

DRAFT
for the Bioethics Commission
Daniel C. Dennett
August 16, 2006
ca 7000 words

How to Protect Human Dignity from Science

Many people fear that science and technology are encroaching on domains of life in a way that undermines human dignity, and they see this as a threat that needs to be resisted vigorously. They are right. There is a real crisis, and it needs our attention now, before irreparable damage is done to the fragile environment of mutually shared beliefs and attitudes on which a precious conception of human dignity does indeed depend for its existence. I will try to show that the problem is real, and that the most widely favored responses to the problem are deeply misguided and bound to fail. There is a solution that has a good chance of success, however, and it employs principles that we already understand and accept in less momentous roles. The solution is natural, reasonable, robust instead of fragile, and it does not require us to try to put the genie of science back in the bottle—a good thing, since that is almost certainly impossible. Science and technology can flourish open-endedly while abiding by restrictive principles that are powerful enough to reassure the anxious and mild enough to secure the unqualified endorsement of all but the most reckless investigators. We can have dignity and science too, but only if we face the conflict with open minds and a sense of common cause.

1. The Problem

Human life, tradition says, is infinitely valuable, and even sacred: not to be tampered with, not to be subjected to “unnatural” procedures, and of course not to be terminated deliberately, except (perhaps) in special cases such as capital punishment or in the waging of a just war. Thou shalt not kill. Human life, science says, is a complex phenomenon admitting of countless degrees and variations, not markedly different from animal life or plant life or bacterial life in most regards, and amenable to countless varieties of extensions, redirections, divisions, and terminations. The questions of when (human) life begins and ends, and of which possible variants “count” as (sacred) human lives in the first place are, according to science, more like the question of the area of a mountain than its altitude above sea level: it all depends on what can only be conventional definitions of the boundary conditions. Science promises—or threatens—to replace the traditional absolutes about the conditions of human life with a host of relativistic complications and the denial of any sharp boundaries on which to hang tradition.

Plato spoke of seeking the universals that “carve Nature at its joints” (*Phaedrus* 265d-266a) and science has given us wonderful taxonomies that do just that. It has identified electrons and protons (which have the mass of 1836 electrons and a positive charge), distinguished the chemical elements from each other, and articulated and largely confirmed a Tree of Life that shows why “creature with a backbone” carves Nature better than “creature with wings.” But the crisp, logical boundaries that science gives us don’t include any joints where tradition demands them. In particular, there is no moment of *ensoulment* to be discovered in the breathtakingly complicated processes that ensue after sperm meets egg and begin producing an embryo (or maybe twins or triplets—when do *they* get their individual souls?), and there is no moment at which the soul leaves the body and human life ends. Moreover, the more we understand, scientifically, about these complexities, the more practical it becomes, technologically, to exploit them in entirely novel ways for which tradition is utterly unprepared: *in vitro* fertilization and cloning, organ harvest and transplant, and, at the end of life, the artificial prolongation of life—of one sort or another—after most if not all the *sacred* aspects of life have ceased. When we start treating living bodies as mother-boards on which to assemble cyborgs, or as spare parts collections to be sold to the highest bidder, where will it all end? It is not as if we could halt the slide by just prohibiting (some of) the technology. Technology may provide the *faits accomplis* that demonstrate beyond all controversy that the science is on the right track, but long before the technology is available, science provides the huge changes in conceptualization, the new vistas on possibility, that will flavor our imaginations henceforth whether or not the possibilities become practical. We are entering a new conceptual world, thanks to science, and it does not harmonize comfortably with our traditional conceptions of our lives and what they mean.¹

In particular, those who fear this swiftly growing scientific vista think that it will destroy something precious and irreplaceable in our traditional scheme, subverting the last presumptions of human specialness which ground—they believe—our world of morality. Oddly enough, not much attention has been paid to the question of exactly how the rise of the scientific vista would subvert these cherished principles—in this regard, it is a close kin to the widespread belief that homosexual marriage would somehow subvert traditional “family values”—but in fact there is a good explanation for this gap in the analysis. The psychologist Philip Tetlock (1999, 2003, 2004) identifies values as *sacred* when they are so important to those who hold them that the very act of considering them is offensive. The comedian Jack Benny was famously stingy—or so he presented himself on radio and television—and one of his best bits was the skit in which a mugger puts a gun in his back and barks “Your money or your life!” Benny just stands there silently. “Your money or your life!” repeats the mugger, with mounting impatience. “I’m thinking, I’m thinking,” Benny replies. This is funny because most of us think that nobody should even think about such a trade-off. Nobody should *have to* think about such a trade-off. It should be unthinkable, a “no-brainer.” Life is sacred, and no amount of money would be a fair exchange for a life, *and if you don’t already know that, what’s wrong with you?* “To transgress this

boundary, to attach a monetary value, to one's friendships, children, or loyalty to one's country, is to disqualify oneself from the accompanying social roles." (Tetlock et al, 2004, p5) That is what makes life a sacred value.

Tetlock and his colleagues have conducted ingenious (and sometimes troubling) experiments in which subjects are obliged to consider "taboo trade-offs," such as whether or not to purchase live human body parts for some worthy end, or whether or not to pay somebody to have a baby that you then raise, or pay somebody to perform your military service. As their model predicts, many subjects exhibit a strong "mere contemplation effect": they feel guilty and sometimes get angry about being lured into even thinking about such dire choices, even when they make all the right choices. When given the opportunity by the experimenters to engage in "moral cleansing" (by volunteering for some relevant community service, for instance) subjects who have had to think about taboo trade-offs are significantly more likely than control subjects to volunteer—for real—for such good deeds. (Control subjects had been asked to think about purely non-sacred trade-offs, such as whether to hire a house-cleaner or buy food instead of something else.)²

So it is not surprising that relatively little attention has been paid to charting the paths by which science and technology might subvert the value of life. If you feel the force of the admonition "Don't even think about it!" you will shun the topic by distracting your own attention from it, if at all possible. I know from experience that some readers of this essay will already be feeling some discomfort and even guilt for allowing themselves to broach these topics at all, so strong is the taboo against thinking the unthinkable, but I urge them to bear with me, since the policy that I will propose may have more going for it than their own.

The fact that the threat has not been well articulated does not mean it is not real and important. Let me try to make it plain by drawing some parallels. Like climate change, the threat is environmental and *global* (which means you can't just move to a different place where the environment hasn't yet been damaged—and time is running out). While global warming threatens to affect many aspects of the *physical* environment—the atmosphere, the flora and fauna, the ice caps and ocean levels—and hence alter our geography in catastrophic ways from which recovery may be difficult or impossible, the threat to human dignity affects many aspects of what we may call the *belief environment*, the manifold of ambient attitudes, presumptions, common expectations—the things that are "taken for granted" by just about everybody—and that just about everybody expects just about everybody to take for granted.

The belief environment plays just as potent a role in human welfare as the physical environment, and in some regards it is both more important and more fragile. Much of this has been well known for centuries, particularly to economists, who have long appreciated the way a currency can become worthless almost overnight, for example, and the way public trust in

financial institutions needs to be preserved as a condition for economic activity in general. Today we confront the appalling societal black holes known as failed states, where the breakdown of law and order makes the restoration of decent life all but impossible. (If you have to pay off the warlords and bribe the judges and tolerate the drug traffic . . . just to keep enough power and water and sanitation going to make life bearable, let alone permit agriculture and commerce to thrive, your chances of long-term success are minimal.) What matters in these terrible conditions is what people in general assume *whether they are right or wrong*. It might in fact be safe for them to venture out and go shopping, or to invest in a clothing factory, or plant their crops, but if they don't, in general, believe that, they cannot resume anything like normal life and rekindle a working society. This creates a belief environment in which there is a powerful incentive for the most virtuous and civic-minded to lie, vigorously, just to preserve what remains of the belief environment. Faced with a deteriorating situation, admitting the truth may only accelerate the decline, while a little creative myth-making might—might—save the day. Not a happy situation.

And this is what people fear might happen if we pursue our current scientific and technological exploration of the boundaries of human life: we will soon find ourselves in a deteriorating situation where people—rightly or wrongly—start jumping to conclusions about the *non*-sanctity of life, the commodification of all aspects of life, and it will be too late to salvage the prevailing attitudes that protect us all from something rather like a failed state, a society in which the sheer security needed for normal interpersonal relations has dissolved, making trust, and respect, and even love, all but impossible. Faced with that dire prospect, it becomes tempting indeed to think of promulgating a holy lie, a myth that might carry us along for long enough to shore up our flagging confidence until we can restore “law and order.”

That is where the doctrine of the soul comes in. People have immortal souls, according to tradition, and that is what makes them so special. Let me put the problem unequivocally: the traditional concept of the soul as an immaterial thinking thing, Descartes's *res cogitans*, the internal locus in each human body of all suffering, and meaning, and decisions, both moral and immoral, has been utterly discredited. Science has banished the soul as firmly as it has banished mermaids, unicorns, and perpetual motion machines. There are no such things. There is no more scientific justification for believing in an immaterial immortal soul than there is for believing that each of your kidneys has a tap-dancing poltergeist living in it. The latter idea is clearly preposterous. Why are we so reluctant to dismiss the former idea? It is obvious that there must be some *non*-scientific motivation for believing in it. It is seen as being needed to play a crucial role in preserving our self-image, our dignity. If we don't have souls, we are *just animals!* (And how could you love, or respect, or grant responsibility to something that was just an animal?)

Doesn't the very meaning of our lives depend on the reality of our immaterial souls? No. We don't need to be made of two fundamentally different kinds of substance, matter and mind-

stuff, to have morally meaningful lives. On the face of it, the idea that all our striving and loving, our yearning and regretting, our hopes and fears, depend on some secret ingredient, some science-proof nugget of specialness that defies the laws of nature, is an almost childish ploy: “Let’s gather up all the wonderfulness of human life and sweep it into the special hidey-hole where science can never get at it!” Although this fortress mentality has a certain medieval charm, looked at in the cold light of day, this idea is transparently desperate, implausible, and risky: putting all your eggs in one basket, and a remarkably vulnerable basket at that. It is vulnerable because it must declare science to be unable to shed any light on the various aspects of human consciousness and human morality at a time when exciting progress is being made on these very issues. One of Aristotle’s few major mistakes was declaring “the heavens” to be made of a different kind of stuff, entirely unlike the matter here on Earth—a tactical error whose brittleness became obvious once Galileo and company began their still-expanding campaign to understand the physics of the cosmos. Clinging similarly to an immaterial concept of a soul at a time when every day brings more understanding of how the material basis of the mind has evolved (and goes on evolving within each brain) is a likely path to obsolescence and extinction.

The alternative is to look to the life sciences for an understanding of what does in fact make us different from other animals, in morally relevant ways. We are the only species with language, and art, and music, and religion, and humor, and the ability to imagine the time before our birth and after our death, and the ability to plan projects that take centuries to unfold, and the ability to create, defend, revise, and live by, codes of conduct, and—sad to say—to wage war on a global scale. The ability of our brains to help us see into the future, thanks to the culture we impart to our young, so far surpasses that of any other species, that it gives us the powers that in turn gives us the responsibilities of moral agents. *Noblesse oblige*. We are the only species that can know enough about the world to be reasonably held responsible for protecting its precious treasures. And who on earth could hold us responsible? Only ourselves. Some other species—the dolphins and the other great apes—exhibit fascinating signs of proto-morality, a capacity to cooperate and to care about others, but we persons are the only animals that can conceive of *the project of leading a good life*. This is *not* a mysterious talent; it can be explained.³

Here I will not attempt to survey the many threads of that still unfolding explanation, but rather to construct and defend a perspective and a set of policies that could protect what needs to be protected as we scramble, with many false steps, towards an appreciation of the foundations of human dignity. Scientists make their mistakes in public, but mostly only other scientists notice them. This topic has such momentous consequences, however, that we can anticipate that public attention—and reaction—will be intense, and could engender runaway misconstruals that could do serious harm to the delicate belief environment in which we (almost) all would like to live.

I have mentioned the analogy with the ominous slide into a failed state; here is a less dire example of the importance of the belief environment, and the way small changes in society can

engender unwanted changes in it. In many parts of rural America people feel comfortable leaving their cars and homes unlocked, day and night, but any country mouse who tries to live this way in the big city soon learns how foolish that amiably trusting policy is. City life is not intolerable, but it is certainly different. Wouldn't it be fine if we could somehow re-engineer the belief environment of cities so that people seldom felt the need to lock up! An all but impossible dream. At the same time, rural America is far from utopia, and is sliding towards urbanity. The felicitous folkways of the countryside can absorb a modest amount of theft and trespass without collapse, but it wouldn't take much to extinguish them forever. Those of us who get to live in this blissfully secure world cherish it, for good reason, and would hate to abandon it, but we also must recognize that any day could be the last day of unlocked doors in our neighborhood, and once the change happened, it would be very hard to change back. That too, is like global climate change; these changes are apt to be irreversible. And unlike global climate change, drawing attention to the prospect may actually hasten it, by kindling and spreading what Douglas Hofstadter once called "reverberant doubt" (Hofstadter, 1983, 1985). The day that our local newspaper begins running a series about what percentage of local people lock their doors under what circumstances is the day that door-locking is apt to become the norm. So those who are in favor of diverting attention from too exhaustive an examination of these delicate topics might have the right idea. This is the chief reason, I think, for the taboo against thinking about sacred values: it can sometimes jeopardize their protected status. But in this case, I think it is already too late to follow the tip-toe approach. There is already a tidal wave of interest in the ways in which the life sciences are illuminating the nature of "the soul" so we had better shift from distraction to concentration, and see what we can make of the belief environment for human dignity and its vulnerabilities.

3. The Solution

How are we to protect the ideal of human dignity from the various incursions of science and technology? The first step in the solution is to notice that the *grounds* for our practices regarding this are not going to be *local* features of particular human lives, but rather more *distributed* in space and time. There is already a clear precedent in our attitude towards human corpses. Even people who believe in immortal immaterial souls don't believe that human "remains" harbor a soul. They think that the soul has departed, and what is left behind is just a body, just unfeeling matter. A corpse can't feel pain, can't suffer, can't be aware of any indignities—and yet still we feel a powerful obligation to handle a corpse with respect, and even with ceremony, and even when nobody else is watching. Why? Because we appreciate, whether acutely or dimly, that how we handle *this* corpse *now* has repercussions for how other people, still alive, will be able to imagine their own demise and its aftermath. Our capacity to imagine the future is both the source of our moral power and a condition of our vulnerability. We cannot help but see all the events in our lives against the backdrop of what Hofstadter (1982, 1985) calls the

implicosphere of readily imaginable alternatives—and the great amplifier of human suffering (and human joy) is our irresistible tendency to anticipate, with dread or delight, what is in store for us.

We live not just in the moment, but in the past and the future as well. Consider the well-known advice given to golfers: *keep your head down* through the whole swing. “Wait a minute,” comes the objection: “that’s got to be voodoo superstition! Once the ball leaves the club head, the position of my head couldn’t possibly affect the trajectory of the ball. This has to be scientifically unsound advice!” Not at all. Since we plan and execute all our actions in an anticipatory belief environment, and have only limited and indirect control over our time-pressured skeletal actions, it can well be the case that the only way to get the part of the golf swing that *does* affect the trajectory of the ball to have the desirable properties is to concentrate on making the later part of it, which indeed could not affect the trajectory, to take on a certain shape. Far from being superstitious, the advice can be seen to follow quite logically from facts we can discover from a careful analysis of the way our nervous systems guides our muscles.

Our respect for corpses provides us with a clear case of a wise practice that does not at all depend on finding, locally, a special (even supernatural) ingredient that justifies or demands this treatment. There are other examples that have the same feature. Nobody has to endorse magical thinking about the gold in Fort Knox to recognize the effect of its (believed in) presence there on the stability of currencies. Symbols play an important role in helping to maintain social equilibria, and we tamper with them at our peril. If we began to adopt the “efficient” policy of disposing of human corpses by putting them in large biodegradable plastic bags to be taken to the landfill along with the rest of the “garbage,” this would flavor our imaginations in ways that would be hard to ignore, and hard to tolerate. No doubt we could get used to it, the same way city folk get used to locking their doors, but we have good reasons for avoiding that path. (Medical schools have learned to be diligent in their maintenance of respect and decorum in the handling of bodies in their teaching and research, for while those who decide to donate their bodies to medicine presumably have come to terms with the imagined prospect of students dissecting and discussing their innards, they have limits on what they find tolerable.)

The same policy and rationale applies to end-of-life decisions. We handle a corpse with decorum even though we *know* it cannot suffer, so we can appreciate the wisdom of extending the same practice to cases where we don’t know. For instance, a person in a persistent vegetative state might be suffering, or might not, but in either case, we have plenty of grounds for adopting a policy that creates a comforting buffer zone that errs on the side of concern. And once again, the long-range effect on community beliefs is just as important as, or even more important than, any locally measurable symptoms of suffering. (In a similar spirit, it is important that wolves and grizzly bears still survive in the wilder regions of our world even if we almost never see them. Just knowing that they are there is a source of wonder and delight, and makes the world a better place. Given our invincible curiosity and penchant for skepticism, we have to keep checking up

on their continued existence, of course, and could not countenance an official myth of their continued presence if they had in fact gone extinct. This too has its implications for our topic.)

What happens when we apply the same principle to the other boundary of human life, its inception? The scientific fact is that there is no good candidate, and there will almost certainly never be a good candidate, for a moment of *ensoulment*, when a mere bundle of living human tissue becomes a person with all the rights and privileges pertaining thereunto. This should not be seen as a sign of the weakness of scientific insight, but rather as a familiar implication of what science has already discovered. One of the fascinating facts about living things is the way they thrive on gradualism. Consider speciation: there are uncounted millions of different species, and each of them had its inception “at some point” in the nearly four billion year history of life on this planet, but there is literally no telling *exactly* when any species came into existence because what counts as speciation is something that only gradually and cumulatively emerges over very many generations. Speciation can only emerge *in the aftermath*. Consider dogs, the millions of members of hundreds of varieties of *Canis familiaris* that populate the world today. As different as these varieties are—think of St. Bernards and Pekinese—they all count as a single species, cross-fertile (with a little mechanical help from their human caretakers) and all readily identifiable as belonging to the same species, descended from wolves, by their highly similar DNA. Might one or more of these varieties or subspecies become a species of its own some day? Absolutely. In fact, every puppy born is a potential founder of a new species, but nothing about that puppy on the day of its birth (or for that matter on any day of its life) could be singled out as the special feature that marked it as the Adam or Eve of a new species. If it dies without issue, it definitely won’t found a new species, but as long as it has offspring that have offspring . . . it might turn out, in the fullness of time, to be a good candidate for the first member of a new species.

Or consider our own species, *Homo sapiens*. Might it divide in two some day? Yes it might, and in fact, it might, in a certain sense, already have happened. Consider two human groups alive today that probably haven’t had any common ancestors in the last thirty thousand years: the Inuit of Cornwallis Island in the Arctic, and the Andaman Islanders living in remarkable isolation in the Indian Ocean. Suppose some global plague sweeps the planet sometime in the next hundred years (far from an impossibility, sad to say), leaving behind only these two small populations. Suppose that over the next five hundred or a thousand years, say, they flourish and come to reinhabit the parts of the world vacated by us—and discover that they are not cross-fertile with the other group! Two species, remarkably similar in appearance, physiology and ancestry, but nevertheless as reproductively isolated as lions are from tigers. When, then, did the speciation occur? Before the dawn of agriculture about ten thousand years ago, or after the birth of the internet? There would be no principled way of saying. We can presume that today, Inuits and Andaman Islanders are cross-fertile, but who knows? The difference between “in principle” reproductive isolation (because of the accumulation of genetic and behavioral differences that make offspring “impossible”) and *de facto* reproductive isolation, which has been already been the case for many thousands of years, is not itself a principled

distinction.

A less striking instance of the same phenomenon of gradualism is *coming of age*, in the sense of being mature enough and well enough informed to be suitable for marriage, or—to take a particularly clear case—to drive a car. It will come as no surprise, I take it, that there is no special moment of *driver-edment*, when a teenager crisply crossed the boundary between being too immature to have the right to apply for a driver's license, and being adult enough to be allowed the freedom of the highway behind the wheel. Some youngsters are manifestly mature enough at fourteen to be reasonable candidates for a driver's license, and others are still so heedless and impulsive at eighteen that one trembles at the prospect of letting them on the road. We have settled (in most jurisdictions) on the policy that age sixteen is a suitable threshold, and what this means is that we simply refuse to consider special pleading on behalf of unusually mature younger people, and also refrain from imposing extra hurdles on those sixteen-year-olds who manage to pass their driving test fair and square in spite of our misgivings about the safety of letting them on the road. In short, we settle on a conventional threshold which we *know* does not mark any special internal mark (brain myelination, IQ, factual knowledge, onset of puberty) but strikes us as a good-enough compromise between freedom and public safety. *And once we settle on it, we stop treating the location of the threshold as a suitable subject for debate.* There are many important controversies to consider and explore, and this isn't one of them. Not as a general rule. Surprising new discoveries may in principle trigger a reconsideration at any time, but we foster a sort of inertia that puts boundary disputes out of bounds for the time being.

Why isn't there constant pressure from fifteen-year-olds to lower the legal driving age? it is not just that they tend not to be a particularly well-organized or articulate constituency. Even they can recognize that soon enough they will be sixteen, and there are better ways to spend their energy than trying to adjust a policy that is, all things considered, quite reasonable. Moreover, there are useful features of the social dynamics that make it systematically difficult for them to mount a campaign for changing the age. We adults have created a tacit scaffolding of presumption, *holding* teenagers responsible before many of them have actually achieved the requisite competence, thereby encouraging them to try to grow into the status we purport to grant them and discouraging any behavior—any action that could be interpreted as throwing a tantrum, for instance—that would undercut their claim to maturity. They are caught in a bind: the more vehemently they protest, the more they cast doubt on the wisdom of their cause. In the vast array of projects that confront them, this is not an appealing choice.

The minimum driving age is not quite a *sacred* value, then, but it shares with sacred values the interesting feature of being considered best left unexamined, by common consensus among a sizable portion of the community. And there is a readily accessible reason for this inertia. We human beings lead lives that cast long beams of anticipation into the foggy future, and we appreciate—implicitly or explicitly—almost any fixed points that can reduce our

uncertainty. Sometimes this is so obvious as to be trivial. Why save money for your children's education if money may not be worth anything in the future? How could you justify going to all the trouble of building a house if you couldn't count on the presumption that you will be able to occupy it without challenge? Law and order are preconditions for the sorts of ambitious life-planning we want to engage in. But we want more than just a strong state apparatus that can be counted on not to be vacillating in its legislation, or whimsical in enforcement. We, as a society, do need to draw some lines—"bright" lines in legalistic jargon—and stick with them. That means not just promulgating them and voting on them, but putting an unequal burden on any second-guessing, so that people can organize their life projects with the reasonable expectation that these are fixed points that aren't going to shift constantly under the pressure of one faction or another. We want there to be an ambient attitude of *mutual* recognition of the stability of the moral—not legal—presumptions that can be taken for granted, something approximating a meta-consensus among those who achieve the initial consensus about the threshold: let's leave well enough alone now that we've fixed it. In a world where every candidate for a bright line of morality is constantly under siege from partisans who would like to change it, one's confidence is shaken that one's everyday conduct is going to be above reproach. (Consider that nowadays in many parts of the world women simply cannot wear fur coats in public with the attitudes their mothers could adopt. Today, wearing a fur coat is making a political statement, and one cannot escape that by simply disavowing the intent. Driving a gas-guzzling SUV carries a similar burden. People may resent the activities of the partisans who have achieved these shifts in opinion even though they may share many of their attitudes about animal rights or energy policy; they have made investments—in all innocence, let us suppose—that now are being disvalued. Had they been able to anticipate this shift in public opinion, they could have spent their money better.)

These observations are not contentious, I think. How, though, can we apply this familiar understanding to the vexing issues surrounding the inception—and manipulation and termination—of human life, and the special status it is supposed to enjoy? By recognizing, first, that we are going to have to walk away from the *traditional* means of securing these boundaries, which are not going to keep on working. They are just too brittle for the twenty-first century. We know too much. Unlike traditional sacred values that depend on widespread acceptance of myths (which even if true, are manifestly unjustifiable—that's why we call them myths rather than common knowledge), we need to foster values that can withstand scrutiny about their own creation. That is to say, we have to become self-conscious about our reliance on such policies, without in the process destroying our faith in them.

4. Belief in belief

We need to appreciate the importance in general of the phenomenon of *belief in belief*.⁴ Consider a few cases that are potent today. Because many of us believe in democracy and

recognize that the security of democracy in the future depends critically on *maintaining the belief* in democracy, we are eager to quote (and quote and quote) Winston Churchill's famous line: "Democracy is the worst form of government except for all the others that have been tried." As stewards of democracy, we are often conflicted, eager to point to flaws that ought to be repaired, while just as eager to reassure people that the flaws are not that bad, that democracy can police itself, so their faith in it is not misplaced.

The same point can be made about science. Since the belief in the integrity of scientific procedures is almost as important as the actual integrity, there is always a tension between a whistle-blower and the authorities, even when they know that they have mistakenly conferred scientific respectability on a fraudulently obtained result. Should they quietly reject the offending work and discreetly dismiss the perpetrator, or make a big stink?⁵

And certainly some of the intense public fascination with celebrity trials is to be explained by the fact that belief in the rule of law is considered to be a vital ingredient in our society, so if famous people are seen to be above the law, this jeopardizes the general trust in the rule of law. Hence we are not just interested in the trial, but in the public reactions to the trial, and the reactions to those reactions, creating a spiraling inflation of media coverage. We who live in democracies have become somewhat obsessed with gauging public opinion on all manner of topics, and for good reason: in a democracy it really matters what the people believe. If the public cannot be mobilized into extended periods of outrage by reports of corruption, or the torturing of prisoners by our agents, for instance, our democratic checks and balances are in jeopardy. In his hopeful book, *Development as Freedom* (1999) and elsewhere (see especially Sen 2003), the Nobel laureate economist Amartya Sen makes the important point that you don't have to win an election to achieve your political aims. Even in shaky democracies, what the leaders believe about the beliefs that prevail in their countries influences what they take their realistic options to be, so belief-maintenance is an important political goal in its own right.

Even more important than political beliefs, in the eyes of many, are what we might call metaphysical beliefs. Nihilism—the belief in nothing—has been seen by many to be a deeply dangerous virus, for obvious reasons. When Friedrich Nietzsche hit upon his idea of the Eternal Recurrence—he thought he had proved that we relive our lives infinitely many times—his first inclination (according to some stories) was that he should kill himself without revealing the proof, in order to spare others from this life-destroying belief.⁶ Belief in the *belief that something matters* is understandably strong and widespread. Belief in free will is another vigorously protected vision, for the same reasons, and those whose investigations seem to others to jeopardize it are sometimes deliberately misrepresented in order to discredit what is seen as a dangerous trend (Dennett, 2003). The physicist Paul Davies (2004) has recently defended the view that belief in free will is so important that it may be “a fiction worth maintaining.” It is

interesting that he doesn't seem to think that his own discovery of the awful truth (what he takes to be the awful truth) incapacitates him morally, but that others, more fragile than he, will need to be protected from it.

This illustrates the ever-present risk of paternalism when belief in belief encounters a threat: we must keep these facts from “the children,” who cannot be expected to deal with them safely. And so people often become systematically disingenuous when defending a value. Being the unwitting or uncaring bearer of good news or bad news is one thing; being the self-appointed champion of an idea is something quite different. Once people start committing themselves (in public, or just in their “hearts”) to particular ideas, a strange dynamic process is brought into being, in which the original commitment gets buried in pearly layers of defensive reaction and meta-reaction. “Personal rules are a *recursive* mechanism; they continually take their own pulse, and if they feel it falter, that very fact will cause further faltering,” the psychiatrist George Ainslie observes, in his remarkable book, *Breakdown of Will* (2001, p88). He describes the dynamic of these processes in terms of competing strategic commitments that can contest for control in an organization—or an individual. Once you start living by a set of explicit rules, the stakes are raised: when you lapse, what should you do? Punish yourself? Forgive yourself? Pretend you didn't notice?

After a lapse, the long range interest is in the awkward position of a country that has threatened to go to war in a particular circumstance that has then occurred. The country wants to avoid war without destroying the credibility of its threat, and may therefore look for ways to be seen as not having detected the circumstance. Your long-range interest will suffer if you catch yourself ignoring a lapse, but perhaps not if you can arrange to ignore it without catching yourself. This arrangement, too, must go undetected, which means that a successful process of ignoring must be among the many mental expedients that arise by trial and error—the ones you keep simply because they make you feel better without your realizing why. (p150)

This idea that there are myths we live by, myths that must not be disturbed at any cost, is always in conflict with our ideal of truth-seeking and truth-telling, sometimes with lamentable results. For example, racism is at long last widely recognized as a great social evil, so many reflective people have come to endorse the second-order belief that *belief in the equality of all people regardless of their race* is to be vigorously fostered. How vigorously? Here people of good will differ sharply. Some believe that belief in racial differences is so pernicious that *even when it is true* it is to be squelched. This has led to some truly unfortunate excesses. For instance, there are clear clinical data about how people of different ethnicity are differently susceptible to disease, or respond differently to various drugs, but such data are considered off limits by some researchers, and some funders of research. This has the perverse effect that strongly indicated avenues of research are deliberately avoided, much to the detriment of the health of the ethnic

groups involved.⁷

Ainslie uncovers strategic belief-maintenance in a wide variety of cherished human practices:

Activities that are spoiled by counting them, or counting on them, have to be undertaken through indirection if they are to stay valuable. For instance, romance undertaken for sex or even ‘to be loved’ is thought of as crass, as are some of the most lucrative professions if undertaken for money, or performance art if done for effect. Too great an awareness of the motivational contingencies for sex, affection, money, or applause spoils the effort, and not only because it undeceives the other people involved. Beliefs about the intrinsic worth of these activities are valued beyond whatever accuracy these beliefs might have, because they promote the needed indirection.” (2005, p649) .

So what sort of equilibrium can we reach? If we want to maintain the momentousness of all decisions about life and death, and take the steps that elevate the decision beyond the practicalities of the moment, we need to secure the appreciation of this very fact, and enliven the imaginations of people so that they can recognize, and avoid wherever possible, and condemn, activities that would tend to erode the public trust in the presuppositions about what is—and should be—unthinkable. A striking instance of failure to appreciate this is the proposal by President Bush to reconsider and unilaterally refine the Geneva Convention’s deliberately vague characterization of torture as “outrages on personal dignity.” By declaring that the United States is eager to be a pioneer in the adjustment of what has heretofore been mutually agreed to be unthinkable, this policy is deeply subversive of international trust, and of national integrity. We as a nation can no longer be plausibly viewed as *above* thinking of arguable exceptions to the sacred value of not torturing people, and this diminishes us in ways that will be difficult if not impossible to repair.

What forces can we hope to direct in our desire to preserve respect for human dignity? Laws *prohibit*; traditions *encourage* and *discourage*, and in the long run, laws are powerless to hold the line unless they are supported by a tradition, by the mutual recognition of most of the people that they preserve conditions that deserve preservation. Global opinion, as we have just seen, cannot be counted on to discourage all acts of degradation of the belief environment, but it can be enhanced by more local traditions. Doctors, for instance, have their proprietary code of ethics, and most of them rightly covet the continuing respect of their colleagues, a motivation intensified by the systems of legal liability and the insurance that has become a prerequisite for practice. Then there are strict liability laws, which target particularly sensitive occupations such as pharmacist and doctor, pre-emptively removing the excuse of ignorance and thereby putting all who occupy these positions on notice that they will be held accountable whether or not they have what otherwise would be a reasonable claim of innocent ignorance. So forewarned, they adjust their standards and projects accordingly, erring on the side of extreme caution and keeping a

healthy distance between themselves and legal consequences. Anyone who attempts to erect such a network of flexible and mutually supporting discouragements of further tampering with traditional ideas about human dignity will fail unless they attend to the carrot as well as the stick. How can we kindle and preserve a sincere *allegiance* to the ideals of human dignity? The same way we foster the love of a democratic and free society: by ensuring that the lives one can live in such a regime are so manifestly better than the available alternatives.

And what of those who are frankly impatient with tradition, and even with the values that tradition endorses? We must recognize that there are a vocal minority of people who profess unworried acceptance of an entirely practical and matter-of-fact approach to life, who scoff at romantic concerns with Frankensteinian visions. Given the presence and articulateness of these proponents, we do well to have a home base that can withstand scrutiny and that is prepared to defend, in terms other than nostalgia, the particular values that we are trying to protect. That is the germ of truth in multiculturalism. We need to articulate these values in open forum. When we attempt this, we need to resist the strong temptation to resort to the old myths, since they are increasingly incredible, and will only foster incredulity and cynicism in those we need to persuade. Tantrums in support of traditional myths will backfire, in other words. Our only chance of preserving a respectable remnant of the tradition is to ensure that the values we defend *deserve* the respect of all.⁸

References

Ainslie, George, 2001, *Breakdown of Will*, Cambridge Univ. Press.

Ainslie, G. (2005) Précis of *Breakdown of Will*. *Behavioral and Brain Sciences* 28(5).

Dennett, Daniel C., 2006, *Breaking the Spell: Religion as a Natural Phenomenon*, New York: Viking Penguin.

Hofstadter, Douglas, 1982, "Metafont, Metamathematics and Metaphysics," in *Visible Language*, August, 1982, reprinted with comments in Hofstadter, 1985, pp290, 595.

Hofstadter, Douglas, 1983, "Dilemmas for Superrational Thinkers, Leading up to a Luring Lottery," *Scientific American*, June, 1983, reprinted with a discussion of reverberant doubt in Hofstadter, 1985, pp 752-755.

Hofstadter, Douglas, 1985, *Metamagical Themas*, New York: Basic Books.

Lewontin, Richard, 2004, “Dishonesty in Science,” *New York Review of Books*, Nov 18, 2004 pp38-40.

Sen, Amartya, 1999, *Development as Freedom*, New York: Knopf.

Sen, Amartya, 2003, “Democracy and Its Global Roots,” *New Republic*, October 6, 2003, pp28-35.

Tetlock, Philip, 1999, “Coping with Trade-offs: Psychological constraints and political implications,” in S. Lupia, M. McCubbins and S. Popkin (eds.), *Political Reasoning and Choice*, Berkeley: Univ. of California Press.

Tetlock, Philip, 2003, “Thinking the unthinkable: sacred values and taboo cognitions,” *Trends in Cognitive Science*, 7, pp320-4.

Tetlock, Philip, McGraw, A. Peter, and Kristel, Orie, 2004, “Proscribed Forms of Social Cognition: Taboo Trade-Offs, Forbidden Base Rates, and Heretical Counterfactuals,” in N. Haslam, ed., *Relational Models Theory: A Contemporary Overview*. Mahway NJ: Erlbaum

¹The philosopher Wilfrid Sellars (1963) distinguished between the *manifest image* of everyday life, with its tables and chairs, trees and rainbows, people and dreams, and the *scientific image* of atoms and particles and waves electromagnetic radiation, and noted that the task of putting these two images into registration is far from straightforward. The dimensions of *meaning*, which reside solely—it seems—in the manifest image, is resistant both to *reduction* (the way chemistry, supposedly, reduces to physics) and to any less demanding sort of unification or coordination with the scientific image. The tension we are exploring here is a particularly vivid and troubling case of the tension between these two images.

²Material in the previous two paragraphs is drawn from Dennett, 2006, p22-3.

³My 2003 book, *Freedom Evolves*, is devoted to an explanation of how our capacity for moral agency evolved and continues to evolve. It begins with a quotation from a 1997 interview with Giulio Giorelli: “*Sì, abbiamo un anima. Ma è fatta di tanti piccoli robot.*—Yes, we have a soul—

but it's made of lots of tiny robots." These "robots" are the mindless swarms of neurons and other cells that cooperate to produce a thinking thing—just not an *immaterial* thinking thing, as Descartes imagined and tradition has tended to suppose.

⁴What follows is drawn, with revisions, from Dennett, 2006, chapter 8.

⁵As Richard Lewontin recently observed, "To survive, science must expose dishonesty, but every such public exposure produces cynicism about the purity and disinterestedness of the institution and provides fuel for ideological anti-rationalism. The revelation that the paradoxical Piltdown Man fossil skull was, in fact, a hoax was a great relief to perplexed paleontologists but a cause for great exultation in Texas tabernacles." (2004, p39)

⁶For a discussion of Nietzsche and his philosophical response to Darwin's theory of evolution by natural selection, see my *Darwin's Dangerous Idea* (1995).

⁷There are significant differences in breast cancer (Li et al, 2003), hypertension and diabetes, alcohol tolerance and many other well-studied conditions. For an overview see Health Sciences Policy, Board (HSP), 2003, *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*

⁸Thanks to Gary Wolf, Tori McGeer and Philip Pettit for asking questions that crystallized my thinking on these topics.