

# Autonomy and self-respect

THOMAS E. HILL, JR.  
*The University of North Carolina at Chapel Hill*

U.C.L.  
INSTITUT SUPERIEUR DE PHILOSOPHIE  
Bibliothèque  
Collège D. Mercier  
rue Cardinal Mercier, 14  
1348 Louvain-la-Neuve

 **CAMBRIDGE**  
UNIVERSITY PRESS

own happiness. After all (as Kant noted) we are usually all too ready to seek our own happiness. An ideal moral legislator, aware of the burden of imposing duties, would see little reason to make such pursuit a matter of duty (as with restraint from killing) even though he counted the happiness of all an ideal end. The most intuitively plausible cases, however, in which one feels a person has a duty to himself seem to be cases in which a person is not respecting himself, not cases in which he is curtailing his happiness. It is tempting to say to the artist, the actress, and the schoolmaster in our examples, "You owe it *to yourself* to change." This is not just an estimate that a change will bring greater happiness, though we may believe that too. It is as if we take a moral interest in persons' setting and living by their own values, even when the values are not required for all. We care for their having the satisfaction of a good opinion of themselves, but not just for the pleasure of it. Perhaps the moral interest is in each person living as an *autonomous* agent, where "autonomy" implies both personal integration and forming values beyond comfort, least resistance, etc.

Thinking of self-respect in the previous way (appreciation of rights), I argued that one should respect oneself as well as every other human being. And, in that sense, absence of self-respect tends to undermine respect for others. But, given the conception of self-respect I have discussed here, one cannot say the same. Perhaps we do believe that everyone should try to respect himself in this sense, but we cannot reasonably ask that everyone respect all others. For respect of the sort we have been considering requires seeing that the respected person has personal standards or ideals and believing that he lives by them. But, unfortunately, not everyone seems to have such standards, and, even more obviously, not everyone lives by the standards he has. The respect in question need not be merited by special talent, achievement, or even ambition; but one must qualify for it by setting and sticking to some personal values. Again, it is not clear that those who lack self-respect of this sort cannot respect others. But what does follow is that one who does not respect himself in this way cannot expect that others will. To ask them to respect him would be to ask them to acknowledge what he himself cannot, namely, that he has and lives by his own (morally permissible) values. As Kant remarked, perhaps too harshly, "One who makes himself a worm cannot complain if people step on him."<sup>8</sup>

<sup>8</sup> *The Doctrine of Virtue* (Part II of the *Metaphysic of Morals*), ed. Mary J. Gregor (Harper Torchbooks, 1964), p. 103 (p. 437 Academy edition). Kant's remark lends itself to my point, but I do not want to argue that it is just what he meant.

## Autonomy and benevolent lies

Often it is easy to see what is wrong with lying.<sup>1</sup> Many lies are vicious: they are meant to hurt, and often do. Other lies are self-serving at the expense of others: they gain something for the liar but are detrimental to those who are deceived. Even well-intentioned lies are sometimes discovered, with consequent damage to valued relationships and to trust and credibility in general. Many lies are violations of professional obligations; others are breaches of promise to particular individuals. But there are also instances in which these explanations do not seem to apply and yet the lie is still not beyond moral question. We feel that there is at least something to be said against lying even then, but it is not obvious what this is. No promises or professional commitments are at stake; no harmful consequences are intended or expected; and yet the lie still seems at least *prima facie* objectionable. We naturally wonder, "why?" To rest the matter with the intuitive remark that truth-telling is always a *prima facie* obligation is hardly satisfying. To say that killing, maiming, and causing pain are *prima facie* wrong may arouse no further questions; but why, one wonders, should truth-telling be viewed this way, especially when a lie seems likely to result in more good than harm?

At least a partial answer, I suggest, is that lies often reflect inadequate respect for the autonomy of the person who is deceived. Unfortunately, though autonomy has been an increasingly popular concept in recent years, there is no uniform understanding about what autonomy is. I hope that by tracing different conceptions of autonomy from their Kantian prototype we can see more clearly, and specifically, the various ways in which lies interfere with autonomy. But the interest in autonomy extends beyond our immediate question about lying: benevolent lies merely illustrate one of many ways in which narrow utilitarian thinking can foster unwarranted interference in others' lives.

My remarks will be divided as follows: *First*, I characterize a special class of benevolent lies which pose the main issue sharply; *second*, I distinguish several conceptions of autonomy with associated moral principles; *third*, I try to explain how, in different ways, these principles oppose benevolent lies; and *finally*, I

<sup>1</sup> I have had the benefit of helpful comments and criticisms from a number of students and colleagues, most notably Robert M. Adams, Gregory Kavka, Warren Quinn, a U.C.L.A. Law and Philosophy discussion group, and colloquium participants at the University of Utah and the University of California, Santa Barbara.

## AUTONOMY AND SELF-RESPECT

comment briefly on how a believer in autonomy might respond to the hedonist contention that it is irrational not to lie if a lie will result in the most favorable pleasure/pain ratio.

My aim is not to find a precise rule or decision procedure for deciding hard cases. In fact I suspect that the search for one would be misguided. Many different considerations oppose most lies, and in difficult situations there are arguments for and against. My object is not to rank these competing factors in importance but rather to articulate one type of consideration that is too often overshadowed. Malice, harm, and breaches of trust are so obvious and so often objections to lies that subtler sorts of objection, though usually present at the same time, may be overlooked. The point of isolating very special examples of benevolent lies is to focus attention on such objections; it is not to deny the importance of other arguments or to articulate considerations that apply only in rare circumstances.

### I

A former teacher related to me the following true story (which I have modified slightly). He had a student who showed in tutorial conversations signs of deep, suicidal depression. The student was later found dead, and the circumstances were such that others could easily have seen his death as accidental. The professor helped to gather up the boy's belongings to return to his mother, and no suicide note was found. But the mother, a devout Roman Catholic, was deeply worried about her son's soul, and she asked the professor point blank whether he had any reason to suspect suicide. The professor, an atheist, wanted to comfort her and so, by a quite deliberate lie, assured her that, as far as he knew, the boy had been in good spirits.

Another true story concerns a doctor who discovered that his mother, a very elderly but happy woman, had extremely advanced atherosclerosis. Her doctor had apparently chosen to treat the problem as best he could without informing the woman how near death she was. The son had no objection to the medical treatment or her doctor's decision to withhold information. Though he thought his mother psychologically and physically capable of handling the truth, he believed that her last days would be happier if she did not know. The problem arose when she asked her son directly, "Do you think the doctor is telling me everything?" The son lied; but since the question concerned his opinion and he had learned of her condition in ways she did not suspect and without anyone else knowing that he knew, he felt confident that she would never discover his lie. He lied to make her more comfortable, and she was in fact happy until her death.

Consider, lastly, a dilemma which could occur even if it has not. Mary has made a painful break from her ex-lover, John, and though pulled towards him, is on the mend. Her roommate is pleased for her, as she knows that John and

## AUTONOMY AND BENEVOLENT LIES

Mary were, and will remain, painfully incompatible. She is fearful, though, that John and Mary will get together again, causing both unnecessary misery before the inevitable final separation. Overhearing John talking with a friend, she learns that John is ready to "start over" if only he receives an encouraging sign; and she expects that Mary, ever the optimist, would give the sign. Later Mary asks the roommate, "Do you think he would want to try again if I asked him?" As an act of kindness, the roommate replies, "No, I am sure he knows it would never work."

These examples illustrate the special sort of benevolent lies I want to consider. The lies are benevolent because they are intended to benefit the person deceived, for no ulterior motives, and they actually succeed in giving comfort without causing pain. Despite the benevolent motives, there is no denying that deliberate lies were told. We are not dealing with examples of mere silence, evasion, ambiguous response, and the like. The lies, moreover, are not designed to protect incompetents from truths beyond their capacities to handle sanely and responsibly. In our sample cases a lie will protect someone from avoidable pain, but it is not needed to prevent serious physical or psychological damage, violent outbursts, gross misperception of reality, and so on.

Our examples also fall outside a range of special problem situations. Some lies, for example, are told in a context where the liar has rather little chance of being believed; but in our cases there is sufficient credibility to make the deception effective. Other lies concern matters which are, intuitively, "none of the business" of the questioner: for example, a lie told to a curious student who asked his teacher about his private sex life. But the questions in our examples are clearly not "out of bounds" in this way. What is asked for is information or opinion about what deeply concerns the questioner's own life. Also the lies in our stories cannot be deemed trivial. Unlike "little white lies," they are about matters of the utmost importance to the deceived: heaven or hell, life or death, reunion or separation from a loved one. Further, our examples concern lies between individuals, not lies from public officials or to institutions, and so certain questions of public responsibility are left aside. Finally, let us imagine that the deceived has not forfeited a right to know, for example, by his own repeated lying or by having a plain intent to misuse the truth.

Lies are often wrong at least in part because they are breaches of a promise to be truthful, but, to simplify matters, let us suppose that there were no such promises in our examples. It is easy to imagine that the professor, the dying woman's son, and the roommate never made an *explicit* promise to tell the truth as, for example, one is required to do before testifying in court. The more difficult matter is to remove the suspicion that they made a tacit or implicit promise to be truthful. W. D. Ross maintained that we make such an implicit promise every time we make an assertion, and so he viewed all lies as breaches of promise. But this position, surely, is implausible. Suppose, for example, two enemies

benevolent  
lies

## AUTONOMY AND SELF-RESPECT

distrust each other, have no desire to be honest with each other, and both know this. As seems to be common in international relations, they tell the truth to each other only when they expect that lying will not give them an advantage. In this situation when one asserts to the other, say, that he has documents damaging to the other's political ambitions, we cannot reasonably interpret this as a promise. Neither person believes that the speaker intends to put himself under obligation. Given their mutual understanding, the speaker cannot seriously intend to lead the other to believe that he is making it a matter of conscience to convey the truth. Furthermore if every assertion amounted to a promise to say what is true, we would not think, as in fact we do, that a lie preceded by an explicit promise to be truthful is usually worse than a lie not preceded by such a promise. There are, of course, implicit promises but it requires more than mere assertion to make one. Suppose, for example, Mary and her roommate had often discussed how they valued each other's honesty and frankness, and each had on other occasions insisted that the other tell the truth, however painful, and neither gave any hint of reservations about giving and counting on complete truthfulness between them. With this special background we might want to say that they had made implicit promises to tell each other the truth. However, to focus attention away from promises, let us suppose that in our examples there were no such special conditions to create implicit promises to be truthful.

Our examples are also meant to minimize the force of utilitarian considerations that so often tell against lying. Most importantly, the lies in our stories are extremely unlikely to be discovered. It is a moralist's fiction that lies can never remain hidden: perhaps a useful fiction, but untrue nonetheless. In each of our examples a person is asked about what he knows or believes, and if he is determined to stand by his response there is no practical way others can find out that he is lying. Even if the student's mother learns that her son committed suicide, she cannot know that the professor lied; the elderly woman can find out that she is seriously ill, but not that her son lied about his opinion; Mary may learn that John is still available, but she has no way of discovering that her roommate knew. There is, of course, always *some* chance, however remote, that those who lie will give themselves away; for example, they may talk in their sleep. If the discovery of the lie would be an utter disaster, then from a utilitarian point of view even this very small risk might not be warranted. But to simplify, let us suppose that in our cases discovery would not be disastrous. The persons deceived, let us say, have an unusually forgiving and trusting nature. If they realized the special circumstances and benevolent intent, they would forgive the lie; and, though disappointed, they would not become unreasonably suspicious and distrustful. Again, typical lies tend to multiply, one lie calling for another and each lie making successive ones easier; but we can imagine this not to be so in our example. Our professor, doctor/son, and roommate, let us suppose, are of

## AUTONOMY AND BENEVOLENT LIES

firm character and would lie only in the special circumstances we have defined, and they do not need an entangled web of further deception to hide the first.

Lies of the sort pictured here are no doubt rare; but, by minimizing the usual considerations of utility and promises, they enable us to focus on other relevant considerations, which may be important in more typical cases as well. In particular, we can reflect on how lies can fail to respect persons' autonomy.

## II

What does it mean to respect the autonomy of a person? Autonomy has been conceived in quite different ways, and there are, accordingly, different principles and ideals associated with autonomy. Since most, if not all, recent conceptions of autonomy have roots in Kant's writing, a useful procedure may be to review Kant's theory of autonomy and then trace out several other conceptions of autonomy to which one is naturally led when one sets aside some of Kant's metaphysical and moral presuppositions.

*Kantian autonomy.* Kant held that autonomy is a property of the *will* of rational beings.<sup>2</sup> To have a will is to be able to cause events in accord with principles. That is, a rational being has a will insofar as he can "make things happen" in a way which makes appropriate explanations like, "He did (caused) it, because it is his principle to . . . (or for the reason that . . .)." To have autonomy it is also necessary that one's will be *free in a negative sense*. This implies that one is capable of causing events without being causally determined to do so. Even more radically, negative freedom implies ability to cause events without being motivated in any way by one's own desires. But since willing requires acting on principles, in order to exercise this capacity one must have some principles to which one is committed not because one desires to follow them or because the principles are expected to lead one to anything one desires (or will desire). All rational beings, Kant argued, have such principles; they are committed to them simply by virtue of being rational. The principles are self-imposed insofar as they stem from one's rational nature rather than from fear of punishment, desire for approval, blind acceptance of tradition, animal instinct, and so on. To have *autonomy* of the will is to be committed to principles in this way and to be able and disposed to follow them.

Kant argued further that the principles to which one is committed by virtue of one's autonomy are the basic principles of morality. This implies that one is under moral obligation to do something if and only if it is required by the principles one accepts for oneself as a rational being free from determining causes

<sup>2</sup> What follows is a brief summary of Kant's views in the *Groundwork of the Metaphysics of Morals*, interpreted in the light of his later writings. Some points are controversial but, I believe, defensible.

## AUTONOMY AND SELF-RESPECT

and independently of all desire. Nothing else can impose moral obligation; tradition, power, ecclesiastical authority, majority opinion, and natural dispositions in themselves have no moral force. Kant believed that he could draw from these ideas more particular moral results which have also been influential in later discussions of autonomy. For example, he maintained that having autonomy is the basis for human dignity, and in particular for the idea that rational nature in every person ought to be treated as unconditionally valuable, above all "price." Other derivative principles include rights to a wide area of individual liberty and insistence on respect for all persons. Given Kant's long-standing reputation for being unduly moralistic,<sup>3</sup> it is somewhat surprising to find that the particular moral principles he sketches in the *Metaphysics of Morals* leave so much room for free choice. There are no ends such as pleasure or Moore's "intrinsic value" which one is morally required to maximize. Within the bounds of basic principles of liberty, respect, limited beneficence, and self-improvement, etc., one may pursue one's own pleasure or others', or whatever else one desires.

These fundamental features of Kant's theory of autonomy are notoriously embedded in a metaphysical framework and surrounded by specific moral opinions which most philosophers today, quite rightly, reject. Kant held, for example, that a person's commitment to moral principles does not take place in space or time and is incapable of empirical explanation. The uncaused decision to follow those principles in particular situations or not also belongs to the "intelligible" or noumenal world which is admitted to be beyond comprehension, even though criminals may be executed for the wrong decision. All of our desires, without distinction, are treated together as "alien" forces, not part of one's "true" self but only of oneself as appearance. Most of Kant's specific principles, such as the prohibition of lying, are held rigoristically; they cannot be overridden, they include within them no built-in exceptions, and yet they are supposed to bind all human beings, at all times and places. But despite these unacceptable features, Kant's theory suggests less encumbered ideals of autonomy which continue to have a wide appeal.

*Sartrean autonomy.* Suppose one is attracted by Kant's idea that one is morally bound by nothing but what one imposes on oneself and also by his denial of determinism regarding human choices, but one cannot accept Kant's noumenal/phenomenal distinction, his moral rigorism, or his belief in principles of conduct which are necessarily rational for everyone. A natural result would be acceptance of what I shall call Sartrean autonomy.<sup>4</sup> To say that persons are autonomous, or

3 Other ways in which Kant opposed moralism I discuss in "Kant's Anti-Moralistic Strain," *Theoria*, Vol. XLIV (1978).

4 This conception is a distillation from some ideas in Sartre, but I would not argue that it fully captures the complexity and variety of Sartre's view. Variations on the theme, with significant differences, can be found in R. M. Hare, *Freedom and Reason*, and H. D. Aiken, *Reason and Conduct*.

## AUTONOMY AND BENEVOLENT LIES

free, in this sense has two main implications: (1) People act as they do because they choose to do so, and their choices are not causally determined; in fact nothing is to be viewed as a force even partially compelling a person to act, and this includes threats from others, one's own desires and emotions, and even so-called compulsions, addictions, and the like. (2) People are morally and rationally free to do as they choose in that there are no objective values, only self-imposed commitments. No general moral principles follow from the contention that persons are autonomous in this sense. A naturally associated principle (which, strictly, does not follow) is that one should not deny anyone's autonomy either by treating him as if he were compelled to act as he does or by moralizing to him as if he were subject to moral constraints not of his own making. Acknowledging persons' autonomy, however, would not require tolerating their behavior or even refraining from value judgments about their choices for, after all, one is free to make and express one's own values and these may prescribe opposition, even violent response.

*Autonomy as a psychological capacity.* Kantian and Sartrean autonomy were supposed to be features of all human beings, but autonomy is often conceived as a characteristic of only mature, reflective persons. Autonomy on this view is not freedom from causal determinism, still less an ability to act independently of desire. It is a capacity and disposition to make choices in a rational manner; and this means choosing in the absence of certain particular attitudes and inner obstacles, such as blind acceptance of tradition and authority, neurotic compulsions, and the like. The paradigms of a person who is nonautonomous in this sense include the child who accepts authority without question, the adolescent who rebels against authority with as little understanding, the traditionalist who will not consider new ways of doing things, the compulsive gambler who cannot stop gambling even though he wants to, and the masochist and the sadist who impulsively hurt themselves or others without any idea why. This conception, like Kant's, treats autonomy as a capacity and disposition to make rational choices free from certain "alien" factors, but now the factors are more narrowly circumscribed. To be autonomous one need not act independently of all causes and desires but only from certain causes and desires which interfere with rational choice. The ideals naturally associated with this conception of autonomy are the development of rational capacities in education, the overcoming of unconscious psychological disabilities through psychotherapy, and the *use* of one's rational capacities in making important choices.

*Autonomy as a right.* Sometimes the assertion that persons are autonomous is an attribution of a right rather than a psychological capacity. To say that persons have autonomy in this sense is not a descriptive statement that they are in fact free from certain influences (such as inner compulsions and slavery to convention) but instead a claim that they ought to be free from certain influences (inappropriate interference by others). Insofar as the idea has roots in Kant, it stems more from

## AUTONOMY AND SELF-RESPECT

his principles of liberty and respect for persons than his metaphysical doctrines. Though other rights have been associated with autonomy, the right I have in mind is a moral right against individuals (not the state) (a) to make one's own decisions about matters deeply affecting one's life, (b) without certain sorts of interference by others, (c) provided certain conditions obtain. The right presupposes a background of other moral rights and legal rights within a just system, which define an area of permissible conduct.

Consider first (a) the area of choice to be protected. In some sense one's "freedom" is limited when others prevent one from beating up children, entering others' houses at will, etc.; but the right of autonomy concerns decisions within an area of morally and legally permissible conduct. Fair competition, when successful, can also "interfere" with what one can choose to do, as when someone else wins a desired spot on a team or buys the last copy of a rare book. But, again autonomy is not a right to be successful in such competitive situations. My "freedom" is also limited when I cannot control others in matters which primarily affect them, e.g., my son's choice of college major, his hair style, and his dating partners. But autonomy is usually thought to be freedom to make one's own decisions about matters which most deeply affect one's own life, not to control others in matters which more seriously affect them.<sup>6</sup> Common examples would be choices about what jobs to take (among those for which one is qualified), what people to associate with (among those willing to reciprocate), what books to read, where to live, etc.

Even within this circumscribed area, the right of autonomy is only a right to make one's choices (b) free from certain interferences by others. Among these interferences are illegitimate threats, manipulations, and blocking or distorting the perception of options. The most obvious sorts of threats which violate a person's right to make his own decisions are threats to do things which would be wrong to do quite aside from considerations of autonomy: for example, threats to kill, maim, spread false rumors, deliberately ruin a career, or disown a child. When a person uses such a threat to control another's decisions, say, about whom to marry or whether to apply for a certain job, then the threat seems especially wrong (more wrong, for example, than if the threats were used to make someone keep a secret not pertaining to his own life). Even threats to do what one has a

5 For some purposes one may well want to claim a right of autonomy against undue interferences by the state, but, while not denied here, this is not especially pertinent to the issue at hand.

6 Two comments may help to prevent misunderstanding. *First*, while the line between matters which deeply affect one's own life and other matters is admittedly imprecise, the principle uses this notion only to characterize roughly the range of cases in question, not to distinguish in general permissible from impermissible interferences. Thus, unlike Mill's liberty principle, the principle does not try to mark off a private sphere in which one's choices do not significantly affect others. What deeply affects my life may deeply affect another's as well. *Second*, the fact that the principle is restricted to matters deeply affecting one's own life does not preclude there being *other* principles opposed to interference in other matters: the right of autonomy is not conceived here as a comprehensive right against all undue interference.

## AUTONOMY AND BENEVOLENT LIES

right to do would be undue interferences with autonomy if they had no other point than to control another's basic life choices. Consider, for example, the classic threat, "If you don't marry me (go to bed with me, etc.), I will marry the first woman who will have me (jump off a cliff, join a monastery, etc.)." Such a threat is obviously designed to control another's decision, not merely to explain one's contingency plans. It would be quite a different situation, and not an illegitimate threat, if at some point in a long relationship a person expressed a genuine conditional intention for which he (or she) had good reasons independently of a desire to control another's choice: for example, "If you do not marry me (go to bed with me, etc.), I intend to look for a new partner."

Manipulation, broadly conceived, can perhaps be understood as intentionally causing or encouraging people to make the decisions one wants them to make by actively promoting their making the decisions in ways that rational persons would not want to make their decisions. Obvious examples would be subliminal advertising, posthypnotic suggestions made to nonconsenting subjects, "brain-washing," getting someone drunk or drugged before a major decision, bribes appealing to a person's weaknesses, playing on a person's neurotic guilt feeling, coloring certain options black by insinuation, and so on. But fully rational persons not only want to make their important decisions free from these subversions of the deliberative process; they also want to have the opportunity to know their options and to reflect on any relevant considerations for and against each option. In short, a rational decision maker wants not only to have a clear head and ability to respond wisely to the choice problems presented to him; he wants also to see the problems and the important facts that bear on them realistically and in perspective. Thus one can also manipulate a person by feeding him information selectively, by covering up pertinent evidence, and by planting false clues in order to give a distorted picture of the problem situation.

Manipulation implies an actual intent to control another person by getting him to make decisions we want him to make; but one can also fail to honor persons' right to make their own decisions when one knowingly and actively interferes with their opportunity to see the significant choices that the circumstances offer, even if such interference is only a foreseen, unintended consequence of what one primarily intends to do. Suppose I liked very much the picture on a poster announcing a scholarship competition for study abroad and so, not caring at all whether anyone might apply for the scholarship or not, I took the poster to decorate my office. Though my intention was not to manipulate anyone's choices, I still knowingly and effectively prevented others from making choices that might significantly affect their lives. Or, again, suppose a father explained to his angry, newly feminist daughter: "When I taught you to be 'feminine,' I never meant to keep you from choosing to be a pilot (jockey, surgeon); I only wanted you to be popular and fit in well." By this he might avoid the charge of *manipulating* her career choice, but not the more general charge that he had

## AUTONOMY AND SELF-RESPECT

failed to honor her right to make such choices without undue interference; at least this is so, if the means by which she was taught to be "feminine" were predictably ones which kept her from realizing that she had choices other than being a housewife, model, etc.

Most of us, surely, believe in a moral right of autonomy only (c) provided certain conditions obtain. Most obviously, the right makes sense only when its possessor has at least a minimum capacity for rational choice. But perfect rationality is not a prerequisite. We do not hold that persons have a right to make their own decisions only so long as they will decide in a perfectly rational way, and that therefore we may interfere with *any* predictably irrational decision. The attitude here is akin to the way many believe we should view autonomous states. At least if they are functioning political/legal systems (not, for example, in anarchy or violent revolution) and they are operating within some minimum standards of justice, it is generally thought inappropriate for other countries to interfere with their "internal decisions," even if likely to be imperfect or stupid. Other nations should not, for example, try to control the outcome of their elections or legislative process by threats, propaganda, or circulating false rumors. This idea of the autonomy of nations is often disregarded, of course, just as the autonomy of individuals is often ignored; but it is still an ideal which many profess.

The right of autonomy of individuals is also commonly understood to be qualified by a *proviso* that interference is not required to avert a major disaster or to prevent the violation of other, more stringent rights. If, for example, the only way to persuade someone to make a decision that will prevent a riot or a series of murders were to make an otherwise impermissible threat or a nonrational appeal to his weaknesses, then surely most would grant that such interference would be justified. Though important, autonomy need not be considered an absolute right.

*Autonomy as capacity for distinctly human values.* A central feature of Kant's theory of autonomy was his belief that human beings can and do value some things in ways that animals do not. Human beings have dignity, Kant thought, because they have distinctively human concerns which elevate them in our estimation above animals. Most of us, no doubt, will disagree with Kant's tendency to identify these special concerns with respect for purely rational moral principles, accepted independently of all desire; but a residue of his view may have wider appeal. Let us say that persons have autonomy in this residual sense if and only if (a) at least some of their values are not simply instinctual responses to an immediate environment, and (b) they value and are disposed to bring about some states of affairs without expecting that these states of affairs will bring them pleasure (or other "good experiences") or prevent pain (or other "bad experiences") for them, or at least they do not value those states of affairs *for the sake of* the pleasure (avoidance of pain, etc.) which they expect will result. To say that

## AUTONOMY AND BENEVOLENT LIES

human beings are autonomous in this sense implies the denial of several forms of psychological hedonism; for autonomous persons not only value and aim for more than their *most* favorable pleasure/pain ratio, they also have some values and aims which are not for the sake of *any* anticipated pleasure (or prevention of pain). What is not denied, however, is the theory that early pleasant and painful experiences are among the causes of our developing the values we have.

That human beings are autonomous in the residual sense just defined seems quite obvious, once this is properly understood. Even setting aside moral values, surely people can and do sometimes care that their children survive and thrive after they, the parents, are dead and incapable of enjoying anything. They also want that people not laugh at them behind their backs, even though they will never know. Not all such values are altruistic, of course; for I might want my enemies to suffer horribly after I am dead. To say that the concerns in these cases are "really" desires *for* one's own good experiences (or avoidance of bad experiences) is surely a mistake. My desire that my children thrive and my enemies suffer after my death is not a desire to have the pleasure of *thinking* that the children will thrive and the enemies suffer. You would not fulfill my desire if, after my death, you tortured my children and rewarded my enemies but had earlier given me the false pleasure of believing that you would do the reverse. What one does not know, perhaps, does not *hurt* one; but sometimes we *want* states of affairs about which we will never know.

What ideals and principles are associated with this conception of autonomy? Much of what we respect and cherish in human beings, which is lacking in (most?) animals, is at least dependent on this human capacity to value more than what is immediately before them and more than agreeable experiences for themselves. More importantly, I believe that many of us at least implicitly accept the following principle: *First*, in dealing with competent human beings who are not violating anyone's rights, one should not presume that they prefer their own comfort (optimal experiences) over other values (such as the welfare of others, the completion of a project, self-awareness, etc.); and, *second*, when one aims to do something *for* others (say, from gratitude, charity, or love), one should not count their comfort (optimal experiences) as more important than values which they sincerely declare to be more important to them, provided at least the declared value is not a violation of others' rights, not the result of momentary impulse, manipulation, obviously false belief, etc. One would violate this principle, for example, if one treated elderly but competent people like small children or pets, caring scrupulously for their comfort and physical well-being but ignoring any desire they might express about what happens in the world outside the confined area which they can experience or check on.

*Autonomy as an ideal rational life.* A person could have autonomy in some of the preceding senses but lack autonomy in other senses. One could, for example, have rational decision-making capacities but still be manipulated by those who



## AUTONOMY AND SELF-RESPECT

control one's information, or the opposite. One could have a right to make one's own decisions but be incapable of caring for anything but maximizing one's pleasures; alternatively, one could have all manner of distinctively human values but lack both the right and the psychological capacities to pursue these values in the ways a rational person would want. These possibilities suggest a final conception of autonomy, which combines several others and adds one further feature.

Let us say that persons have autonomy, or live autonomously, in a final sense if the following is true: (1) They have the psychological capacities for rational decision making which are associated with autonomy; (2) they actually use these capacities when they face important choice situations; (3) they have the right of autonomy discussed previously, i.e., a right to make morally and legally permissible decisions about matters deeply affecting their own lives free from threats and manipulation by others; (4) other people actually respect this right as well as their other rights; (5) they are able and disposed to have distinctly human values; (6) others respect this capacity by not presuming that they value only good experiences for themselves and by not counting their comfort as more important than their declared values; and, finally, (7) they have ample opportunities to make use of these conditions in living a life over which they have a high degree of control.

This last point requires special explanation. One might at first think it sufficient for an ideal rational life to have the capacities, rights, and good treatment from others indicated in (1)–(6); but further reflection quickly shows otherwise. Even if (1)–(6) were true, people could still find themselves unable to make use of these favorable conditions for any of several reasons. For example, though rationally disposed to make the best of their situation and unhindered by threats and manipulation by others, they might be severely confined in the choices they could make by widespread poverty, disease, overpopulation, and absence of technology and culture. Even if it is no one's fault, when one has to labor in the fields all day to survive, one has little opportunity to live *as* a rational person controlling his life. The choice to labor may be perfectly rational, of course; but it may be almost the only rational choice one has a chance to make. Harsh conditions also restrict the range of morally permissible choices: one cannot do philosophy if one must mine coal to feed one's children. Opportunities to live an ideally rational life may be further restricted by pointless role-expectations, conformist attitudes, and the lack of what Mill called "experiments in living." And even though one may be able to select from many brands of soup and cosmetics, if communal values are lost in a capitalistic society then more significant options are effectively closed. Finally, and significantly for present purposes, opportunities for rational, self-controlled living are restricted when one does not know the realities of one's choice situations. I could be able and eager to seek information, to reflect critically, to be on guard against manipulation

## AUTONOMY AND BENEVOLENT LIES

and neurotic patterns, and to decide rationally on the basis of my beliefs, but still I would not really be in rational control of my life if my beliefs about my situation were drastically mistaken. Opportunities to learn the relevant facts are also needed.

### III

How might the principles and ideals of autonomy we have considered help to explain the intuitive feeling that, even in our special cases, benevolent lies are to some degree objectionable? I will pass over Kant's conception of autonomy as a respected ancestor of later conceptions but not itself a viable option. I also cannot accept the Sartrean conception, and in any case it offers little help with our problem. The only relevant principle suggested by the Sartrean conception is that one should not lie in "bad faith," that is, pretending to oneself, when one "really" knows better, that a lie is necessary because the truth would *make* someone do something undesirable or because comforting someone when it causes no harm is an *objective* duty. The ideal of developing the psychological capacities associated with autonomy may give some reason to hesitate to tell lies to protect people from painful realities, but not a reason that applies in all cases. Probably, as a rule, having to face unpleasant truths about matters deeply affecting one's life helps one to develop the capacity for mature, reflective decision making. If so, there would be a general presumption against benevolent lies, even if it would not always be persuasive as, for example, when we are dealing with the very elderly whose capacities have presumably already been developed as much as they will be.

If we believe in the *right* of autonomy, however, we have more reason to object to benevolent lies. This is most obvious in our example of the roommate lying to keep her friend from reuniting with her ex-lover. The roommate manipulates her friend's decision (to call or not to call her "ex") by actively concealing pertinent information. If we accept the right of autonomy, this could only be justified if the reunion would have been so great a disaster that the right is overridden. In other cases the right of autonomy may be violated but in a less obvious way. The professor and the doctor/son, for example, did not lie in order to control the decisions of the people they deceived; they only wanted to spare them avoidable pain. Nevertheless, there were important, life-altering decisions which the deceived might have made if they had not been deprived of relevant information; and surely the professor and the doctor/son knew this. They knowingly prevented certain options presented by the real situation from ever being faced by the people they deceived: to pray or not, and, if so, how; to continue life as usual or to reorder one's priorities; to face death and tragedy stoically or to be open in a new way with friends.

Someone may object as follows: "Sometimes benevolent lies interfere with life-

GE 11 Conclusions

9 April 1971

Autonomous



altering decisions, but not always; often benevolent lies merely keep people from suffering unnecessarily because of something which they can do nothing about. When, for example, a widow demands to know whether her husband suffered when he was killed in the war, there is little she can *do* if she is told truthfully that he died in horrible agony. And similarly, if the suicide's mother had been bedridden and terminally ill, the professor's lie would not have interfered with any important decisions."

The appropriate response, I think, is this: Benevolent lies do not necessarily or always violate the right of autonomy, but we should not be hasty in concluding that a particular lie does not concern any significant decisions. Good novelists and biographers know what philosophers too easily forget, namely, that the most important decisions in life are not always about external behavior, about what to *do* in the public world. How we face death, family tragedy, our own successes and failures, and the way others treat us, is partly a matter of decision, as Sartreans knew but exaggerated. Even *whether* to see a situation as success or failure, tragic or routine, is not simply a matter of perception of fact. We can also interfere with these life-altering decisions, or prevent a person from facing them, by keeping certain truths from him — even if he is immobile for the rest of his life.<sup>7</sup>

Consider next the principles associated with autonomy as a capacity for distinctly human values. Their implications for benevolent lies depend upon what we know about the preferences of the person to be deceived. Suppose, first, that we have no reason to doubt that the questioner wants an honest answer. His question is in effect an expression of a desire to know the truth. To give him less because we want to spare him pain would be to count his comfort more important than what he himself professes to value more and so would be contrary to our principles.

Sometimes, of course, people ask questions wanting to be reassured rather than to learn the truth. What should we do if we have indirect evidence that the questioner does not really want to know? Much depends, I think, on the nature and strength of the evidence. Suppose, for example, the evidence is rather evenly mixed: the person often shrinks from painful realities but, on the other hand, he asked in a serious tone, he never said in advance not to reveal the sort of fact in question, and the truth is not outside the range of answers he could anticipate. Often when we are in doubt whether a person really prefers what he professes, we can remove the uncertainty by asking further questions; but the peculiarity

7 Several have suggested to me that opposition to lying in these cases stems from the judgment that knowing the truth, or facing tragic realities, is intrinsically valuable regardless of the pain it causes; but I suspect that theories (such as G. E. Moore's) which make it a duty to promote an objective intrinsic value will repeatedly call for interference with autonomy. Robert Adams suggested that an ideal of autonomy might include "*living one's own life*," e.g., experiencing the tragic realities actually surrounding one, quite aside from opportunities to make *decisions*, rational or otherwise; but I think that autonomy is so closely associated with the idea of "*self-governing*" that his ideal is probably better classified under some other conception.

of the dilemma of the would-be benevolent liar is that he cannot resolve the uncertainty this way. To ask, "Would you *really* prefer the truth even though it will hurt?" is in effect to give away the answer. When faced with such mixed evidence and unresolvable uncertainty, one guided by our principles of autonomy would, I believe, again be disposed to tell the truth; for respecting a person's capacity for distinctly human values implies that, other things equal, it is worse to presume that someone prefers comfort to some other declared value than to presume the opposite.<sup>8</sup>

If there were definitive evidence that the questioner preferred not to learn the painful truth, then autonomy as a capacity for distinctly human values would not be relevant. This would be the case if, for example, the questioner had explicitly requested in advance not to be told the truth in specified circumstances, and then, later, those circumstances arose and ample evidence indicated that he had not changed his mind.

Such cases, however, are probably rare. Normally even if a person has previously asked not to be told the truth, his subsequent question raises legitimate doubts about his current preferences. Suppose the earlier request was not made in anticipation of a period of incompetence — like Ulysses' request to his crew before facing the Sirens ("Don't listen to what I say later"). Then the would-be liar is apparently faced with two conflicting requests: an earlier request for deception, and a later request for truth. Unless there are independent reasons for discounting the latter, or for not treating the later question as a request for truth, then one might argue that respect for autonomy gives precedence to the more recent request. Other things equal, we respect a person's autonomy more by allowing changes of mind, honoring what he *does* profess to value over what he *did* profess to value.

The many-sided *ideal* of autonomous living will usually give further reason for hesitating to tell benevolent lies. Even if benevolent lies do not violate a *right*, they still deprive people of a realistic picture of their situation. Insofar as having such a realistic picture is needed for genuine rational control over one's life, to that extent the benevolent liar fails to promote an ideal end.<sup>9</sup>

It may be objected that this argument supports the desirability of volunteering the truth just as much as it supports the desirability of not actively depriving

8 This may seem strange if one supposes (mistakenly) that we should give people what they want — truth or comfort, whichever they prefer. But the principle in question was in fact rooted in a different idea, namely, that persons are to be respected for their distinctly human (e.g., non-hedonistic) values. From this point of view, given uncertainty, it is worse to err in supposing that they prefer comfort to truth than to err in the opposite direction.

9 It may be argued, rightly, that sometimes benevolent lies may promote the ideal of autonomous living in other respects. This might be so if, for example, coping with a painful truth, about which little could be done, would so preoccupy a person that other important aspects of life would be comparatively neglected. Sometimes, perhaps, too much information can also interfere with rational decision making.

## AUTONOMY AND SELF-RESPECT

someone of the truth; and yet, it might be said, it is counterintuitive to suppose that we have as much reason to volunteer painful truths as to tell them when directly asked. The ideal does give reason to volunteer the truth, I think, but there are also reasons why lying in response to a direct question is worse than merely not volunteering the truth. There is a general presumption that one should not cause avoidable pain to others, but this presumption is at least partially set aside when the person requests the painful treatment for the sake of something he wants: e.g., painful medical tests. Thus, although there is a general presumption against expressing truths which cause pain, this presumption is at least partially set aside when a competent person asks for truth; but the presumption is not set aside when one simply volunteers the truth without being asked. Thus, though the ideal of autonomy gives some reason for volunteering painful information about someone's life, the case for volunteering is not as strong as the case for telling the truth when asked.

Another objection might be this: "Sometimes we need to lie in order to increase the chances that a person will make his own decisions (and so live autonomously). For example, when my son asked me where I wanted him to go to college, I lied, telling him that I did not care. Actually I wanted very much for him to go where I went; but I figured that he could make up his own mind better if I kept my preference to myself."

The objection points to a practical problem difficult to resolve in real cases, but it does not, I think, show that the ideal of autonomy unequivocally recommends lying even in the example just presented. *One* aspect of the ideal, to be sure, was encouraging people to make their important decisions in a rational way free from inner psychological obstacles such as neurotic need for a father's approval. Thus, if the son in our example was so dominated by his father's opinions that he could not make a rational choice once his father expressed his desires, then one aspect of the ideal of autonomy would urge the father to hide his opinion. But let us suppose, as in our previous examples, the person deceived is rationally competent with respect to his choice problem and so is not a slave to his father's wishes. In this case another aspect of the ideal of autonomy would urge the father to express his wishes: he should make clear both that he prefers his son to go to his old college and also that he wants his son to decide on the basis of what he, the son, most wants. This puts the pertinent facts on the table, giving the son an opportunity he would have otherwise lacked, namely, to choose whether to give weight to his father's wishes or not and, if so, which wish to count more important. By lying, the father would have helped the son make a self-interested choice; but, as we have seen, one's autonomous choice is not always self-interested. To "make up one's own mind" is not necessarily to decide without regard for others' wishes but to decide maturely in the light of the facts about the situation.

So far we have considered ways in which principles and ideals of autonomy

## AUTONOMY AND BENEVOLENT LIES

help to explain why we view even benevolent lies as to some degree objectionable; but we also have intuitive opinions about which sort of lies (or deceptions) are worse than others. Let us consider, then, whether considerations of autonomy help to explain these intuitions as well.

To consider several factors together, I suppose it is commonly accepted that deceptive responses to questions are worse, other things equal, when (a) the response is a direct lie rather than a merely evasive, misleading, or deceptively ambiguous response, (b) the person deceived trusts the deceiver and was encouraged to do so, and (c) the lie concerns the life of the deceived rather than matters only remotely touching him. The lies of the roommate and the doctor/son described earlier exemplify the first sort. An example of the second, less significant sort of deception might be this: A person asks me, simply from curiosity, "Do you know whether so-and-so is gay?" and, though I know, I answer, "How would I know?"

Now utilitarians will have familiar explanations why the first sort of lie is regarded as more serious than the second; but it is worth noting that our principles and ideals of autonomy provide an alternative, or additional, explanation. In brief, one's opportunity to live in rational control of one's life is increased when there are people one can unmistakably identify as prepared to give straight, honest answers to direct pointed questions. If one does not want to know, one can refrain from asking; if the first answer is evasive or ambiguous, suggesting a reluctance on the other's part to reveal the truth, then one can choose to put the question again more pointedly or to back off; and if one does insist ("I want a straight, honest answer!"), then, while allowing for honest errors, one can make important decisions with more confidence that one understands the real situation. To live in a world without people we can rely on in this way would be to live in a world in which we have less control over our lives. Utilitarians often stress the unpleasantness that results when lies which violate trust become discovered, and for this reason our examples were designed to minimize the risk of discovery. But now it emerges that ideals of autonomy not only oppose undiscoverable benevolent lies; they also oppose lies which risk discovery of a breach of trust, for discovery of such lies encourages us to be distrustful and suspicious and so less able to make use of even the honest answers trustworthy persons give us; and this limits our opportunities for rational control over our lives.

## IV

These conclusions, of course, are both hypothetical and intuitive: that is, the argument has been that if one accepts certain principles of autonomy, then one has reasons to refrain from benevolent lies. But imagine now an objection from a normative hedonist unwilling to rest the issue on intuitive principles. He argues

that, intuitions aside, it is *irrational* to prefer truth to comfort, unless having the truth would maximize one's pleasure in the long run. Thus, he continues, when one aims to be benevolent towards another, it is *irrational* to give him the truth if a lie will contribute more to his total satisfaction.

The objection rests on the common, but mistaken, assumption that, at least when free from moral constraints, a fully rational person would always aim for his most favorable pleasure/pain ratio. But why so? As we have seen, people do in fact have (nonmoral) concerns independent of any anticipated good experiences. Some, perhaps, make maximum pleasure their goal; and others do not. What determines whether one is rational is not, by itself, the content of one's aims, but how they are arrived at, how they fit into one's life plan, etc. More plausible than the hedonist's conception of rationality, I think, is that of John Rawls, who defines ideal rationality, roughly, as satisfying certain "counting principles" (means-end efficiency, inclusion, etc.) and then deciding in light of full information about one's desires, circumstances, etc. Given this conception and the falsity of *psychological* hedonism (i.e., that all seek only to maximize their pleasure), then the rational life will be different for different people. For some, maybe, it will be predominantly pursuit of pleasure; but, unless we suppose that all non-hedonistic desires would extinguish when exposed to more information, for many the rational life will include pursuit of other values, such as truth, independently of their payoff in personal satisfaction.

The principles of autonomy which we have considered, though still un-unified in a general theory, point toward a conception of morality quite different in spirit from familiar forms of utilitarianism, hedonistic and otherwise. The latter start with views about what is intrinsically valuable as an end, and then define morality, in one way or another, as what promotes this end. A theory of autonomy, following Kant, Rawls, and others, would first define principles for moral institutions and personal interactions, leaving each person, within these constraints, the freedom to choose and pursue whatever ends they will. Such a theory would not oppose benevolent lies on the ground that truth-telling will maximize some intrinsic value other than pleasure (e.g., self-awareness); rather, it would encourage truthfulness as, in general, a way of respecting people as free to choose their own ends.

## The importance of autonomy

For many years we have been hearing that *autonomy* is important.<sup>1</sup> Immanuel Kant held that autonomy is the foundation of human dignity and the source of all morality; and contemporary philosophers dissatisfied with utilitarianism are developing a variety of new theories that, they often say, are inspired by Kant. Autonomy has been heralded as an essential aim of education; and feminist philosophers have championed women's rights under the name of autonomy.<sup>2</sup> Oppressive political regimes are opposed on the grounds that they deny individual autonomy; and respect for the autonomy of patients is a recurrent theme in the rapidly expanding literature on medical ethics. Autonomy is a byword for those who oppose conventional and authoritative ethics; and for some existentialists, recognition of individual autonomy is apparently a reason for denying that there are objective moral standards. Both new rights theorists and the modern social contract theorists maintain that their theories best affirm autonomy.<sup>3</sup> Finally, and not least in their esteem for autonomy, well known psychologists speak of autonomy as the highest stage of moral development.

Recently, however, the importance of autonomy has been questioned from a variety of sources. Utilitarian critics have struck back at the neo-Kantians, and a group of moral philosophers, sometimes labeled "personalists," have challenged the Kantian ideal that we should be moved by regard for impartial principles rather than concern for particular individuals.<sup>4</sup> A "different voice" is being heard, emphasizing aspects of morality too often ignored in the persistent praise of autonomy. Some suggest that, far from being the source and highest development of morality, autonomy may be the special ideal of a particular dominant group, and in fact an ideal which serves to reinforce old patterns of oppression.<sup>5</sup> If so,

- 1 This paper was written for presentation to students and faculty at Ripon College as a part of a colloquium on "Autonomy and Caring," which was inspired by Carol Gilligan's *In A Different Voice*. Thanks are due to the participants at that conference, and especially to Robert Hannaford, for their comments.
- 2 A noteworthy example of the latter is Sharon Bishop Hill's "Self-determination and Autonomy" in Richard Wasserstrom's *Today's Moral Problems* (New York: Macmillan, Third Edition, 1985).
- 3 See, for example, John Rawls's *A Theory of Justice* (Cambridge: Harvard University Press, 1971) and Robert Nozick's *Anarchy, State and Utopia* (New York: Basic Books, Inc., 1974).
- 4 Notably among these are Lawrence Blum, whose *Friendship, Altruism and Morality* (London, Boston and Henley: Routledge & Kegan Paul, 1980) has stimulated much useful discussion.
- 5 This remark applies only to those who try to draw certain moral and political conclusions from Carol Gilligan's research. In her book, Gilligan herself claims only to have uncovered tendencies in male and female populations and not to have established any ideological generalizations.