Paternalism and partial autonomy

Onora O’Neill  Department of Philosophy, Essex University

Author’s abstract

A contrast is often drawn between standard adult capacities for autonomy, which allow informed consent to be given or withheld, and patients’ reduced capacities, which demand paternalistic treatment. But patients may not be radically different from the rest of us, in that all human capacities for autonomous action are limited. An adequate account of paternalism and the role that consent and respect for persons can play in medical and other practice has to be developed within an ethical theory that does not impose an idealised picture of unlimited autonomy but allows for the variable and partial character of actual human autonomy.

Autonomous action, understood literally, is self-legislated action. It is the action of agents who can understand and choose what they do. When cognitive or volitional capacities, or both, are lacking or impaired, autonomous action is reduced or impossible. Autonomy is lacking or incomplete for parts of all lives (infancy, early childhood), for further parts of some lives (unconsciousness, senility, some illness and mental disturbance) and throughout some lives (severe retardation). Since illness often damages autonomy, concern to respect it does not seem a promising fundamental principle for medical ethics. Medical concern would be strangely inadequate if it did not extend to those with incomplete autonomy. Concern for patients’ well-being is generally thought a more plausible fundamental principle for medical ethics.

But it also commonly thought implausible to make beneficence the only fundamental aim of medical practice, since it would then be irrelevant to medical treatment whether patients possessed standard autonomy, impaired autonomy or no capacity for autonomous action. All patients, from infants to the most autonomous, would be treated in ways judged likely to benefit them. Medical practice would be thorough and through paternalistic, and would treat patients as persons only if beneficence so required.

Recurrent debates about paternalism in medical ethics show that the aim of subordinating concern for autonomy to beneficence remains controversial. The group of notions invoked in these debates – autonomy, paternalism, consent, respect for persons, and treating others as persons – are quite differently articulated in different ethical theories. A consideration of various ways in which they can be articulated casts some light on issues that lie behind discussions of medical paternalism.

1. Paternalism and autonomy in result-oriented ethics

Most consequentialist moral reasoning does not take patients’ autonomy as a fundamental constraint on medical practice. Utilitarian moral reasoning takes the production of welfare or well-being (variously construed) as the criterion of right action. Only when respect for patients’ autonomy (fortuitously maximises welfare) is it morally required. Paternalism is not morally wrong; but some acts which attempt to maximise welfare by disregarding autonomy will be wrong if in fact non-paternalistic action (such as showing respect for others or seeking their consent to action undertaken) would have maximised welfare. Only some ‘ideal’ form of consequentialism, which took the maintenance of autonomy as an independent value, could regard the subordination of autonomy to beneficence as wrong. In utilitarian ethical thinking, autonomy is of marginal ethical importance, and paternalism only misplaced when it reflects miscalculation of benefits.

This unambiguous picture is easily lost sight of because of an historical accident. A classical and still highly influential utilitarian discussion of autonomy, and paternalism is John Stuart Mill’s On Liberty (1). Mill believed both that each person is the best judge of his or her own happiness and that autonomous pursuit of goals is a major source of happiness, so his thought happiness could seldom be maximised by any action which thwarted or disregarded others’ goals, or took over securing them. Paternalists, on this view, have benevolent motives but don’t achieve beneficent results. They miscalculate.

Mill’s claims are empirically dubious. Probably many people would be happier under beneficent
policies even when these reduce the scope for autonomous action. Some find autonomous pursuit of goals more a source of frustration and anxiety than of satisfaction. In particular, many patients want relief from hard decisions and the burden of autonomy. Even when they don’t want decisions made for them they may be unable to make them, or to make them well. The central place Mill assigns autonomy is something of an anomaly in result-oriented ethical thought (2). It is open to challenge and shows Mill’s problem in reconciling liberty with utility rather than any success in showing their coincidence.

2. Paternalism and autonomy in action-oriented ethics

Autonomy can have a more central place only in an entirely different framework of thought. Within a moral theory which centres on action rather than on results, the preconditions of agency will be fundamental. Since autonomy, of some degree, is a presupposition of agency, an action-centred ethic, whether its fundamental moral category is that of human rights, or of principles of obligation or of moral worth, must make the autonomy of agents basic rather than derivative moral concern. This concern may be expressed as concern not to use others, but to respect them or ‘treat them as persons’, or to secure their consent and avoid all (including paternalistic) coercion.

A central difficulty for all such theories is once again empirical. It is obvious enough that some human beings lack cognitive and volitional capacities that would warrant thinking of them as autonomous. But where autonomous action is ruled out what can be the moral ground for insisting on respect or support for human autonomy (3)? The question is sharply pointed for medical ethics since patients standardly have reduced cognitive and volitional capacities.

Yet most patients have some capacities for agency. Their impairments undercut some but not all possibilities for action. Hence agent-centred moral theories may be relevant to medical ethics, but only if based on an accurate view of human autonomy. The central tradition of debate in agent-centred ethics has not been helpful here because it has tended to take an abstract and inaccurate view of human autonomy. The history of these discussions is revealing.

Enlightenment political theory and especially Locke’s writings are classical sources of arguments against paternalism and for respect for human autonomy. Here the consent of citizens to their governments is held to legitimate government action. In consenting citizens become, in part, the authors of government action: the notion of the sovereignty of the people can be understood as the claim that they have consented to, and so authorised, the laws by which they are ruled. In obeying such laws they are not mere subjects but retain their autonomy.

This picture invited, and got, a tough focus on the question ‘What constitutes consent?’ An early and perennial debate was whether consent has to be express – explicitly declared in speech or writing – or can be tacit – merely a matter of going along with arrangements. In a political context the debate is whether legitimate government must have explicit allegiance, or whether, for example, continued residence can legitimatize government action. A parallel debate in medical ethics asks whether legitimate medical intervention requires explicit consent, recorded by the patient’s signing of consent forms, or whether placing oneself in the hands of the doctor constitutes consent to whatever the doctor does, provided it accords with the standards of the medical profession (4).

The underlying picture of human choice and action invoked by those who advocate the ‘informed consent’ account of human autonomy is appropriate to a contractual model of human relations. Just as parties to commercial contracts consent to specific action by others, and have legal redress when this is not forthcoming, and citizens consent to limited government action (and may seek redress when this is exceeded), so patients consent to specified medical procedures (and have cause for grievance if their doctors do otherwise). Those who argue that informed consent criteria are not appropriate in medical practice sometimes explicitly reject the intrusion of commercial and contractual standards in medical care.

The contractual picture of human relations is clearly particularly questionable in medicine. We may think that citizenship and commerce are areas where we are autonomous decision-makers, enjoying what Mill would have called ‘the maturity of our faculties’. In these areas, if anywhere, we come close to being fully rational decision-makers. Various well-known idealisations of human rationality – ‘rational economic man’, ‘consenting adults’, ‘cosmopolitan citizens’, ‘rational choosers’ – may seem tolerable approximations. But the notion that we could be ‘ideal rational patients’ cannot stand up to a moment’s scrutiny. This suggests that we cannot plausibly extend the enlightenment model of legitimating consent to medical contexts. Where autonomy is standardly reduced, paternalism must it seems be permissible; opposition to medical paternalism appears to reflect an abstract and inaccurate view of human consent which is irrelevant in medical contexts.

3. The opacity of consent: a reversal of perspective

However, the same picture might be seen from quite a different perspective. Human autonomy is limited and precarious in many contexts, and the consent given to others’ actions and projects is standardly selective and incomplete. All consent is consent to some proposed action or project under certain descriptions. When we consent to an action or project we often do not consent even to its logical implications or to its likely results (let alone its actual results), nor to its unavoidable
corollaries and presuppositions. Put more technically, consenting (like other propositional attitudes) is opaque. When we consent we do not necessarily 'see through' to the implications of what we consent to and consent to these also. When a patient consents to an operation he or she will often be unaware of further implications or results of that which is consented to. Risks may not be understood and post-operative expectations may be vague. But the opacity of patients' consent is not radically different from the opacity of all human consenting. Even in the most 'transparent', highly-regulated, contractual arrangements, consent reaches only a certain distance. This is recognised whenever contracts are voided because of cognitive or volitional disability, or because the expectations of the 'reasonable man' about the further implications of some activity do not hold up. Medical cases may then be not so much anomalies, with which consent theory cannot adequately deal, as revealing cases which highlight typical limits of human autonomy and consent (5).

Yet most discussions of consent theory point in the other direction. The limitations of actual human autonomy aren't taken as constraints on working out the determinate implications of respect for autonomy in actual contexts, but often as aberrations from ideally autonomous choosing. The rhetoric of the liberal tradition shows this clearly. Although it is accepted that we are discussing the autonomy of 'finite rational beings', finitude of all sorts is constantly forgotten in favour of loftier and more abstract perspectives.

4. Actual consent and 'ideal' consent

There are advantages to starting with these idealised abstractions rather than the messy incompleteness of human autonomy as it is actually exercised. Debates on consent theory often shift from concern with dubious consent actually given by some agent to a proposed activity or arrangement to concern with consent that would hypothetically be given by an ideally autonomous (rational and free) agent faced with that proposal. This shift to hypothetical consent allows us to treat the peculiar impairments of autonomy which affect us when ill as irrelevant: we can still ask what the (admittedly hypothetical) ideally autonomous patient would consent to. This line of thought curiously allows us to combine ostensibly concern for human autonomy with paternalistic medical practice. Having reasoned that some procedure would be consented to by ideally autonomous patients we may then feel its imposition on actual patients of imperfect autonomy is warranted. But by shifting focus from what has (actually) been consented to, to what would (ideally) be consented to, we replace concern for others' autonomy with concern for the autonomy of hypothetical, idealised agents. This is not a convincing account of what it is not to use others, but rather to treat them as persons (6).

If we don't replace concern for actual autonomy with concern for idealised autonomy, we need to say something definite about when actual consent is genuine and significant and when it is either spurious or misleading, and so unable to legitimise whatever is ostensibly consented to. Instead of facing the sharp outlines of idealised, hypothetical conceptions of human choosing we may have to look at messy actual choosing. However, we don't need to draw a sharp boundary between genuine, morally significant consent and spurious, impaired consent which does not legitimate. For the whole point of concern for autonomy and hence for genuine consent is that it is not up to the initiator of action to choose what to impose: it is up to those affected to choose whether to accept or reject proposals that are made. To respect other's autonomy requires that we make consent possible for them (7), taking account of whatever partial autonomy they may have. Medical practice respects patients' autonomy when it allows patients as they actually are to refuse or accept what is proposed to them. Of course, some impairments prevent refusal or acceptance. The comatose and various others have to be treated paternalistically. But many patients can understand and refuse or accept what is proposed over a considerable range. Given some capacities for autonomous action, whatever can be made comprehensible to and refusible by patients, can be treated as subject to their consent – or refusal. This may require doctors and others to avoid haste and pressure and to counteract the intimidation of unfamiliar, technically bewildering and socially insecure medical environments. Without such care in imparting information and proposing treatment the 'consent' patients give to their treatment will lack their autonomous character which would show that they have not been treated paternalistically but rather as persons.

5. 'Informed consent' and legitimating consent

There is a long-standing temptation, both in medical ethics and beyond, to find ways in which consent procedures can be formalised and the avoidance of paternalism guaranteed and routinised. But if the ways in which human autonomy is limited are highly varied, it is not likely that any set procedures can guarantee that consent has been given. Early European colonialists who 'negotiated treaties' by which barely literate native peoples without knowledge of European moral and legal traditions 'consented' to sales of land cession of sovereignty reveal only that the colonisers had slight respect for others' autonomy. Medical practice which relies on procedures such as routine signing of 'consent forms' may meet conditions for avoiding litigation, but does not show concern for human autonomy as it actually exists. Such procedures are particularly disreputable given our knowledge of the difficulties even the most autonomous have for assimilating distressing information or making unfamiliar and hard decisions.

Serious respect for autonomy, in its varied, limited forms, demands rather that patients' refusal of consent, at least to fundamental aspects of treatment, be made possible. The onus on practitioners is to see
that patients, as they actually are, understand what they can about the basics of their diagnosis and the proposed treatment, and are secure enough to refuse the treatment or to insist on changes. If the proposal is accessible and refusible for an actual patient, then (but only then) can silence or going along with it reasonably be construed as consent. The notions of seeking consent and respecting autonomy are brought into disrepute when the ‘consent’ obtained does not genuinely reflect the patient’s response to proposed treatment.

6. Partial autonomy, coercion and deception

Once we focus on the limited autonomy of actual patients it becomes clear that consent to all aspects and descriptions of proposed treatment is neither possible nor required. Only the ideally, unrestrictedly autonomous could offer such consent. In human contexts, whether medical or political, the most that we can ask for is consent to the more fundamental aspects of proposed policies, practices and actions. Patients can no more be asked to consent to every aspect of treatment than citizens can be asked to consent to every act of government. Respect for autonomy requires that consent be possible to fundamental aspects of actions and proposals, but allows that consent to trivial and ancillary aspects of action and proposals may be absent or impossible.

Treatment undertaken without consent when a patient could have reached his or her own decisions if approached with care and respect may fail in many ways. In the most serious cases the action undertaken uses patients as tools or instruments. Here the problem is not just that some partially autonomous patient couldn’t (or didn’t) consent, but that the treatment precluded consent even for ideally autonomous patients. Where a medical proposal hinges fundamentally on coercion or deception, not even the most rational and independent can dissent, or consent. Deceivers don’t reveal their fundamental proposal or action; coercers may make their proposal plain enough but rob anyone of choice between consent or dissent. In deception ‘consent’ is spurious because cognitive conditions for consent are not met: in coercion ‘consent’ is spurious because volitional conditions for consent are not met.

However, some non-fundamental aspects of treatment to which consent has been given may have to include elements of deception or coercion. Use of placebos or of reassuring but inaccurate accounts of expected pain might sometimes be non-fundamental but indispensable and so permissible deceptions (8). Restraint of a patient during a painful procedure might be a non-fundamental but indispensable and so permissible coercion. But using patients as unwitting experimental subjects or concealing fundamental aspects of their illness or prognosis or treatment from them, or imposing medical treatment and ignoring or preventing its refusal, would always use patients, and so fail to respect autonomy. At best such imposed treatment might, if benevolent, constitute impermissible paternalism; at worst, if non-benevolent, it might constitute assault or torture.

7. Partial autonomy, manipulation and paternalism

Use of patients is an extreme failure to respect autonomy; it precludes the consent even of the ideally autonomous, let alone of those with cognitive or volitional impairments. Respect for partial autonomy would also require medical practice to avoid treatment which, though refusible by the ideally autonomous, would not be refusible by a particular patient in his or her present condition. Various forms of manipulation and of questionable paternalism fail to meet these requirements. Patients are manipulated if they are ‘made offers they cannot refuse’, given their actual cognitive and volitional capacities. For example, patients who think they may be denied further care or discharged without recourse if they refuse proposed treatment may be unable to refuse it. To ensure that ‘consent’ is not manipulated available alternatives may have to be spelled out and refusal of treatment shown to be a genuine option. ‘Consent’ which is achieved by relying on misleading or alarmist descriptions of prognosis or uninformative accounts of treatment and alternatives does not show genuine respect. Only patients who are quite unable to understand or decide need complete paternalist protection. When there is a relationship of unequal power, knowledge or dependence, as so often between patients and doctors, avoiding manipulation and unacceptable paternalism demands a lot.

Avoiding unacceptable paternalism demands similar care. Manipulators use knowledge of others and their weaknesses to impose their own goals; paternalists may not recognise either others’ goals, or that they are others’ goals. Patients, like anyone with limited understanding and capacity to choose, may be helped by advice and information, and may need help to achieve their aims. But if it is not the patients’ but others’ aims which determine the limits and goals of medical intervention, the intervention (even if neither deceptive nor coercive) will be unacceptably paternalistic. Handicapped patients whose ways of life are determined by others may not be deceived or coerced – but they may be unable to refuse what others think appropriate for them. This means that patients’ own goals, medical and non-medical, and their plans for achieving these, are constraints on any medical practice which respects patients’ autonomy. Since return to health is often central to patients’ plans, this constraint may require little divergence from the treatment that paternalistic medical practice would select, except that patients would have to be party to fundamental features of their treatment. But where patients’ goals differ from doctors’ goals – perhaps they value quality of life or avoiding pain or dependence more than the doctor would – respect for the patient
requires that these goals not be overridden or replaced by ones the patient does not share, and that the patient’s own part in achieving them not be set aside.

Debates on medical paternalism often assume that the goals of medical action can be determined independently of patients’ goals. But in action-oriented ethical thinking morally required goals are not given independently of agents’ goals. Paternalism in this perspective is simply the imposition of others’ goals, (perhaps those of doctors, nursing homes or relatives) on patients. These goals too must be taken into account if we are to respect the autonomy of doctors, nursing homes and relatives. But imposing their goals on patients capable of some autonomy does not respect patients. The contextually-sensitive, action-oriented framework discussed here does not reinstate a contractual or consumer-sovereignty picture of medical practice, in which avoiding deceit and coercion is all that respect requires. On the contrary, it insists that judgements of human autonomy must be contextual, and that what it takes to respect human autonomy will vary with context. When patients’ partial autonomy constrains medical practice, respect for patients may demand action which avoids not only deceit and coercion but also manipulation and paternalism; but where autonomy is absent there is no requirement that it be respected.

8. Respecting limited autonomy

Medical paternalism has been considered within three frameworks. Within a result-oriented framework of the standard utilitarian type it is not only permissible but required that concern for human autonomy be subordinated to concern for total welfare. Within an action-oriented framework that relies on an abstract, ‘idealising’ account of human autonomy, medical practice is too readily construed as ruling out all paternalism and permitting only treatment that would be consented to by ‘idealised’ autonomous agents. Within an action-oriented framework that takes account of the partial character of human autonomy we can sketch patterns of reasoning which draw boundaries in given contexts between permissible and impermissible forms of paternalism. This account yields no formula, such as the requirement to avoid coercion and deception may be thought to yield for abstract approaches. But the inadequacies of that formula for guiding action when impairment is severe speak in favour of a more accurate and contextual view of human autonomy.

By trying to incorporate concern for actual, partial capacities for autonomous action into an account of respect for patients and medical paternalism we find that we are left without a single boundary-line between acceptable and unacceptable medical practice. What we have are patterns of reasoning which yield different answers for different patients and for different proposals for treatment. One patient can indeed be expected to come to an informed and autonomous (if idiosyncratic) decision; another may be too confused to take in what his options are. A third may be able to understand the issues but too dependent or too distraught to make decisions. Attempts to provide uniform guidelines for treating patients as persons respecting their autonomy and avoiding unacceptable medical paternalism are bound to be insensitive to the radical differences of capacity of different patients.

A theory of respect for patients must rely heavily and crucially on actual medical judgements to assess patient’s current capacities to absorb and act on information given in various ways. But it does not follow that ‘professional judgement’ or ‘current medical standards’ alone can provide appropriate criteria for treating patients as persons. For if these do not take the varying ways in which patients can exercise autonomy as constraints on permissible treatment, they may institutionalise unjustifiable paternalism. Professional judgement determines what constitutes respect for patients only when guided by concern to communicate effectively what patients can understand and to respect the decisions that they can make.

9. Issues and contexts

Sections 1, 2 and 3 above discussed some ways in which treatment of autonomy, paternalism and respect for patients are articulated in result-oriented ethics and action-oriented approaches which take an abstract view of cognition and volition, and hence of autonomy. The alternative account proposed in sections 4 to 7 that only consideration of the determinate cognitive and volitional capacities and incapacities of particular patients at particular times provides a framework for working out boundaries of permissible medical paternalism. If such judgements are contextual, there is no way to demarcate unacceptable paternalism in the abstract. The following headings only point to contexts in which these issues arise and have to be resolved.

Which resolutions are justifiable will depend not only on following a certain pattern of reasoning but on the capacities for autonomous action particular patients have at the relevant time.

A. TEMPORARILY IMPAIRED CAPACITY FOR AUTONOMY

If respect for autonomy is morally fundamental, restoring (some) capacities is morally fundamental. Survival is necessary for such restoration; but not sufficient. If patients’ autonomy constrains practice, survival can never be foregone in favour of autonomy, but it is an open question whether survival with no or greatly reduced capacities for autonomy can be a permissible goal. Risky surgery may sometimes reasonably be imposed for the sake of restoring capacities, even when mere survival would be sure without surgery.

Temporary loss of autonomy offers grounds for paternalistic intervention to restore autonomy – but not for all paternalistic interventions. It might be better for an unconscious sportsman if advantage were
taken of his temporary incapacity to perform some non-urgent operation or to make some non-medical intervention in his affairs. But if restoration of autonomy is likely, an action-oriented ethic offers no ground for such paternalism.

B. LONG TERM OR PERMANENT IMPAIRMENT OF AUTONOMY

This is the standard situation of children, and so the original context of paternalism. Those with long and debilitating illnesses, physical as well as mental, may suffer very varied impairments of autonomy. Hence consideration of parental paternalism may illuminate these cases. While the law has to fix an age to end minority, parents have to adapt their action to a constantly altering set of capacities for autonomous action. Choices which cannot be made at one stage can at another; autonomy develops in one area of life and lags in another (9). Unfortunately, medical trajectories may not be towards fuller capacities. Medical and other decisions may then have to be to some extent imposed. But there is no general reason to think that those who are unable to make some decisions are unable to make any decisions, and even when full return of capacities is unlikely, patients, like children, may gain in autonomy when an optimistic view is taken of their capacities.

C. PERMANENT LOSS OF AUTONOMY

Here decisions have (eventually) to be made that go beyond what is needed for restoring (some) autonomy. Sometimes medical staff and relatives may be able to make some use of a notion of hypothetical consent. But what they are likely to be asking is not ‘What would the ideally autonomous choose in this situation?’, but rather ‘What would this patient have chosen in this situation?’ If this can be answered, it may be possible to maintain elements of respect for the particular patient as he or she was in former times. But usually this provides only vague indications for medical or other treatment, and respect for absent autonomy can be at best vestigial.

D. LIFELONG INCAPACITY FOR AUTONOMY

For those who never had or will have even slight capacities for autonomous action the notion of respect is vacuous. There is no answer to the hypothetical question ‘What would he or she have chosen if able to do so?’ and the hypothetical question ‘What would the ideally autonomous choose in this situation?’ may have no determinate answer. Here, unavoidably, paternalism must govern medical practice indefinitely and the main questions that arise concern the appropriate division of authority to make paternalistic decisions between relatives and medical staff and legally appointed guardians.

References and notes


(4) Here US and British practice differ. US legislation and debates often stress the need to secure informed consent from patients (or their guardians). Cf. discussions and bibliography in Veatch R M. Case studies in medical ethics. Cambridge Mass: Harvard University Press, 1977. British law holds that ‘what information should be disclosed, and how and when, is very much a matter of professional judgement’, and that ‘there is no ground in English law for extending the limited doctrine of informed consent outside the field of property rights’. See Sidaway v Board of Governors of the Bethlem Royal Hospital and the Maudsley Hospital and Others, Law Report, The Times, 1984 Feb 24. However, medical paternalism may be more practised in the US than it is praised by those who write on medical ethics. See Buchanan A E Medical paternalism. Philosophy and public affairs 1978; 7: 371–390, and reprinted in Sartorius, see reference (2).

(5) For further comments on the limitations of ‘normal’ abilities see Wilker D. Paternalism and the mildly retarded. Reprinted in Sartorius, see reference (2).


(8) Bok S. Lying: moral choice in public and private life. New York: Harvester Press, Random House, 1978: 234. Bok points out that sometimes the use of placebos may be more than ancillary, (61–68), and also discusses fundamental forms of deception such as hiding from the patient that the illness is terminal. On the latter point see also KUBLER-ROSS E. On death and dying. New York: Macmillan 1969, and the bibliography in Veatch, reference (4).