

Dans « *Sanity and the metaphysics of responsibility*¹ », Susan Wolf examine certaines théories de l'autonomie et de la responsabilité — notamment celles mises de l'avant par Frankfurt, Watson ou Taylor — qu'elle regroupe sous le terme commun de « théories du moi profond » (*deep-self view*). D'emblée, Wolf reconnaît que ces théories visent juste en exigeant, pour l'autonomie et la responsabilité, que le moi superficiel soit placé sous le contrôle d'un moi plus profond. Par contre, Wolf n'hésite pas à reprocher aux théories du moi profond leur incomplétude. Bien sûr, les actions de l'agent autonome et responsable doivent être motivées par ses valeurs et volontés profondes, et non par de simples désirs superficiels. Mais cela ne suffit pas. Encore faut-il que l'agent soit pourvu de santé d'esprit (*sanity*) pour qu'on soit véritablement en présence de responsabilité. Ainsi, en surajoutant le concept de santé d'esprit aux théories du moi profond, Susan Wolf fixe des exigences plus élevées en ce qui a trait à la définition de la responsabilité.

Plus spécifiquement, Wolf définit la santé d'esprit comme une compétence, en grande partie caractérisée par la juste saisie de contenus normatifs donnés. En ce sens, elle affirme le caractère hautement substantif de sa théorie et prend ses distances face aux aspects plus procéduraux des théories du moi profond. En fait, elle laisse entendre qu'une relation de biconditionnalité serait à l'œuvre entre la *compétence* de l'agent et la saisie du *contenu* normatif : non seulement la saisie du contenu normatif impliquerait la compétence de l'agent, mais, en outre, un agent compétent saurait nécessairement saisir ledit contenu. Il ressort de cette proposition biconditionnelle que quiconque n'ayant pas su saisir un contenu normatif donné serait jugé incompétent.

Donc : si l'agent doit saisir un contenu normatif donné pour être jugé sain d'esprit, si par ailleurs la santé d'esprit est une condition nécessaire à sa responsabilité, alors on doit exonérer de toute responsabilité l'auteur d'un comportement moralement répréhensible, lorsque la faute morale qu'on lui attribue repose sur la non-saisie du contenu normatif.

Parce qu'elle tend à exonérer les agents de leur responsabilité envers les actes qu'ils ont commis, sous prétexte de leur incompétence, la théorie de Wolf pourrait à juste titre être qualifiée de paternaliste. Gerald Dworkin² établit cinq grands axes à partir desquels on peut caractériser les diverses instances de paternalisme; un de ces axes permet de distinguer un paternalisme léger — en vertu duquel l'intervention vis-à-vis un agent est justifiée seulement dans les cas où il y a lieu de croire que celui-ci est incompétent —, d'un paternalisme plus lourd, qui autorise une intervention même dans les cas où l'agent est jugé pleinement compétent. Par ailleurs, même si dans la plupart des cas, le paternalisme consiste à intervenir pour limiter la *liberté d'action* de l'agent, au nom de son bien, de sa protection ou de son intérêt, il existe néanmoins certaines situations où l'attitude paternaliste se traduit plutôt par une intervention visant à limiter la *responsabilité* d'un agent par rapport à une action déjà commise. À ce titre, on peut évoquer certains cas de paternalisme judiciaire, lors desquels des peines moins lourdes qu'à l'ordinaire sont prononcées si les contrevenants sont mineurs, atteints d'une déficience intellectuelle, etc. Ainsi, un « paternalisme à l'égard de la responsabilité » pourra juger qu'un agent n'est pas pleinement responsable de certains de ses actes, s'il n'était pas pleinement compétent au moment de les poser. Contrairement au paternalisme à l'égard de la liberté, qui intervient *a priori*, c'est-à-dire avant que l'agent ne passe à l'action, le paternalisme à l'égard de la responsabilité intervient *a posteriori*, par rapport à des actions déjà commises. Dans les deux cas cependant, c'est le même constat d'incompétence de l'agent qui justifie l'intervention.

Formellement, la théorie de Susan Wolf fait preuve d'un paternalisme léger à l'égard de la responsabilité, au sens où seuls les agents jugés incompétents se voient exonérés de leur responsabilité relativement aux comportements dont ils sont les auteurs. Toutefois, dans la mesure où une définition substantive et très pointue de la compétence restreint la classe des agents compétents pour étendre celle des agents jugés incompétents, on peut penser que la théorie de Wolf est en fait beaucoup plus paternaliste qu'il n'y paraît au premier abord.

¹ Susan WOLF, « *Sanity and the metaphysics of responsibility* », dans John Christman (dir.), *The Inner Citadel: Essays on Individual Autonomy*, New York, Oxford University Press, 1989, p. 46-62.

² Gerald DWORKIN, « *Paternalism* », *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (dir.), 2005, <http://plato.stanford.edu/archives/win2005/entries/paternalism/>; Gerald DWORKIN, « *Moral Paternalism* », *Law and Philosophy*, vol. 24, 2005, p. 305-319.