

Flashlights, Mirrors, Real Brains and Willpower: Steering Ourselves to Steer Our Future

by Neil Bates

nfb222@email.vccs.edu

Abstract: The challenges facing humanity require not just action, but better understanding and transformation of the human mind. It is more important to find out why so many have trouble with lifestyle and cooperative issues (obesity, lack of sleep, employment and economic problems, increasing controversies and tensions between groups, etc.) than it is to design ever more clever cell phones and "pads" and so forth. We examine the problem of flawed thinking as both a factor reducing current well-being and advancement, as well as being a hindrance to human improvement and the forming of better minds and responses. It is argued that mechanistic models of consciousness and choice are inadequate. Appreciating that we are more than computing machines will lead to improved modes of thinking and behaving suitable for preparing and sustaining a better future, as well as inspiring us to make the vital effort.

Note: small portions of this article are adapted from "Descartes Was Wrong" by Neil Bates, in *M-Tides, the Newsletter of Tidewater Mensa*; (February, 2014) and various unpublished manuscripts by the author.

Who, Not Just What

The question of how humanity should steer the future begins with: who are we really, and how have and how shall we steer ourselves? This is no silly self-reference loop. We face greater challenges today for several reasons. One, our ability to affect our environment or harm each other reaches to orders of magnitude beyond that of past eras. Also, it seems that polarization increased despite hopes after civil rights and independence movements, the erstwhile rise of the middle class, global interdependence, and the presumed end of the Cold War and the great dichotomy posed by energetic Communism: there is more religious and regional strife, disparity of income, and friction between those of different ideological and cultural bent. Hence it is even harder to resolve any "how" question, when philosophical tribes want to answer in such different ways, even within the same civil community. No "end of history" has arrived, although our history is at risk of no longer being a wide-ranging progression of greater well being and personal development. (It is still of course to some extent at risk of ending in the bad sense.) Our response balances on the fulcrum of the question: must we be more, not just do more, to guide ourselves into a good future?

Here I argue that it will not be enough to take ourselves for granted, "as is" (also, as we *think* we are ...) and simply ask: what should we, as we are, do? Facing and creating our future cannot just be about technology and changes to the world outside the skin. We need to better understand and express ourselves, the better to be worthy and effective beings occupying the future along with our posterity and many other species and resources and wonders. We also must consider changing ourselves and our descendants to be more capable and more compatible with each other.

Approaching and implementing such a personal and interpersonal transformation requires both the best use of interdisciplinary engagement as well as the highest ethical standards and commitment to the common good. Otherwise it could be just an arms race, a race to out-evolve other nations or to field victorious cyborgs. We must utilize the best of biology, medicine, math, cybernetics, physics, engineering; psychology and philosophy and all the other soft sciences and practical skills of accomplisners everywhere. (For a recent review of the prospects and limitations of some "soft" approaches, see Carr [1].)

Of course, we have been altering our own and others' selves for millennia. Enculturation is modification, so is body building etc. We have also learned much about the human body and mind. So, here we ask about real insight and real change: deep understanding (which should come first), and making our underlying selves different. Furthermore, those of us proposing any such efforts must realize that not everyone will agree on whether and which such changes can or should be made. It is therefore perhaps more helpful to touch on some basic and "non-partisan" (politically at least) arguments supporting the idea that 1. there is a problem with how humans think, 2. there are reasons to believe that we have the power within us to transcend our current state, in outline and avoiding dramatic and ethically controversial methods such as genetic engineering, selective breeding, neurological implants or drug alterations, and so on. I see this as much more than a dry exercise in improved brain "power" or enhancing sympathy, although both of those would be helpful. I think we need to both inspire people and empower them. First, can we discover that humans are more than machines? Are consciousness and free will real in some sense, and can we better understand our true selves and act more effectively as a result?

What's Wrong? A Big Little Story ... (or Two or ...)

As a boy, growing up during the Cold War and a hot one in Vietnam, it was particularly galling to know I and many more could die because one megatribe did not get along with another one thousands of miles away. Yes of course this needs to get better, and many people are trying. But what showed me there was something wrong with society's utilitarian practices, was the humble little encounter of trying flashlights (the older kind) and thinking about how they work. I understood the basic concept that the curved reflector behind the bulb gathered light and directed it into a beam. The beam was not perfectly straight, or it would provide too small a spot of light at distances. All sensible so far, but I was annoyed by the uneven illumination spot from most of

them. There was either a darker spot in the middle, or (sometimes a little of both) helical-style or less regular patterns of brighter light. The latter was clearly an imaging (if imperfectly) of the filament, its structure known to me by reading and magnification. Wouldn't a soft, evenly glowing spot make more sense? As I read in depth, I learned about parabolic mirrors and how they were used to focus light into a beam, and for telescopes. Considering their ability to bring parallel light to a point, the latter use made sense in particular. Also I could see the wonderful results in my own Newtonian reflector.

However, I came to realize the perfection of the parabolic reflector was the wrong choice for flashlights. The same sharp focus wonderful for forming images of distant craters and stars, also projected detail outwards. Hence, the dark spot (from nearly parallel rays being obstructed in the middle by the bulb) and the filamentary highlights (like through a telescope in reverse, even if the filament was not exactly at the focus.) A better choice might be a spherical mirror, with the bulb closer than the center of curvature. (Faceted mirrors are sometimes really used.) This would direct some peripheral rays toward the center and fill in the variations to some extent, as well as providing more helpful "finder rays" at the periphery. After all, we aren't designing a laser beam, but useful household illumination. For example, E. Edmund Ellion came up with a better idea [2]. In any case, the choice that was wrong and harder to manufacture became the standard, instead of better alternatives. So I wondered, why almost everyone made flashlights that way.

Yes, this is petty isn't it, hardly earth shaking. Yet it's a commonplace example of how people really don't think at their best, because habits and deceiving impulses are getting in the way. What has gone wrong? I think the problem here was the appeal of the ideal, without context and thinking through. "This is the 'best' reflector, so we will use it." And I'm not sure how conscious this thought really is. The crappy donut of light is clearly an example of the saying, "the perfect is the enemy of the good." It can also be the enemy of the sensible... People lack an appreciation of paradox and irony. A website [3] evaluating different reflector shapes illustrates the misdirection perfectly, fixating on perfection of parallelism for the "ideal" flashlight rather than best use. Much has been written in recent decades on fundamental flaws in human thought. One notable example is by Ariely [4]. Scientists and science students are susceptible [5], not just the general public.

This is a helpful if embarrassing effort, and transcends merely cataloguing the various fallacies that are possible. We are finding out more and more, and in detail, how and why our thinking gets into odd ruts and dead ends. This knowledge does provide direct hints about better thinking process, but it doesn't "inspire" us to aspire, to a better condition. Apathy and seductive motivation, not just lack of technical savvy; is a limitation. Otherwise, teachings about fallacies (done for millennia) would have already made for much less confusion and error.

It's not right to say "people are actually stupid," either. The world is subtle. Consider the

treachery of the common *flat* mirror. I ask a "simple" question: "do plane mirrors reverse left and right?" Think you know? Let's think it through. First, printing in a book is certainly seen "backwards" in the mirror, as is the lettering of things seen in a rear-view mirror (hence AMBULANCE is reversed and the letters are painted as \exists ) So that means yes? At a college reunion I titillated and took aback a few alums by working this question. If I got the usual "yes" answer, I said: wait a minute. If you wave your right hand in front of the mirror, you see waving motion to your right, and so on. Hmm. Then the embarrassing shocker: imagine holding up a pane of *glass* between you and the mirror. Write "hello" on the glass with dark grease paint. In the image (i.e., on your retinas, where "the image" really starts), the letters will have to match up to their positions, as will each point ... indeed, they will thus be oriented the normal way as well! The lettering will not be reversed! So why is some lettering "reversed"? The following made a lady groan: lettering in books is on an opaque surface. If you hold the book to yourself normally, light from the pages can't reach the mirror. For us to "see the page in the mirror," it must face the mirror instead. Now we have it. *You* are the one that literally reverses left and right, by turning the pages around with your own hands! And *you* are turned around relative to the ambulance behind you. It's not the mirror's fault.

Yet did I trick these poor marks a bit? Perhaps. The issue is not this simple. The mirror world is certainly reversed front to back. Hence, if you consider left-to-right to be an x axis and down-to-up to be a y axis, then forward-to-behind can be a z axis (as per increasing positive.) The mirror image inverts the z axis. So here's the ironic twist, pun fully intended: it could be argued, that reversal changes the standard of "left" and "right" for the mirror image. If the image was considered "real," then your reflection regards the corresponding hand "on the right" as a "left" hand, and vice versa. Now what? It seems to depend on the context of what we meant by the question, as well as a metaphysical decision about whether to reify a sort of Platonic other world behind the mirror.

But wait, there's more. By contrast, if we turn around an actual structure to reverse its front and back, its left and right portions *would* literally reverse as well. Turning around a coordinate system makes its x-axis reverse sign too, not just the z-axis. But we found that a mirror image is not turned around that way, and facing sides continue to correspond. So why can't I imagine that "in the mirror world" the x-axis, not just the y-axis, has the sign of numbers showing just as labels would be reflected. That is, positive and negative would still point in the same directions. Then left and right remain face to face if we retain the standard: "my right is the direction of positive x, when z increases behind me." So perhaps we should say after all that left and right are not reversed by or in the mirror, in any sense. (In which case, my original tease is simply correct.) Actually I'm not really sure. On top of all that, specific handedness cannot even be defined in purely descriptive terms. One needs a literal example to distinguish beyond relative comparison ("Ozma problem." Hat tip to the late Martin Gardner [6] for often bringing up such flummoxes about

mirrors, symmetry etc.) How can such a simple query end up no clearer than mud?

Ironically it *may* be easier to deal with apparently complex problems than apparently simple things like mirrors and nothingness. One of the great problems facing society is finding enough employment for the population as businesses try to operate with fewer and fewer work hours. Usually unemployment is tackled through interest rates, government spending or projects, complex incentives, calls for more training etc. Yet perhaps once again, the simple answer waits for assumptions and habits to move aside. Economists will need to evaluate the aptness of this proposal, but the following serves for illustration. Suppose that for business tax purposes, money paid to employees counted *as if* say 1.2x greater an expense. That is, make it a superdeductible. The multiplier could be tweaked to count only the first million earned, etc., or how about tied to whatever was taxable for social security. If need be, other expenses could be made infradeductible (say, 0.9x); perhaps other adjustments would be needed. Logically, this should encourage hiring because it would cost less than otherwise to pay employees compared to say, for robots. It's reasonably simple and less politicizeable because "pay" is a broad category and not picky selection. It's a win-win because it makes it easier for employers to do what we're wanting them to do. It appears it logically *should* work well.

I don't know how well this idea would actually work, but I am curious why it isn't at least tried or talked about more. That is the issue most germane to our query. Have models shown it wouldn't work as well as the proposal comes across, or is it another one of those ideas that is smothered under "but *this* is the way we've done things?" How rational is current clunky tax policy? I suspect not much, so is reform a mental bog like handness through mirrors? Or, would academics given free reign have long ago perfected something that the political process inevitably poisons? Is this new proposal the fix, or is it another flashlight with a parabolic reflector? We need to dig deeper into why so many things are either built or done the wrong way or become conceptual morasses for us, and why many good ideas are resisted. We need to know why we don't conceive and implement more good solutions, and teach our children directly to accomplish them. This means, in detail, and the reasons for our failures that involve limitations of accomplishment rather than of good will. We already appreciate that beings evolved to deal with survival in savannahs and forests can already consider themselves very lucky to do calculus. And isn't that already ironic and mystifying too, that such complex operations and subtle concepts of infinitesimals, imaginary numbers and infinite sets of various kinds are manipulated with such apparent ease, even though we can't literally handle them at all? Why is this, while we are often stumped about "the obvious"? Roger Penrose [7] has written well of the oddity of our great abstract abilities and what they might mean for our not being mere automatons.

More than just "working smarter, not harder," we need a science of "mistakeology" and we need to teach mistake avoidance. Our educational institutions are not tackling this, but continue to act

as if imparting knowledge and positive specific competencies is enough. Yes, critical thinking and review of formal logical fallacies, scientific method etc. is helpful and more teaching of it should be done. But we need to train minds to think in intrinsically less fallacious ways, to be more creative, to be less intimidated by practical custom so they can say things like "flashlights are stupid" and more often come up with something better. (Think also of jury duty!) If indeed schools are mostly designed to create "organization men and women", they need to create challengers who will actually make their organizations more productive and competitive. But this requires specific understanding of better thinking processes, not just promoting a nebulous rebellious or artistically creative streak.

Thinking We Are More Than Machines, And Doing More

The Human Brain Project [8] is a massive effort underway to model the human brain by computational technologies, and the the BRAIN Initiative [9] is similar. They are implicitly based on the assumption that computationalism is essentially correct: that is, our minds are basically like computers running programs (although more like "machine code" than C++ at the neural level.) This is the perspective that natural intelligence is the same as artificial intelligence. However, I want to plead for the limitations of such AI efforts to model our brains. I also suggest that believing our minds are more than just computers can inspire us, as well as encourage more effective behavior such as the application of "will power." I present a rather novel argument on behalf of our not being machines, to join other efforts such as the famous Chinese Room, zombies, gradual replacement of neurons with circuitry, etc. The debate about the nature of the mind goes on and on, with each side rather attached to their positions and classic memes. AI advocates point to a sense of basic necessity and the way our brains are built: interconnected cells that are like switching stations. Of course, that doesn't demonstrate that their mode of functioning is like formal programming, this being a sub-debate apart from more radical distinctions.

Was Descartes wrong? René Descartes' famous dictum "I think, therefore I am" (*cogito ergo sum*) seems as self-evident as it gets. If you think, then you must exist, true? However, as Bill Clinton reminded us, "it depends upon what the meaning of the word 'is' is." "Existence" has long been a mystery in philosophy. Here I mean the what of it, not the why. It is utterly simple (so it seems) and yet we can't really describe or explain it. As Kant noted, it is not a trait like being tall or sharp-edged. We can't divide lions into those that exist and those that don't. Or can we? The conundrum of existence is related to the claim that our minds are basically just powerful computers. I say, our minds are more than that, and the difference is intimately connected to the very nature of existence.

There is a challenging proposal going around that our universe is literally made of math. Physicist Max Tegmark's new book *Our Mathematical Universe* [10] argues that this world of apparent

objects is of the same nature as things like perfect circles and collections of points on graphs. Yes it is not "perfect" because the mathematical array forming it is more complex and messy, but not from matter being a kind of "more real" essence compared to conceptual entities like a dodecahedron. Tegmark imagines that all possible mathematical structures are in fact real, and some of them are "worlds" like our own. It is somewhat like the reification of the conceptual plan for a video game. Indeed, all possible configurations of things and rules and happenings have to exist! We are just one of infinite worlds of maybe, all in something called "the ultimate ensemble." This "mathematical universe hypothesis" is of course a form of Platonism. It is a more formal version of David Lewis' concept [11] of modal realism, asserting that all possible worlds are equivalently existent.

Modal realism ran early into the criticism that there is more variety in disorder than in order, and thus our Bayesian expectation is to be in a "rubbishy" world poised on the brink of chaos. Lewis' reply that we cannot compare infinite sets of the same cardinality provides no positive basis to expect an orderly universe, and there are defenses of the original complaint. Note also, Hume's criticism that "laws" are post facto descriptions of order and what actually "makes" things happen the way they do? For example, all the mathematical structures representing the "following of the rules of chess" simply are chess games by definition, and all those structures in which those rules are broken, simply aren't. There are clearly many more of the latter than the former (since these are discrete configurations.)

So how does that relate to how our minds work? Well, suppose that a mind really works like a computer. That means, all it does is describable in terms of math. If that is so, then its operations don't depend on whether MUH is true or not. All your brain could know is the abstractions represented by its computational activity. The same computations happen anyway in either a purely abstract sense, or in a world that is "made of math", as would happen in a special material world. Mathematics is defined and operates only on its own ideals, even though it seems to us to represent "realness" and not just pure form. Hence there would be no way for your brain to detect its substantive existence in a material world that was not just an abstract model world. Descartes was wrong because a mathematical brain in any of the countless structure worlds of the UE, would be able to "think" in the sense of having analytical processes. But I don't think a mere "math brain" could have real feelings: love, nausea, itches and pains, delight, experiences of pretty color sensations, and above all: the basic "sense" of being alive and real.

Most of us already think we aren't just machines because we have feelings, which we want to believe computers can't have. But hard-core believers in functionalism and computational intelligence do think that feelings come from such activity. However, my argument shows that if we are CI minds, then we can't have a "feeling of really existing" that shows we are more than just Platonic mathematical constructions. Such thinkers will ironically have to either accept MUH

and deny materialism, or else give up on the idea of CI in order to keep the idea that matter is special and more than math. Materialism is actually a sort of mysticism, since matter has to transcend pure abstraction and thus pure logic.

I think the world and our minds, are both more than just math, and neither can be completely described by math. The world is not a machine, and we are not computers. Just consider the mysterious character of quantum mechanics, with its apparently genuine randomness. Mathematics is of course deterministic, being driven by logical necessity. (Entities like "random variables" just describe a field of possible outcomes, they do not actually *produce* varying results as do real random number generators we must physically build.) Despite pretensions, the enigma of Schrödinger's Cat is not really resolved by current attempts to bring back a sort of determinism through continued Schrödinger evolution [12]. If we know that we and our world are "really real" and not just mathematical structures, our minds must intimately connect to the ground of being that constitutes the universe. Otherwise, brain processes would not be able to access that fundamental fact and reality - they would operate "above" (perhaps "below" is a better metaphor) that level and not be able to have awareness of it. That means that consciousness and "being" are the same essence, as the great mystics have experienced. *Sentio ergo sum* - I feel, therefore I am.

There is no specific ability clearly conferred by appreciating this point. However, it can serve to stimulate renewed interest in contemplative exercises, meditation, development of intuitive faculties and empathy, and greater respect for our world as if, at least, it is a sort of living being. Perhaps "inspiration" can motivate us as well as providing a hint to look for special features of our minds, if indeed they are intimately connected to physical ground rather than being "substrate independent" processes. We may finally resolve whether psychic abilities really exist, and why. All this is rather speculative of course, but it deserves attention. Recent research on microtubules (see review [13]) has revived the idea that special quantum processes - non deterministic and therefore credibly involved in the noncomputational awareness that I have argued for here - are a significant or at least relevant aspect of neural activity. Also, recent findings on myelin distribution [14] show more gapping than previously thought - which may allow more diffuse, parallel-type interactions and not just discrete synaptic switching. These may also enable the exercise of "will", a concept defended below.

"Free will" is the other sort of transcendent property we experience of our minds, and is even more controversial than a special status for conscious awareness. Even though determinism was technically overthrown by quantum randomness, the relevance of this for choice behavior has remained unclear. Many authors have renewed the attack on the concept of free will on various grounds, ranging from basic causal processes, to the influence of unconscious motivations (having been more revealed by recent research), the measurement of unconscious electrical changes before the subject claims to have decided to act, etc. However, here I review my own

fundamental argument [15] supporting the claim of some kind of special, global executive function or "will" not only in humans, but higher animals in general. There are various theories of behavior of course, but all that deny an effective executive "homunculus" (if there is such, it need not be immaterial) rely on spontaneous coordination of patterns of brain activity. One example is "pandemonium." These conceive that if you have been doing a certain activity for awhile, like raking leaves or playing the piano, the patterns corresponding to creating that behavior have built up in your brain and are expressing as your continued application of that activity.

But what if you think to yourself, "right now I will stop whatever I am doing, and sit perfectly still." Or, more complex, suddenly change to a completely different activity. How could a maelstrom of activity with no intrinsic holism to it, turn on a dime like that without some global supervisory metaprocess? Whatever enables a sudden yet smooth "freeze" or shift, and then resumption "when we choose to" indicates global correlation of neural activity. I don't see how signals that are simply on their own, as it were, could make such a change. At the least, I for one, would expect that whatever initiated the choice (whatever its nature) to suddenly stop, would be fighting much inertia and confused independence that would result in a fitful and prolong putting on of brakes. Furthermore, unlike a crude freeze or crash of a computer it can be so easily resumed: play the piano for 20 minutes, freeze in place for four, then start playing again. It is so easy, the experience is just as the traditional concept of an overarching "will" suggests. Sure, I can't prove this, any more than those of more reductionist conventional views can directly simulate and demonstrate the validity of their models either. However, it does knock against the plausibility of reductionistic, bottom-up theories of choice, and it intuition-pumps and gives hope to the traditional view that some sort of whole person is involved in making choices.

Yes, we still face well-known facts such as the anticipatory potentials, although indulging the whim of sporadically picking some irrelevant act is hardly the classic idea of weighing a course of action. However, in a 2000 phone conversation with the late distinguished choice researcher Benjamin Libet, he confirmed that the clearest support for active choice is "free won't": our ability to stop from doing things we'd otherwise do. This is of course the traditional role of "will power." It is something we need to understand and enable if we are to enhance people's abilities to avoid harmful habits, overeating, antisocial behavior and so on. Various research as reviewed by Kelly McGonigal [16] and various investigations by or on behalf of Alfred Mele of Florida State University [17] have supported some traditional views of willpower. For example, those who believe that we have free will are more able to stop themselves from indulging in some habit they are struggling against. Also, exercising self control builds its effectiveness, as though it were a muscle. Just put aside debates over what it really is. Adding willpower training to our educational process will create healthier and better behaved adults, more able to take care of themselves and coexist with each other. We can only steer the future if we can better steer our own selves, and we will only passionately care if we think we are truly alive.

References:

- [1] Nicholas Carr, "The Limits of Social Engineering," <http://www.technologyreview.com/review/526561/the-limits-of-social-engineering> (April 16, 2014.)
- [2] Danie F. Cuff, "Tech Notes; Flashlights With a Twist", <http://www.nytimes.com/1993/06/13/business/tech-notes-flashlights-with-a-twist.html> (published June 13, 1993.)
- [3] J. Schattman, "Parabolic Reflectors and the Ideal Flashlight" (<http://www.maplesoft.com/applications/view.aspx?SID=5523&view=html>, November 28, 2007.)
- [4] Dan Ariely, *Predictably Irrational, Revised and Expanded Edition: The Hidden Forces That Shape Our Decisions* (Harper Perennial, New York, 2010.)
- [5] *Scientific American*, <http://www.scientificamerican.com/article/your-scientific-reasoning-more-flawed-than-you-think> (Aug 21, 2012.)
- [6] Martin Gardner, *The New Ambidextrous Universe: Symmetry and Asymmetry from Mirror Reflections to Superstrings: Third Revised Edition* (Dover Publications, Mineola, 2005.)
- [7] Roger Penrose, *Shadows of the Mind* (Oxford University Press, Oxford, 1994.)
- [8] The Human Brain Project: <https://www.humanbrainproject.eu>.
- [9] The BRAIN Initiative: <http://www.whitehouse.gov/the-press-office/2013/04/02/fact-sheet-brain-initiative>
- [10] Max Tegmark, *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality* (Knopf, New York, 2014.)
- [11] David Lewis, *On the Plurality of Worlds* (Oxford: Blackwell Publishers, Oxford, 1986.)
- [12] Neil Bates, "Our Non-Deterministic Reality Is Neither Digital Nor Analog: Experimental Tests Can Show That Decoherence Fails to Resolve the Measurement Problem" (<http://fqxi.org/community/forum/topic/949>, posted 02/17/2011.)
- [13] Elsevier, "Discovery of quantum vibrations in 'microtubules' inside brain neurons supports controversial theory of consciousness", <http://www.sciencedaily.com/releases/2014/01/140116085105.htm> (January 16, 2014.)
- [14] News Staff, "Myelin Finding Turns Neuroscience On Its Head," http://www.science20.com/news_articles/myelin_finding_turns_neuroscience_on_its_head-134541 (April 2014.)
- [15] Neil Bates, "Chosen Sudden Behavioral Suppression: The Key to Understanding Free Will?", Poster Presentation, *Toward a Science of Consciousness* (Conference), Tucson AZ (2000).
- [16] Kelly McGonigal, *The Willpower Instinct* (Penguin, New York, 2011.)
- [17] "Free Will: Empirical and Philosophical Investigations," <https://www.templeton.org/what-we-fund/grants/free-will-empirical-and-philosophical-investigations> (Copyright © 2009 John Templeton Foundation.)