The Assent of Man: 
Mental Representation 
and the Control of Belief

DANIEL T. GILBERT


An old story has it that a middle-aged businessman from Cleveland traveled to Tibet and climbed its highest mountain so that he could put a question to the venerable yogi who lived at the peak. After weeks of torturous climbing the businessman finally reached the summit—cold, hungry, and bruised. He found the yogi meditating in a cave and fell at his feet. "Great master," said the businessman, "I've come such a long way to find you. Please tell me, what is life?" The yogi turned to the man. "I have spent 40 years contemplating this very question, my son, and I have concluded that life is a fountain." The businessman looked aghast. "What? I've come all this way, through ice and snow, without food or water, and all you can tell me is that life's a fountain?" A disappointed expression came over the yogi's face. "Damn!" he said. "You mean life's not a fountain?"

What makes this story funny (and if it did not seem funny, then just play along) is the fact that when the businessman objected to having his question answered metaphorically, the yogi assumed that he had chosen the wrong object of comparison. Yet, the particular object was not the source of the businessman's anguish—it did not matter to him whether life was a fountain, a flower, or a large cheese pizza. Rather, he objected to having a serious question answered with a metaphor of any kind. After

*This chapter was written while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences. The fellowship was made possible by the John D. and Catherine T. MacArthur Foundation and by a Research Scientist Development Award from the National Institute of Mental Health (1-K02-MH00939-01). Research reported herein was funded by a grant from the National Science Foundation (BNS-8819836). The generous support of all these institutions is gratefully acknowledged. The chapter was improved by the thoughtful comments of Mark Bickhard, Pat Malone, and Romin Tafarodi; by the patient philosophical advice of Sissela Bok, Bill Lycan, Ruth Millikan, and Red Watson; and by the thorough and insightful critiques of Susan Fiske and Robin Vallacher.
such an arduous ascent, the businessman felt that he surely deserved to know what life was, and not just what life was like.

Most of us are like a businessman from Cleveland in that a metaphor may strike us as no answer at all. And yet, metaphors are the only vehicle by which psychologists can understand and explain the unobservable. No one can actually see the mind, and thus psychologists have traditionally described its features, their relations, and their functions, in terms of the common physical systems that everyone seems to understand. Metaphors are powerful devices for understanding that which otherwise eludes us, but the trouble is that we occasionally forget that we are using metaphors and we begin to confuse the objects that are being compared. Life may be a fountain, but we would be justifiedly concerned about the philosopher who wanted to discover whether life's spigot took a number six or number eight washer. That philosopher would have forgotten that he or she was using a metaphor and would have been fooled into asking questions that are meaningfully asked about fountains but not about life (see Weiner, 1991).

The metaphors for mind are numerous (Lakoff & Johnson, 1980; Leary, 1990; Roediger, 1980), but perhaps the most enduring has been the mind is a warehouse of information. Tulving (1983, p. 5) attributes this metaphor to the nobleman Gregor von Feinaigle, who explicitly described the mind as a warehouse in 1807, but warehouse-like models of the mind are usually traced, in one form or another, to Plato (see Bruner & Feldman, 1990; Jaynes, 1976; Kellner & Roediger, 1984; Leary, 1990; Roediger, 1980).

In its simplest form, the warehouse metaphor has it that the things we know (which we could call mental representations or ideas or memory traces, or any of a dozen other names) are indeed things, and that they are stored in an orderly fashion in a place called the mind (or some subdomain, such as the memory). When we want to use these things (for thinking or judging or talking or remembering), we go to the place where the requisite knowledge is located and retrieve it. Despite the fact that such stories have enough loops in their logic to make Möbius dizzy (e.g., if “we” are the warehouse, then how do “we” go into the warehouse to retrieve things without turning ourselves inside out?), they do provide the undergirding for most current psychological theorizing. The warehouse metaphor becomes much more sophisticated in sophisticated hands of course, but its basic features persist.

Damn! You mean the mind’s not a warehouse? Of course it is, and of course it isn’t. The metaphor is only an attempt to help us think about the mind and its workings, and this it has done quite well. But there may also be ways in which it has caused us to ask questions about number-eight washers, and moreover, ways in which it has obscured other conceptualizations of mental life. Despite all appearances, this is not a chapter about metaphors or fountains, but a chapter about the psychology of belief—specifically, about how our beliefs control us and how we control them. And the main task of this chapter is to argue that there is another way to think about the mind, and that even if the warehouse metaphor turns out to be the best one, it is at least not the only one.

When Philosophers Collide

All science, it is said, begins with taxonomy. When good thinkers set out to understand new things, they usually begin by parsing the object of their inquiry into manageable bits. Linnaeus parsed living forms into their various kingdoms and phyla; Aristotle parsed physical reality into matter and forces; Freud parsed the mind into conscious and unconscious domains. So when Rene Descartes decided to try to understand the nature of everything, he predictably began by parsing everything into a few tractable chunks. In so doing, he made several mistakes, some of them so grand that even today the word Cartesian is often followed in conversation by the word dualism, which together have come to mean really bad parsing.

Despite differences of opinion about the details, every citizen of the seventeenth century knew that God had made the physical universe, and therefore everyone agreed that God and it should be parsed first—that is, everyone except a young Jew from Holland named Baruch de Spinoza. When Spinoza (or, as he would later call himself, Benedict Spinoza) was all of 23 years old, he informed the rabbis of Amsterdam that contrary to Jewish doctrine, God was not a deity who dwelled outside the universe and watched it as a child might watch an ant farm, but that instead, the universe and all its laws were God and thus there was no outside from which to watch it. For which the rabbis promptly invited Spinoza to leave Judaism and never come back. Spinoza's excommunication was both cruel and unusual, but it also meant that he had little more to lose by challenging a variety of bad parsings. So he rented an attic, took up lens grinding, and in the corollary and scholium to Proposition 49 of Part II of his masterwork, Ethics, he pieced together two concepts that Descartes had parsed. Those two concepts were intellect and will, and although psychologists tend not to use these words quite so much anymore, they surely use the concepts—and they use them as though they were the very different things that Descartes claimed them to be.

Here is what Spinoza had to say about them. Here is his entire conclusion: “Will and intellect are one and the same thing.” Spinoza once argued that if philosophers would just say things succinctly and in ordinary language, then ordinary people would understand them and the philosopher would not have to say them all over again. It is probably fair to say that Spinoza was a bit of an optimist.

The Doctrine of Separate Mental Acts

Descartes' position is both a product and progenitor of the warehouse metaphor of mind. That metaphor is so deeply embedded in modern theorizing that it is usually quite easy for twentieth-century psychologists to understand what Descartes was talking about. In essence, Descartes suggested that the mind performs two mental acts when it encounters a proposition. First, it understands the proposition. Today we might say that the

1I will use the word proposition in its common sense to mean a sentence that asserts that
mind represents the meaning or content of the proposition in a symbol-based system or "language of thought" (Fodor, 1975). Second, the mind assesses the veracity of the proposition that it has represented. That is, it somehow decides if the proposition that it has just understood is true or not. Just how this is done is no trivial matter, but we can pretend for the moment that assessment merely involves comparing the proposition to other propositions that are known to be true (a method called coherenstism, which will become important later). Thus, according to Descartes, minds first represent the meaning of propositions and then later, as a separate mental act, assess the veracity of those representations. Only when a proposition is first represented and then assessed as true is it said to be believed.

Why did Descartes believe that believing involved these two steps? First, because it feels that way. Most of us have the experience of understanding what is said to us, pondering for a moment, and then judging the validity of the assertion on its merits. Descartes's theory captures the phenomenology of believing quite well. Second, politically shrewd thinkers of the seventeenth century generally put psychology in the service of theology. Descartes knew that people often believe what is not so, and thus he required a psychological model that would hold people (and not their Creator) responsible for mental errors. So Descartes argued that God had bestowed the intellect or understanding (e.g., the capacity to turn auditory vibrations into meaningful mental representations of propositions), but that each individual had been given free will with which to judge the merits of the propositions they understood. God had assembled the hardware of human cognition so that people would understand what other people meant when they said things like "kumquats are tastier than roofing materials," but people themselves had to decide whether such things were true or not. In essence, the Doctrine of Separate Mental Acts was Descartes's way of making sure that God was not to blame when someone decided to eat shingles.

Descartes's theory of mind, then, has two simple and important axioms: (1) the mental acts of understanding and believing ideas are separate and sequential, and (2) people have no control over how and what they understand, but they are perfectly free to believe or disbelieve what they please. Of course, modern psychologists know that just because a mental process feels a certain way does not mean that it really is that way. And they also know that sound psychological theorizing can only happen when psychologists are free to challenge the orthodoxy of their day. Spinoza knew both of these things a long time ago. As such, he felt free to suggest that the parsing of intellect and will, of passive representation and active assessment, of mere understanding and partisan subscription, was one of the magnificently wrong ideas that God apparently did not force Descartes to accept.

---

The Doctrine of Empowered Ideas

Spinoza's theory is not grounded in a familiar spatial metaphor, and thus it can at first be a bit difficult to understand. According to Spinoza, the Cartesian Doctrine of Separate Mental Acts was an erroneous bit of psychology that grew out of a fundamentally defective metaphor for mind. For most modern psychologists, the truly interesting questions about the warehouse metaphor have to do with how information is stored, how it is retrieved, and so on. But the metaphor's most basic assumption is that knowledge is indeed something that can be stored, retrieved, activated, or otherwise operated on. And this, Spinoza argued, was the erroneous assumption that led Descartes mistakenly to parse the intellect and the will. As Spinoza (1677/1982, p. 97) noted of the Cartesianists: "They look on ideas as dumb pictures on a tablet, and misled by this preconception they fail to see that an idea, insofar as it is an idea, involves affirmation or negation." The word dumb here does not, of course, mean stupid, but mute. A better rendering might even have been impotent.

Ideas certainly are things in the sense that when we say "I have an idea" we are talking about some physical activity in our brains that occurs as a result of our commerce with the world. That much any materialist (which Spinoza was not) admits. But if one dwells on the "thingness" of ideas for long, one slips into the bad habit of thinking of ideas as having another property of things, namely impotence. Most things, from salad tongs to bicycles, are in and of themselves impotent, and can only have an effect or "make a difference" when they are used. When things are not being used they are stored impotently (in a drawer or a garage, for instance) and thus cannot have effects (such as lettuce-moving or pedestrian-ramming). According to Descartes, ideas are, but they do not do. As such, the Cartesian mind requires a faculty that can do something with the impotent products of the intellect—a mechanism that can somehow energize or empower a "mere idea" and turn it into a belief or "an idea that controls action." For Descartes, that mechanism was the will.

Spinoza did not subscribe to the thinly impotent of ideas. He argued that empowering mechanisms such as the will are unnecessary theoretical baggage because ideas are already empowered and thus do not need a fictitious will to serve as their engine. Descartes described a belief as a special sort of idea, one that had been transformed or given the nod of assent, so to speak. But Spinoza thought that belief was a mere description of what any mental representation does to the behavioral propensities of the system in which it exists. For Spinoza, "every idea is intrinsically belief-like" (Bennett, 1984, p. 164), and thus to have a subcategory of ideas called beliefs is much like having a subcategory of living persons called breathers. All mental representations contribute to the behavioral propensities of the person, and this is all we mean when we say that a person "has a belief" or "considers an idea true" (see Braithwaite, 1932; Price, 1954, 1969; Ryle, 1949). In short, the warehouse metaphor makes it possible for us to imagine an idea that is impotent—to imagine an idea that is not a belief. And that, Spinoza argued, is the metaphor's problem.
If ideas are not impotent things in a warehouse, then what are they? How are we to construe Spinoza’s Doctrine of Empowered Ideas? We may be tempted to make a friendly amendment to the warehouse metaphor, that is, to concede to Spinoza that ideas are “things that can somehow do” and promise ourselves that we will later figure out what that somehow is. But this is precisely the wrong move. If we imagine ideas as things that do (e.g., automatic, battery-powered salad tongs), then we preserve the possibility of imagining other ideas as things that don’t. And this leaves us right back where we started. If we want to be rid of Descartes’s notion that ideas are things that don’t, then we cannot even allow the possibility of things that do. Instead, we must recognize that ideas are not things that do because ideas are not things at all—they are states.

When a system encounters information, that information changes the state of the system, which leads the system to behave in certain ways. Aspects of this state can be described in propositional terms, just as we can say that the mechanical state of an open thermostat is a representation of the proposition “The room is too cold” (see Dennett, 1989, pp. 13–42). We might even go on to say that the thermostat believes the room is cold and there would be nothing philosophically embarrassing about saying so. Indeed, anyone who describes his or her office computer as “knowing where to look for a file” or an automobile as “not wanting to start” realizes that intentional descriptions are extremely useful ways to talk about complex systems whose actual mechanical states are very difficult to discuss. But these descriptions should not lead us to think that thermostats can represent the proposition “The room is too cold” and then do nothing about it. The representation of that proposition is not an impotent picture in a thermostat’s mental warehouse, but an occasionally useful and occasionally misleading description of the internal state of the thermostat that leads it to behave as it does. The thermostat will turn up the temperature when it believes the room is too cold and turn it down when it believes the room is too hot. And this is fine. But would we really want to claim that the thermostat “has” both ideas in its “mental warehouse,” and that at different times of year it believes one idea but not the other? Surely not. Rather, we might want to say that the thermostat believes the room is cold until the room is warm, at which time the thermostat believes something else.

If we want to try out the Spinozan way of thinking about belief and free ourselves of the warehouse metaphor, we can begin with a simple bit of linguistic self-deception.2 Rather than saying that a person believes a proposition or that a person has a belief, we might try saying that propositions induce states of belief. Most representational systems can be described in terms of dynamic states or static mental tokens: Anything that can be represented by a universal Turing machine (which represents an idea as a symbol or collection of symbols) can probably be represented by a neural network (which represents an idea as the state of a hidden layer of neurons), and vice versa. So, one can usually translate a concept from Thing Talk to State Talk and back again. But new talk can promote new ways of thinking. And even though saying that understanding a proposition puts a person in a state of belief is not really different from saying that a person who understands a proposition has a belief about it, the former phrase does not bestow an accidental thingness on the mental representation. When we eliminate such thingness, we can no longer imagine a representation that is impotent. And when we can no longer imagine a representation that is impotent, then we can no longer appreciate the distinction between ideas and beliefs. And when we can no longer appreciate this distinction, we are Spinozans.

The Concept of Analytic Work

The first Spinozan principle, then, is that encounters with propositions cause people to enter a state of belief such that they are prepared to act as though the proposition were true. Such states are inevitable consequences of understanding a proposition—in fact, they are what we mean by the phrase “understanding a proposition.” A proposition alters a system, restructures a system, tunes a system for action. If Thing Talk is more comforting, we can say that understanding a proposition involves generating a potent representation or empowered idea. Now these are rather striking claims, and they will be evaluated with appropriate skepticism shortly. But before doing so, we must note that this claim is still only half the story. The Doctrine of Empowered Ideas describes what propositions initially do to people, but not what people subsequently do to themselves.

Even if we grant for a moment that people are forced into states of belief by the propositions they encounter, we all recognize that they do not simply remain in those states until assaulted by another proposition. Indeed, if a friend tells me “Spinoza was a Baptist,” I may enter a state of belief such that I am momentarily prepared to act as though this proposition were true (e.g., if I met Spinoza in that moment I might tell him how much I admire Bill Moyers or vice versa). Nonetheless, I do not remain a drooling imbecile of this particular variety for very long. Almost instantly I recognize that my friend’s statement conflicts with what I know of Spinoza from a dozen different authoritative sources, all of which report him to hav: been a Jew, and I tell my friend that she is mistaken. Exactly what happened in those few moments between my initial belief in my friend’s declaration and my ultimate belief in my friend’s pitiful ignorance?

Two rejectable accounts. Philosophers have traditionally answered this question in one of two ways, neither of which has ever satisfied them. One maneuver is to say that between the initial and final states of belief I simply “changed my mind.” Unfortunately, not only does such a statement say very little, but it also suggests a freedom of will that would not have pleased Spinoza, who argued (much as B. F. Skinner would three centuries later) that free will is an illusion perpetrated by ignorance of true causes:

---

2Technically, a Spinozan position embraces Spinoza’s entire philosophical system, whereas a Spinozistic position merely derives from that system. Surely Spinoza did not make (nor would he have necessarily endorsed) all the points I will make in his name. Unfortunately, Cartesian is the only common adjectival form of Descartes’s name, and thus in the interest of symmetry I will refer to the present position as Spinozan although it is merely Spinozistic.
Men are deceived in thinking themselves free, a belief that consists only in this, that they are conscious of their actions and ignorant of the causes by which they are determined ... As to their saying that human actions depend on the will, these are mere words without any corresponding idea. For none of them knows what the will is and how it moves the body, and those who boast otherwise and make up stories of dwelling places and habitations of the soul provoke either ridicule or disgust. (Spinoza, 1677/1982, p. 86)

Avoiding Spinoza’s ridicule is not, of course, modern psychology’s primary mission, but there are plenty of other good reasons why we should not be satisfied to say that people decide to change their beliefs. For one, it doesn’t fit all the facts. William James crystallized several centuries of debate when he asked, “Can we, by any effort of our will, or by any strength of wish that it were true, believe ourselves well and about when we are roaring with rheumatism in bed?” (1897, p. 16). The inescapable answer, James concluded, is no. Only a lunatic can believe himself healthy when he is obviously ill, female when he is obviously male, happy when he is obviously sad. Given all the will in the world, the sane individual cannot simply adjust her brain and merrily believe as she pleases. Not only is it directly impossible to change our beliefs at will, but it is also impossible for good reason. Beliefs are what bind us to the reality beyond our skins, and any organism that could simply undo these bindings and substitute a more gratifying version of reality would not be long for this world. Why gather food when one can believe oneself fed? Why procreate when fantasy feels just dandy? To the extent that the Willful Change of Belief account is an account at all, it just doesn’t work.  

If people do not change their beliefs, then perhaps beliefs change themselves. We could describe the moments between my initial and ultimate beliefs about my friend’s statement by saying that, as a general principle of human psychology, belief states evolve toward coherence. Coherence is a property of a state that measures how well all the many propositional descriptions of that state fit together. The thermostat’s state could be described as “The room is too cold” or “I do not feel warm enough” or any of a dozen other ways, all of which are coherent or noncontradictory. But a belief state that can be described as “Spinoza was a Baptist” and also as “Spinoza was a Jew” is an incoherent state because these descriptions are mutually exclusive propositions. From Boltzman machines to one-celled organisms, many natural and artificial systems organize themselves into stable states after perturbations have moved them into unstable states, and we could suppose that psychological systems move toward stable and co-

herent belief states after propositions induce unstable and incoherent belief states.

Rather than explicate the Evolution of Belief States account, however, it is probably best to dismiss it outright. The idea that belief states naturally evolve toward coherence does indeed eliminate the need for a concept of free will, but it eliminates it all too thoroughly. This account fails to provide any role for the individual believer who becomes a mere bystander to the drama being played out in her head—a spectator to the waxing and waning of belief states that simply evolve in accordance with some unwritten rule. This does more than offend our sense of propriety—it actually leaves some serious questions unanswered. For example, if belief states naturally evolve toward coherence, then why do so many people seem to have so many incoherent beliefs? And shouldn’t two people who encounter all the same information come away believing precisely the same (coherent) things?

Two Features of Work

Neither the Willful Change of Belief account nor the Evolution of Belief States account is entirely satisfactory. Happily, there is an account that does not allow people to believe whatever they like (which the Willful Change of Belief account unfortunately does), yet does provide a role for individuals in the revision of their own beliefs (which the Evolution of Belief States account unfortunately does not). What happened in the moment between my understanding of my friend’s idiotic statement about Spinoza the Baptist and my realization that it was, in fact, idiotic? Work happened. Analytic work. By analyzing the proposition, I worked my way out of one state of belief (the one induced by my friend’s statement) and into another. The concept of work is useful because it has two connotations, one more noticeable than the next. First and most noticeably, work is an act that requires time and energy. It is not “a mere word without a corresponding idea,” but rather a measurable activity with measurable consequences; as such, a psychologist with tools of sufficient sensitivity should be able to tell whether work has taken place.

Second and less noticeably, the concept of analytic work enables people to have a limited hand in the management of their own minds, a kind of unfree will. The decision to work may lie (at least in part) with the worker, but the results of that work are highly constrained by the rules of the system into which the worker pours forth her effort. One may or may not decide to spend 15 minutes randomly pushing the keys marked 3 and 9 and + on a pocket calculator, but one cannot achieve any result one desires through such work. The work cannot result in a negative sum, much less in harvested corn or a painted barn. One may or may not choose to pedal a bicycle, but the result of the pedaling is constrained by the nature of the chain assembly; it will not result in levitation, pregnancy, or the liberation of Kuwait. The point, then, is that a person may choose to do work or not to do work, but one cannot decide what the result of that work will be. The result is determined by the rules of the system in which the work is done.

We may think of belief change as the product of work whose result
is constrained, not by the nature of a chain assembly or a printed circuit board, but by the rules of propositional analysis—the laws of deductive and inductive logic that all people carry about with them. In State Talk we would say that work is vital to the transmutation of belief states, but that there are inviolable rules that dictate what sort of new state an old state can be transmuted into. In Thing Talk we would say that the mind has a Belief Reviser, which is a device that compares beliefs to others and, based strictly on its rules for weighing evidence, decides which are true. But that processor will not work unless the individual decides to exert effort—someone has to turn the crank to make the wheels of the Belief Reviser spin, so to speak. The individual may or may not choose to supply the energy necessary for the analytic work to be performed, but that person cannot decide what the consequences of that analytic work will be.

It is very important to note that the individual cannot outsmart this arrangement by previewing the results of analytic work. One cannot think, “Hmnm. If I do analytic work I will have to conclude that Spinoza was a Jew rather than a Baptist, and given my desire to remain friends with Bill Moyers, I don’t really want to believe that Spinoza was a Jew. So I’m not going to do the work and thus I will retain the belief that the proposition ‘Spinoza was a Baptist’ originally induced.” A person cannot do this because to preview the result of the work is to do the work. There is no difference between analyzing the assertion that Spinoza was a Baptist and determining what the results of that analysis would be were it carried out, because the analysis must be carried out to determine what carrying it out would reveal. A wife cannot say, “I would hate to believe that my husband was unfaithful. So I’m going to find out whether he is having an affair, and if he is, then I’m not going to spy on him because I would never want to believe he betrayed me.” To foresee is to see, and once analytic work is performed, the cat is out of the bag.

A third feature of work. Work is metaphorically useful in that it connotes both measurable activity and unfree will (or “light determinism”). Individuals may put forth effort and thereby play a role in the revision of their own beliefs—that is, they can decide whether or not to transmute them—but that role is not a determining role because they cannot decide what the beliefs will be transmuted into. But work has another attribute that is also metaphorically useful: It is graduated. Walking and jogging both require work, but they require different gradations or amounts. An easy task requires less effort than a difficult one, and if beliefs are revised via constrained analytic work, we might expect some revisions to require

more work than others. What feature of a belief state determines how much work is required to transmute it? Coherence!

This may be illustrated by examining the simple case of disputing a verbal claim. If someone says to us, “You had less sleep last Wednesday than last Thursday,” we will dispute that claim if it conflicts with something else we know (e.g., that Wednesday was poker night). But the amount of work required to dispute the claim is nontrivial, because the beliefs to which we must compare the claim (and this is pure Thing Talk) are relatively few and obscure. Most of us cannot instantly reply, “Not so! I slept twenty-seven minutes longer on Thursday than on Wednesday.” Instead, we furrow our brows, recollect the events of those days, do a bit of mathematics, and refute the claim. On the other hand, if someone tells us, “You had no sleep on Wednesday” or “You had no teeth on Wednesday” or even “Wednesday didn’t happen last week but it will happen twice tomorrow,” the refutation of the claim is quick and virtually effortless. In fact, it is so easy that it may feel as though it required no work at all.

And so it is with verbal claims that one may easily dispute a claim that contradicts a vast number of facts, and more effortfully dispute a claim that contradicts only a few facts. As every debater knows, the more coherent a claim, the harder it is to dispute. Just as coherent claims are difficult to dispute, the coherent belief states induced by these claims should be difficult to transmute. If we subscribe to the Doctrine of Empowered Ideas, then the initial claim about Wednesday providing less sleep than Thursday induced in us a relatively coherent state of belief, whereas the claim about nonexistent teeth and nonexistent days produced relatively incoherent states of belief. By analogy to verbal claims, we would expect the latter to be more easily changed than the former.

SPINOZA DOES PSYCHOLOGY

Psychologists tend to find philosophical talk sort of sexy. For the first five minutes. Then they start to wonder why no one is measuring anything. The previous discussion of Cartesian and Spinozan epistemology measured nothing at all, but it did suggest something. It suggested that we can at least conceive of a representational system that, by its very nature, is forced to believe every proposition it encounters and is able to alter that belief through work that is constrained by a set of fundamental rules. Indeed, the Doctrine of Empowered Ideas and the Concept of Analytic Work can be mated to produce a testable theory of belief. That theory suggests that (1) propositions automatically induce belief states that (2) can be changed by analytic work. Although (3) the person can decide whether to do the work, he or she (4) cannot foresee or determine the result of the work. Ultimately, (5) the amount of work required depends on the coherence of the belief state (that is, the more coherent the state, the more work is required to change it).

The system described by this theory is a very weird system indeed. It is a system that is initially and inexorably gullible and only secondarily and
energy to assess its merits. Figure 4.1 presents this notion visually. As the left side of Figure 4.1 shows, when Cartesian systems cannot get going, so to speak, they retain their understandings but do not achieve beliefs. Notice that this is not what happens to Spinoza systems, which cannot be said to "achieve" beliefs so much as to have beliefs thrust upon them. When Spinoza systems encounter false propositions in the world, they are immediately put into a state of false belief that they then must work their way out of. If the work cannot be done, then the Spinoza system is stuck with its incorrect belief. In short, a lack of energy should stop Cartesian systems from achieving false beliefs and should stop Spinoza systems from shedding them.

This simple difference between the way these two hypothetical systems react to a loss of analytic energy is quite important, because it means that we can take the energy out of real systems (e.g., people) and watch to see which of the two hypothetical systems their behavior most closely resembles. (For a similar approach to a rather different problem, see Rock and Nijhawan, 1989.) If people are Spinoza systems, then de-energization (that is, any manipulation that reduces the capacity to perform analytic work on a proposition) should cause them to retain their initial beliefs in both true and false propositions. If people are Cartesians, on the other hand, then de-energization should prevent them from achieving beliefs of any kind. If one wants to discover whether people are Spinoza or Cartesian systems, one reasonable first step is to de-energize a few.

![Diagram](image-url)
The Paralysis of Analysis

Typically, psychologists diminish a person's capacity to perform a task by requiring the person to perform an additional task at the same time. The assumption is that if the additional task is reasonably demanding, then the person cannot spend as much energy on the original task (see Baddeley & Hitch, 1974; Kahneman, 1973; Kantowitz, 1974). Just such a technique was used by Festinger and Maccoby (1964), who were the first to find that people become particularly gullible when they are busy performing extra tasks (but see Sidis, 1898, for a prescient discussion). In the decades since that discovery, a middle-sized research tradition has grown up around this topic. In general, subjects are exposed to persuasive communications (e.g., speeches, editorials, television commercials), and some subset are required to perform an additional demanding activity (e.g., listening for certain tones, counting characters in an alphanumeric display, and so on). Ultimately, all subjects are asked to report their beliefs on the topic of the communication. Virtually all such studies find that subjects who perform an extra task are more deeply persuaded by the communications to which they are exposed than are subjects who perform no such task (see Baron, Baron, & Miller, 1973; also DePaulo, Stone, & Lassiter, 1985; and Gilbert, 1989, for parallel findings in other areas).

Why does this happen? The mechanism that mediates the effect was described by Petty, Wells, and Brock (1976), who found that extra tasks facilitate persuasion by preventing people from performing the analytic work necessary to reject the communications they hear. Petty and Cacioppo (1986) incorporated this finding into a powerful and general model of persuasion processes, and noted that “distraction disrupts the thoughts that would normally be elicited by a message” (p. 141). But a moment’s reflection reveals that there is something important hidden in this argument. The disruption of the analytic work that would normally lead to the rejection of a false proposition should not induce the acceptance of that proposition if, as the Cartesian hypothesis argues, acceptance and rejection are products of analytic work. Indeed, if Petty and Cacioppo’s model is to make sense, it must assume that acceptance is the person’s initial response to a false message, and that this response is then modified by analytic work that distraction may undermine.6

Is there any evidence to support such an assumption? Several studies by Gilbert, Krull, and Malone (1990) suggest that there is. The idea in the first of these studies was to show subjects some true and some false propositions and then de-energize subjects before they could assess the proposition. To prevent subjects from rushing ahead and assessing the propositions before the experimenter had a chance to de-energize them, the experimenter controlled the timing of the assessment process by showing subjects propositions that they could not assess without the experimenter’s help. Specifically, each proposition contained a word that was ostensibly in the Hopi language (e.g., a monoshina is a bat’s eyebrow) and, because none of the subjects knew Hopi, they had to wait until the experimenter told them whether the proposition was true or false. The experimenter explained that he was hoping to model language acquisition in natural contexts (e.g., when a person moves to a foreign country), and that in such contexts people are often exposed to propositions that they later learn to be true or false. As such, subjects were led to expect that some of the propositions they would see in the experiment would turn out to be false and some would turn out to be true, and they expected to learn which of these was the case only after they had read the proposition. This, of course, meant that subjects could not transmute their initial belief states until the experimenter gave them this crucial information.

After reading the proposition, subjects were indeed told whether the preceding proposition had been true or false. But on some occasions, subjects also performed an additional task: They listened for a specific tone to which they had to respond with a rapid button press. The tone sounded very shortly after the subject learned whether the proposition was true or false. The idea, of course, was that the sudden imposition of the tone task would de-energize subjects; that is, it would cause them to spend some energy on the tone task, and thus have less energy with which to transmute their beliefs in propositions that they had just been told were false. When the subject responded to the tone, the next proposition was shown, which meant that the subject did not have an opportunity to go back and reconsider the preceding proposition.

When subjects were later asked about their beliefs, we expected them to mistake false propositions for true ones—and we expected this to be especially likely when the assessment of the proposition was impaired by a de-energizing tone task. As Figure 4.2 shows, the de-energizing manipulation had a very unusual effect on subjects’ beliefs in the propositions they encountered. When a subject was required to respond to a tone after learning that a proposition was false, that subject tended to believe the proposition was true (see the False as True column to the far right of the graph). Interestingly though, when the subject was required to respond to the tone after learning that proposition was true, the subject was not particularly inclined to consider the proposition false (the slight difference in the third pair of columns is not statistically reliable—that is, it might well have happened by sheer chance). The tone task, then, caused subjects to believe the propositions they had just read. When that proposition just so happened to be true, then the subject’s accuracy was increased by the tone task; when the proposition just so happened to be false, the subject’s accuracy was decreased. Subsequent analyses of reaction times ruled out the possibility that these results were caused by a guessing bias (i.e., a tendency to say, but not really believe, that propositions are true).

This study shows that de-energization increases belief in false prop-
ositions—a quintessentially Spinozan effect. Of course, inducing belief in propositions about Hopi vocabulary may not be a particularly compelling demonstration, but Gilbert, Tafarodi, and Malone (1992) showed that this dull phenomenon can, in another context, have very powerful consequences. In one study, subjects read crime reports with the goal of sentencing the perpetrators to prison. There were two oddities about the reports and the conditions under which they were read. First, subjects were told that each report contained some false propositions, and that these propositions would be printed in red, whereas the rest of the report would be printed in black. As it turned out, the false propositions in one report just so happened to exacerbate the severity of the crime (e.g., “Kevin leaned closer and threatened to sexually assault the clerk”), whereas the false propositions in another report just so happened to attenuate the severity of the crime (e.g., “Tom said he was ashamed of what he had done”). The second oddity was this: The reports were shown on a television monitor and each sentence was made to “crawl” across the screen, much like an emergency weather bulletin. Below the text of the report was a series of similarly crawling digits, and some subjects were instructed to search for the digit 5 whenever it appeared, as well as to read the crime reports aloud. This visual-detection task, like the tone task in the previous study, was meant to de-energize subjects.

At the end of the session, subjects recommended what they thought were fair sentences for the various perpetrators of the crimes. Memory tests showed that de-energized subjects were more likely than control subjects to believe the false propositions they had seen. Consequently, as Figure 4.3 shows, de-energized subjects were more willing to forgive the perpetrator when those false propositions attenuated the severity of the crime, but were more eager to nail the perpetrator to the wall when the false propositions exacerbated the severity of the crime. This is precisely what a Spinozan system should do. In the act of understanding the false assertion (“Kevin leaned closer and threatened to sexually assault the clerk”) a Spinozan system must believe the assertion. And if number 5’s are popping up all over the place and demanding detection, then the system may not have enough energy to perform analytic work (in this case, the very simple work of “Oh, it’s printed in red and so it is false”) and thus to change its belief about the assertion. As the results show, counting a few 5’s caused subjects to sentence perpetrators to nearly 60 percent more jail time for

The Habit of Supposing

It certainly seems as though subjects in these experiments initially believed the propositions they encountered. But why did they do so? One possibility is that these subjects are Spinozan systems who simply had no other option. Another is that these subjects were merely supposing that the propositions they encountered were true, in the belief that they could always change their minds later if they found out otherwise (cf. Kohler, 1991). In other words, might initial belief in a proposition be a strategy that even a non-Spinozan system could temporarily adopt? Perhaps. But notice that there were no good reasons for subjects to adopt such strategies in either of the experiments just described. In one case they knew that there would be approximately equal numbers of true and false propositions about the Hopi language, and thus they should have realized that a “suppose it is true” strategy would yield no better result than a “suppose it is false” strategy. In the other case, subjects could see right away that a red proposition about a particular perpetrator was false, and thus should have realized that a “suppose it is true” strategy would be positively counterproductive.

Nonetheless, one might argue that people are just so used to supposing this way in everyday life (where true propositions surely outnumber false ones by a longshot), that they have gotten into the habit of supposing, and that this habit is a hard one to break. The difference between having “a habit of supposing” and having a representational system that, by its very nature, represents propositions as true, is subtle but real. If people are merely in the habit of supposing, then there should be some way to get them out of the habit. There should be some circumstance in which any reasonable person would suspend this habit for a moment if he could. What might those circumstances be? Two come quickly to mind.

First, a reasonable person would not adopt a “suppose it is true” strategy when he knows full well and before the fact that he is about to hear something false. Imagine that a friend says, “I’m trying out one of those psychology tricks from the back of the cereal box, and in ten seconds I’m going to tell you something that is patently false and then ask you if you believe it.” If one can actually control one’s supposing habits, then such a circumstance would be a most propitious time to take the reins. Second, a reasonable person would not adopt a “suppose it is true” strategy when that person was entirely uninterested in the veracity of what he was hearing. If a friend loses her eyeglasses and says, “Do me a favor and tell me whether that coffee mug says America Is God’s Country or Antarctica Is God’s Country” one can certainly help one’s friend by reading her the words on the mug. But it would be rather pointless to suppose those words were true—or false, or kind of true, or anything else. The veracity of the inscription is simply irrelevant when one is asked merely to articulate strings of consonants and vowels for a far-sighted friend.

Both of these circumstances have been modeled in the laboratory. In one study (Gilbert, Krull, & Malone, 1990), female subjects were told that they would be trained to detect deceit in a man’s smile, and that the initial stage of their training would involve viewing a series of slides of smiling men. Subjects were informed that some of these smiling men were actually quite happy, and others were quite sad and were smiling only because they had been asked to do so for a picture. In other words, each of the faces was an icon for the assertion “I am happy,” but only some of those assertions were true. (Of course, none of the males were actually particularly happy or sad when they posed for the pictures.) The experimenter promised to tell subjects which faces were expressing true happiness and which were expressing false happiness as subjects watched the slides, but explained that sometimes he would divulge this information before the subject saw the face and sometimes afterwards. (Ostensibly this was part of the complicated and scientific deceit-detection training procedure that was far too complicated and scientific to explain to college students.)

As they watched the slides, some subjects were asked to perform a second task that demanded some energy. Specifically, they were told to listen for the presence of a tone, to determine whether the tone was a high-pitched or low-pitched tone, and quickly to press a button that recorded their determination. As in the Hopi vocabulary experiment, we expected this task to decrease the energy that subjects could devote to unbelieving the propositions that they initially believed, and thus cause them to mistake false smiles for true ones. Of course, if the supposing hypothesis was correct, then we expected this to happen only when subjects assessed the veracity of the smile after seeing it. When they knew ahead of time that they were about to see a false smile, then the supposing hypothesis suggests that they should abandon their strategy and simply disbelieve the smile from the start.

As Figure 4.4 shows, subjects were more likely to mistake a false smile for a true one when they were de-energized by the tone-discrimination task (though, just as the Spinozan hypothesis predicts, the tone task did not cause the opposite kind of mistake—the slight increase shown in the graph is not statistically reliable). What Figure 4.4 does not show is how the timing of the assessment information affected this particular error. Figure 4.4 does not show this because there is nothing to be shown. Analyses revealed that the timing made absolutely no difference whatsoever—not even a marginal one. Subjects seemed to begin by believing the false facial prop-
they had finished following the computer's instruction. Subjects were told that whenever they were instructed merely to read a sentence, their reading speed would be measured by the computer, and thus they should read as quickly as possible in order to "score well."

The trick in this study was that all the propositions about glarks were shown twice—never back to back, but usually only about ten seconds apart. Sometimes a subject was asked to assess the proposition both times; sometimes the subject was asked to read it quickly the first time and assess it the second time; sometimes the subject was asked to assess the proposition the first time and then read it quickly later, and so on. Of particular interest were the false propositions that were first read quickly and then, a few moments later, assessed. If subjects could not help but believe the false propositions they comprehended on the reading trials, then they should have been particularly likely to report believing those false propositions when they were later asked to assess them.

As Figure 4.5 shows, this is precisely what happened. A comparison of the first and third sets of columns in the figure shows that merely reading a false proposition increased the likelihood that subjects would later believe

In some sense, the example with the far-sighted friend's coffee mug is even more compelling. What happens when people try their hardest to understand and to do nothing more? Can they help but believe what they understand? Gilbert, Krull, and Malone (1990) taught a group of subjects all about an imaginary animal called a glark. Subjects learned about glarkian morphology, mating habits, social structure, and so on. After subjects had become glark experts, a computer showed them a series of propositions about glarks, some of which were true and some of which were false. Subjects were told that prior to seeing each proposition, the computer would tell them whether to assess the proposition (that is, determine whether it was true of glarks) or simply to read the proposition as quickly as possible and then move along. In either case, subjects were told to press a key when
Propylactic strategies. If Spinozian systems must believe everything they understand and, must understand everything to which they can be exposed, then there is no escape from the necessity of having beliefs, no matter how false. And if, in order to comply with this requirement, we must control our beliefs, then this control is exercised on the level of our own minds, by not only controlling our own, but also modifying our level of control of our belief systems. The only way to avoid this is to reduce the number of beliefs we hold, or to reduce the number of events that can influence our beliefs. The only way to do this is to reduce the number of events that can influence our beliefs. The only way to do this is to reduce the number of events that can influence our beliefs.

The Urge to Hide
Most of us find it difficult to believe that we must believe everything that we understand, and that we cannot control our beliefs. But the fact that we cannot control our beliefs does not mean that...
others feel compelled to avoid (see Tedford, 1985)." In the laboratory, as well as in everyday life, people will go to great lengths to avoid hearing what they consider to be false utterances about religious, political, and financial matters (Frey, 1986), and they will even avoid false propositions when those false propositions are excessively positive descriptions of themselves (Swann, 1987; Swann, Hixon, Stein-Seroussi, & Gilbert, 1990). But avoiding exposure to false propositions is not in and of itself a behavior unique to Spinozan systems. There are at least two reasons why a Cartesian system might also avoid such exposure.

First, there is only so much time in the world, and when one spends time with false propositions, one passes up opportunities to meet true ones. If both Cartesian and Spinozan systems want to have correct beliefs, they should both have a demonstrable preference for exposure to true rather than false propositions. Second, both Cartesian and Spinozan systems should be particularly concerned about encountering a false proposition when they are uninformed. Neither system should care to bump into a false proposition that it has no evidentiary grounds for rejecting. Parents, for instance, often worry that their kindergarteners will be exposed to ideas that they are too young to reject (e.g., "People of this race are much better than people of that race" or "Your mommy sent me to pick you up from school, so get in the car"). By "too young," the parent usually means that the child does not yet have enough true beliefs to allow the younger to identify the false proposition as such (i.e., to measure its coherence). For both these reasons, a system's tendency to engage in Exposure Control does not imply that the system is Spinozan.

Nonetheless, there is a circumstance under which Cartesian and Spinozan systems ought to differ in their tendencies to use Exposure Control. When a Spinozan system is de-energized it becomes gullible; that is, it cannot Unbelieve very well. As such, de-energized Spinozan systems should be particularly likely to choose Exposure Control as a method of protecting themselves against false beliefs. Cartesian systems, however, should not be so affected by de-energization. As Descartes (1644/1984, p. 205) noted, "That we have power... to give or withhold our assent at will, is so evident that it must be counted among the first and most common notions that are innate in us." As such, a Cartesian system that wants to control its belief in a false proposition need only sit back, relax, and withhold its assent.

"In 1919, Oliver Wendell Holmes wrote an opinion for a unanimous Supreme Court (Schenck vs. United States, 249 U.S. 47) in which he argued that "The most stringent protection of free speech would not protect a man in falsely shouting fire in a theater, and causing a panic. It does not even protect a man from an injunction against uttering words that may have all the effect of force..." That last phrase is an excellent description of how utterances affect Spinozan systems. Justice Holmes, however, showed his Cartesian stripes when in 1925 he wrote an opinion for the minority (Gillow vs. New York, 268 U.S. 652) in which he speculated that "Every idea is an incitement. It offers itself for belief, and, if believed, it is acted on unless some other belief outweighs it, or some failure of energy stills the movement at its birth." If ideas merely offer themselves for belief (as he argues in Gillow), then how can words have all the effect of force (as he argues in Schenck)? If nothing else, these landmark cases should remind us that seemingly exoteric psychological issues may play pivotal roles in the fate of nations.

Because without willful assent a proposition can become an idea but never a belief. Indeed, if a Cartesian system had to be de-energized in the presence of either true or false information, we might even expect it to prefer the latter. After all, for a Cartesian system to be de-energized is for it to miss the opportunity to acquire beliefs of any kind, and surely it is worse to miss the opportunity to acquire true beliefs than to miss the opportunity to reject false ones.

**Informational refugees.** Does de-energization lead to an increased likelihood of Exposure Control? This would, on the face of it, seem a rather easy hypothesis to test. One might give subjects the opportunity to encounter true or false propositions, and then see if de-energized subjects show an enhanced preference for the former. If people are Spinozan systems, de-energization should cause them to eschew the impaired strategy (Unbelieving) and to pursue the more reliable one (Exposure Control). If people are Cartesian systems, de-energization should prohibit them from achieving beliefs in any proposition they encounter, which is exactly what happens when they fail to encounter the proposition in the first place. As such, de-energized Cartesian systems should not be more inclined than energetic Cartesian systems to avoid false propositions.

Unfortunately, there are two complexities that make the waters a bit muddier than this simple experiment suggests. First, we can pretend (as we have thus far) that people want to believe only what is true, but when we are done pretending we will have to acknowledge this as fiction. People certainly do want to have true and correct beliefs—there are strong evolutionary, social, and moral pressures that encourage such desires—but they also want their beliefs to have other attributes as well. Who, for example, would not want to believe that others find him utterly brilliant and admirable? Indeed, people want their beliefs to be gratifying as well as true. For those who are brilliant and admirable a single belief may have both attributes, but the rest of us usually have to choose. We can either face the music or we can pretend that the orchestra is playing our tune, but not both (see Swann, Pelham, & Krull, 1989). In any given situation, then, it is extremely difficult to know what sorts of beliefs—false ones, uncomfortable ones, or some other one—people will use Exposure Control to guard against.

Second, even if we knew that a particular person was, on a particular occasion, in hot pursuit of the truth and nothing but the truth, we would not necessarily be able to predict which of two propositions that person would most want to encounter or avoid. The problem is that when we say a person seeks true information we really mean that the person seeks information that he or she considers true. A truth-loving Klansman who is offered a choice between a book by a white supremacist and Volume 6 of the *Encyclopedia Britannica* may choose the former because, from the Klansman's perspective, it contains the higher proportion of true statements. Indeed, as we have seen, subjective truth is largely a matter of coherence: Statements that complement (rather than contradict) what one already believes are likely to be seen as true. Thus, we cannot expect even the most
earnest seekers of truth necessarily to expose themselves to objectively or consensually true propositions; rather, we should probably expect such individuals to expose themselves to familiar propositions.

All of this is to say that it can be difficult to know whether de-energization increases the probability that people will use Exposure Control to manage their beliefs, because (1) people do not always want to have true beliefs and thus should not always seek exposure to true propositions, and (2) even when people do want true beliefs, the propositions to which they seek exposure may not be those that an impartial observer would label as true. These complexities mean that great care must be taken when examining the effects of de-energization on a person's tendency to use Exposure Control, but they do not rule out the possibility of such an experiment. Although no one has yet de-energized subjects in a laboratory and measured their tendency to avoid information, social psychologists who study Cognitive Dissonance theory have performed studies that make some contact with this question.

Festinger's (1957) theory of cognitive dissonance essentially suggests that people want to believe propositions that justify their decisions (e.g., the proposition "Festinger is a brilliant psychologist" should prove justifying for any graduate student who endured his demanding tutelage). As such, people should be especially likely to seek contact with such propositions (e.g., to read homages to Festinger, to seek relations with other Festingerians) and to avoid contact with the proposition's logical opposite (e.g., to shun Festinger's critics and dismiss their critiques). Unlike most other predictions of dissonance theory, this "selective exposure hypothesis" proved a beautiful loser. In study after study, subjects were induced to make a decision and then offered a choice between justifying (consonant) and nonjustifying (dissonant) information. In study after study, subjects showed no reliable preference for the former. Indeed, Freedman and Sears (1965) reviewed a decade of discouraging research and concluded that there simply was no empirical support for the selective exposure hypothesis (see also McGuire, 1968).

But in the nearly 30 years since the death knell sounded, dissonance theory has matured considerably (Cooper & Fazio, 1984; Steele, 1988; Wicklund & Brehm, 1976) and the selective exposure hypothesis has done a rather convincing impression of Lazarus: It is now clear that subjects in laboratory experiments do engage in selective exposure, but only under some circumstances, and current research strives to determine just what those circumstances might be. Frey's (1986) recent survey of the literature concluded that selective exposure to justifying propositions is least likely to occur "when the cognitive system is so stable that the dissonant information can be easily integrated (for example, by refuting it)" (p. 73). In other words, when systems can ward off the effects of dissonant or nonjustifying or otherwise unwanted information, they do not resort to Exposure Control.

Klein-hesselink and Edwards (1975) used Brock and Ballou's (1967) "de-jamming" paradigm to demonstrate just this effect. In their study, subjects were played a tape recording of a speech that was badly muffled by white noise. By pressing a button, subjects could eliminate the white noise for a few seconds (or, in Cold War Talk, "de-jam the broadcast"), and the extent to which they did so served as an index of their decision to expose themselves to the propositions contained in the speech. Some subjects were strongly in favor of the legalization of marijuana and others were strongly opposed, but regardless of their opinions, all were played a tape of a pro-marijuana speech. This meant that the pro-marijuana subjects were exposed to propositions that they considered true, and the anti-marijuana subjects were exposed to propositions that they considered false. In addition, some subjects were played speeches whose propositions were logically weak (e.g., "Marijuana should be legal because it makes rock-and-roll musicians feel less paranoid"), and others were played speeches whose propositions were logically strong (e.g., "Marijuana should be legalized because prohibitions against popular drugs never work"). As the left side of Figure 4.6 shows, when subjects were exposed to weak propositions, they exhibited a very mild tendency to avoid true more than false propositions; but as the right side of Figure 4.6 shows, when exposed to strong propositions, subjects displayed a profound tendency to avoid false propositions more than true ones (see also Lowin, 1967).

With a few assumptions, this study can be understood in Spinozan

![De-Jamming Rate](chart.png)

**Figure 4.6**

- **False Propositions**
- **True Propositions**
terms. When we say that a system is de-energized, we do not, of course, mean that every last ounce of \textit{élan vital} has been sucked from its narrow so that it can do nothing but lie in a pitiful heap. Rather, we only mean that the system no longer has \textit{sufficient} energy to perform a specific task. If one wanted to be sure that a mountaineer lacked sufficient energy to scale Mt. Ranier, one could either reduce the energy the mountaineer has or increase the energy the mountaineer needs. That is, one could fatigue the climber (e.g., by making her climb with one hand tied to her waist) or one could steepen the ascent. Similarly, if one wishes to be sure that a subject has insufficient energy to unbelieve a proposition, one can decrease the amount of energy that the subject can direct to analytic work (e.g., by giving the subject an extra task to perform) or one can increase the amount of analytic work required to unbelieve the particular proposition. Gradation is, of course, the "third nice feature" of Analytic Work and, as noted earlier, more analytic work is required to transmute a coherent than an incoherent belief state. It is not unreasonable to suspect that strong and weak propositions induce just those states, respectively. In Spinozan terms, then, subjects in this study were particularly likely to engage in Exposure Control when they lacked the energy necessary to unbelieve a false proposition—which is something Spinozan systems, but not Cartesian systems, should do.

### DANGEROUS IDEAS

In the winter of 1633, aged and infirm, Galileo Galilei was carried on a litter from Florence to Rome so that he could kneel before Pope Urban VIII and disavow "the false doctrine" that "the sun is the center of the world and immobile and that the earth is not the center and moves" (Boorstin, 1983, p. 326). Galileo was not the first, nor would he be the last, to suffer for having had a dangerous idea. From Socrates to Spinoza to Salmand Rushdie, the history of human civilization is a litany of those who have been executed, excommunicated, and exiled for sharing their thoughts—litanies of vicious criminals whose unforgivable transgressions consisted mainly of inkning paper and vibrating air. Why are we so frightened and moved and offended by mere ideas? After all, if an idea is "obviously" false—as virtually anyone with a lick of sense in the seventeenth century knew Galileo's idea to be—then one need not believe it. Right?

Baruch de Spinoza was a child of the Marranos—a community of Spanish and Portuguese Jews who had been forcefully converted to Catholicism during the various Inquisitions but who continued to practice Judaism in secret—and therefore he personally confronted this problem of the control of belief (Yovel, 1989). Can a Jew really be forced to embrace the Christian faith? Can a philosopher really refrain from believing, as Descartes insisted he must, when there is no evidence bearing on the matter at hand? Can any person really choose to believe or choose to doubt? Spinoza thought not. He urged us to reject the warehouse metaphor and the illusion of mere understanding that its language fosters, and to realize

that the mental representation of an idea is not a mute picture that we store and inspect, but rather a state of the system that represents it. A belief is a behavior-inducing state that initially represents any proposition encountered in the environment; only later, in a lawful and physically demanding way, may the system change its state to one that reflects its better judgment. The heretic who sets the air astir with dangerous ideas does, in fact, create other heretics, if only for a moment, because the idea replicates in the fertile medium of other minds and is instantly empowered to control their actions. Only with effort and knowledge and logic can the innocence of the hearer be restored.

As strange as these suggestions may seem, there are a variety of laboratory situations in which people do precisely what we would expect of Spinozan systems. When de-energized, people become particularly likely to believe what they hear and, as such, do their best to avoid hearing that which they do not want to believe. But laboratory demonstrations of these facts merely bolster the intuitions that most of us developed long ago through experience: Is there one among us who seriously doubts that people are gullible when stressed or distracted, and that the more stressed and distracted they are, the more they avoid situations in which their gullibility puts them in peril? The experiments described herein do not suggest that we alter our beliefs about what people do, but rather that we consider altering our beliefs about why and how they do it.

The word \textit{consider} is important here. This chapter has not been an attempt to prove that human beings are Spinozan rather than Cartesian systems, because with a few more boxes and arrows the Cartesian model could probably be made to account for any result predicted by the Spinozan model. Psychological models can usually be rescued from the facts by additional machinery. Instead, this chapter has been an attempt to sketch an alternative to a metaphor that has managed to guide a great deal of thinking about belief. At most, this chapter suggests that we might try the idea that people are Spinozan systems. If we do, we may find that when the details are worked out, Spinoza's model is a sophisticated and heuristic way to construe belief. We may find that it is sophisticated in some ways but simple-minded in others. Or we may find that it is merely useless, untestable, or wrong. But for the moment, the weight of evidence suggests that modern psychology should consider offering Spinoza the fair trial that the rabbis of Amsterdam never gave him.

### REFERENCES


