la consommation. Lorsque la préférence pour le présent des épargnants diminue, ils réduisent leur consommation présente et augmentent leur épargne. Cela entraîne une augmentation de l'offre de fonds prêtables, qui a à son tour pour conséquences une augmentation de l'investissement et une baisse du taux d'intérêt d'équilibre (voir figure 2). Cette hausse de l'investissement se manifeste, dans le cadre de la théorie autrichienne du capital, par une réallocation des facteurs de production dans des processus dont la durée moyenne est plus longue qu'auparavant. En d'autres termes, la structure de production « s'allonge » : par application de la célèbre théorie böhm-bawerkienne du détour de production, la structure devient de ce fait plus productive en termes réels, ce qui signifie que la quantité et/ou la qualité du flux annuel des biens de consommation s'est accrue, ou en d'autres termes qu'une croissance est survenue (Skousen 1990, p. 233-241, Garrison 2001, p. 61-67).

²² Notons au passage cette différence majeure avec la théorie macroéconomique « standard », dans laquelle c'est le *niveau général des prix* qui régule l'équilibre entre offre globale et demande globale. Skousen (1991, chap. 7) propose une intéressante critique, dans une perspective « autrichienne », de la théorie standard de l'offre et de la demande globales.

²³ « The loan rate that clears [the market for loanable funds] is vital in translating the intertemporal consumption preferences of income earners into intertemporal production plans of the business community. »

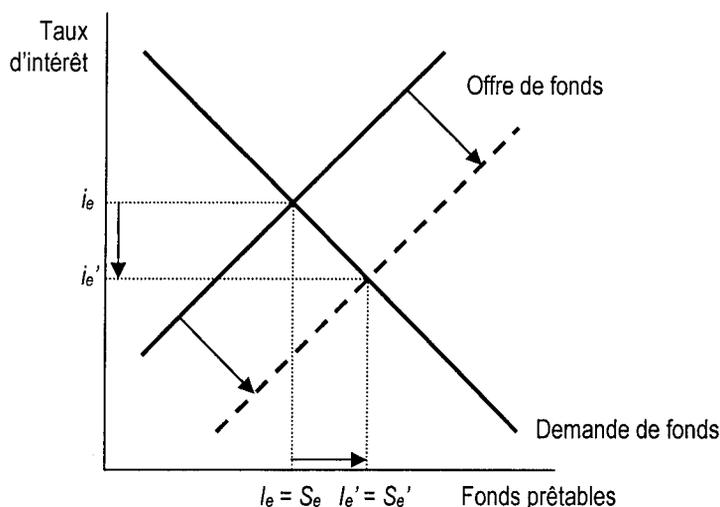


Figure 2. Conséquences d'une baisse de la préférence pour le présent des épargnants

Le cycle. Dans la théorie autrichienne du cycle, une nouvelle quantité de monnaie ΔM entrant dans le système économique par le marché des fonds prêtables (par exemple suite à une politique d'*open market* du gouvernement) peut avoir un effet déstabilisateur. L'offre de fonds prêtables augmente, comme dans le cas de la croissance, mais la hausse de l'investissement et la baisse du taux d'intérêt qui en résultent ne correspondent pas à un changement des préférences intertemporelles des agents : elles sont la conséquence d'une *épargne forcée*, comme le montre la figure 3 (O'Driscoll 1977, p. 74, Skousen 1990, p. 285-286, Horwitz 2000, p. 113, Garrison 2001, p. 69). Au taux d'intérêt artificiellement abaissé i' , les épargnants souhaitent épargner un montant S' , mais l'investissement vaut I' , d'où une « épargne forcée » ($= I' - S'$). Dans ces conditions, un allongement de la structure de production survient, qui correspond à la phase de boom. Mais contrairement au cas de la croissance, cet allongement n'est pas durable. Lorsque les travailleurs récupèrent un surplus de monnaie sous forme de salaires accrus, ils choisissent de consommer et non d'investir. Dès que l'augmentation de la quantité de monnaie s'arrête ou même ralentit, la consommation remplace l'épargne forcée et une partie des investissements qui avaient été consacrés à l'allongement de la structure n'est plus rentable. La liquidation de ces investissements – qui induit du chômage et des pertes en capital – constitue la crise.

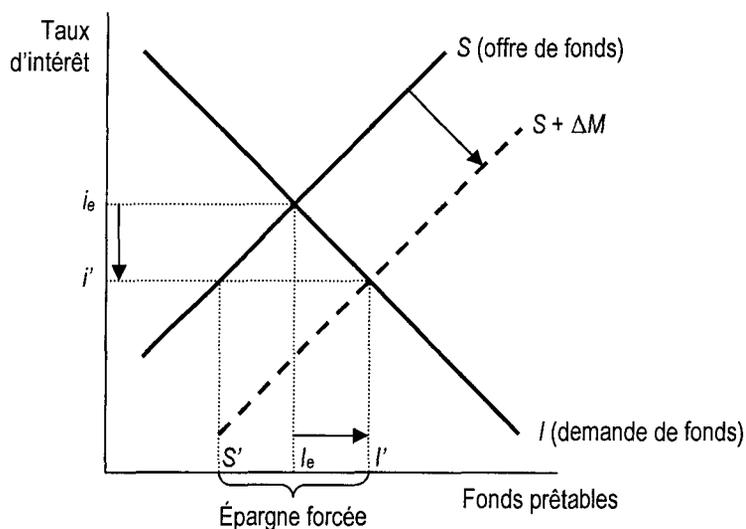


Figure 3. Conséquences d'une augmentation de la quantité de monnaie

2. LES LIMITES DE LA THÉORIE DU MARCHÉ DES FONDS PRÊTABLES

Cette section présente, à l'encontre de la théorie du marché des fonds prêtables, une série de critiques qui reposent toutes sur la même base, à savoir la distinction entre taux d'intérêt *originnaire* et taux d'intérêt *contractuel*. Cette distinction a été clairement formulée par Böhm-Bawerk dès les premières pages de son célèbre traité *Capital et Intérêt* (1884[1959], p. 6). Le taux d'intérêt originnaire est l'écart relatif qui apparaît à l'équilibre final entre le prix d'un bien produit et le prix des facteurs qui ont servi à le produire. En dehors de l'équilibre final, dans le monde réel caractérisé par la présence d'une incertitude radicale, l'écart entre les prix du produit et de ses facteurs est la somme de quatre composantes hétérogènes qui sont l'intérêt originnaire, le salaire de l'entrepreneur (le cas échéant), le profit/perte pur (dû à l'incertitude), et la composante de prix due aux changements dans l'offre ou la demande de monnaie (Mises 1949, chap. 19-20). Quant à l'intérêt contractuel, il est bien sûr celui qui apparaît sur le marché des prêts en monnaie, et qui est expliqué par la théorie du marché des fonds prêtables.

2.1 La théorie du marché des fonds prêtables ne peut pas expliquer l'intérêt originnaire

Il doit d'emblée être bien clair que les économistes de référence de l'école autrichienne, comme Mises, Hayek et Rothbard, utilisent la théorie du marché des fonds prêtables²⁴. Mais ils restreignent son utilisation aux situations de déséquilibre, et en particulier à l'analyse du cycle telle qu'elle est illustrée à la figure 3. L'impression très nette qui se dégage de leurs travaux est cependant que cette théorie ne peut pas constituer *la* théorie fondamentale de l'intérêt. La raison en est simple. Le phénomène primordial à expliquer est celui de l'intérêt

²⁴ Mises l'emploie implicitement dans le passage où il explique que lorsqu'une quantité supplémentaire de monnaie fiduciaire est offerte sur le marché des prêts, « Additional loans can be placed only at a lower gross market rate » (1949, p. 549). Rothbard illustre ainsi ce processus : « Suppose, now, that the banks in the economy expand credit and increase the money supply by 2 000 000 ounces. What are the consequences ? The new money is loaned to businesses. These businesses, now able to acquire the money at a lower rate of interest, enter the capital goods' and original factors' market to bid resources away from the other firms » (1962, p. 855). Hayek (1941, p. 363-365), de son côté, utilise explicitement la théorie du marché des fonds prêtables, mais dans une version qui lui est propre et qui est assez différente de celle qui a été présentée plus haut.

originnaire. Or ce phénomène existe indépendamment de – et préalablement à – l’existence d’un marché des capitaux, et il ne peut donc pas relever du principe explicatif de la théorie du marché de fonds prêtables. S’il n’y avait pas une tendance préalable pour les prix des produits à dépasser les prix de leurs facteurs, alors les entreprises ne pourraient pas emprunter des fonds car elles n’auraient pas ensuite les moyens de payer les intérêts dus²⁵. Lorsque Hayek (1941) présente de façon approfondie sa théorie du capital et de l’intérêt, il procède en une succession logique de trois étapes. D’abord, il élabore une théorie du capital et de l’intérêt dans un système centralement planifié, puis il passe à un système capitaliste (concurrentiel) mais sans monnaie. Ce n’est que dans un troisième temps qu’il aborde le cas d’un système capitaliste monétaire et qu’il présente alors sa version de la théorie du marché des fonds prêtables :

« Nous pourrions concevoir une société où le prêt monétaire (tout au moins le prêt à intérêt) serait interdit et où néanmoins, aussi longtemps qu’existerait la possibilité de répartir les investissements au moyen de partenariats, de participation à l’actionnariat, etc., le taux de rendement sur l’investissement serait uniforme à travers le système.

Le taux de rendement sur l’investissement tel que déterminé par la relation entre les prix des biens de capitaux et des biens de consommation est donc antérieur à, et en principe indépendant de, l’intérêt sur les prêts en monnaie, même si, bien sûr, lorsque les prêts monétaires sont possibles le taux d’intérêt sur ces prêts monétaires tendra à correspondre au taux de rendement sur les autres investissements » (1941, p. 266).

Fetter est un économiste américain aujourd’hui quelque peu oublié, mais qui a développé les aspects subjectivistes de la théorie de l’intérêt de Böhm-Bawerk et a eu de ce point de vue une influence majeure sur von Mises. Il explique très clairement que le taux « contractuel » résulte de forces plus fondamentales que celles qui sont à l’œuvre sur le seul marché des capitaux. Son argumentation est de nature historique plutôt que logique :

« L’intérêt contractuel (en pourcentage) est un phénomène relativement superficiel. Il est aussi, dans l’histoire économique, un phénomène très récent à grande échelle. Il est apparu à la suite de l’utilisation de la monnaie et d’un régime de prix monétaires. Si, donc, une relation causale peut être repérée historiquement, l’intérêt [contractuel] apparaît comme un effet plutôt que comme une cause des prix. Il est impossible de concevoir qu’un taux d’intérêt général, exprimé en pourcentage du principal, puisse précéder le système des prix ; alors qu’il est possible de se représenter, comme antérieur à la fois à l’intérêt et aux prix monétaires, un système de prix de troc, non monétaires, impliquant des taux d’échange entre les marchandises (1927, p. 275-276).

Le système des prix, même dans ses formes les plus anciennes et les plus rudimentaires, incorpore une dimension temporelle. Les biens futurs subissent une décote de valeur par rapport aux biens présents. Cette décote provient selon Fetter de la préférence pour le présent et constitue le phénomène de l’intérêt originnaire. La décote se manifeste par le fait que les prix des facteurs de production (biens présents) ont tendance à être inférieurs au prix de leurs produits (biens futurs), ou encore par le fait que les usages plus lointains des biens durables va-

²⁵ Comme le remarque Hülsmann (2002, p. 99) : « the role of credit markets is a derived one. [...] The question is merely why the demand for interest payments can be satisfied. Why do creditors obtain these payments ? The answer is that these interest payments derive from the price spread that firms realize as part and parcel of their market activities ; entrepreneurs can pay interest to their creditors because, from the very outset, they exclusively engage in money-productive ventures. »

lent moins que leurs usages plus proches²⁶ :

« l'escompte sur les biens et usages futurs, et la prime aux biens et usages présents, doit pénétrer dans tous les recoins d'un système des prix et intervenir dans presque tous les prix, indépendamment de l'utilisation du crédit sous une forme ou une autre, sans même parler du prêt monétaire à intérêt » (1927, p. 280).

Le point de vue de Mises est tout à fait similaire à celui de Fetter :

« L'intérêt originaire n'est pas un prix déterminé sur le marché par l'interaction de la demande et l'offre de capitaux ou de biens de capitaux. Son niveau ne dépend pas de l'étendue de cette demande et de cette offre [...] Le marché des prêts ne détermine pas le taux d'intérêt. Il ajuste le taux d'intérêt sur les prêts au taux d'intérêt originaire tel qu'il se manifeste dans l'escompte sur les biens futurs » (1949, p. 523-524).

Quant à Rothbard, il estime que « du point de vue de l'analyse fondamentale, il n'est pas du tout besoin qu'il existe un marché des fonds prêtables » :

« Pour le mettre en évidence, considérons ce que serait la situation des affaires en l'absence d'un marché des prêts aux producteurs. Que faut-il pour cela ? Les individus épargnent, consommant moins que leur revenu. Ensuite, ils *investissent directement* cette épargne dans la structure de production, l'incitation à investir provenant du taux de rendement – de l'écart de prix – sur l'investissement [...] Il n'y a, dans ce cas, aucun marché des prêts aux producteurs. Il n'y a pas de prêt provenant d'un groupe d'épargnants vers un autre groupe d'investisseurs. Et il est clair que le taux d'intérêt dans la structure de production existe encore ; il est déterminé par des facteurs qui n'ont rien à voir avec les considérations habituelles des économistes sur le marché des prêts aux producteurs » (1962, p. 364).

Chez tous ces auteurs de l'école autrichienne, la théorie du marché des fonds prêtables ne peut en aucun cas prétendre constituer la théorie de base de l'intérêt. La critique qu'ils adressent à cette théorie s'applique tout autant – et peut-être même plus encore – à la conception macroéconomique « standard » de Mankiw (1998, p. 897), selon laquelle la théorie du marché des fonds prêtables explique l'investissement et le taux d'intérêt uniquement à l'équilibre de long terme. Or, à l'équilibre de long terme, l'incertitude a disparu et les capitalistes peuvent investir leurs fonds eux-mêmes. Ils n'ont besoin ni des banques ni des entrepreneurs, et le marché des fonds prêtables n'a donc aucune raison de subsister. Et si ce marché n'existe pas, il ne peut à l'évidence pas déterminer l'investissement et le taux d'intérêt. Mankiw restreint l'applicabilité de la théorie du marché des fonds prêtables à l'équilibre de long terme, c'est-à-dire précisément au cas où elle est, nous semble-t-il, le plus inapplicable !

²⁶ Si un bien a une durée de vie de 10 ans et un loyer annuel de 1 000 €, sa valeur totale actuelle – valeur en capital – ne sera pas égale à 10 000 € mais à une somme plus faible (indépendamment de tout problème d'incertitude), ce qui prouve qu'un usage futur tend à être évalué au-dessous d'un usage présent. Fetter précise bien que ce phénomène de capitalisation est *antérieur* à l'existence d'un marché monétaire, et ne présuppose donc pas du tout l'existence d'un taux d'intérêt contractuel. La relation causale joue selon lui en sens inverse : c'est parce que les usages futurs tendent à être sous-évalués par rapport aux usages présents qu'un taux d'intérêt positif va apparaître dès lors que seront pratiqués les prêts et emprunts de sommes de monnaie – dès lors qu'apparaîtra un marché monétaire.

2.2 Coordination et croissance

Commençons par la notion de coordination. Skousen, on l'a vu, décrit le taux d'intérêt monétaire comme le « régulateur » de l'économie, qui permet aux producteurs d'allouer correctement les ressources entre la production des biens de consommation et celle des biens intermédiaires. Garrison évoque lui aussi le rôle du marché des fonds prêtables dans la détermination de la répartition que souhaitent les agents entre consommation et investissement. Bien que ces considérations soient en elles-mêmes tout à fait acceptables, elles font perdre de vue un aspect essentiel de la tradition autrichienne, à savoir que la production s'effectue par étapes successives. L'équilibre macroéconomique autrichien n'est donc pas un équilibre simple qui s'établirait entre la consommation d'un côté et la production de l'autre, mais un équilibre plus complexe qui s'instaure entre *toutes* les étapes de la structure de production.

Une brève présentation de la théorie autrichienne du capital est ici indispensable. Prenons un exemple aussi simplifié que possible, inspiré des travaux fondateurs de Hayek (1931[1935])²⁷. Supposons que la production s'effectue par un processus qui dure quatre ans et qui peut être subdivisé en quatre étapes d'une année chacune. Ce raisonnement peut être illustré en disant que l'étape 1 est la récupération de minerais, l'étape 2 la production de métaux à partir des minerais, l'étape 3 la fabrication de machines à l'aide des métaux, et l'étape 4 la production des biens de consommation grâce aux machines. La production est bien sûr « synchronisée » en ce sens que l'on n'a pas besoin d'attendre quatre ans avant de pouvoir consommer les biens produits : quatre processus identiques se déroulent simultanément, avec un décalage d'un an, de sorte que l'un d'eux arrive à son terme chaque année et permet de disposer annuellement des fruits d'un processus de production qui a duré quatre ans. La figure 4 montre les quatre processus dans leur déroulement temporel, ainsi que la structure de production de l'année courante. Il est bien évident que la structure de production d'une économie développée est beaucoup plus complexe, avec des étapes plus nombreuses et de durées diverses, mais ce schéma très simple va suffire à illustrer notre propos.

²⁷ Voir aussi Rothbard (1962), Skousen (1990), et pour un exposé plus technique Fillieule (2005).

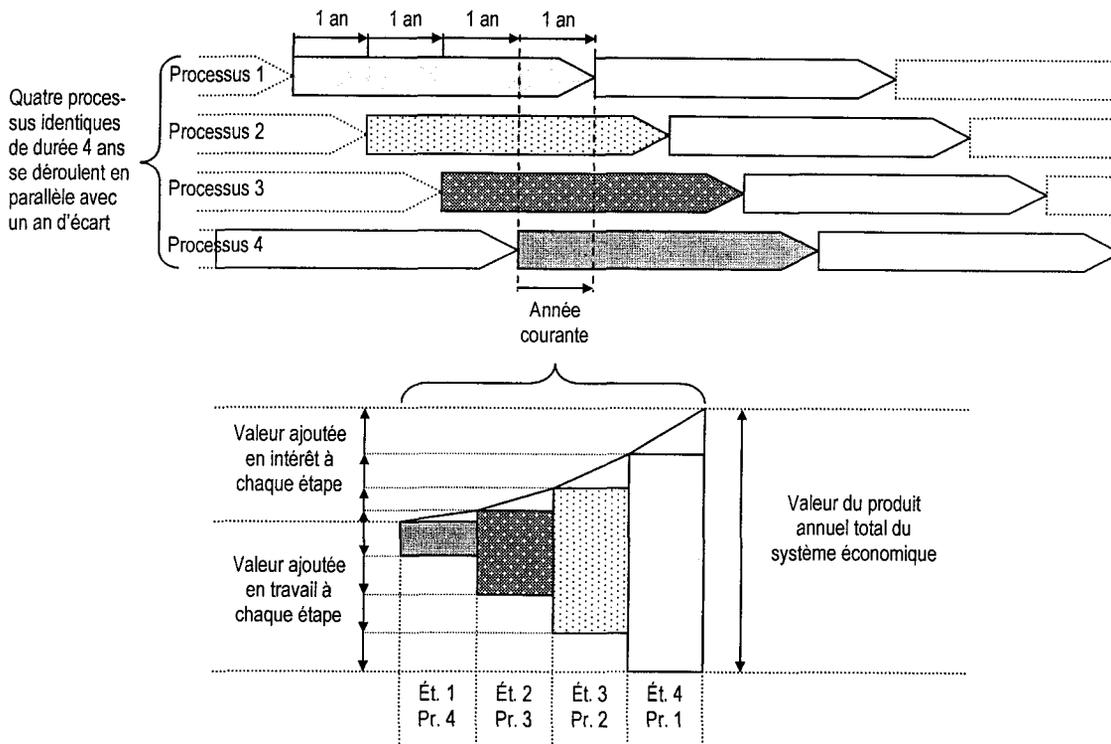


Figure 4. Synchronisation des processus et structure de production hayekienne

La structure de production de l'année courante (figure 4) permet de visualiser les quatre processus qui se déroulent simultanément. Compte tenu de leurs décalages temporels, chacun se situe à une étape différente : le processus 1 arrive à son terme (étape 4), le processus 4 commence (étape 1), et les processus 2 et 3 occupent les positions intermédiaires (respectivement à l'étape 3 et l'étape 2). En d'autres termes, lors de l'année courante : pendant que les travailleurs de l'étape 1 récupèrent les minerais (début du processus 4), ceux de l'étape 2 produisent les métaux (grâce aux minerais récupérés l'année précédente dans le processus 3), ceux de l'étape 3 fabriquent les machines (grâce aux métaux produits l'année précédente dans le processus 2) et enfin ceux de l'étape 4 produisent les biens de consommation (grâce aux machines produites l'année précédente dans le processus 1). Les segments horizontaux de la structure représentent le *temps* : en l'occurrence, chaque étape dure une année. Les segments verticaux représentent les *prix* (les valeurs) des facteurs et des produits de chacune des quatre étapes : comme on le voit sur la figure 4, à chaque étape de la valeur s'ajoute en travail dans la partie inférieure de la structure (rectangles), et de la valeur s'ajoute en intérêt originaire dans la partie supérieure (triangles). La valeur du produit total annuel est la somme des valeurs ajoutées – en travail et en intérêt originaire – à toutes les étapes.

Au terme de cette présentation, il est clair que dans la perspective autrichienne l'équilibre macroéconomique est un équilibre « intertemporel » en ce sens qu'il requiert un taux d'intérêt originaire uniforme à *toutes* les étapes de production. Lorsque les taux sont différents d'une étape à l'autre, la convergence vers l'équilibre ne se réduit évidemment pas à l'équilibrage du marché des fonds prêtables. Il s'agit d'un processus plus complexe qui égalise les taux d'intérêt originaires par la réallocation des facteurs de production des étapes où le taux est plus faible vers celles où le taux est plus élevé. Si par exemple le taux d'intérêt originaire est plus important à une certaine étape *E* qu'aux autres étapes, les entrepreneurs vont investir davantage à cette étape pour bénéficier de ce taux avantageux. Ce faisant ils vont tendre à réduire sa profitabilité puisque d'un côté les coûts de production vont tendre à augmenter

(puisque la demande de facteurs à l'étape E augmente), et de l'autre le revenu agrégé de la production de l'étape E va tendre à baisser (la demande de facteurs baisse aux autres étapes, et en particulier à l'étape $E + 1$ qui suit l'étape E ; or la demande de biens de capitaux de l'étape $E + 1$ constitue justement le revenu de l'étape E). Les tendances à la baisse du revenu et à la hausse des coûts impliquent que le taux d'intérêt originaire (écart relatif du revenu et des coûts) de l'étape E tend à baisser et à revenir au niveau des taux des autres étapes (qui, de leur côté, tendent à augmenter pour des raisons symétriques). C'est donc bien le système des prix de marché – prix des facteurs et prix des produits – qui régule les taux d'intérêt des différentes étapes et promeut la coordination macroéconomique. Le taux d'intérêt du marché monétaire est un indicateur beaucoup trop agrégé, beaucoup trop synthétique pour assurer cette coordination dans une perspective autrichienne.

Passons au cas de la croissance. Dans le paradigme autrichien, la croissance par intensification capitaliste consiste en un allongement de la structure de production (Böhm-Bawerk 1889[1959], Hayek 1931[1935], Rothbard 1962). Cet allongement résulte d'une baisse de la demande nominale de biens de consommation et d'une augmentation de la demande de facteurs de production. Or ce double mouvement va avoir pour effet de réduire les taux originaires des étapes proches de la consommation et d'accroître les taux des étapes plus éloignées (voir Fillieule 2005 pour une présentation formelle). L'égalisation de ces taux va, comme précédemment, s'opérer par une réallocation des facteurs de production des étapes proches de la consommation (devenues relativement moins rentables) vers les étapes plus éloignées (devenues relativement plus rentables). À nouveau, dans ce cas, ce qui se passe sur le marché des fonds prêtables ne peut suffire à rendre compte de ces déformations de la structure de production. Il faut recourir à une désagrégation aux différentes étapes et raisonner en terme d'intérêt originaire en s'appuyant sur le fonctionnement du système des prix.

En fin de compte, la théorie du marché des fonds prêtables ne permet pas d'analyser les processus de coordination et de croissance avec le degré de finesse requis par le paradigme autrichien. Il n'est donc pas étonnant qu'elle ne soit pas utilisée pour traiter ces thèmes par les auteurs classiques de l'école autrichienne.

3. LA THÉORIE DE LA CONSOMMATION NETTE (REISMAN)

Les travaux de Reisman (1996, chap. 16-17) sont intéressants pour notre discussion car ils montrent en quoi peut consister une théorie macroéconomique de l'intérêt qui à la fois se situe dans un cadre autrichien (miseso-hayekien, si l'on peut dire) et se passe de toute référence au marché des fonds prêtables. Reisman a formulé une théorie de l'intérêt très détaillée et très approfondie. Son texte sur le sujet équivaut à un livre de 300 pages, contre une vingtaine de pages tout au plus dans les ouvrages de Skousen ou de Garrison. Faute de place, nous devons laisser de côté tous les approfondissements pour nous concentrer sur l'essentiel. D'emblée, Reisman se démarque nettement de la théorie du marché des fonds prêtables²⁸. Il se situe dans la lignée subjectiviste de Mises, en ce sens que pour lui la « cause ultime » de l'existence de l'intérêt se trouve dans la préférence pour le présent, mais il développe principalement les aspects macroéconomiques de la question de l'intérêt en s'appuyant sur la notion hayekienne

²⁸ « In no sense is the rate of profit determined by the supply of and demand for "capital". Rather, it has everything to do with the difference between two demands : the demand for factors of production by business – that is, the demand for capital goods and producer's labor – which shows up as costs of production to be deducted from sales revenues, and the greater demand for the products of business, which constitutes business sales revenues. Profit, in essence, is the difference between these two demands" (Reisman 1996, p. 721). L'expression « demande » est utilisée ici, non pas au sens néo-classique d'une fonction ou d'une courbe reliant des prix et des quantités, mais au sens classique d'une *dépense*. Reisman emploie le terme « profit » pour désigner l'*intérêt originaire*, c'est-à-dire la différence de prix entre les produits et les facteurs qui ont servi à les produire.

de structure de production. Il commence par analyser une structure de production à l'équilibre statique, et montre que l'intérêt originaire agrégé est égal à ce qu'il nomme la « consommation nette » (c'est-à-dire la consommation des capitalistes), puis il examine les effets sur la valeur de l'intérêt originaire d'un investissement net. Sa théorie nous paraît à plusieurs égards supérieure à celle du marché des fonds prêtables, nous expliquerons pourquoi à la fin de cette section.

3.1 Structure de production, intérêt agrégé et « consommation nette »

Reisman raisonne sur la base d'un modèle simple de structure de production représenté à la figure 5. Au cours de chaque année n se déroule un processus de production dans lequel la quantité K_n de biens de capitaux est combinée à la quantité L_n de travail pour produire : d'une part, la quantité C_{n+1} de biens de consommation qui sera consommée tout au long de l'année suivante (l'année $n+1$), et d'autre part, la quantité K_{n+1} de biens de capitaux qui sera utilisée au cours de l'année $n+1$ (et servira à produire les biens de consommation et les biens de capitaux de l'année $n+2$). Les biens de capitaux K_n sont entièrement détruits lors de leur utilisation pendant l'année n , et il en va de même pour K_{n+1} lors de l'année $n+1$, et ainsi de suite. Cette hypothèse est assez artificielle mais elle facilite beaucoup le calcul des coûts de production et donc de l'intérêt originaire.

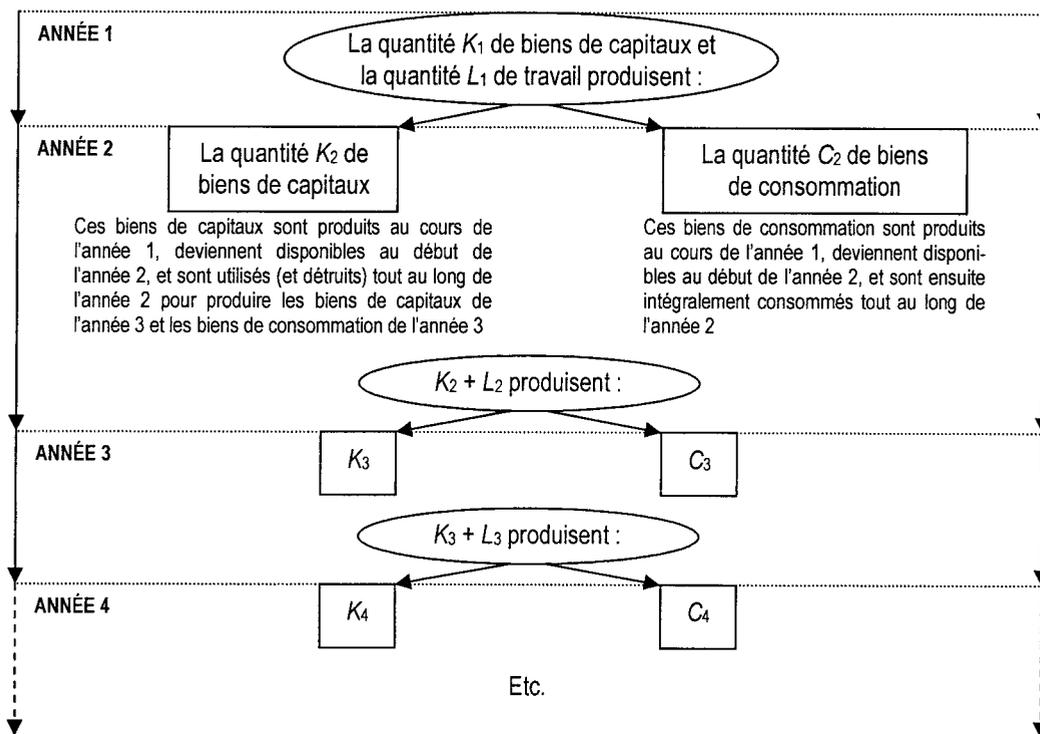


Figure 5. Le déroulement temporel du processus de production (d'après Reisman 1996, p. 729)

Reisman attribue ensuite des valeurs numériques aux différents types de revenus et de dépenses annuels de la production. Les producteurs vendent deux types de biens, des biens de consommation et des biens de capitaux. Reisman suppose que le revenu annuel de la vente des biens de consommation est de 500 unités monétaires, et que celui des biens de capitaux est lui aussi égal à 500. La production rapporte donc chaque année un revenu brut de 1 000 unités monétaires. Les producteurs effectuent deux types de dépenses, des dépenses en salai-

res (= 300 unités monétaires par an) et des dépenses en biens de capitaux (= 500 puisque dans un système statique la dépense annuelle en biens de capitaux est égale au revenu annuel de leur vente). Si l'on suppose que le système économique se reproduit à l'identique d'une année sur l'autre, on obtient la succession de revenus et de dépenses représentée dans le tableau 1.

Tableau 1. Revenus, dépenses et intérêt originaire annuels dans une structure de production à l'équilibre statique (d'après Reisman 1996, p. 730)

Année	Revenus annuels de la production			Dépenses annuelles de production (coûts pris en compte l'année suivante)			Consommation nette	Intérêt originaire
	Vente des biens de consom.	Vente des biens de capitaux	TOTAL	Salaires	Achats de biens de capitaux	TOTAL		
1	$R_{C1} = 500$	$R_{K1} = 500$	$R_1 = 1\ 000$	$D_{S1} = 300$	$D_{K1} = 500$	$D_1 = 800$	$R_{C1} - D_{S1} = 200$	~
2	$R_{C2} = 500$	$R_{K2} = 500$	$R_2 = 1\ 000$	$D_{S2} = 300$	$D_{K2} = 500$	$D_2 = 800$	$R_{C2} - D_{S2} = 200$	$R_2 - D_1 = 200$
3	$R_{C3} = 500$	$R_{K3} = 500$	$R_3 = 1\ 000$	$D_{S3} = 300$	$D_{K3} = 500$	$D_3 = 800$	$R_{C3} - D_{S3} = 200$	$R_3 - D_2 = 200$
4	$R_{C4} = 500$	$R_{K4} = 500$	$R_4 = 1\ 000$	$D_{S4} = 300$	$D_{K4} = 500$	$D_4 = 800$	$R_{C4} - D_{S4} = 200$	$R_4 - D_3 = 200$
...
n	$R_{Cn} = 500$	$R_{Kn} = 500$	$R_n = 1\ 000$	$D_{Sn} = 300$	$D_{Kn} = 500$	$D_n = 800$	$R_{Cn} - D_{Sn} = 200$	$R_n - D_{n-1} = 200$

La *consommation nette* est définie par Reisman comme la dépense annuelle totale de consommation, nette de toute dépense productive. Lorsqu'un travailleur dépense son salaire pour consommer, sa consommation ne fait pas partie de la consommation nette puisqu'elle correspond à une dépense productive salariale effectuée au préalable par l'entreprise où il travaille. La consommation des capitalistes, en revanche, s'effectue sur la base des revenus de la production et non pas sur la base de ses coûts²⁹. Ainsi : « au fond, en allant à l'essentiel, la consommation nette est la *dépense de consommation des hommes d'affaires et des capitalistes*, en incluant parmi ces derniers les créiteurs des entreprises aussi bien que les actionnaires » (Reisman 1996, p. 725).

Dans le tableau 1, la consommation nette est calculée par la formule $R_C - D_S$: revenu de la vente des biens de consommation de l'année courante moins dépense salariale de l'année courante ($500 - 300 = 200$). Compte tenu de l'hypothèse simplificatrice selon laquelle les biens de capitaux produits au cours d'une année sont entièrement détruits par leur utilisation lors de l'année suivante, l'intérêt originaire est facile à calculer : aux revenus de la production vendue lors de l'année courante (année n) doivent être retranchés les coûts de production, ces coûts étant intégralement encourus lors de l'année précédente ($n - 1$). L'intérêt originaire est donc calculé par la formule $R_n - D_{n-1}$: revenu total de la production lors de l'année n moins dépenses totales de production lors de l'année $n - 1$ ($1\ 000_n - 800_{n-1} = 200_n$). L'intérêt originaire agrégé est donc bien égal à la consommation nette, et cette égalité paraît à Reisman suffisamment importante pour qu'il donne son nom à sa théorie macroéconomique de l'intérêt.

3.2 La « cause ultime » de l'intérêt originaire : la préférence pour le présent

Reisman écrit que la consommation nette est un « déterminant » (*determinant*, p. 723), une

²⁹ Rappelons que la conception autrichienne traditionnelle de l'intérêt, à laquelle Reisman se rattache, est celle de l'intérêt *originaire*. Dans cette conception, l'intérêt est considéré comme un *revenu* et non pas (surtout pas, pourrait-on même dire) comme un coût de production.

« source » (*source*, p. 726) de l'intérêt originaire. Mais l'emploi de ces termes est quelque peu abusif car la relation entre la consommation nette et l'intérêt originaire agrégé est plutôt de l'ordre d'une égalité comptable que d'une causalité. Reisman lui-même reconnaît que la relation de causalité se situe à un autre niveau :

« La consommation nette n'est pas une cause ultime du profit [de l'intérêt originaire]. Elle reflète à son tour l'opération de la préférence temporelle. [...] la préférence temporelle est le fait que, toutes choses égales par ailleurs, les gens évaluent plus haut la satisfaction de leurs besoins dans le présent que dans le futur, et dans le futur plus proche que dans le futur plus éloigné » (1996, p. 743).

Il se place ici explicitement dans la tradition de von Mises (1949), qui a été le premier à énoncer le principe de la préférence pour le présent tel qu'il est formulé dans cette citation.

La relation entre la préférence pour le présent et l'intérêt originaire est illustrée dans le tableau 2. Une baisse de la préférence pour le présent implique que les agents ont tendance à épargner davantage et à consommer moins. En d'autres termes, la dépense en facteurs de production augmente (l'investissement s'accroît) et la dépense de consommation diminue. Ceci entraîne, toutes choses égales par ailleurs, une baisse de la consommation nette et donc une baisse de l'intérêt originaire agrégé. Dans le tableau 2, et par rapport aux données du tableau 1, les capitalistes ont réduit leurs dépenses de consommation de 200 à 100 unités monétaires par an. Ils ont augmenté d'autant leurs dépenses d'investissement, les ventes/achats annuelles de biens de capitaux passant de 500 à 600³⁰. La consommation nette et l'intérêt originaire ne valent plus que 100 (au lieu de 200 précédemment). Si l'on calcule le taux d'intérêt originaire comme la différence relative entre le revenu et les coûts de la production³¹, alors il vaut $(1\ 000 - 800)/800 = 25\ %$ pour le tableau 1 et il vaut $(1\ 000 - 900)/900 = 11\ %$ pour le tableau 2. La baisse de la préférence pour le présent entraîne donc une hausse de l'investissement (de 800 à 900) et une baisse du taux d'intérêt (de 25 % à 11 %). Si, à l'inverse, on avait supposé que la préférence pour le présent avait augmenté, alors la dépense productive aurait diminué et la consommation augmenté, conduisant à une baisse de l'investissement et une hausse de l'intérêt agrégé et du taux d'intérêt. La théorie de Reisman aboutit ici à des résultats similaires à ceux que l'on obtiendrait dans la théorie du marché des fonds prêtables et qui sont illustrés dans la figure 2.

³⁰ Reisman se place sous l'hypothèse que la dépense totale annuelle reste constante, qu'il appelle « l'hypothèse de monnaie invariable ».

³¹ Notons que ce n'est pas ainsi que Reisman calcule le taux d'intérêt. Les taux qu'il obtient sont systématiquement plus faibles car il intègre au capital, en plus des dépenses productives, les sommes de monnaie détenues par les capitalistes (le dénominateur est donc plus élevé et le taux résultant est plus faible).

Tableau 2. Revenus, dépenses et intérêt originaire annuels, pour un degré de préférence pour le présent plus faible que dans le tableau 1 (d'après Reisman 1996, p. 736)

Année	Revenus annuels de la production			Dépenses annuelles de production (coûts pris en compte l'année suivante)			Consommation Nette	Intérêt originaire
	Vente des biens de consom. R_C	Vente des biens de capitaux R_K	TOTAL $R = R_C + R_K$	Salaires D_S	Achats de biens de capitaux D_K	TOTAL $D = D_S + D_K$		
1	400	600	$R_1 = 1\ 000$	300	600	$D_1 = 900$	100	~
2	400	600	$R_2 = 1\ 000$	300	600	$D_2 = 900$	100	$R_2 - D_1 = 100$
3	400	600	$R_3 = 1\ 000$	300	600	$D_3 = 900$	100	$R_3 - D_2 = 100$
4	400	600	$R_4 = 1\ 000$	300	600	$D_4 = 900$	100	$R_4 - D_3 = 100$
...
n	400	600	$R_n = 1\ 000$	300	600	$D_n = 900$	100	$R_n - D_{n-1} = 100$

3.3 Les effets d'un investissement net

La seconde relation majeure mise en évidence par Reisman, après l'égalité entre intérêt agrégé et consommation nette, concerne l'effet sur l'intérêt d'un investissement net. Il montre que lorsque survient un investissement net, l'intérêt agrégé est égal à la somme de la consommation nette et de l'investissement net. Cette relation est illustrée par le tableau 3.

Tableau 3. Investissement net et intérêt originaire (d'après Reisman 1996, p. 745)

Année	Revenus annuels de la production			Dépenses annuelles de production (coûts pris en compte l'année suivante)			Consommation nette $R_C - D_S$	Investissement net ΔI	Intérêt originaire
	Vente des biens de consom. R_C	Vente des biens de capitaux R_K	TOTAL $R = R_C + R_K$	Salaires D_S	Achats de biens de capitaux D_K	TOTAL $D = D_S + D_K$			
1	500	500	$R_1 = 1\ 000$	300	500	$D_1 = 800$	200	0	$R_1 - D_0 = 200$
2	500	500	$R_2 = 1\ 000$	300	500	$D_2 = 800$	200	0	$R_2 - D_1 = 200$
3	400	600	$R_3 = 1\ 000$	300	600	$D_3 = 900$	100	100	$R_3 - D_2 = 200$
4	400	600	$R_4 = 1\ 000$	300	600	$D_4 = 900$	100	0	$R_4 - D_3 = 100$
5	400	600	$R_5 = 1\ 000$	300	600	$D_5 = 900$	100	0	$R_5 - D_4 = 100$
...
n	400	600	$R_n = 1\ 000$	300	600	$D_n = 900$	100	0	$R_n - D_{n-1} = 100$

Lors de l'année 3 survient un investissement net : les capitalistes décident d'accroître leurs investissements au détriment de leur consommation (leur préférence pour le présent a baissé). Plus précisément, ils renoncent à 100 unités monétaires de consommation pour dépenser 100 unités supplémentaires en investissement (biens de capitaux). L'intérêt originaire agrégé de l'année 3 est égal au revenu de la production de l'année 3 (qui n'a pas changé puisque la

baisse du revenu de la vente des biens de consommation est compensée par la hausse du revenu de la vente des biens de capitaux) net des coûts de production encourus l'année précédente. L'intérêt originaire est donc égal à : $R_3 - D_2 = 1\ 000 - 800 = 200$. Or la consommation nette de l'année 3 est en baisse par rapport à l'année 2 puisque la dépense de consommation a baissé (= 400) alors que la masse salariale est restée constante (= 300). La consommation nette vaut donc : $400 - 300 = 100$. Dans ces conditions, l'intérêt originaire de l'année 3 (= 200) n'est plus égal à la consommation nette, mais à la somme de la consommation nette de la *nouvelle* structure de production (= 100) et de l'investissement net (= 100). Lors de l'année suivante (année 4), puis des années ultérieures, il n'y a plus d'investissement net et l'intérêt (= 100) est à nouveau égal à la consommation nette (= 100). Le tableau 3, tout en illustrant l'effet d'un investissement net, permet de voir comment évoluent les variables macroéconomiques nominales lors de la phase de transition entre la structure décrite par le tableau 1 et celle décrite par le tableau 2 (dans laquelle la préférence pour le présent a baissé)³².

3.4 Quelques éléments de comparaison entre la théorie de Reisman et la théorie du marché des fonds prêtables

La théorie du marché des fonds prêtables possède à première vue un avantage par rapport à celle de Reisman. Elle propose en effet un *principe de détermination* du taux d'intérêt et de l'investissement d'équilibre, qui est tout simplement le point de croisement des courbes d'offre et de demande de fonds prêtables. La théorie de Reisman, en revanche, ne possède pas un tel principe de détermination. Cette dernière repose bien sur une relation de causalité, qui est celle entre préférence pour le présent et intérêt, mais cette relation est mise en évidence de façon seulement indirecte, par le biais de la relation entre les *variations* de ces variables. Comme on l'a vu dans les tableaux 2 et 3, la baisse de la préférence pour le présent entraîne dans la théorie de Reisman une baisse de l'intérêt et une hausse de l'investissement. Mais cette théorie ne spécifie aucun principe de détermination conduisant, pour des préférences intertemporelles données des agents, de ces préférences vers le taux d'intérêt et l'investissement d'équilibre. Elle se contente de montrer que si la préférence pour le présent diminue, alors l'intérêt baisse et l'investissement augmente (et inversement). De ce point de vue, Reisman se rattache bien à la tradition de von Mises puisque ce dernier mettait lui aussi en évidence la relation entre préférences intertemporelles et taux d'intérêt par un raisonnement portant sur leurs variations relatives³³. Cet avantage de la théorie du marché des fonds prêtables est pourtant très discutable car, comme on l'a dit plus haut, à l'équilibre de long terme les capitalistes n'ont pas besoin d'intermédiaires. Ils investissent directement leurs fonds dans leurs entreprises, et il n'existe donc pas de marché des capitaux. Le principe de détermination qui est à la base de la théorie du marché des fonds prêtables (croisement des courbes d'offre et de demande de capitaux) a donc tendance à se dissoudre lorsque le système converge vers son équilibre final, et l'avantage qu'il confère à cette théorie par rapport à celle de Reisman apparaît en fin de compte comme purement superficiel. En ne présupposant pas l'existence d'un marché des capitaux et en raisonnant directement sur les prix des biens, c'est-

³² La troisième et dernière relation majeure analysée par Reisman concerne l'ajout d'une nouvelle quantité de monnaie. Le résultat qu'il met ici en évidence est plus conventionnel : « The rate of increase in the quantity of money and, accordingly, the volume of spending in the economic system, tends to cause an *approximately equivalent increase in the rate of profit* » (1996, p. 763). Nous ne développerons pas ce point.

³³ Mises (1949, p. 524) évoque les deux cas polaires, celui où les gens penseraient que la fin du monde est toute proche (les facteurs de production perdraient alors leur valeur et le taux d'intérêt originaire tendrait vers l'infini puisque la valeur des facteurs qui constitue son dénominateur tendrait vers zéro) et celui où la préférence pour le présent disparaîtrait (ce qui entraînerait une disparition de la consommation et une annulation du taux d'intérêt puisque son numérateur tendrait lui-même vers zéro). Rothbard (1962, p. 380-381) utilise des illustrations similaires.

à-dire sur l'intérêt originaire, la théorie de Reisman se situe à un niveau plus fondamental que celle du marché des fonds prêtables et s'avère de ce point de vue plus satisfaisante que cette dernière.

Les autres éléments tendant à conforter la supériorité de la théorie de Reisman nous semblent être les suivants.

(a) La théorie de Reisman donne accès aux variations de l'intérêt agrégé, et pas seulement aux variations du taux. Lorsque le taux d'intérêt et l'investissement varient en sens inverse (par exemple lorsque l'offre de fonds varie à demande constante), la théorie du marché des fonds prêtables ne peut pas se prononcer sur le sens de la variation de l'intérêt agrégé, à moins de faire une hypothèse sur l'élasticité de la demande de fonds.

(b) La théorie de Reisman permet de distinguer les aspects temporaires et les aspects définitifs de certains changements. Un investissement net, par exemple, accroît d'autant l'intérêt agrégé pendant l'année courante (effet temporaire), puis conduit à une baisse de l'intérêt (effet de long terme). La théorie du marché des fonds prêtables est incapable de mettre en évidence la relation temporaire entre intérêt agrégé et investissement (ou désinvestissement) net.

(c) La théorie du marché des fonds prêtables peut être interprétée en termes nominaux, comme le font les Autrichiens (en particulier Garrison 2001 et Murphy 2003), ou en termes réels, comme le fait Mankiw (1998). Ces deux possibilités d'interprétation peuvent entraîner des confusions qui n'existent pas chez Reisman puisque sa théorie est très clairement, sans la moindre ambiguïté possible, une théorie portant sur les variables nominales.

(d) La théorie de Reisman s'appuie sur une représentation de la structure de production, c'est-à-dire sur une image globale des interactions entre les grandes variables macroéconomiques, et en particulier entre les dépenses de consommation et d'investissement. Elle conduit donc à faire des raisonnements systématiques : dès qu'un revenu/dépense varie, il faut se demander d'où provient cette variation et quels sont les autres revenus/dépenses affectés dans la structure. Dans la théorie du marché des fonds prêtables, il existe le risque de séparer ce qui se passe sur ce marché de ce qui se passe dans le reste du système économique. Les économistes autrichiens évitent de tomber dans ce travers, et Garrison, en particulier, a développé un modèle destiné à tenir compte des interdépendances macroéconomiques (Garrison 2001, p. 50). Un exemple emprunté à Mankiw permet néanmoins d'illustrer la différence entre les deux démarches. D'après Mankiw (1998, p. 693-694), si le gouvernement décide d'instaurer un impôt sur les investissements, cela aura pour effet de réduire la demande de fonds prêtables à offre constante, et donc de diminuer l'investissement et le taux d'intérêt. Plaçons-nous maintenant dans le cadre de la théorie de Reisman. Le raisonnement que l'on doit faire est alors tout à fait différent. Il faut d'abord se demander à quoi va servir l'impôt ainsi prélevé par l'État. Il importe surtout de savoir si cet impôt va servir à financer un investissement (dans une entreprise concurrentielle visant à maximiser son profit) ou une consommation. Nous supposons qu'il va servir à financer des traitements de fonctionnaires, traitements qui seront à leur tour intégralement dépensés en consommation. Il s'agit ensuite de déterminer comment les capitalistes vont réagir à la mise en place de cet impôt. À un extrême, ils pourraient maintenir à l'identique leur dépense d'investissement et *payer la totalité de l'impôt sur leur dépense de consommation*. Dans ce cas, l'investissement ne changerait pas (cette conclusion contredit celle de Mankiw), ni non plus l'intérêt agrégé avant impôt. À un autre extrême, les capitalistes pourraient maintenir à l'identique leur dépense de consommation et *payer l'intégralité de l'impôt sur leurs dépenses d'investissement*. Dans ce cas, la dépense de consommation des fonctionnaires se substituerait à la dépense d'investissement ; la part « productive » de l'investissement (part consacrée à l'achat des facteurs de production) se réduirait, alors que l'intérêt originaire agrégé ne changerait pas. De fait, la réaction des capitalistes va se situer entre ces deux extrêmes : l'impôt supplémentaire va pour une part se substituer à la consommation des capitalistes (consommation nette) et pour une autre part se substi-

tuer à leur dépense d'investissement. L'intérêt agrégé avant impôt ne changera donc pas, mais l'investissement « productif » se trouvera partiellement réduit. Il est clair que le raisonnement fondé sur la théorie de Reisman incite à davantage de rigueur que celui – trop étroit – fondé sur la théorie du marché des fonds prêtables.

CONCLUSION

La question de la détermination de l'intérêt ne nous paraît pas obtenir de réponse pleinement satisfaisante dans la macroéconomie autrichienne contemporaine. La théorie du marché des fonds prêtables pose des difficultés majeures. Elle ne permet en effet pas de rendre compte du phénomène de l'intérêt originaire, ni des mécanismes de coordination intertemporelle et de croissance économique tels qu'ils sont conçus dans le paradigme autrichien. Il est donc assez surprenant que des macroéconomistes contemporains de l'école autrichienne aient exclusivement recours à cette théorie pour expliquer le phénomène de l'intérêt. Elle n'a en outre rien de spécifiquement « autrichien » puisqu'elle constitue aujourd'hui l'un des piliers de la macroéconomie « standard ». La théorie de la consommation nette de Reisman entre beaucoup mieux dans le cadre du paradigme autrichien dans la mesure où elle vise à expliquer l'intérêt originaire, et non pas seulement l'intérêt monétaire qui apparaît sur le marché des capitaux. Cette dernière pose pourtant elle aussi une difficulté qui tient au fait qu'elle ne propose aucun principe de détermination de l'équilibre macroéconomique. Cet équilibre est pré-supposé dans les illustrations de Reisman, sans que ses caractéristiques aient été explicitement indiquées. Reisman ouvre cependant une voie qui mérite d'être explorée et approfondie car elle pourrait permettre de doter la macroéconomie autrichienne d'une théorie de l'intérêt solide et convaincante.

Bibliographie

- Böhm-Bawerk, Eugen von (1884[1959]) *History and Critique of Interest Theories*, trad. de la 4^e éd., South Holland, Libertarian Press.
- Böhm-Bawerk, Eugen von (1889[1959]) *Positive Theory of Capital*, trad. de la 4^e éd., South Holland, Libertarian Press.
- Fetter, Frank A. (1927) « Interest Theory and Price Movements », *American Economic Review*, suppl., vol. 17, repris dans Fetter (1977, p. 260-316).
- Fetter, Frank A. (1977) *Capital, Interest, and Rent. Essays in the Theory of Distribution*, Kansas City, Sheed Andrews and McMeel.
- Fillieule, Renaud (2005) « A Formal Model of the Hayekian Structure of Production », article non publié, 26 p.
- Garrison, Roger (1979) « In Defense of the Misesian Theory of Interest », *Journal of Libertarian Studies*, vol. 3(2).
- Garrison, Roger (2001) *Time and Money. The Macroeconomics of the Capital Structure*, Londres, Routledge.
- Hayek, Friedrich A. (1931[1935]) *Prices and Production*, 2^e éd., Londres, Routledge.
- Hayek, Friedrich A. (1941) *The Pure Theory of Capital*, Chicago, The University of Chicago Press.
- Horwitz, Steven (2000) *Microfoundations and Macroeconomics. An Austrian Perspective*, Londres, Routledge.
- Hülsmann, Jörg Guido (2002) « A Theory of Interest », *The Quarterly Journal of Austrian Economics*, vol. 5, p. 77-110.

- Mankiw, N. Gregory (1998) *Principes de l'économie*, Paris, Economica.
- Mises, Ludwig von (1949) *Human Action. A Treatise on Economics*, réimpression de la 1^{re} éd., Auburn, The Ludwig von Mises Institute, 1998.
- Murphy, Robert (2003) *Unanticipated Intertemporal Change in Theories of Interest*, unpublished Ph.D. dissertation, Department of Economics, New York University.
- O'Driscoll, Gerald P. (1977) *Economics as a Coordination Problem. The Contributions of Friedrich A. Hayek*, Kansas City, Sheed Andrews and McMeel.
- O'Driscoll, Gerald P. et Rizzo, Mario J. (1985) *The Economics of Time and Ignorance*, Oxford, Basil Blackwell.
- Reisman, George (1996) *Capitalism. A Treatise on Economics*, Ottawa, Jameson Books.
- Rothbard, Murray N. (1962) *Man, Economy, and State. A Treatise on Economic Principles*, 2 vol., Princeton, Van Nostrand.
- Skousen, Mark (1990) *The Structure of Production*, New York, New York University Press.
- Skousen, Mark (1991) *Economics on Trial. Lies, Myths, and Realities*, New York, Irwin.

