

DE PIAGET À HABERMAS ET RAWLS: LES PROBLÈMES DE LA RECONSTRUCTION RATIONNELLE DU JUGEMENT MORAL CHEZ KOHLBERG

Freddy Mortier

1. *Introduction*

La psychologie est un terrain de rencontre choisi entre la philosophie et les sciences humaines. Le psychologue qui entend étudier le comportement moral, et plus particulièrement les compétences que les sujets mettent en oeuvre pour résoudre des problèmes moraux, se tourne communément vers la philosophie. Il est apparu que les théories morales, et plus particulièrement la méta-éthique, sont un outil heuristique indispensable aux pénibles travaux de la classification et de la reconstruction des principes de l'intuition des sujets.¹ D'autre part, les psychologues sont amenés à adresser des questions spécifiquement philosophiques à leur propre discipline.

A ce sujet, dans les sciences de la morale deux positions s'affrontent. La première avance que la psychologie morale, bien que présupposant des partis pris philosophiques quant à la définition de la morale, s'occupe en fait de la description et de l'explication de l'expérience et du comportement moral. Les tenants de cette position admettent souvent qu'il faut d'abord clarifier la signification du concept de 'morale', mais que l'apport de la philosophie s'arrête là. Plus précisément, l'étude de la morale ne présuppose pas de parti pris normatif. On peut bien admettre que les sciences humaines ne sont pas 'neutres' au sens large: elles mettent en oeuvre des catégories et des schèmes de pensée (paradigmes, analogies,

¹ Voir p.ex. Th. Wren (ed.), *The Moral Domain. Essays in the Ongoing Discussion Between Moral Philosophy and the Social Sciences*. Cambridge, Mass., MIT-Press, 1990.

métaphores, etc.) imprègnés d'évaluations cognitives. Comme toute pensée, la pensée psychologique est essentiellement découpage et réorganisation. Elle reflète des jugements et des décisions, qui groupent et sélectionnent à partir de critères le plus souvent implicites et donc 'non neutre'. Mais la psychologie ou la sociologie morale n'impliquent pas nécessairement des jugements de valeur sur les intuitions et les comportements étudiés. Les sciences peuvent faire 'l'époché' de la question de la validité des positions éthiques exprimées par le sujet. La science de la morale n'est pas elle-même fondée sur une morale. Ses descriptions ne sont pas des prescriptions déguisées ou avouées, ou, du moins, elles ne le sont pas nécessairement. Le chercheur est en mesure de transcender le point de vue du participant dans la vie morale et de se placer au point de vue de l'observateur.²

Le seconde position s'oppose terme-à-terme à la première. Les sciences de la morale, et plus précisément celles qui traitent de sa genèse, doivent être en mesure de se faire une idée sur le degré de rationalité spécifiquement morale des jugements ou comportements étudiés. Ainsi donc, la description-même du développement moral est logiquement impossible sans avoir recours à une conception morale normative. Cette nécessité découle de ce que l'expérience morale vécue des sujets implique la référence implicite ou explicite à des critères de rationalité éthique: le chercheur ne peut se dispenser de juger de ces critères s'il veut objectiver l'expérience morale proprement dite. La position du participant est donc inhérente à la tentative d'objectivation scientifique. Cette position est intimement liée à une certaine conception des conditions qui façonnent la compréhension de l'expérience morale. Ces conditions engageraient le

² N. Luhmann, 'Soziologie der Moral', in : N. Luhmann, S.H. Pfürtnner, *Theorietechnik und Moral*, Frankfurt/M., Suhrkamp, 1978, pp. 8 - 116; B. Puka, 'Toward the Redevelopment of Kohlberg's Theory', in: W. Kurtines, J. Gewirtz, *Handbook of Moral Behavior and Development. Vol. I: Theory*. Hillsdale, NJ, L. Erlbaum, 1991, pp. 373-94; A. Bandura, *Social Foundations of Thought and Action. A Social Cognitive Theory*. Englewood Cliffs, Prentice-Hall, 1986; L. Boltanski et L. Thévenot, *De la justification*. Paris, Gallimard, 1991; A. Blasi, 'Psychologische oder philosophische Definition der Moral. Schädliche Einflüsse der Philosophie auf die Moralpsychologie', in: W. Edelstein, G. Nunner-Winkler (Hrsg.), *Zur Bestimmung der Moral. Philosophische und Moralphilosophische Beiträge zur Moralforschung*. Frankfurt/M., Suhrkamp, 1986, pp. 55-85.

chercheur activement.³

La discussion entre ces deux partis est fréquemment brouillé par une autre, qui, bien que proche de la première, n'en est pas identique. Il s'agit du débat entre le réductionnisme et le non-réductionnisme, et donc de la question du langage à employer afin de décrire et d'expliquer l'expérience morale. Le langage réducteur présuppose que la référence aux points de vue à partir desquels le sujet formule ses jugements n'est pas intrinsèquement liée à la description ou à l'explication du réel. La façon dont le sujet vit phénoménalement sa moralité serait irrélevante. Le langage non-réducteur, par contre, présuppose que la compréhension par le chercheur des prétentions chez le sujet quant à la validité de son jugement moral est d'une importance fondamentale à la description et à l'explication du réel.

La position que j'appellerai 'normative' envers les sciences de la morale se double donc parfois d'une critique du sociologisme et du psychologisme réducteurs: faute d'objectiver les prétentions subjectives en regard de la validité normative, les réductionnistes sont condamnés à, comme l'exprime Genard, une 'psychologie de la non-maîtrise'⁴ et à l'utilisation de notions pseudo-causales, comme la 'socialisation'. En fait la position 'neutre' est prise aussi bien par des réductionnistes (comme Bandura) que par des 'phénoménologues' (Blasi, Puka, Boltanski et Thévenot).

Dans cet article, je me propose d'exposer et d'examiner les arguments qui ont conduit Lawrence Kohlberg à embrasser la position 'normative', à partir d'une herméneutique dite 'objective' proche de et largement inspirée par Jürgen Habermas. Depuis 1958 jusqu'à sa mort en 1987, Lawrence Kohlberg a élaboré et raffiné une psychologie morale reposant sur une réflexion philosophique poussée. Son structuralisme cognitif est aux origines d'une vaste entreprise de recherches sur la morale qui est loin d'aboutir.

³ J.-L. Genard, *Sociologie de l'éthique*. Paris, Editions L'Harmattan, 1992; J. Habermas, *Moralbewusstsein und kommunikatives Handeln*. Frankfurt/M., Suhrkamp, 1983; Z. Bauman, 'Towards a Sociological Theory of Morality', in: Idem, *Modernity and the Holocaust*. Cambridge, Polity Press, 1989, pp. 169-200; L. Kohlberg, *Essays on Moral Development, Vol II: The Psychology of Moral Development. The Nature and Validity of Moral Stages*. San Francisco, Harper & Row, 1984.

⁴ J.-L. Genard, *op. cit.*, p. 41.

Je tenterai de démontrer que la théorie des stades moraux de Kohlberg combine deux approches partiellement dissociables: une théorie 'forte' inspirée par des philosophies morales du type déontologique et universaliste; et une théorie 'faible', compatible avec un éventail de types de philosophies morales. Pour des raisons épistémologiques, seule la seconde me semble soutenable. Elle implique l'adoption de la position 'neutre'.

2. La théorie des stades de Kohlberg

Selon Kohlberg, le développement du jugement moral chez l'individu comprendrait six stades successifs. Tous les individus, dans toutes les cultures, passeraient par la même séquence de réorganisation du point de vue global à partir duquel ils envisagent les problèmes moraux.⁵ Les stades auxquels s'intéresse Kohlberg ont les quatre propriétés suivantes:

- 1) Les stades impliquent une différence qualitative entre des structures (modes de raisonnement) qui servent une fonction fondamentale invariante (par exemple, l'intelligence, le jugement moral).
- 2) Ces structures différentes se succèdent invariablement, bien qu'elles puissent être parcourues plus ou moins vite et que très peu de personnes n'atteignent le point final.
- 3) Les structures forment chacune un tout intégré et global, une organisation de la pensée, agençant de façon cohérente les raisonnements moraux du sujet.
- 4) Ces structures sont intégrées hiérarchiquement: les stades supérieurs servent mieux la fonction de base que les stades inférieurs parce qu'ils sont plus différenciés et intégrés.⁶

Les stades sont groupés en trois niveaux, qui se distinguent entre eux par la perspective morale prise par le sujet. Au niveau préconventionnel, le sujet juge plus ou moins égocentriquement et ne tient pas encore compte

⁵ L. Kohlberg, *op. cit.*, p. 286.

⁶ A. Colby, L. Kohlberg, *The Measurement of Moral Judgment. Vol I: Theoretical Foundations and Research Validation*. Cambridge, University Press, 1987, pp. 6-7.

des intérêts de communautés particulières. Au niveau conventionnel par contre, il juge de la validité de règles morales à partir de l'identification avec une communauté. Au niveau postconventionnel enfin, les règles du groupe ou de la société toute entière sont jugées elles-mêmes du point de vue d'un sujet se plaçant en dehors des communautés actuellement existantes et s'efforçant de construire des règles qui devraient idéalement régir les rapports sociaux. A l'intérieur de chaque niveau, des caractéristiques plus fines de la perspective sociale permettent de distinguer deux stades. Le premier, le stade de la moralité hétéronome, est défini par cela que l'individu ne tient pas compte des intérêts des autres et n'essaie donc même pas de les relater dans une perspective englobante. De là découlent une définition particulière de ce qui est 'juste' et des raisons pour agir 'justement': il faut obéir aux autorités supérieures et éviter les punitions. Le stade de l'individualisme et de l'échange instrumental a ses origines dans une perspective qui intègre la compréhension que chacun a de ses propres intérêts et que le 'juste' est donc relatif. Il faut promouvoir ses propres intérêts, tout en reconnaissant les intérêts des autres. Il s'agit donc d'une sorte d'égoïsme à terme plus long. Au troisième stade, le sujet juge du point de vue d'une petite communauté, dont il partage les émotions, règles et attentes (la famille, le groupe des copains, ...). Il faut donc s'efforcer d'être 'bien' à ses propres yeux et à ceux des autres, et supporter les rapports réciproques et les normes correspondantes (comme la loyauté, la confiance, le respect, la gratitude). Au stade suivant, le sujet prend le point de vue du système qui définit les rôles sociaux et les règles communes. Est 'juste' ce qui contribue à la conservation du système, ou, autrement dit, ce qui correspond aux obligations liées à la position occupée dans le système. Au cinquième stade, l'on prend la perspective 'hors-société': l'organisation sociale est elle-même jugée du point de vue des droits et des valeurs qu'elle est censée protéger et promouvoir. L'intégration des intérêts divers est opérée par le biais de mécanismes formels (une convention, un contrat, l'impartialité objective, des procédures justes). Le 'juste' consiste généralement à respecter les valeurs et les règles que le groupe a librement convenues, parce qu'elles découlent d'un contrat social. De cette façon l'utilité générale est protégée. Néanmoins, quelques valeurs, comme la vie et la liberté, sont fondamentales et non-annulables. Au dernier stade, le sujet prend le point de vue proprement moral (*moral point of view*). Sa perspective est celle d'un sujet rationnel quelconque qui reconnaît la nature de la moralité, autre-

ment dit, que les personnes sont des fins en soi et doivent être traitées comme telles. De là l'idée qu'il faut suivre des principes moraux universellement acceptables, qui sont prioritaires aux lois et contrats particuliers.

Quelques remarques s'imposent. Premièrement, à chaque stade Kohlberg distingue clairement la perspective sociale sous-jacente, la définition de la 'justesse' correspondante et les raisons invoquées pour 'agir juste'. L'attribution d'un certain stade repose sur une analyse poussée des normes et des principes invoqués dans le raisonnement justificatoire du sujet. Le schéma est accompagné d'un manuel de score, qui montre comment en arriver correctement (c'est-à-dire de façon standardisé) à une attribution de stade (voir plus loin). Secondement, bien que la procédure de distinction des stades passe par la perspective sociale, Kohlberg prétend que la transition entre les stades n'est pas déterminé par la dynamique génétique de la prise de perspective, mais par une logique proprement morale. Le principe en est à chercher, entre autres, dans la théorie de la justice de John Rawls. Je reviendrai sur ce point. En outre, Kohlberg prétend que la supériorité des stades ultérieurs sur les stades antérieurs n'est pas seulement d'ordre fonctionnel, c'est-à-dire reflétant une conception plus réaliste des rapports de coopération entre les agents sociaux. Il s'agirait encore d'un progrès de nature spécifiquement morale: les jugements rendus à partir du point de vue des stades supérieurs relèvent d'un progrès dans l'ordre de la *validité* normative. En avançant cette thèse, Kohlberg ne se rend pas simplement à 'l'argument fallacieux naturaliste' de Moore, bien qu'à un moment donné il ait cru le faire légitimement.⁷ Kohlberg soutient que des jugements de validité sont inhérents à la position occupée par le chercheur dans le processus de recherche. La conduite-même de la recherche psychologique présupposerait la validité de certaines conceptions morales. L'entreprise d'autofondation normative de la psychologie morale passe donc par la position dite 'normative' envers les sciences de la morale.

Remarque finale: il faut bien se garder de ne voir dans la théorie de Kohlberg qu'une rationalisation d'une conception particulière de la morale

⁷ L. Kohlberg, 'From is to ought. How to commit the naturalistic fallacy and get away with it in the study of moral development', in: Idem: *Essays in Moral Development, Vol. I: The Philosophy of Moral Development. Moral Stages and the Idea of Justice*. San Francisco, Harper & Row, 1981, pp. 101-189.

qui serait typique des sociétés occidentales, plus précisément de la théorie politique libérale et du kantianisme, et de l'envoyer sans ambages à la poubelle. L'étude comparative a démontré que, bien qu'il y ait bon nombre d'idiosyncrasies occidentales dans la séquence proposée par Kohlberg, elle peut être modifiée et utilisée à grand profit dans l'étude comparative de la morale.⁸

Je m'attacherai à analyser l'argumentation de Kohlberg de plus près. On voit bien ce que sa théorie des stades moraux peut avoir de séduisant (pour certains): la coïncidence entre ses prétentions à la vérité et à la validité. Nous voilà donc débarrassés de certaines questions encombrantes et voilà la psychologie transformée en véritable pierre de touche philosophique.

3. *L'héritage de Piaget*

L'argumentation de Kohlberg est complexe. Pour bien la comprendre, il s'agira d'examiner critiqueusement ses bases théoriques et les problématiques dont elles relèvent. Il est devenu commun de placer Kohlberg dans la lignée de Jean Piaget. Et en effet, la thèse de doctorat du premier a été conçue par son auteur comme une continuation des recherches que Piaget a rapportées dans son ouvrage pionnier.⁹ Interrogeant des enfants entre 4 et 13 ans, Piaget avait cru pouvoir distinguer deux 'phases' ou 'stades' distincts dans leurs valorisations: à une morale hétéronome succéderait une morale autonome. La première serait vécue par l'enfant comme un code à prendre littéralement, parce que d'origine transcendante et de caractère immuable. Le mal et le bien contenus dans une action particulière sont déterminés, de façon tout à fait extérieure, par le rapport de conformité du comportement à la règle. La morale autonome au contraire se caractériserait par l'idée que les règles de la morale sont conçues souverainement par le groupe. Les règles sont changeables parce qu'elles sont le résultat de la délibération rationnelle qui prolonge la coopération

⁸ Voir p.ex. Hing Keung Ma, 'The moral judgment development of the Chinese people: a theoretical model', *Philosophica*, 49 (1992, 1) (n° spécial: F. Mortier (ed.), *Philosophical Perspectives in Moral Psychology and Education*), pp. 55-82.

⁹ J. Piaget, *Le jugement moral chez l'enfant*. Paris, PUF, 1932.

entre les membres du groupe. A la notion de la responsabilité objective succède celle de la responsabilité subjective.

La transition de la morale hétéronome à la morale autonome ne se produirait pas automatiquement. La raison en est que la succession des phases de la moralité dépend décisivement de l'environnement social des enfants: la morale hétéronome appartient au monde de la contrainte; la morale autonome au monde de la coopération égalitaire.

Creusons un peu la problématique de Piaget, puisqu'elle a été mal comprise par Kohlberg. La problématique qui hante Piaget dans son ouvrage est d'expliquer comment s'opère le passage d'une structure à une autre. On verra que chez Kohlberg, l'accent est mis davantage sur la description des propriétés formelles des stades et de leur filiation logique. Mais il est clair qu'expliquer une transition est autre chose qu'en décrire les étapes.

Sans qu'il y ait opposition entre les deux démarches, elles conduisent à des écarts plus ou moins systématiques entre les significations des concepts clefs de respectivement Piaget et Kohlberg ('logique', 'cognitif', 'stade', etc.)

Le jugement moral date des débuts de Piaget. Je concède que je cours le risque de commettre un anachronisme, mais il me semble que l'épistémologie génétique plus tardive peut aider à clarifier les idées de Piaget. Dans l'*Introduction*¹⁰ il explique que l'explication causale offre peu d'appui dans l'étude de l'action intentionnelle. La coordination neuromusculaire par exemple ne peut pas expliquer la cohérence interne d'actions pareilles; elle ne s'explique qu'à condition de considérer les rapports internes entre les actions extériorisées et les actions intériorisées. Or, ces rapports sont de nature *logique*. Par exemple, les émotions ne sont pas seulement liées entre elles de façon causale, mais encore et surtout de façon 'interne'. Si je veux A, alors il me faut bien désirer B, si B est le moyen d'obtenir A. De cette façon, les désirs définissent des chaînes d'implications entre des 'valorisations'. Ces implications n'enlèvent rien à l'aspect causal de la vie affective, elles y ajoutent.

Ces implications, poursuit Piaget, peuvent en outre être plus ou moins structurées du point de vue strictement logique. Ainsi, il y a des

¹⁰ J. Piaget, *Introduction à l'épistémologie génétique*. Tomes I-III. Paris, PUF, 1950, III, pp. 149-52.

chaînes implicatives de nature temporaire et locale: les préférences, désirs, inclinations, émotions de plaisir, de chagrin, etc. Il y en a d'autres qui sont réglées par des normes permanentes: ce sont les valeurs normées: les passages entre les valorisations sont eux-mêmes déterminés par des implications de *second* ordre. Il est clair, selon lui, que l'explication psychologique ne peut être efficace qu'à condition de partir de l'action causale pour en aboutir en fin de compte aux systèmes des implications.¹¹

Il est intéressant de considérer l'exposé de Piaget du point de vue de la valorisation. Au niveau le plus causal, il paraît que la valorisation est guidée par des pulsions biologiques, qui n'ont pas de composante spécifiquement psychologique. Au niveau intermédiaire, celui des systèmes d'implications momentanées et locales, la direction de l'action intentionnelle est bien déterminée par des schèmes spécifiquement psychologiques. Mais la cohérence interne des valorisations n'est pas déterminée elle-même par la logique des transitions entre ces valorisations. Ce n'est qu'au niveau supérieur de la régulation normative que les valorisations forment un système. Ce système, fermé sur lui-même, est devenu autonome, il trouve sa loi en soi, sans être déterminé par des facteurs externes, comme la biologie, les inclinations et émotions immédiates, etc. Bref, il y a autorégulation. On peut donc conclure que la hiérarchie des chaînes implicatives reflète l'autorégulation croissante du système: le domaine moral se constitue progressivement en un système fermé d'implications entre les valorisations. C'est ainsi qu'il faut comprendre, dans le domaine moral, le mécanisme de l'équilibration. Quand le système de valorisation est soumis à l'action causale et infra-normative, il y aura déséquilibre, ce qui se traduira par une tendance à l'autorégulation progressive, la morale se fermant sur elle-même.

Or, dans *Le jugement moral chez l'enfant*, la notion de l'équilibre est elle-aussi intimement liée à celle de l'autorégulation. C'est ce qui ressort de la citation suivante:

'... , un acte intelligent ne saurait être qualifié de logique, et un trait de sensibilité de moral, qu'à partir du moment où certaines normes impriment à de telles matières une structure donnée et des règles

¹¹ Piaget, *Introduction*, III, p. 232.

d'équilibre. La logique n'est pas coextensive à l'intelligence, mais consiste en *l'ensemble des règles de contrôle* dont use l'intelligence pour se diriger elle-même. La morale joue un rôle analogue à l'égard de la vie affective. ..., (C)e sont les personnes *extérieures* qui canalisent les sentiments élémentaires de l'enfant, et non ceux-ci qui tendent eux-mêmes à *se régulariser de l'intérieur*.¹²

On retrouve les trois modes de régulation dans l'exposé du *Jugement moral chez l'enfant*. En effet, l'égoïsme infantin fait figure d'alogisme moral; tantôt l'affectivité prédomine sur l'objectivité, tantôt l'activité propre sur celle des autres. Les sentiments qui surgissent dans son esprit lui apparaissent comme ayant de la valeur, sans qu'il faille les soumettre à une évaluation ultérieure. Ce mode correspond à la causalité bio-fysiologique. Puis, les rapports de respect unilatéral et de contrainte contribuent à la constitution d'un premier type de contrôle logique et moral: 'la pensée cesse d'affirmer simplement ce qui lui plaît pour se conformer à l'opinion de l'entourage'.¹³ L'enfant conçoit qu'il y a des conduites bonnes et mauvaises, comme il y a des affirmations vraies et fausses, mais la raison n'est pas encore en mesure de contrôler l'accord ou le désaccord de ces jugements avec 'la réalité'. Ce mode correspond à la régulation non-normée. Finalement, grâce à la coopération, qui exige discussion entre égaux et dont naît l'attitude critique, l'obligation est elle-même soumise à une norme, à savoir celle de la réciprocité entre égaux.¹⁴ Piaget ne dit pas grand chose sur cette norme idéale régissant la conscience autonome. A part des caractéristiques de cette norme déductibles de la manière dont il traite de l'équité (dans la partie consacrée à la justice), il n'en dit que le suivant:

'La morale de la conscience autonome ne tend pas à soumettre les personnalités à des règles communes en leur contenu même: elle se borne à obliger les individus à 'se situer' les uns par rapport aux autres, sans que les lois de perspective résultant de cette réciprocité

¹² Piaget, *Le jugement moral*, p. 323. C'est moi qui souligne.

¹³ *Ibid.*, p. 325.

¹⁴ *Ibid.*, pp. 323-28.

suppriment les points de vue particuliers'.¹⁵

Il me semble donc que pour Piaget le développement moral est lié à la régulation de deuxième ordre de la vie affective. Il s'agit d'une sorte de cybernétique de la vie affective. La conscience morale (au sens de *Gewissen*) n'est autre chose qu'une prise de conscience (au sens de *Bewusstwerdung*) de la pratique morale, corrigée par la norme de la réciprocité, qui oblige les individus à se situer les uns par rapport aux autres (stades de réflexion).

4. *L'apport néo-kantien*

C'est là qu'intervient Kohlberg. Bien qu'il ne se rendit pas compte du fait que Piaget envisageait le développement moral comme une sorte d'affectivité de deuxième ordre, par détours et partant d'observations cliniques, il en est venu à expliciter les 'lois de perspective' dont parle Piaget.

Dans ses recherches, Kohlberg utilisa des dilemmes hypothétiques (comme le fameux dilemme de Heinz, dont la femme mourante d'un cancer pourrait être sauvé par un médicament particulier. Le pharmacien demande un prix exorbitant pour son invention, que Heinz ne réussit pas à payer. Le pharmacien refuse de le donner. Heinz se voit donc placé devant le dilemme de cambrioler l'officine du pharmacien ou de laisser mourir sa femme.)

Essayant de classer selon la dichotomie piagétienne les réponses à des dilemmes pareils d'une centaine de jeunes garçons âgés de 10 à 16 ans, il se rendit vite compte qu'elle ne suffisait pas à couvrir toutes les alternatives. Par exemple, la soi-disante autonomie morale des garçons de 10 ans ne lui parût pas reposer sur le respect intrinsèque de l'autre, mais sur une réciprocité calculée, sur un égocentrisme différé. Ainsi, élaborant le schème piagétien, Kohlberg formula une séquence de six 'types idéaux', qui lui permit de rendre compte de toutes des réponses

¹⁵ Ibid., p. 322.

contenues dans le matériau clinique.¹⁶

Toutefois, dans cette première version, la description des stades moraux est bien plus basée sur une certaine théorie morale que sur la recherche de structures opératoires *organisant* le point de vue moral. Plus particulièrement, à cette époque-là Kohlberg s'inspira principalement de la philosophie néo-kantienne de Richard Hare¹⁷, donc de l'analyse logique du langage moral quotidien. Comme Kohlberg l'admettait plus tard, il presupposa que le noyau du développement moral est à chercher dans la compréhension de la spécificité des obligations et droits moraux.¹⁸ Ainsi donc, les stades se distinguent entre eux par une compréhension plus adéquate des différences entre le bien technique, le bien prudentiel et le bien proprement moral.

Plus particulièrement, selon Hare, le jugement proprement moral se distingue par (1) sa prescriptivité (c'est une prescription d'agir de certaine façon), (2) son universalité (ce qui signifie que nous nous contredisons quand nous jugeons de façon différente en deux situations que nous reconnaissons pour être identiques quant à leurs propriétés descriptives universelles.¹⁹

Bref, le jugement moral est une prescription à prétention universelle. Cette prétention à l'universalité implique entre autres choses que nous sommes tenus à l'application universelle de la règle, même quand, dans la situation problématique, nous changeons de places avec d'autres sujets.²⁰ Finalement, le jugement moral se caractérise par sa primauté, ce qui signifie que quand il est en collision avec d'autres prescriptions,

¹⁶ Cfr. L. Kohlberg, 'Moral development', *International Encyclopedia of the Social Sciences*. Crowell, Collier & McMillan, 1968, p. 489. C'est une version 'primitive' de la séquence décrite plus haut.

¹⁷ Plus particulièrement R.M. Hare, *The Language of Morals*. Oxford, University Press, 1952; R.M. Hare, *Freedom and Reason*. Oxford, University Press, 1963. Pour les explications de Kohlberg à ce sujet, voir L. Kohlberg, Ch. Levine, A. Hewer, *Moral Stages: A Current Formulation and a Response to Critics*. Basel, Karger, 1983, p. 67.

¹⁸ Kohlberg e.a., *Moral Stages*, p. 18.

¹⁹ R.M. Hare, *Moral Thinking. Its Levels, Method and Point*. Oxford, University Press, 1981, p. 21.

²⁰ Hare, *op. cit.*, p. 89.

il a toujours la priorité.²¹

Ainsi, pour Kohlberg, les stades moraux se définissent par une différenciation croissante entre d'une part les jugements moraux universels et prescriptifs et d'autre part les jugements esthétiques, prudentiels, etc. En somme, au fur et à mesure que le sujet approche le stade ultime, il apprend à juger d'un point de vue spécifiquement moral, qui est en même temps un point de vue impartial. On le voit, subrepticement Kohlberg a remplacé la question de la 'logique interne' qui est à chercher au niveau de la causalité sociale qui pousse les structures à changer dans un sens donné, par la question de la logique du langage moral.

Néanmoins, l'approche cybernétique, bien qu' à cette époque-là ne jouant aucun rôle dans la distinction des stades, n'est pas absente des préoccupations de Kohlberg. Dès l'abord il partait de la question comment l'enfant arrive à coordonner sa propre perspective et celle des autres sur une situation et un problème moral donnés. En effet, il pensait que la dynamique de la progression est régie par une sophistication cognitive croissante des aptitudes du soi à s'objectiver à partir du point de vue de l'autre concret et fonctionnel (l'autre dans un rôle social donné). Jouer un rôle dans la famille, un groupe d'égaux, une institution particulière, présuppose implicitement que le sujet prenne le rôle des autres vis-à-vis de lui-même et vis-à-vis d'autres sujets impliqués dans la situation. Elaborant les travaux de J.M. Baldwin et G.H. Mead, il postulait que la capacité de s'envisager du point de vue de l'autre est médiée d'une part par des actes mentaux de réciprocité et de complémentarité et d'autre part par des actes et des attitudes d'identification, de solidarité et d'imitation. Intimement associés au développement du langage et du symbolisme, ces 'opérations' seraient à la base de toutes les institutions sociales. Ainsi, ces dernières représentent un enchevêtrement complexe d'attentes communes ou complémentaires.²² On peut considérer cette approche comme une première tentative embryonnaire de préciser les 'lois de perspective' piagésiennes, qui obligent les individus à se situer les uns par rapport aux autres.

Dans les versions plus tardives de la théorie, les stades eux-mêmes sont distingués part rapport à la 'perspective sociale' sous-jacente et non

²¹ Idem, p. 56.

²² L. Kohlberg, 'Moral Development', *Encyclopedia*, p. 471.

par rapport aux orientations morales déduites de la théorie de Hare. Les raisonnements d'un locuteur sont assignés à un certain stade par l'application d'une procédure constante. Prenons comme exemple le dilemme de Heinz:

Question : 'Dans cette situation la loi et la vie entrent en conflit. Comment pourrait-on résoudre le cas, prenant en compte les meilleurs arguments pour chacun des choix?'

Réponse: Eh bien, la loi sert l'humanité et aide à maintenir l'ordre, mais d'où vient-elle? Elle provient de la valeur fondamentale accordée à la vie ... la loi est faite pour protéger la vie. En ce cas-ci, la loi ne protège pas la vie, mais un droit subordonné, et devrait être violée'.

D'abord l'interprète cherche quel *choix* a été fait par le locuteur (la vie, dans la réponse du sujet). Ensuite il cherche à déterminer la *norme* soutenant la justification du locuteur (la vie, encore, puisque le sujet justifie sa réponse en première instance en se référant à la valeur fondamentale de la vie). Puis, il cherche l'*élément* (ou le principe) supportant en deuxième ordre le choix de la norme (en ce cas-ci: 'Protéger la dignité et l'autonomie humaine'; les conditions nécessaires et suffisantes pour la distinction des éléments sont clairement indiqués dans le manuel sous la forme de listes d'indicateurs critiques). Finalement la combinaison choix-norme-élément est relaté à la perspective socio-morale sous-jacente (dans l'exemple retenu: le jugement exprime une hiérarchie rationnelle de droits et de valeurs. L'idée que la finalité idéale de la loi est la protection des droits humains, p.e. le droit à la vie, permet de juger de la moralité de lois particulières. Cette idée présuppose que le sujet se place non au point de vue d'un groupe, ou d'un individu isolé, mais à celui d'un sujet prenant en compte les intérêts bien compris de tous les autres, et qui essaie de les coordonner dans une trame idéale qui peut ou ne peut pas exister dans les faits. L'idée que la protection de droits individuels est le but final idéal du fonctionnement social est à distinguer de l'idée voisine, au stade 4, de l'existence d'une hiérarchie morale reconnue par la société ou par la religion. Là, c'est la perspective du groupe actuel qui est prise. Ici, l'accent est mis sur la hiérarchie logique que des personnes rationnelles

sauraient accepter. Classification: stade 5).²³

En résumé: dans la pensée de Kohlberg, deux conceptions des stades moraux se confrontent: d'une part, une conception centrée sur la logique du langage moral (c'est-à-dire le jugement moral universel, objectif, et impartial); d'autre part une conception centrée sur les 'lois de perspective', c'est-à-dire sur la problématique de la coordination des points de vue des agents sociaux. La première conception est soutenue par des philosophies morales. Comme nous le verrons, Kohlberg n'a pas cessé d'intégrer ces philosophies, en plus particulièrement celle de John Rawls, dans sa psychologie génétique. La seconde conception est originaire de théories psychologiques de la construction des points de vues des agents sociaux. Chez Kohlberg, ces deux pistes ont partie liée. Mais je crois qu'il lui faudrait choisir, comme elles relèvent de deux façons incompatibles de concevoir des 'stades' moraux. C'est ce qui ressortira de l'analyse suivante.

5. La notion de 'stade'

Il convient d'abord d'analyser de plus près les présuppositions de la notion générale de stade telle qu'elle apparaît dans les sciences humaines. L'idée de stade comprend les propriétés suivantes: il faut qu'il y ait 1) une succession d'états qui 2) représentent dans leur succession une continuité en voie d'amplification, et 3) ou chaque état forme un tout relativement stable, caractérisé par l'émergence de capacités ou de régulations nouvelles, inexistantes au niveau antérieur. La possibilité même de la distinction de stades présuppose forcément au moins un critère évaluatif permettant de comparer deux états différents du 'système'.

Je ne crois pas que l'existence de 'stades' en ce sens général puisse être contestée. Ainsi, dans le système concentrationnaire nazi, le complexe d'Auschwitz continue et amplifie un état embryonnaire représenté par le KZ-Börgermoor. Il s'agit là d'une succession de deux états, qui représente une continuité en voie d'amplification. Auschwitz-Birkenau en particulier disposait de 'capacités nouvelles' inexistantes à Börgermoor.

²³ L'exemple est tiré des pages 53-54 du manuel: A. Colby, L. Kohlberg, *The Measurement of Moral Development. Vol. II: Standard Issue Scoring Manual*. Cambridge, University Press, 1987.

Entre les deux états il y a bien eu 'développement', mesuré au critères de la capacité d'internement, de la polyfonctionnalité des camps, du nombre de catégories d'internés, etc. La seule stabilité temporelle définie par rapport à un critère donné suffit à démarquer les 'touts' requis.

Venons-en à la psychologie. La notion de stade a été, et est toujours très populaires en psychologie génétique: elle sert à indiquer des capacités nouvelles acquises par un organisme/sujet quelconque, lui permettant de mieux s'adapter aux ressources et défis d'un environnement donné. Le critère évaluatif impliqué est ordinairement d'ordre technique: il s'agit d'un critère permettant d'évaluer la progression du sujet dans ses efforts d'adaptation.

Comme la notion de stade est principalement utilisée à propos de processus de maturation psychosociaux (la 'maturation' de la personnalité, de l'intelligence, du jugement moral, de la conduite morale, de la cognition sociale, etc.) il peut se dégager l'impression que l'idée de 'développement' implique le même jugement moral positif qui serait impliqué dans l'idée de 'maturation'. En fait, aucune des deux idées n'implique ce jugement. L'exemple du système concentrationnaire nazi nous montre que le critère évaluatif n'est pas forcément d'ordre moral. Il suffit que d'une façon ou d'une autre il y ait continuité aussi bien qu'amplification d'une certaine dimension. Le développement peut très bien représenter une progression dans le mal moral.

Parcourant la littérature psychologique, on trouve plusieurs types de stades qui répondent aux critères retenus. Voici une brève liste d'exemples, qui ne prétend nullement à l'exhaustivité.

1) Entre les stades il y a des rapports cognitifs et/ou pratiques de complexité croissante, soit:

A. des rapports d'aptitudes à leurs présuppositions: ainsi, Albert Bandura argumente que les régularités génétiques observées par Kohlberg et Piaget peuvent être expliquées par ceci que la plupart des activités cognitives comprennent des séquences hiérarchiques. Les capacités plus complexes résultent de l'intégration de plusieurs capacités plus simples. Un écrivain ne saurait guère écrire des essais sans maîtriser le vocabulaire, la construction de phrases, etc. Le développement cognitif peut dès lors être considéré du point de vue des séquences inhérentes aux struc-

tures des tâches que les sujets s'efforcent à maîtriser.²⁴

Il s'agit bien d'une conception permettant de parler de stades, le critère comparatif étant formulé à partir de la maîtrise d'une capacité finale. L'intégration de plusieurs capacités pré-existantes fournit les 'capacités nouvelles' requises pour la reconnaissance de stades.

B. des rapports du simple au réflexif: à partir de critiques adressés à Kohlberg, Gibbs a élaboré une 'théorie constructiviste' du développement moral. L'idée fondamentale en est que le développement ne résulte ni de l'internalisation directe des normes socio-morales, ni d'une logique constructiviste quasi-autonome. L'acquisition de l'idée de la conservation de la quantité d'un fluide, par exemple, peut résulter de la confrontation entre deux enfants, dont l'un croit que, de deux récipients, le plus grand contient plus de liquide que le plus petit et dont l'autre croit que l'inverse soit vrai. Du conflit entre les convictions contradictoires peut résulter une conscience nouvelle. Il s'agit bien là de la construction d'une conscience 'de novo', qui résulte indirectement de l'internalisation.²⁵

Les stades distingués par Gibbs résultent de la réflexion du sujet sur ses expériences morales et représentent le progrès dans la 'maturité réflexive'.²⁶ Jane Loevinger fournit un autre exemple de stades réflexifs. Selon elle, le développement de l'ego passe par une succession de prises de conscience de la psychologie du soi.²⁷ Le point final idéal de cette progression est la possession d'un soi intégré. Ces stades réflexifs, on le voit, répondent bien aux propriétés mentionnées.

²⁴ A. Bandura, *op. cit.*, p. 485; voir aussi D. H. Feldman, *Beyond Universals in Cognitive Development*. Norwood, NJ, Ablex, 1980. Donald Campbell croit que, dans ce sens-ci, des stades pourraient être distingués non seulement dans la psychologie cognitive individuelle, mais aussi dans les sciences. Cfr. W. Callebaut (ed.), *Taking the Naturalistic Turn, or, How Real Philosophy of Science is Done*. Chicago, University Press, 1993, pp. 57-58.

²⁵ J.C. Gibbs, K.S. Basinger, D. Fuller, *Moral Maturity: Measuring the Development of Sociomoral Reflection*. Hillsdale, NJ, Lawrence Erlbaum, 1992, p. 14.

²⁶ J.C. Gibbs, K.F. Widaman, *Social Intelligence. Measuring the Development of Sociomoral Reflection*. Englewood Cliffs, NJ, Prentice-Hall, 1982, p. 23, p. 25.

²⁷ J. Loevinger, *Ego development. Conceptions and Theories*. San Francisco, Jossey-Bass, 1976.

2) Un rapport fonctionnel: Les stades du développement de l'identité de Erik Erikson fournissent un autre exemple. Reprenant l'idée des stades épigénétiques en biologie, il distinguait des stades psychosociaux. A travers des crises d'identité le sujet acquerrait un sens du soi, qui lui permettrait de mieux s'adapter aux arrangements institutionnels et aux rôles sociaux qu'ils confronte au cours de sa trajectoire biographique.²⁸ Un autre exemple sont les stades psychosexuels de Freud.

Ces stades-ci sont déterminés par la confrontation à l'environnement: une certaine idée de ce qui constitue le fonctionnement optimal permet de déterminer l'objectif final de la progression.²⁹

3) Un rapport logique, dont la description est possible sans référence aucune aux contenus des structures. Un exemple en sont les stades théorético-cognitifs de Piaget, dont le point final consiste en l'équilibration optimale d'une structure d'opérations. Ces structures peuvent être décrites par des modèles mathématiques et logiques. Ces stades sont liés entre-eux par le mécanisme de l'équilibrage et donc par une logique interne. Les stades du développement moral de Kohlberg prétendent appartenir à cette dernière catégorie. Il les appelle, pour les distinguer des autres (qu'il confond en une seule catégorie), des stades 'durs', les opposant aux stades 'mous'.^{30 31}

Les structures du développement moral chez Piaget ressemblent plus aux stades réflexifs de Gibbs qu'aux stades cognitifs du Piaget tardif. La morale de l'autonomie résulte de l'intériorisation indirecte, puisqu'elle est induite par la coopération et en constitue une prise de conscience corrigée par la norme de la réciprocité.

En rangeant sa théorie du développement moral dans la catégorie des stades logiques, Kohlberg n'a pas choisi la voie la plus évidente. Le critère de 'totalité' requis est le plus fort de tous, en ce qu'il ne présuppose pas la simple stabilité (comme le type 1), ni une intégration dispa-

²⁸ E. Erikson, *Childhood and Society*. New York, Norton, 1963.

²⁹ Voir également au sujet de ces stades la conversation fictive entre Kohlberg, Piaget et Erikson dans J. Fowler, *Stages of Faith*. San Francisco, Harper, 1981, pp. 41-51.

³⁰ Kohlberg, *The Psychology of Moral Development*, p. 235 sq.

³¹ Les auteurs cités ne se limitent pas forcément au type de stades sous lequel je les ai mentionnés. Il s'agit ici d'explicitier quelques possibilités théoriques.

rate (comme le type 2), mais une globalité fortement homogène, c'est-à-dire une structure formelle d'opérations morales qui dirige les façons dont le sujet cherche à résoudre des problèmes moraux. Tandis que les deux premiers types permettent qu'il y ait perte de capacités (les stades d'Erikson, par exemple, se remplacent ou s'enchevêtrent, plutôt que de s'intégrer sans perte), les stades 'durs' exigent la différenciation et intégration croissante de capacités orientées vers une seule fonction constante (dans le cas présent, le jugement moral).

En fait, on constate que chez Kohlberg il y a une confusion entre deux conceptions des stades. Il reprend l'idée fondamentale de Piaget que les lois de perspective distinguent les différents stades entre-eux. Dans cette approche, les stades sont démarqués les uns des autres par la complexité croissante de la coordination des différents rôles sociaux impliqués.

Or, comme je l'ai indiqué, il s'agit là d'une conception des stades qui est moins exigeante — du point de vue théorique aussi bien qu'empirique — que celle des stades 'durs'. S'il est probable que des stades de type 1 et 2 'existent' (encore faudrait-il préciser de quelle manière, ce qui est loin d'être facile), le problème fondamental se pose s'il en va de même pour les stades de type 3. Assumant que Piaget a réussi à découvrir de tels stades dans la pensée orientée sur le monde physique, on est bien obligé de constater que ses 'structures' ont un sens mathématique et logique précis: il s'agit d'approximations successives de structures de groupes et de 'groupements'. Le groupe INRC, par exemple, forme une structure algébrique aisément reconnaissable comme telle. Mais quelle pourrait être la structure correspondante régissant la filiation des stades moraux? C'est là que, chez Kohlberg, intervient la philosophie morale analytique. Il prétend que cette 'logique' est fournie par la philosophie néo-kantienne, qui lui permettrait de considérer la filiation des structures comme régie par une logique des 'opérations morales'. Regardons de plus près comment Kohlberg assimile continûment la piste logique avec la piste psychologique.

6. La question de la supériorité des stades ultérieurs

A la question en quoi consiste la 'supériorité' des stades plus avancés sur les stades moins avancés, Kohlberg avance trois réponses: leur supériorité

est de nature 1) cognitive; 2) ethico-formelle; 3) ethico-substantielle. Examinons d'abord les deux premiers points.

La supériorité cognitive est indiquée, selon Kohlberg, par le caractère 'dur' des stades: ils représentent des façons plus adéquates d'adaptation aux réalités de l'interaction sociale. Selon lui, cette affirmation est équivalente à l'affirmation qu'il représentent une progression de complexité croissante des 'prises de rôle'. Or, nous avons vu que la présupposition de l'existence de stades cognitifs 'durs' est logiquement plus forte que la présupposition de l'existence de stades 'réflexifs'.

La seconde sorte de supériorité consiste en ceci qu'à mesure que le sujet progresse, le jugement moral approche de plus en plus de ce qui constitue la spécificité du jugement moral *proprement dit*: c'est-à-dire la primauté, l'universalité et la prescriptivité.

Or, il me semble que ce ne soit pas vraiment le cas que la 'logique de la prise de rôle social' conduise automatiquement à l'épuration du jugement moral dans le sens indiqué par Kohlberg. Ce sens serait celui également indiqué par les théories déontologiques ('les morales de l'obligation', comme celles de Brandt, Rawls, Frankena, Harsanyi et Hare³²). Comme je l'ai montré, Kohlberg n'a pas clairement démontré que cette route déontologique seule prolonge celle de la coordination des perspectives. Il présuppose d'emblée que ces deux processus sont identiques. Il y a lieu d'en douter. Même les versions conceptuellement les plus primitives de l'utilitarisme, qui sont considérées par Kohlberg, du point de vue des stades 'durs', comme des formes inférieures de théorie morale, impliquent le respect du principe de l'égalité de considération des intérêts des personnes, groupes ou institutions concernés³³ et donc une prise de per-

³² R.B. Brandt, *Value and Obligation: Systematic Readings in Ethics*. New York, Harcourt, 1962; W.K. Frankena, *Ethics*. Englewood, N.J., Prentice Hall, 1973; J.C. Harsanyi, 'Morality and the theory of rational behaviour', in: A. Sen, B. Williams, *Utilitarianism and Beyond*. Cambridge, University Press, 1982, pp. 39-62; J. Rawls, *A Theory of Justice*. Oxford, University Press, 1971. Il s'agit-là d'un ensemble de théories fort diverses, allant du reconstructionisme politique de Rawls à l'utilitarisme neo-kantien de Hare, passant par les théories de la justice dérivant de théories du choix rationnel. Selon Kohlberg, elles offrent toutes une description partielle du stade 6.

³³ Il suffit de prendre un ouvrage quelconque écrit par un utilitariste contemporain pour comprendre que c'est bien le cas. Voir p.ex. P. Singer, *Practical Ethics*. Cambridge, University Press, 1979.

spective 'hors-société'. La logique des lois de perspective soustend aussi bien des théories du type conséquentialiste que celles du type déontologique. Il faut donc bien distinguer le processus de clôture cybernétique des prises de perspective *en général*, d'une *certaine façon* de concevoir cette clôture, c'est-à-dire d'une façon uniquement déontologique.

Néanmoins, trompé par cette présupposition fallacieuse, Kohlberg a délibérément restreint sa théorie du développement moral à la pensée de justice seule. Non seulement a-t-il relégué la plupart des morales utilitaristes au second rang, encore a-t-il explicitement dévalué les éthiques téléologiques en leur assignant une place parmi les 'stades mous', c'est-à-dire les stades fonctionnels.³⁴ Or, il n'y a pas lieu de penser que les théories déontologiques reflètent des stades plus 'durs' que les autres.

7. *L'herméneutique objective de Habermas*

Kohlberg avance aussi que les stades ultérieurs transcendent les stades antérieurs parce que les jugements moraux que fondent les premiers sont plus justifiables d'un point de vue normatif que ceux fondés par les stades inférieurs. C'est là que le psychologue se convertit en philosophe non seulement méta-éthique (qui analyse l'usage commun des mots, concepts et notions moraux) mais en éthicien proprement dit. A vrai dire, la position que Kohlberg a prise en ce respect a changé fondamentalement au cours des années. Dans son article provocateur 'From is to ought'³⁵ il avait d'abord avancé, de façon d'ailleurs assez ambiguë, que la théorie scientifique de la manière dont les personnes en viennent à parcourir *en réalité* la succession des stades et de préférer les raisonnements caractéristiques des stades supérieurs, est équivalente à une théorie morale qui

³⁴ Cfr. les stades de la foi, Fowler, *op. cit.*; les théories psychologiques du développement des éthiques de la sollicitude et de la responsabilité, p.ex. C. Gilligan, *In a Different Voice. Psychological Theory and Women's Development*. Cambridge, Mass., Harvard University Press, 1982 et N. Noddings, *Caring. A Feminine Approach to Ethics and Moral Education*. Berkeley, University of California Press, 1984. Cette dernière range l'éthique de Martin Buber dans la catégorie des éthiques de la sollicitude. On peut donc supposer que la psychologie impliquée par l'éthique de Lévinas appartiendrait elle-aussi, selon Kohlberg, à la logique des stades 'mous'.

³⁵ Cité plus haut.

explique pourquoi ces personnes *doivent* (moralement parlé) préférer les raisonnements des stades supérieurs.

Selon Jürgen Habermas, Kohlberg, dans l'article mentionné, propose encore une autre conception de la relation entre philosophie et psychologie morale, sans la distinguer clairement de la première. Dans la conception alternative, il n'est pas question de l'identité entre les approches philosophique et psychologique, mais de leur complémentarité. Dans cette conception la psychologie peut vérifier si, phénoménalement parlé, les conceptions de la philosophie morale coïncident avec les données psychologiques. Mais la psychologie ne peut faire davantage: elle ne peut *justifier* la conception morale particulière qui se retrouve dans les données psychologiques. L'hypothèse qu'un stade ultérieur est supérieur à un stade antérieur est implicitement basée sur une théorie morale philosophique qui justifie les prétentions à la validité de jugements moraux en ses propres termes normatifs.³⁶ Kohlberg a ultérieurement reconnu que Habermas avait raison et déclaré que le progrès moral n'est pas seulement de nature fonctionnelle, comme il en était chez Piaget, mais encore de nature spécifiquement morale. Cette démonstration étant une question relevant de l'ordre de la *validité* normative, elle n'appartiendrait pas au domaine de la psychologie, qui s'occupe uniquement de prétentions à la *vérité*. Pour bien comprendre ce point, il convient de l'étudier de plus près l'argumentation de Habermas (Kohlberg l'a également faite sienne).

On sait combien Habermas était prédisposé à 'découvrir' cette ambiguïté dans le programme kohlbergien: dans la théorie des actes de langage de Habermas, les prétentions à la vérité, à la validité et à l'authenticité sont rigoureusement différenciées à partir des domaines de la réalité auxquels les actes de langage ont rapport (resp. le monde externe; le monde social; le monde interne), des modes de communications qu'ils visent (les modes cognitif, interactif et expressif) et des fonctions qu'ils remplissent (représentation de faits, instauration de relations interpersonnelles légitimes, révélation de la subjectivité du sujet parlant).³⁷

Dans ce schème tripartite, la psychologie a pour objectif la représen-

³⁶ J. Habermas, *op. cit.*, p. 48.

³⁷ J. Habermas, *op. cit.*, p. 33; voir également, pour un tableau plus sophistiqué, J. Habermas, *Theorie des kommunikativen Handelns*. Bd. 1. Frankfurt/M., Suhrkamp, 1981, p. 45.

tation du monde des relations interpersonnelles légitimes, et cela à partir de l'explicitation des visées communicatives du sujet parlant. Pour Habermas, le chercheur se trouve par conséquent dans la position de l'interprète qui doit s'efforcer de comprendre les énoncés communicatives du locuteur (le sujet interrogé). Son attitude ne consiste pas tellement en ce qu'il essaie simplement de constater ce qui est le cas (un rapport objectif à son objet de recherche) qu'en ce qu'il s'efforce, en tant que participant, de comprendre le sens des paroles de son interlocuteur (rapport *performatif*). Or, continue-t-il, le rapport performatif se caractérise par ceci que le locuteur est pris dans l'attente du consentement ou du rejet de la part du destinataire. Ainsi, il y a une orientation *réciproque* sur les prétentions à la validité des énoncés (validité descriptive, normative ou expressive).³⁸ Le chercheur perd son privilège d'observateur pour entrer dans le rôle de participant. En fait, l'interprète se retrouve, tout comme l'interlocuteur, à tour de rôle dans la position de la première, de la seconde et de la troisième personne.

Le problème épistémologique central de la psychologie morale réside dans la traduction des compréhensions du 'je', produites par un nombre de changements de rôle, en langage de troisième personne. Il s'ensuit que n'importe quelle science sociale est posée devant le problème particulier que le sens n'est pas simplement 'présent' dans le mode 'cognitif', mais que l'interprète cherche en vérité à 'objectiver' des significations 'données' par les interlocuteurs conjoints, lui-même compris, dans des processus de communication. Tout ceci semble condamner les sciences sociales, soit à une herméneutique radicale (impossibilité tout court de l'objectivation scientifique, la position de R. Rorty), soit à une herméneutique reconstructiviste, qui cherche à dissocier, par le biais d'une théorie empathique de l'entendement, d'une part le sens donné par le sujet parlant à ses paroles et d'autre part la condition herméneutique de l'interprète lui-même (position de Dilthey, selon Habermas). Habermas signale une troisième possibilité, celle d'une herméneutique 'objective'.³⁹

Afin de clarifier cette position, Habermas propose une métaphore assez proche de la situation du chercheur essayant de comprendre 'objectivement' des sujets placés dans un contexte clinique d'interview: celle

³⁸ J. Habermas, *Moralbewusstsein und kommunikatives Handeln*, p. 34.

³⁹ J. Habermas, *op. cit.*, pp. 36-38.

de l'interprète qui cherche à reconstruire la signification d'un texte. A partir de l'expérience faite par l'interprète que les énoncés de l'auteur sont souvent difficiles à comprendre (bien qu'il croyait auparavant avoir compris le texte), celui-là est amené à conjecturer que le contexte dans lequel il a placé le texte n'est pas adéquat. Il lui faut alors réviser sa compréhension première. Ce n'est que dans la mesure où l'interprète arrive à découvrir les *raisons* qui font apparaître les énoncés de l'auteur comme *rationnels*, qu'il arrive à comprendre les intentions premières de l'auteur. Ainsi donc, l'interprète comprend le texte vraiment à partir du moment où il en arrive à avancer lui-même les raisons que l'auteur aurait pu avancer dans la situation originaire, mais qu'il a laissé dans l'ombre de l'évidence.

A partir de ce point-là, Habermas cherche à nous convaincre que comprendre un texte comme éclaircir le sens d'un énoncé obscur, présuppose, de la part de l'interprète, une norme rationnelle, permettant de juger des raisons de l'auteur. D'une part, ces raisons ne peuvent être comprises comme telles qu'à condition que l'interprète cherche, au moins implicitement, à les évaluer, la compréhension impliquant l'évaluation. D'autre part, l'élucidation des énoncés obscurs exige une explication, de la part de l'interprète, pourquoi les raisons données ou présupposées par l'auteur ne sauraient plus convaincre. Bref, une interprétation est forcément une interprétation rationnelle, ce qui signifie, entre autre, que l'interprète, en évaluant les raisons de l'auteur, doit présupposer la validité de normes engageant aussi bien lui-même et son public actuel que l'auteur et ses contemporains.⁴⁰

Il est évident que la possibilité même de cette 'herméneutique objective' est étroitement liée au programme habermasien de la justification d'une éthique discursive. Selon lui, la coordination des attentes des participants dans un échange social peut être réalisée de deux façons diamétralement opposées: soit elle est réalisée de façon stratégique (par le biais des contraintes du marché et de l'état), soit de façon communicative (par le biais de la motivation rationnelle du consentement des participants). A ces deux types correspondent en gros deux types de rationalité: la première orientée sur l'efficacité, la seconde sur l'entendement ou le consensus. Il est vrai que, historiquement parlé, le premier type de

⁴⁰ Idem, p. 40.

rationalité a tendance à se superposer au second. N'empêche que l'action communicative obéit à une logique propre qui peut être réperée en examinant les présupposés implicites des communications orientées sur l'entendement intersubjectif (bien qu'il soit vrai que, dans le domaine de la reproduction symbolique du monde social, la rationalité spécifiquement stratégique est présente tout comme la rationalité communicative). Or, la place centrale de l'action communicative dans le 'Lebenswelt' (monde vécu) paraît en ce que, graduellement, elle réussit à s'institutionnaliser dans ce Lebenswelt et à s'imposer sous forme de compétences communicatives des agents sociaux. Ainsi est fondée une sorte de spirale de la rationalité proprement communicative: les vestiges de la rationalité stratégique sont progressivement 'problématisés', à partir d'un arrière-fond d'évidences qui forment le 'contexte' du 'texte' problématique (pour nous en tenir au langage herméneutique).

Il est clair que pour Habermas, un aspect de cette rationalisation progressive des compétences communicatives est à chercher dans les compétences morales des agents sociaux. Ceci explique son intérêt pour Kohlberg: pour le premier il est évident que le développement moral reflète les tendances profondes de la rationalité communicative vers l'idéal d'une 'communication exempte de pouvoir' (herrschaftsfreie Kommunikation).⁴¹ Le stade ultime du développement moral ne saurait être autre qu'une éthique universelle du discours (universelle Sprachethik).⁴² Cette éthique explicite 'ce que parler moralement veut dire moralement parlé': elle systematise une pragmatique qui sous-tend l'activité de ceux qui parlent morale et qu'ils seraient obligés à reconnaître et à accepter en tant que système de procédures régissant les conditions idéales du

⁴¹ J. Habermas, 'Moral Development and Ego Identity', in : J. Habermas, *Communication and the Evolution of Society*. Boston, Beacon Press, 1979, pp. 69-94.

⁴² Dans l'article 'Moral Development and Ego Identity', pp. 89-90, Habermas introduit cette éthique comme un septième stade kohlbergien. Après une réaction de Kohlberg et ses collaborateurs (*Moral Stages*, p. 164), il a reconnu l'identité profonde entre ce septième stade et les stades cinq et six de Kohlberg lui-même. En réalité, Habermas voit ses stades, à la façon de Gibbs, comme des stades de réflexion, de justification des principes qui orientent le jugement moral, (voir *Moralbewusstsein*, pp. 184-87). Pour Habermas il ne s'agit donc pas de stades 'durs'. Ceci est confirmé par Kohlberg (*Moral stages*, p. 162). La problématique de Habermas, écrit-il, n'est pas de trouver des analogies entre les rapports interpersonnels et les opérations logiques de Piaget.

consensus, si seulement les agents en étaient conscients et se limitaient à la logique propre de leurs activités (il s'agit de procédures comme les suivantes: 'ne vous contredites pas vous-même', 'Appliquez un prédicat F que vous appliquez dans une situation x, sur les situations y, z, ... si elles se ressemblent dans tous les aspects relevants', etc.).

Kohlberg a reconnu ces propres thèses dans ces idées de Habermas. Il estime que le chercheur qui conduit l'interview clinique et doit classifier les réponses, se trouve dans la position de l'interprète devant un texte. Il est obligé à collectionner des données en posant des questions précises, stimulant des réponses, qui lui permettent de pénétrer la vision du sujet interrogé jusqu'au point où le chercheur peut générer lui-même celles que son interlocuteur aurait donné, si demandé. Bref, il s'agit de décrire le sens trouvé par le sujet 'dans le monde' à partir du point de vue où se place ce sujet lui-même. L'interprétation du texte (la conduite de l'interview et la classification des réponses) ne se déroule pas aveuglément: elle se déroule à partir de catégories philosophiques du sens communes au sujet interrogé et au chercheur. La possibilité même de cette entente découle du fait que le chercheur a lui-même parcouru les stades et sait en reproduire les visions et divisions pertinentes.

Jusqu'ici cette approche ne va pas au-delà de l'herméneutique dite reconstructiviste. Ce qui la range, selon Kohlberg, dans la famille 'objective', c'est le manuel de score des réponses. Selon lui, ce manuel ne cherche pas simplement à classifier les réponses des sujets et les sujets eux-mêmes, mais présuppose une position envers le plus ou le moins de rationalité des énoncés interprétés. Les attributions de stades ne sont pas 'neutres', mais impliquent une évaluation morale. Ainsi, Kohlberg estime que sa théorie offre une 'reconstruction rationnelle' du développement moral, et ceci parce que 1) la description présuppose une logique interne dans le développement moral, à l'aide de 2) la norme idéale du stade six, qui philosophiquement représente la forme la plus adéquate de la pensée de justice.⁴³

⁴³ Kohlberg e.a., *Moral Stages*, p. 14.

8. *Que penser de cette herméneutique?*

Examinons d'abord l'hypothèse qu'une approche herméneutique, bien comprise, devrait être du type objectif. Concédon's que la compréhension d'un texte présuppose une évaluation des raisons données par le locuteur. S'ensuit-il pour autant que cette évaluation repose nécessairement sur une éthique discursive implicite dans la communication morale ou sur une théorie morale qui systématise implicitement la globalité des sens moraux actuellement présents? Est-ce qu'il faut vraiment partir de raisons qui sauraient convaincre un auditoire universel de la validité des prétentions normatives, sous peine de tomber dans une herméneutique radicale (qui, je le rappelle, conteste la possibilité même de la validité théorique et objective, cfr. Rorty; l'autre position, l'herméneutique reconstructiviste est rejeté sans plus par Habermas: lire Gadamer suffirait à entrevoir l'absurdité profonde de l'entreprise). Est-ce qu'on a vraiment besoin de cette lumière universelle chez l'interprète pour tirer de l'ombre les évidences non problématisés par le locuteur critique ou par le sujet moral interrogé?

Chez Habermas, dont les vues ont été profondément influencées par K.-O. Apel⁴⁴, toute la conception de l'herméneutique rationnelle s'orientant sur une éthique discursive universelle et implicite tombe ou tient avec la distinction première entre l'action stratégique et l'action communicative. L'action stratégique dans le domaine symbolique y est d'emblée mise sous le signe de la *médiation* des formes proprement socio-culturelles du 'Lebenswelt'; elle y représente l'intrusion des formes de la rationalité des *systèmes*, colonisant la rationalité proprement (c'est-à-dire idéalement) communicative du 'Lebenswelt'.

Il me paraît très probable que cette dichotomie entre l'action communicative et l'action stratégique est trop rigide pour rendre compte de la psychologie phénoménale de la construction d'un monde commun à deux ou plusieurs locuteurs. Ainsi, le psychologue M. Miller a fait des recherches empiriques sur le rôle de l'argumentation dans la construction par des coopérants d'un monde commun. Il a conclu que des sujets cherchant coopérativement à définir et à résoudre des problèmes définis

⁴⁴ Pour un exposé récent, voir K.-O. Apel, *Diskurs und Verantwortung. Das Problem der Übergang zur postkonventionellen Moral*. Frankfurt/M., Suhrkamp, 1990.

en arrivent à construire des procédures pour la résolution de conflits entre les perspectives des participants, mais sans que disparaissent pour autant les dissensions. En fait, il y a même co-construction entre d'une part des opinions différenciant clairement les participants et d'autre part des procédures visant à réduire les désaccords.⁴⁵ Même la constatation de désaccords entre les participants aboutit à une compréhension meilleure du monde commun: bien que localement, les participants sont contraints de prendre de la distance vis-à-vis de leurs propres opinions et à les réviser, il reste tout aussi vrai que la constatation de désaccords insurmontables contribue à la construction d'un monde commun. Bien qu'il n'y ait pas consensus, ils font l'apprentissage de l'existence de divergences dans un monde commun.

Ces résultats contredisent bien la conception de Habermas. Appliqué à la situation interprète/texte, le processus indiqué par Miller signifie que, même si la compréhension du texte implique un jugement sur les raisons de l'auteur, elle ne présuppose pas l'horizon lointain d'une communauté transcendentale régnée par le consensus. Il peut aussi y avoir co-construction d'un monde commun entre le locuteur et l'interprète, dans lequel la différence des perspectives est maintenue: l'interprète cherche à clarifier la position du locuteur en établissant les différences entre sa propre perspective et celle du locuteur, tout en s'efforçant de comprendre quelles raisons le locuteur aurait pu avancer, la différence de perspective étant explicitée. Tout cela se passe sans qu'il y ait une 'norme rationnelle' qui embrasse les deux types de raisons et qui permet de décider de leur validité respective. La logique sous-jacente n'a pas deux valeurs de vérité (valide et non-valide), mais trois (valide, non-valide, indécis). Par exemple, l'interprète pourrait chercher à vérifier si les différences de perspective coïncident en gros avec celles entre une morale déontologique et une morale perfectionniste (de type aristotélien). Cela suffit parfaitement à la compréhension du locuteur à condition que l'interprète a objectivé les différences de type entre ces deux morales. S'il y a reconstruction, elle ne présuppose point la présence, chez l'interprète, d'un idéal de rationalité engageant les deux partis.

La question se présente un peu autrement chez Kohlberg: ceci peut

⁴⁵ Voir R. Rogoff, *Apprenticeship in Thinking. Cognitive Development in Social Context*. New York, 1990, pp. 178-80.

étonner, mais le manuel de score ne donne aucune indication comment reconnaître la présence de raisonnements du stade 6: aucun des sujets de Kohlberg, ni dans le Etats Unis, ni en Turquie ou Israel, a fait preuve de raisonner au stade ultime. En fait, Kohlberg admet que le stade ultime est une construction purement théorique, qui ne sert aucunement à ordonner les données empiriques:

'Nous continuons à postuler (l'existence du stade 6) parce que nous concevons notre théorie comme une tentative de reconstruire rationnellement l'ontogenèse de la pensée de justice, une entreprise qui requiert un stade ultime afin de définir la nature et le point final du type de développement que nous étudions'.⁴⁶

Cette postulation, continue-t-il, est indispensable à la *définition* de l'objet de recherche. Or, cet objet de recherche ce sont les 'stades durs' dans le raisonnement sur la justice, reposant sur des opérations de réciprocité, d'égalité et d'équité. Ces opérations sont des 'analogues', dans le domaine de l'interaction interpersonnelle, des opérations logiques de Piaget dans le domaine cognitif.⁴⁷

Plus spécifiquement, la théorie de la justice de John Rawls est à la psychologie morale de Kohlberg, ce que le groupe INRC est à la théorie du développement théorético-cognitif de Piaget. La théorie de John Rawls offre une description partielle du stade ultime. Non que la théorie de Rawls est la théorie ultime, que tous les philosophes devraient reconnaître comme la théorie morale la plus adéquate. Celle-ci 'formalise' plutôt le point optimal d'équilibration des 'opérations de justice', en ce qu'elle permet de comprendre la forme que prend la 'réversibilité' des opérations de justice dans le raisonnement moral.⁴⁸ En effet, la conception du voile d'ignorance de Rawls⁴⁹ permet de définir en quel sens un jugement mo-

⁴⁶ Kohlberg e.a., *Moral Stages*, p. 61.

⁴⁷ *Ibid.*, p. 162.

⁴⁸ *Ibid.*, p. 61.

⁴⁹ Rawls se demande quels principes de justice seraient choisis dans une négociation par des agents qui seraient ignorants de leur position dans la société ou même de leur place dans la distribution des talents et des capacités. Afin de trouver des principes de justice valides, il s'imagine donc que les agents se trouvent sous un voile d'ignorance.

ral adéquat doit être 'réversible': on doit être prêt à changer de place avec les autres. Le jugement moral qui obéit aux conditions stipulées sous le voile d'ignorance et aux principes de justice qui en découlent, est optimalement équilibré.⁵⁰ Ainsi, le principe rawlsien de la différence, qui stipule que des inégalités dans la distribution des biens fondamentaux sont justifiées si et seulement si elles sont à l'avantage des plus désavantagés, représente l'équilibre optimal entre les opérations d'égalité et d'équité. Les conceptions dont partent les philosophies morales alternatives, comme l'utilitarisme, ne permettent pas de définir le point d'équilibre optimal, ce qui n'implique pas que ces philosophies sont inférieures du point de vue philosophique.⁵¹

En fin de compte donc, c'est l'idée qu'il y ait des stades 'durs' dans le développement moral, qui pousse Kohlberg à avancer son herméneutique objectiviste. Le raisonnement est simplement circulaire (et il ne s'agit pas du cercle herméneutique, mais d'un raisonnement circulaire tout court): l'idée qu'il y ait herméneutique plutôt que simple classification découle de la présupposition qu'il y a des stades 'durs'; et inversement, l'idée qu'il y ait des stades durs est fondée sur la présupposition que l'herméneutique a bien besoin d'un point d'ancrage final (le stade six).

9. Conclusions

Cette observation ne réduit pas l'oeuvre de Kohlberg en miettes. En fait, examinant le manuel de plus près, on voit bien que la soi-disante approche herméneutique ne joue aucun rôle dans la construction des stades à partir des réponses des locuteurs. Kohlberg a construit le manuel mentionné, qui formule des règles d'interprétation strictes, afin d'éviter que l'interprète ne projette librement ses propres points de vue sur les réponses des répondants. En explicitant ces règles, Kohlberg a espéré pouvoir

⁵⁰ L. Kohlberg, 'Justice as Reversibility', in: *Essays on Moral Development*, vol. I, pp. 190-226, plus particulièrement pp. 194-201.

⁵¹ En fait, Kohlberg n'est pas tout-à-fait clair sur ce point: puisqu'il prétend aussi que le stade 6 est moralement supérieur au stade 5, et puisqu'il argumente que la position utilitariste représente la perspective morale du stade 5, il s'ensuivrait que les morales déontologiques universalistes sont moralement supérieures à certaines philosophies utilitaristes (pas celles de Harsanyi et Hare, par exemple).

combiner le souci de compréhension profonde avec les exigences scientifiques de fiabilité et de validité des procédures d'interprétation. Or, il me semble que les procédures retenues n'impliquent nulle référence au stade ultime dans le sens rawlsien.

Le chercheur kohlbergien est confronté au protocole d'une conversation à demi structurée qu'il doit réduire à des unités plus petites. Les réponses à un dilemme donné peuvent contenir plusieurs de ces unités. Ceci ne doit pas trop nous occuper ici: Kohlberg a formulé des règles précises — et d'ailleurs discutables — pour réduire cette pluralité à une assignation de stade univoque. Il s'agit de la procédure de décomposition des réponses en des unités de choix x norme x élément x niveau socio-moral (voir plus haut).

Le manuel lui-même a été construit à partir de sept 'cas de construction', qui ont été choisis arbitrairement dans une population plus large de sujets interrogés à six moments différents de leur vie.⁵² Dans les 42 protocoles Kohlberg et son équipe ont su reconnaître en tout 12 normes et 17 éléments. Les 17 éléments sont répartis en plusieurs groupes, dont certains représentent des théories morales. Ainsi, il y a un groupe qui rassemble les éléments référant à la justice comme 'fairness' (p.e. réciprocité et mérite positif; équité et justice procédurale; ...). Un autre groupe, intitulé 'conséquences promouvant un idéal ou une harmonie', rassemble des éléments qui rappellent les éthiques de la vertu (p.ex.: 'maintenance de la personnalité', 'maintenance du respect de soi-même, promotion d'un idéal social ou de l'harmonie, protection de la dignité humaine et de l'autonomie). Un troisième groupe réunit les éléments utilitaristes (bonnes/mauvaises conséquences pour l'individu ou le groupe), etc. En réalité, les principaux types de théorie morale sont représentés: les théories déontologiques, téléologiques, conséquentialistes et les théories morales construites sur le modèle du droit naturel.

Il est remarquable que l'on trouve dans le manuel des exemples de raisonnements au stade ultime appartenant à tous ces groupes. Nulle part il ne faut supposer que les théories déontologiques représentent en quelque sorte la 'vérité ultime' des autres types de théories. La procédure 'choix x norme x élément' opère une classification pure et simple, partant bien entendu d'une interprétation des énoncés du locuteur. Rien n'indique

⁵² Colby, Kohlberg, *The Measurement of Moral Judgment*, Vol. I, p. 40.

que l'interprète s'en remet à la 'norme de rationalité' afin de comprendre le sens donné par le sujet à ses mots. En vérité, il les juge non à partir de la théorie de la justice de Rawls, mais à partir d'une classification préétablie, dans laquelle la théorie de Rawls n'est même pas représentée.⁵³ Encore ne faut-il pas supposer que l'interprète juge les énoncés du locuteur à partir d'un point de vue normatif. L'assignation de stades est faite en examinant la perspective sociale qui sous-tend les raisonnements du sujet (la théorie 'faible' des stades), bien que, occasionnellement, Kohlberg base les assignations de stade sur les modalités déontologiques des énoncés du sujet (la théorie 'forte').⁵⁴

Revenons-en, pour conclure, aux positions dites 'normative' et 'neutre' envers la psychologie morale. Kohlberg écrit: 'Non seulement les jugements moraux que nous étudions sont des jugements normatifs. Encore, notre théorie est elle-même de nature normative. Notre méthode et notre théorie présupposent une prise de position normative envers le plus ou le moins de rationalité des jugements interprétés'.⁵⁵ Je me suis efforcé de montrer que, dans ce cas précis, cette thèse repose sur une confusion entre une théorie 'forte' et une théorie 'faible' des stades. La théorie 'faible' seule semble suffire aux recherches entreprises par Kohlberg. Elle peut très bien faire *l'époque* de la question de la validité des jugements émis par le locuteur.

Néanmoins, il est également clair que la psychologie morale ne peut se dispenser d'objectiver les prétentions à la validité des agents moraux, faute de quoi elle serait indubitablement condamnée à une 'psychologie de la non-maîtrise'. Une sorte d'herméneutique est donc requise. Je me suis tenu assez loin de la question de quel type d'herméneutique il pourrait s'agir. En tout cas, ce ne saurait être l'herméneutique 'objective' imaginée par Habermas et Kohlberg. La question nous conduit droit au

⁵³ Ce point est également fait par B. Puka, *op. cit.*, qui relate que A. Colby a confirmé dans une conversation que la théorie de Rawls n'a joué aucun rôle dans la construction du manuel.

⁵⁴ Ainsi, dans A. Colby, L. Kohlberg, *op. cit.*, Vol II, p. 381, il croit que certain raisonnement appartient au stade 4/5 parce que le sujet considère qu'il est *permis* de violer des lois si l'on est prêt à en subir les conséquences. Le sujet ne comprend *pas encore* que parfois il est *obligatoire* de les violer. Des lapsus pareils sont rares dans le manuel.

⁵⁵ L. Kohlberg e.a., *Moral Stages*, p. 14.

coeur du problème des rapports entre les intuitions des agents, leurs expressions et systématisations dans la philosophie morale 'savante' et les traditions historiques qui relient la *folk psychology* (comme on dit en Anglais) et la philosophie morale.

Université de Gand