Reprinted from AMERICAN PSYCHOLOGIST, Vol. 46, No. 2, February 1991
Printed in U. S. A.

How Mental Systems Believe

Daniel T. Gilbert *University of Texas at Austin*

Is there a difference between believing and merely understanding an idea? Descartes thought so. He considered the acceptance and rejection of an idea to be alternative outcomes of an effortful assessment process that occurs subsequent to the automatic comprehension of that idea. This article examined Spinoza's alternative suggestion that (a) the acceptance of an idea is part of the automatic comprehension of that idea and (b) the rejection of an idea occurs subsequent to, and more effortfully than, its acceptance. In this view, the mental representation of abstract ideas is quite similar to the mental representation of physical objects: People believe in the ideas they comprehend, as quickly and automatically as they believe in the objects they see. Research in social and cognitive psychology suggests that Spinoza's model may be a more accurate account of human belief than is that of Descartes.

Though Truth and Falsehood bee Neare twins, yet Truth a little elder is.

—John Donne, 1635/1930, p. 129

Everyone knows that understanding is one thing and believing is another, that people can consider ideas without considering them so, and that one must have an idea before one can determine its merit. "Everyone knows the difference . . . between supposing a proposition and acquiescing in its truth" (James, 1890, p. 283). Nonetheless, this article suggests that what everyone knows may be, at least in part, wrong. It will be argued that the comprehension and acceptance of ideas are not clearly separable psychological acts, but rather that comprehension includes acceptance of that which is comprehended.

The Process of Belief

Components of Believing

"Believing," wrote Bertrand Russell (1921), "seems the most mental thing we do." (p. 231) Indeed, the problem of belief—what it is, how it happens, and what it does—constitutes "the central problem in the analysis of mind" (Russell, p. 231). Exactly what does it mean to say that we believe something? What do we mean when we say we believe that bald men are especially licentious, that sharks can swim backwards, or that love is blind and justice is not? The answers to these questions are so complex and varied that no brief discussion can treat them fairly. Nonetheless, in all the epistemological talk that philosophers and psychologists have produced over the years, one point of consensus seems to emerge: Beliefs, in the broad and colloquial sense, involve both the *mental representation* and the *positive assessment* of meaningful information (for a review, see Bogdan, 1986).

A proposition is believed when the proposition's meaning is represented, coded, or symbolized in a mental system and when that symbolic representation is treated as if it was true. It makes little sense to say that we believe something of which we have absolutely no knowledge or which we, through our words and actions, acknowledge to be false. Thus, to believe that armadillos have four legs, for example, requires that one have the information *armadillos have four legs* coded and stored in one's mental system and that one behave as though the information correctly characterizes some real zoological state of affairs. Whereas the representation component of believing refers merely to the existence of meaningful information within a mental system, the assessment component refers to the relation between that information and other information that already exists within the system.[1]

The Cartesian Procedure

When we think about how the human mind believes, it is only natural that we should immediately think about how it *ought* to believe. Because we cannot actually observe the processes by which comprehension and assessment occur, we tend to ask the next best question: How *should* they occur? What is the most logical way for these things to happen? How would we design a device to accomplish them? One piece of the answer to such questions is obvious and has been obvious for so many centuries

[1] Such a "coherence theory" of truth is obviously inadequate: One may not define the veracity of a proposition exclusively in terms of its fit with other propositions (see Tarski, 1969). Nonetheless, propositions do (somehow) get assessed, and because the actual mechanics of the assessment procedure are not crucial to the present discussion, no attempt to define its necessary and sufficient conditions will be made.

The writing of this article was supported by National Science Foundation Grants BNS-8605443 and BNS-8819836 to Daniel T. Gilbert.

So many people have provided kind and insightful comments on versions of this article that I will surely fail to thank them all. Among them, then, are John Bargh, Roy Baumeister, Mark Bickhard, David Cohen, Fred Dretske, Russ Fazio, Susan Fiske, Don Foss, Windflower Gilbert, Nancy Hazen-Swann, Marcia Johnson, Ned Jones, Stan Klein, Doug Krull, Mark Lepper, James Liu, Charlie Lord, Pat Malone, Brett Pelham, Mike Ross, Johnathan Schooler, Dave Schneider, Jim Sherman, Janet Spence, Bill Swann, Romin Tafarodi, Elizabeth Tighe, Yaacov Trope, Barbara Tversky, Dan Wegner, Rich Wenzlaff, Tim Wilson, Jason Young, the students in Susan Fiske's and Shelley Taylor's seminars on social cognition, and several anonymous reviewers.

Correspondence concerning this article should be addressed to Daniel Gilbert, Department of Psychology, 330 Mezes Hall, University of Texas at Austin, Austin, TX 78712. Electronic mail may be sent to PSHU666@UTXVM.BITNET.

 0003-066X/91/$2.00
Vol. 46, No. 2, 107-119

that very few thinkers have challenged it: Obviously, one must comprehend an idea before one can assess it. Put another way, before one can consider an idea true, one must consider it truly.

This simple notion—that comprehension precedes and is separate from assessment—seems more than merely agreeable and innocent: It seems to have the force of logical necessity. Borrowing liberally from the Stoic philosophers, René Descartes was the first modern thinker to formalize this notion by partitioning the mind into relatively active (controlled) and passive (automatic) domains. Comprehension, he claimed, was passive: Ideas impressed themselves upon the mind as physical objects might upon soft wax, and just as wax did nothing to receive the object's impression, so the mind did nothing to comprehend the world around it.

Although having ideas was effortless and automatic, accepting or rejecting those ideas was not. Descartes considered the assessment of an idea's veracity to be the operation of the *voluntas*—the active, conscious, willful force of the psyche. This principle reigned axiomatic in Descartes's philosophy: "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" (1644/1984, p. 205), and "All that the intellect does is to enable me to perceive, without affirming or denying anything, the ideas which are subjects for possible judgments" (1641/1984, p. 39; see also Bennett, 1984, pp. 159–167; Lehrer, 1983; Nuchelmans, 1983, p. 45; Price, 1969, pp. 221–240; Russell, 1921, pp. 247–249).

Although his fellow philosophers often took exception with Descartes's epistemology, they generally accepted the basic distinction between comprehension and assessment on which it was based. And indeed, that distinction continues to pervade contemporary views of how mental systems should and do operate. Computers, for example, employ languages and architectures that enable us to distinguish clearly between the mere representation and the subsequent assessment of information. A machine may represent the proposition *armadillos deplore raisin cookies* before or even without assessing the veracity of that proposition. Information can exist inside a machine in an unassessed form, and it can do so because the machine's human designers have generally considered this the most logical and efficient way for machines to think. Computers, then, are good examples of what one might call Cartesian belief systems—devices in which the operations of comprehension and assessment are as separate and sequential as everybody knows they should be.

The Spinozan Procedure

That is, *almost* everybody. Baruch Spinoza was a part-time lens grinder who never saw a digital computer, but it is clear that he did not think the human mind worked like one. In the 49th proposition of Part 2 of his *Ethics* (1677/1982), Spinoza put forth a thesis that, on the face of it, seems simply preposterous: He dismissed Descartes's distinction between comprehension and assessment and insisted instead that comprehending and accepting were, in fact, the same operation. Spinoza argued that to comprehend a proposition, a person had implicitly to accept that proposition; only later, if the person realized that this proposition conflicted with some other, might he or she change his or her mind. As William James (1890) would later characterize Spinoza's position, "All propositions, whether attributive or existential, are believed through the very fact of being conceived" (p. 290). Indeed, Spinoza argued that one need not even distinguish between mere ideas and beliefs—because all ideas *are* beliefs (see Bennett, 1984).

It is critical to note that this claim does not mean that persons are incapable of representing falsehoods and are thus doomed to believe forever that which they comprehend: Upon hearing that *armadillos make excellent appetizers,* few of us feel compelled to pack our cutlery and head for Texas. Indeed, having comprehended and accepted an idea, Spinoza considered persons entirely free either to *unaccept* or to *certify* it. Spinoza's thesis simply implies that unacceptance is a secondary psychological act in which the initial accepting that invariably accompanies comprehension is subsequently undone. Disbelief is by no means an impossibility in Spinoza's scheme; rather, it is merely a deliberate revision of belief.

The Mechanics of Belief

Virtually all current and classical theories of mental representation presume that once the truth value of a proposition is assessed, the mental representation of that proposition must somehow be altered or "tagged" to indicate that truth value—otherwise we would have to reassess the validity of our knowledge each time we deployed it. The differences between the Spinozan and Cartesian procedures may be framed as a disagreement not about the nature of mental representation itself, but about the nature of this tagging system. A familiar metaphor may serve to illustrate the key elements of this division.

Spinozan and Cartesian Tagging Systems

Imagine a library of a few million volumes, of which only a small number are fiction. There are (at least) two reasonable methods by which one could tag the spines of books so that fiction could be distinguished from nonfiction at a glance. One method would be to paste a red tag on each volume of fiction and a blue tag on each volume of nonfiction. Another method would be to tag the fiction and leave the nonfiction untagged. Either of these systems would accomplish the goal of allowing a librarian to distinguish fiction from nonfiction without necessitating that he or she actually reread the book each time such a discrimination needed to be made.

It is only a mild oversimplification to say that Descartes considered the mind to be a library of ideas that used something akin to the red–blue tag system. A new book (new information) appeared in the library (was represented in the mind), its contents were read (assessed), and the book was then tagged (recoded or rerepresented)

as either fiction (false) or nonfiction (true). New books (unassessed ideas) lacked a tag, of course, and thus were not identifiable as either fiction or nonfiction until they had been read. Such new and unread books were "merely" represented in the library.

Spinoza, however, argued that the mind was more like a library that used a tagged–untagged system. In Spinoza's view, books were represented before they were assessed; but because of the particular tagging system that was used to denote the outcome of that assessment, a new book that appeared without a tag looked exactly like a work of nonfiction. In a Spinozan library, a book's spine *always* announced its contents; no book could be "merely" represented in the library, because the absence of a tag was itself informative (or misinformative) about the content of the book. Analogously, ideas whose truth had been ascertained through a rational assessment process were represented in the mind in precisely the same way as were ideas that had simply been comprehended; only ideas that were judged to be false were unaccepted, or given a special tag (cf. Wegner, Coulton, & Wenzlaff, 1985; see also Clark & Chase, 1972).

Perhaps the most striking feature of these different systems is how similar they are. Despite the fact that the two libraries use somewhat different tagging systems, both ultimately accomplish the same end. If, for example, the Cartesian and Spinozan procedures (shown in Figure 1) were implemented on a pair of hypothetical machines that had unlimited processing resources (i.e., time, energy, and data) then the two machines should end up believing precisely the same things, and the particular method by which each achieved its beliefs would be largely inconsequential.

The rub is that no mental system, whether natural or artificial, has unlimited processing resources. On the contrary, real mental systems must often operate under imperfect conditions in which multiple tasks vie for access to a finite pool of resources (Baddeley & Hitch, 1974; Kahneman, 1973; Norman & Bobrow, 1975), and when this happens, the procedures these systems execute tend to break down in predictable ways. As it turns out, Spinozan and Cartesian procedures break down quite differently, and thus have very different consequences for the mental systems that employ them. In reality, any pair of machines that implemented the Spinozan and Cartesian belief procedures would, after some time, come to have very different beliefs. But before examining the particular ways in which these systems react to stress, it is necessary to discuss briefly a more general principle of systems breakdown.

The Principle of Premature Output

Perhaps the most fundamental feature of mental systems is that they are modular: Several different subsystems or modules work in a serial sequence to accomplish the information-processing tasks that presumably underlie human actions as diverse as raising an eyebrow, watching a wrestling match, and composing a sonata (e.g., Fodor, 1983; Garfield, 1987; Gazzaniga, 1985; Minsky, 1985). Even connectionist or parallel distributed processing models incorporate these serial components (Rumelhart, Smolensky, McClelland, & Hinton, 1986, p. 38). Modular systems, then, pass information from one module to another, and each module modifies the information in some way before passing it along to the next. Virtually all information-processing schemes—from those that describe language comprehension (Just & Carpenter, 1977) and visual pattern recognition (Marr, 1982) to those that model more general problem-solving procedures (Newell & Simon, 1972)—contain this information-passing feature. Surely, a flowchart with one box and no arrows would be a singularly useless sort of map.

All of this information-passing is not without beginning or end. Processing is initiated by the introduction of information into a modular system, that information is passed between and modified by various modules, and finally an informational product is outputted. In some systems, outputting can occur only after the last module has received some information as input. By way of illustration, imagine a very simple, boxlike machine whose input is a fuzzy, white tennis ball. The ball rolls into the machine, is shaved, dyed lime green, tattooed with a trademark, and then rolled out of the machine and into its retail canister. Such a machine is an example of a single exit system: The tennis ball can only emerge after it has been modified by, and passed to, each of the shaving, dying, and tattooing modules in sequence. Until the tennis ball is received by the tattooing module, it cannot exit the machine.

In general, however, sophisticated systems allow informational products to exit at a variety of stages (see Norman & Bobrow's, 1975, principle of continually available output). Imagine a different contraption: A pancake-making machine whose input is an irregular-shaped pancake. This irregular pancake is modified by the first module, a square-cutting device, which cuts a square out of the irregular pancake's center. This square

Figure 1
Two Conceptual Stages of the Spinozan and Cartesian Belief Procedures

	SPINOZAN PROCEDURE	CARTESIAN PROCEDURE
REPRESENTATION STAGE	Comprehension & Acceptance	Comprehension
	↓	↓
ASSESSMENT STAGE	Certification OR Unacceptance	Acceptance OR Rejection

pancake can either exit into a blue box or it can be passed along to a second module, a circle-cutting device, which cuts a circle from the square pancake and then drops a circular pancake into a red box. Such a multiple exit system would be capable of producing output at either of its two processing stages, and would thus be capable of producing a variety of different products.

One interesting feature of a multiple exit system is this: Because resources are required to send the information (or, in the preceding case, a pancake) from one module to the next, any sort of resource depletion will cause premature output (see Norman & Bobrow's, 1975, principle of graceful degradation; Tversky & Kahneman's, 1974, anchor-adjust heuristic). In other words, when multiple exit systems break down they generally do so by outputting the product of an early module. The pancake-making machine is a good example. A resource shortage can cause the machine to produce a square pancake (a product of the first module) but not a circular pancake (a product of the second module), because circular pancakes depend on the prior manufacture of square pancakes but not vice versa. It is worth noting that this tendency toward premature output will occur when the two modules perform equally complex tasks (i.e., when they require equal resources); to the extent that the second module performs a more complex task than the first, this tendency will be even further pronounced.

This tendency can be codified as a general principle of systems breakdown: *When stressed, modular information-passing systems with multiple exit capabilities will often show a bias toward prematurely outputting the products of early modules.* This principle enables very clear predictions about the unique ways in which Spinozan and Cartesian systems should break down and, therefore, about what these systems should ultimately come to believe. A stressed or resource-depleted system should represent propositions (a product of the first module) but should occasionally be unable to assess those representations (a product of the second module). This means that resource depletion should prevent a Cartesian system from either accepting or rejecting the propositions that it merely comprehends, whereas it should prevent a Spinozan system from unaccepting or certifying the propositions that it both comprehends and accepts.

The Breakdown of Belief

The principle of premature output suggests that the way a mental system behaves under stressful conditions may be a good clue to the sort of belief procedure it employs. In the language of the library metaphor, one can imagine placing a book on the library's shelves and stressing the librarian so that he or she is unable to assess the book's contents and tag its spine. If this were done, then the librarian's response to a subsequent inquiry about the book would reveal a great deal about the library's tagging system. If the library used the red–blue (Cartesian) system, then the librarian would claim not to know the book's contents; but if the library used the tagged–untagged (Spinozan) system, then the librarian would show a profound tendency to mistake such new and unassessed books for nonfiction. The translation of this metaphor should be clear: When a multiple-exit Spinozan system is unable to assess the veracity of information, it should consider the information true. In short, the two procedures shown in Figure 1 will produce identical products when run to completion, but very different products when truncated. With this fact in mind, it is possible to begin asking the question toward which this discussion has been driving: Are human beings Cartesian or Spinozan systems? A preliminary answer to this question may be found by examining systems in which the belief process has not yet grown up and by examining systems in which the belief process has been broken down.

The Construction of Doubt

In *The Analysis of Mind* (1921), Bertrand Russell specifically denied Spinoza's position, but he did so with more than a hint of reluctance: "There is a great deal to be said in favour of this view," wrote Russell, "and I have some hesitation in regarding it as inadequate" (p. 249). Indeed, after denouncing the Spinozan hypothesis, Russell proceeded to marshal evidence in its support. In particular, he noted that the hypothesis "is recommended by the way in which it accords with mental development. Doubt, suspense of judgment and disbelief all seem later and more complex than a wholly unreflecting assent" (p. 249).

Indeed, thinkers of all stripe have suggested that doubt is less quickly and easily acquired than belief (see Flavell, 1985; Lundholm, 1936; Piaget, 1962). As early as 1859, Bain contrasted the "primitive credulity" of human beings with their "acquired skepticism" and noted that "we begin by believing everything; whatever is, is true" (p. 511). Similar sentiments were expressed by such eminent contemporaries as William James (1890, p. 318), G. F. Stout (1909a, pp. 111–114; 1909b, pp. 258–260), and William McDougall (1923), who noted, "It is only gradually that we learn to suspend judgment, to doubt" (p. 375). But it was the Scottish philosopher, Thomas Reid (1764/1895), who made this point most explicitly:

> If credulity were the effect of reasoning and experience, it must grow up and gather strength, in the same proportion as reason and experience do. But if it is the gift of Nature, it will be strongest in childhood, and limited and restrained by experience; and the most superficial view of human life shews, that the last is really the case, and not the first. (p. 197)

Two observations provide some empirical support for these contentions. First, the ability to deny propositions (i.e., truth-functional negation) is, in fact, one of the last linguistic abilities to emerge in childhood (Bloom, 1970; Pea, 1980). Although very young children may use the word *no* to indicate rejection or nonexistence, the denial function of the word is not mastered until quite a bit later. To the extent that linguistic competences reflect underlying cognitive capabilities, this suggests that disbelieving propositions may be more demanding than be-

lieving them. Second, young children (who lack truth-functional negation) are particularly suggestible—that is, they are generally prone to accept propositions uncritically (see Ceci, Ross, & Toglia, 1987; cf. Loftus & Davies, 1984). Although such hypersuggestibility is surely exacerbated by the child's inexperience and powerlessness (Dodd & Bradshaw, 1980; Messerschmidt, 1933), young children are more suggestible than older children even when such factors are controlled (Ceci et al., 1987).

Frankly, the ingenuousness of children may seem too self-evident to warrant extended remark. Yet, a moment's reflection reveals that this self-evident fact does not fit well with Descartes's hypothesis. If acceptance and rejection are merely two sides of the same psychological coin, as the Cartesian perspective asserts, then we should expect children to learn to believe and disbelieve at about the same time and with equal ease. Yet, even a "most superficial view" reveals that this simply isn't so. Children are especially credulous, especially gullible, especially prone toward acceptance and belief—as if they accepted as effortlessly as they comprehended but had yet to master the intricacies of doubt. In short, human children do precisely what one would expect of immature Spinozan (but not Cartesian) systems.

The Destruction of Doubt

The ontogeny of belief is at least consistent with the idea that unacceptance is a more difficult operation than is acceptance. Not only does doubt seem to be the last operation to emerge, but it also seems to be the first to disappear. Investigations of the experimentally induced breakdown of belief offer clear and convincing evidence in favor of this view.

Persuasive propositions. Throughout the course of history, governments have routinely practiced the art of changing minds. Many of the techniques they have developed trade on the assumption that beliefs are most easily instilled when the believer's cognitive resources have been depleted. Political prisoners, for example, are commonly deprived of sleep because it is thought that mental fatigue will facilitate their rapid indoctrination. As one victim of the Maoist thought-reform programs recalled of his own "reeducation" in China, "You are annihilated, exhausted, you can't control yourself, or remember what you said two minutes before. You feel that all is lost. From that moment the judge is the real master of you. *You accept anything he says* [italics added]" (Lifton, 1961, p. 23).

The prisoner's acceptance is not, of course, limited to those propositions put forth by interrogators or judges. Indeed, a ubiquitous feature of belief-induction programs is the forced confession, a remarkably effective technique that requires resource-depleted prisoners to tell themselves the propositions that their captors wish them to embrace. Prisoners are coerced to write and recite the appropriate political doctrines and, eventually, even the most intransigent prisoner begins to experience some difficulty doubting his or her own words (Lifton, 1961, pp. 38–64; see also Bem, 1966).

The facilitating effects of resource depletion on persuasion are, of course, more than anecdotal. One common method of depleting processing resources in the laboratory requires subjects to perform several tasks simultaneously. Logic dictates that when a person performs two or more resource-consuming tasks at the same time, then he or she must devote a smaller portion of his or her total processing capacity to each of the tasks. In this vein, Festinger and Maccoby (1964) demonstrated that subjects who listened to a counterattitudinal communication while attending to an irrelevant stimulus were particularly likely to accept the propositions they comprehended (see Baron, Baron, & Miller, 1973, for a review of such studies). More exacting research has replicated this phenomenon and specified its mediators: When resource-depleted persons are exposed to doubtful propositions (i.e., propositions that they normally would disbelieve), their ability to reject those propositions is markedly reduced (see Petty & Cacioppo, 1986, for a review).

Once again, an obvious fact of psychological life deserves close scrutiny. Why should resource depletion facilitate believing? Why does the fatigued or distracted individual seem especially susceptible to persuasion? Current models of persuasion (e.g., Chaiken, 1987; Petty & Cacioppo, 1986) have suggested that resource depletion specifically disables cognitive elaboration—the tendency for persons to attempt logical rejections of the doubtful propositions they comprehend. But why should impairment of such a mechanism necessarily engender acceptance? Indeed, if rejection efforts were to follow on the heels of "mere" comprehension (i.e., representation of meaning without acceptance), then the disabling of these efforts should leave the person without a belief of any sort. Such a person should understand what he or she has heard, but should not be particularly inclined toward or against it. In short, resource-depleted Cartesians should be uncertain, uncommitted, but not persuaded.

It would seem, then, that for models of persuasion to make sense, they must implicitly assume that acceptance occurs prior to or more easily than rejection, or both, and that as a result, this initial acceptance remains even when subsequent attempts at rejection are experimentally impaired. This assumption is, of course, the very heart of the Spinozan hypothesis: When one disables the assessment mechanism that follows Spinozan comprehension, one should, in fact, find a person with a belief in the comprehended proposition. One should find a person who has been artificially reduced to a state of primitive credulity, a state of judgmental innocence in which everything that is, is true. And this is what one finds.[2]

[2] I avoid a discussion of how resource-depleted persons react to true propositions. As Figure 1 shows, the Spinozan procedure certifies (or further accepts) true propositions that are initially accepted at the comprehension stage. This suggests that resource-depleted persons should believe a true proposition less than should normal persons. In fact, just such an effect has been documented by Petty, Wells, and Brock (1976). It is important to note that in these studies, resource depletion did not decrease belief in true propositions; rather, it prevented belief from increasing as much as it otherwise would have. I thank Richard Petty for making the raw data from these studies available to me for reanalysis.

Autobiographical propositions. When an acquaintance says that all Democrats are atheists or that squash is the sport of kings, one may determine not only whether these contentions are true, but also whether the acquaintance believes them to be true. It may be, for example, that the acquaintance is merely trying to get a rise out of a nearby liberal or a raise out of a squash-playing superior, and that he or she does not really believe either of these claims. In effect, when a speaker asserts that "This is so," he or she also asserts (quite independently) that "I *believe* that this is so." Such autobiographical assertions are not limited to verbal declarations of belief; virtually any human action can be construed as inherently containing a propositional description of the actor (e.g., a smile says "I am happy," a vote says "I am liberal," and so on).

How do people react to others' verbal and behavioral self-descriptions? After decades of research activity, both the lie-detection and attribution literatures have independently concluded that people are particularly prone to accept the autobiographical propositions implicit (or explicit) in others' words and deeds (for reviews of these literatures see, respectively, Zuckerman, Depaulo, & Rosenthal, 1981, and Jones, 1979). What makes this phenomenon so intriguing is that people accept these assertions even when they know full well that the assertions stand an excellent chance of being wrong. So, for example, when subjects listen to a person read aloud a position statement (e.g., "I am in favor of the federal protection of armadillos"), they generally assume that the reader's autobiographical claim is true—despite the fact that it was delivered at the behest of a powerful authority. This robust tendency—variously known as the *fundamental attribution error* (Ross, 1977), the *correspondence bias* (Gilbert & Jones, 1986), or *the truthfulness bias* (Zuckerman et al.)—is precisely the sort that a resource-depleted Spinozan system should display. A Spinozan system should accept another's implicit self-description as part of the comprehension of the other's action and should subsequently unaccept that self-description only if the system has both the time and energy necessary to do so. Gilbert, Pelham, and Krull (1988) have shown that this tendency is, in fact, exacerbated by the experimentally induced depletion of resources (see also Gilbert, 1989; Gilbert, Krull, & Pelham, 1988; Gilbert & Osborne, 1989).

The Language of Doubt

Despite their considerable differences, both the Cartesian and Spinozan positions contend that belief begins with the comprehension of propositions. How people comprehend propositions is, of course, a primary province of psycholinguistic research, and several of the phenomena in this area speak to the present concerns.

Components of propositions. A fundamental assumption of psycholinguistic research is that "complexity in thought tends to be reflected in complexity of expression" (Clark & Clark, 1977, p. 523) and, conversely, that "simplicity in cognition is reflected in simplicity of (linguistic) form" (Lakoff, 1987, p. 60; cf. Geach, 1980). The *markedness* of a word is usually considered the clearest index of such linguistic complexity, and it can be determined by two simple criteria (Greenberg, 1966). First, unmarked words (such as *happy*) tend to have fewer morphemes than do their marked counterparts (such as *unhappy*). Second, whereas unmarked words may be used neutrally, marked words generally cannot. For example, the neutral inquiry "How *happy* are you?" presupposes no particular degree of happiness on the part of the listener, whereas the inquiry "How *unhappy* are you?" clearly reveals the speaker's prior belief that the listener is, in fact, sad. This is because *happy* (unmarked) and *unhappy* (marked) are thought of as degrees of *happiness* rather than degrees of *unhappiness,* just as *long* (unmarked) and *short* (marked) are considered degrees of *length* rather than of *shortness.* Both marked and unmarked words describe a particular point on a dimension, but in addition, unmarked words may be used neutrally to describe the dimension itself.

Unmarked terms are thought to describe operations that are more conceptually basic than their marked counterparts. The Spinozan hypothesis states that unacceptance is a more complex operation than is acceptance and, interestingly enough, the English words that indicate the acceptance of ideas are generally unmarked with respect to their antonyms. Thus, people speak of propositions as *acceptable* and *unacceptable,* but (unless one is a neologizing psychologist) not as *rejectable* and *unrejectable.* One's statements may be *true* or *untrue,* but they may not be *false* and *unfalse.* People hope their ideas are *correct, accurate,* and *credible* rather than *incorrect, inaccurate,* and *incredible,* but they cannot grammatically wish to be *unwrong.* Indeed, people even speak of *belief* and *disbelief* more naturally than they speak of *doubt* and *undoubt.* To the extent that one's words for mental processes do reveal something about the processes themselves, the structure of the English lexicon suggests (as did Spinoza, who wrote in Latin) that the rejection of false ideas is more complex than the acceptance of true ones.

Verification of propositions. These observations about the structure of words are complemented by studies of how words are understood. Research on sentence verification (i.e., how people decide if a sentence correctly characterizes its referents) has shown that people are generally quicker to assess the veridicality of true than false affirmative sentences (Gough, 1965). Clark and Chase (1972) have explained this finding with a "paramorphic" model (i.e., a model that produces the same output as a human being, but does not necessarily mimic the methods by which the human being generated that output. See also Carpenter & Just, 1975; Trabasso, Rollins, & Shaughnessey, 1971). According to Clark and Chase's model, when people process assertions (a) they mentally represent both the assertion (e.g., the sentence *an armadillo's eyes are above its nose*) and the state-of-affairs it purports to describe (e.g., a picture of an armadillo), and (b) they then compare the two representations and compute a "truth index." This simple model makes very

clear predictions about how long it should take a person to verify a variety of logically identical but syntactically different sentences, and research has shown that these predictions are excellent approximations of empirical data. Yet, what enables this model to work so well is the asymmetrical fashion in which the comparison stage is assumed to be executed:

> First, people start with the truth index set to *true* . . . (and) second, they compare the two representations. If the two match in every respect, the truth index is left alone. If the two do not match . . . (then) the truth index is changed to its opposite, *false.* (Clark & Clark, 1977, p. 103)

In other words, all sentences are initially coded as true, and the comparison stage leads to further mental work only when the sentence turns out to be false. This feature of the Clark and Chase model is what enables it to predict listeners' reaction times so well. Spinoza would simply have added that this model is not paramorphic; it is perfectly descriptive.

Self-assessing propositions. There is something unusual about linguistic denials. *Homer's armadillo is not male* differs from *Homer's armadillo is female* in more than superficial structure; indeed, the word *not* seems to be of a different "logical type" than the other words in the sentence (Russell, 1919, pp. 131–143; Tarski, 1969) and is probably best thought of as a metalinguistic instruction to the listener to reject the affirmative proposition that the remainder of the sentence puts forth. Horn (1989) neatly characterized this view: "Every negative statement presupposes a corresponding affirmative . . . but not vice versa. Negation is consequently a second-order affirmation: Negative statements are about positive statements, while affirmatives are directly about the world" (p. 3). Indeed, a linguistic denial is "like an affirmative supposition and its cancellation all rolled into one" (Clark & Clark, 1977, p. 98).

All of this means that to comprehend a denial (*armadillos are not herbivorous*), a listener must first comprehend the core assertion (*armadillos are herbivorous*) and then reject it. As Russell (1948, p. 122) noted, "When I say truly 'This is not blue' there is, on the subjective side, consideration of 'This is blue,' followed by rejection" (p. 122). Tesniere (1959) made the same point: "Before denying the contents of a sentence, the mind must first affirm it" (p. 225).[3] A number of experimental studies have substantiated the psychological reality of such philosophical contentions (see chap. 3 of Horn, 1989, for a review).

If denials are self-assessing propositions in that they offer both a core assertion and an instruction to reject that assertion once it is comprehended, then the Spinozan hypothesis makes a unique prediction about them. Because denials require the initial comprehension of that which is being denied and because Spinozan comprehension always entails acceptance of that which is comprehended, then a Spinozan listener who comprehends a denial should momentarily believe the very state of affairs that he or she is being instructed not to believe—more so than should someone who is hearing nothing at all. So, for example, upon learning that *Homer's armadillo is not aroused,* a Spinozan listener should comprehend (and thus accept) the core assertion *Homer's armadillo is aroused,* and should only then be able to finish the processing of this self-assessing proposition by rejecting the core assertion. This means that Spinozan listeners should occasionally end up believing the very assertions they hear denied!

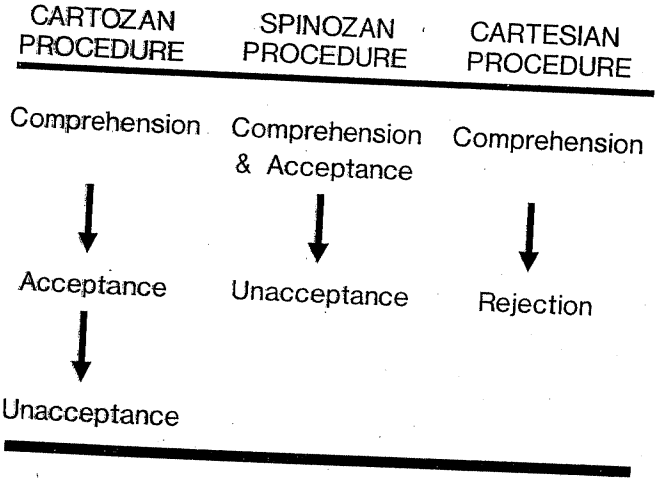
In a series of investigations, Wegner, Wenzlaff, Kerker, and Beattie (1981) showed that subjects who read denials such as *Bob Talbert not linked to Mafia* were in fact left with more negative impressions of the fictitious Talbert than were subjects who read a neutral assertion such as *Bob Talbert celebrates birthday.* If this tendency for people to believe denied information is, in fact, a result of their use of the Spinozan procedure, then it should be exacerbated by resource depletion. Gilbert, Krull, and Malone (1990) asked subjects to learn a fictitious vocabulary by reading assertions (e.g., *a monishna is an armadillo*) whose veracity was either affirmed or denied by the subsequent presentation of an assessment word. On some trials, subjects' resources were momentarily depleted by having them quickly identify a musical tone that followed the assessment word. Whereas a Cartesian learner should have been able to comprehend each assertion and then wait for the assessment word to determine whether to accept or reject that assertion, a Spinozan learner should not have had that prerogative. A Spinozan learner should have accepted the assertion as it was comprehended and then, if that assertion was denied, unaccepted it. Indeed, a unique pattern of errors emerged: Resource depletion did not cause subjects to believe that affirmed propositions were false, but it did cause them to believe that denied propositions were true. This is precisely what one would expect to happen if unacceptance was disrupted by the musical tone task. The Cartesian model predicts no such asymmetry in response to resource depletion during assessment.

The Control of Belief

Up to this point, evidence for the Spinozan procedure has consisted largely of evidence for the precedence of acceptance over rejection. Clearly, acceptance is not merely the flip side of rejection. However, the Spinozan hypothesis is really composed of two separable claims—the first concerning the temporal precedence of acceptance over rejection (the *asymmetry hypothesis*), and the second concerning the unity of comprehension and acceptance (the *unity hypothesis*). These hypotheses are conceivably independent. It is possible, for example, to imagine a system in which acceptance occurs prior to rejection (as it does in the Spinozan system) and yet re-

[3] I thank Benne Willerman for translating this passage from the French.

Figure 2
Comparison of Cartozan, Spinozan, and Cartesian Belief Procedures



mains distinct and separate from comprehension (as it does in the Cartesian system).

A hybrid "Cartozan" processor of this sort is illustrated in Figure 2. Such a system would display virtually all of the effects documented hitherto. Specifically, the principle of premature output would cause such a system to show a bias toward accepting whatever propositions it assessed; however, the Cartozan system would not necessarily assess those propositions it comprehended! In short, the fact that assessment is asymmetric does not mean that acceptance is an automatic consequence of comprehension. As such, the research discussed thus far may impugn the Cartesian model, but by and large it supports the Cartozan and Spinozan models equally well. The question then remains: Can *mere* comprehension occur? That is, can one understand a proposition without representing it as true?

According to some theorists, the answer is by definition *no*. Comprehension is the process by which the meaning of a proposition is mentally represented, and many theorists have found it difficult to define this process without reference to the veracity of the proposition. Thus, Johnson-Laird (1988) noted that to comprehend a proposition one must "imagine how the world should be granted its truth" (p. 110), and Dowty, Wall, and Peters (1981) argued that "to give the meaning of a sentence is to specify its truth conditions" (p. 4). Rips and Marcus (1977) were even more explicit in claiming that the comprehension of a sentence involves "creating a temporary context in which the sentence is true" (p. 192). Indeed, many of us feel that the meaning of an aphorism such as *oppression is the mother of liberty* is not clear until we have represented the aphorism in a way that renders it true ("Ah yes—the yearning for freedom becomes acute when freedom is denied"). In some sense, to generate a proposition's meaning is to consider it so.

Nonetheless, other theorists have offered definitions that preserve the possibility of mere comprehension. For example, Bransford and Johnson (1973) argued that comprehension "may involve options such as whether or not to judge the truth value of a statement or presuppose its truth value" (p. 433). Such statements suggest that comprehension without acceptance is, at the very least, a viable cognitive option. If this is so, then one would expect people who are faced with erroneous, suspect, unsubstantiated, or otherwise invalid information to exercise such options; specifically, they should be able to comprehend the invalid information without accepting it.

How *do* people deal with invalid information? Research shows that people often use information that they clearly recognize as invalid. Ross, Lepper, and Hubbard (1975) gave subjects feedback about their performance on an ability-linked task (namely, the discrimination of actual from bogus suicide notes) and later confessed that the feedback had been unrelated to the subject's actual performance. Despite the experimenter's confession, subjects persisted in believing the invalid feedback. Similar results have been documented with many different kinds of invalid information (e.g., Anderson, Lepper, & Ross, 1980; Valins, 1972; Walster, Berscheid, Abrahams, & Aronson, 1967). Interestingly, one of the few effective methods for ameliorating belief in such false propositions is to ask subjects to comprehend the false proposition's logical opposite (Anderson, 1982).

These findings fit with a larger body of research that suggests that people are particularly poor at ignoring, forgetting, rejecting, or otherwise failing to believe that which they have comprehended (Bjork, 1972; Schul & Burnstein, 1985; Wyer & Budesheim, 1987; Wyer & Unverzagt, 1985). Yet, in virtually all such investigations, subjects have used information whose invalidity was revealed only *after* it was comprehended. Cartozan, Spinozan, and even Cartesian systems are all capable of displaying such effects. Spinozan systems are unique, however, in that they should occasionally believe and use information even when they have *prior* knowledge of the information's invalidity.

Wegner et al.'s (1985) extension of the Ross et al. (1975) study suggested that being forewarned in this way is not necessarily to be forearmed. Rather than confessing to subjects *after* giving them phony feedback on a suicide-note detection task, Wegner et al. (1985) told subjects about the invalidity of the feedback *before* the subjects received it. Despite this clear forewarning, subjects believed the arbitrary feedback they were given (cf. Fleming & Arrowood, 1979; Kiesler & Kiesler, 1964). Similarly, Gilbert et al. (1990) showed subjects a series of smiling faces (i.e., icons of the proposition *I am happy*) and told subjects that each face was expressing either true or false happiness. On some trials subjects received these assessments before seeing the face, and on other trials they received the assessment only after seeing the face. As the Spinozan hypothesis predicted, resource depletion caused subjects to misidentify false faces as true, but not vice versa. Moreover, the timing of the assessment made absolutely no difference whatsoever. Resource-depleted subjects mistook false faces for true ones, and they were

just as likely to do so when informed of the face's falseness *before* seeing it as when informed only after seeing it. In both of these experiments, then, subjects were apparently unable to merely comprehend; rather, they seemed to accept invalid information whose invalidity was known prior to its comprehension. A Spinozan, but not a Cartesian or Cartozan, system should behave in just this way.

It is important to note that in each of the aforementioned studies, subjects were notified that they were to receive invalid information, and this presumably motivated them to exercise their cognitive option and withhold acceptance if they could. Gilbert et al. (1990) took a more direct approach to this problem. Subjects first learned about an imaginary animal called a *glark,* and then tried to decide if some new propositions about glarks were true or false. As they did this, subjects were occasionally asked to quickly read a statement about glarks, but not to assess it. Thus, rather than motivating subjects to withhold acceptance by presenting them with invalid information, subjects in this experiment were asked to merely read a few statements—and a premium was put on the speed with which they could do so. Reaction-time data indicated that subjects responded to this premium and did not assess the propositions that they were instructed merely to read. Yet, when later asked about the veracity of the statements they had merely read, subjects tended to regard them as true (cf. Tousignant, Hall, & Loftus, 1986). These data suggest that subjects were unable to represent the statements in a truth-neutral fashion, even when directly motivated to do so.

The Suspension of Belief

The unity hypothesis is clearly the more difficult of the two Spinozan postulates. The hypothesis states that acceptance is part of comprehension, and therefore when one represents the meaning of a proposition, that proposition is immediately believed. But how can one demonstrate this hypothesis? There are, in general, two strategies. First, one can show that when people comprehend information that they are motivated not to believe (either because they know the information is false or because they have been given reason not to linger), they act as though the information was true anyway. This would suggest that people are unable to decouple acceptance and comprehension, even when it would be propitious to do so. Evidence of this sort was just discussed, and although it is not voluminous, what evidence there is tends to support the Spinozan hypothesis.

A second strategy is to show that when people comprehend information whose truth value cannot be determined via assessment, they act as though they believe the information to be true. For example, upon hearing that *armadillos relish imported cheese,* most people will find in memory neither evidence for nor evidence against this proposition and thus may simply wish to reserve judgment about the matter. If assessment revealed no evidentiary basis for the acceptance or rejection of a proposition, then both Cartesian and Cartozan systems should be able to hold that proposition in abeyance—to suspend both acceptance and rejection—until decisive evidence can be acquired. A Spinozan system, of course, cannot enjoy this particular convenience. As Spinoza (1677/1982) himself argued:

> Experience seems to tell us most indisputably that we are able to suspend judgment so as not to assent to things that we perceive . . . (and) that the will, that is, the faculty of assenting, is free, and different from the faculty of understanding . . . (but) I reply by denying that we have free power to suspend judgment. (pp. 97 and 99)

The unity hypothesis, then, suggests that a mental system *must* believe that which it comprehends, and that it may not abstain. If people are Cartesian or Cartozan systems, then they should be able to avoid both accepting and rejecting when an "internal audit" provides inconclusive evidence. If, on the other hand, people are Spinozan systems, then they should retain their original acceptance of a proposition when the results of a subsequent assessment do not provide information with which to unaccept the proposition. So what happens when people comprehend propositions whose veracity they cannot evaluate?

Researchers who have investigated the ways in which human beings entertain "mere possibilities" have documented a robust tendency for people to seek information that confirms those possibilities (Snyder & Swann, 1978; see Klayman & Ha, 1987, for a review). Although this is not an optimal strategy under most circumstances, there is an exception: A confirmatory search is reasonable when one already believes the hypothesis to be tested because, under such circumstances, confirmatory searches usually prove more informative than do other kinds of searches (Trope & Bassock, 1982). For example, it is quite reasonable to ask questions about a person's extroversion if one already has reason to suspect that the person is an extrovert because (a) such questions can help one determine just *how* extroverted the person is, and (b) interrogating an extrovert about the details of his or her shyness is generally a poor way to gain information. In short, confirmatory information is subjectively more informative than neutral information. It is interesting to note in this regard that when Swann and Giuliano (1987) asked people to entertain a hypothesis and then to rate the diagnosticity of relevant information, they found "direct support for the notion that simply entertaining a belief elevates the perceived informativeness of evidence that may confirm that belief" (p. 522). This suggests that in the process of entertaining a hypothesis, subjects may actually embrace it—thus raising the perceived diagnosticity of confirmatory evidence.

If people seek confirmatory evidence because they embrace the hypotheses they are testing (and not because they are simply inept at testing them), then this tendency should be annulled when people are led to entertain both a hypothesis and its contrary. If, for example, a person is asked merely to comprehend the proposition *John's armadillo is an introvert* before testing the proposition *John's armadillo is an extrovert,* then the person should

not conduct a confirmatory search. Lord, Lepper, and Preston (1984) have demonstrated just such a phenomenon. When subjects were led to think about another person as an introvert before testing the possibility that the person was an extrovert, they showed no tendency to conduct a confirmatory search. In fact, this simple manipulation was far more effective than was a straightforward appeal to hypothesis testers to be fair and accurate.

Conclusions

The foregoing review of research on human belief attests to the truth of Bain's (1859) dictum: "The great master fallacy of the human mind is believing too much" (p. 513). People are especially prone to accept as true the things they hear and see—but why is this so? The explanation examined here is that people are Spinozan systems that, when faced with shortages of time, energy, or conclusive evidence, may fail to unaccept the ideas that they involuntarily accept during comprehension. Evidence for the asymmetry hypothesis (which distinguishes the Spinozan and Cartozan models from the Cartesian model) is quite strong: Acceptance and rejection are not merely alternative outcomes of a single assessment process, but rather, acceptance is psychologically prior. Evidence for the unity hypothesis (which distinguishes the Spinozan from the Cartozan model) is somewhat harder to come by, but at least it is consistent: It is not unreasonable to suggest that the temporary acceptance of a proposition is part of the nonvoluntary process of comprehension itself.

Yet, even if one endorses the Spinozan hypothesis, there is no denying that this method of accepting and then unaccepting false ideas is a somewhat awkward one, in terms of both its procedural redundancy and its susceptibility to errors. Why do human beings use this ungainly belief procedure rather than its esthetically superior alternatives?

The Origin of Belief

"My thinking," wrote William James (1890), "is first and last and always for the sake of my doing" (p. 333). What James said of cognition is even truer of perception. Organisms are immersed in a world of activity, and the perceptual system enables them to play their parts quickly and well. One of the ways in which the system accomplishes this end is by using the outputs of early stages of processing to guide urgent action. Organisms immediately believe what they see and only question their percepts subsequently and occasionally (see Bargh, 1989). Perception can afford to work this way because the correlation between a perceptual representation (i.e., one's mental image of an object) and the presence of that object is nearly perfect. Organisms need not question percepts, because percepts are for the most part faithful representations of reality.

Perception, then, is quintessentially Spinozan, and it would probably be fair to say that this article has been a meditation on one simple suggestion: *As perception construes objects, so cognition construes ideas.* In both cases, the representation of a stimulus (an object or idea) is *believed*—that is, empowered to guide behavior as if it were true—prior to a rational analysis of the representation's accuracy. But perception is not merely a convenient metaphor for cognition; indeed, the two processes may be similar because one (namely, the propositional system of representation that underlies cognition) is an evolutionarily outgrowth of the other (namely, the imaginal system of representation that underlies perception). If this were so, then the cognitive system (which represents ideas *about* reality) could be expected to retain vestiges of the perceptual system (which represents features *of* reality), for example, the tendency initially to treat all representations as if they were true. In other words, a newly evolved cognitive system might treat socially communicated propositions as if they were visually transmitted images, believing what it comprehends (*Life is a bowl of cherries*) just as immediately and thoroughly as it believes what it sees (*This is a bowl of cherries*). Perhaps it is not entirely absurd to think of human understanding as a sensory system that uses the propositional assertions of others as data—a kind of "vicarious observation" that is "the first and greatest human device for stepping up the observational intake" (Quine & Ullian, 1978, p. 51).

Of course, a system that believes its representations prior to assessing them (e.g., the perceptual system) can only work if those representations are largely accurate. It is certainly true that human societies impose a moral onus on lying (Bok, 1978), which results in a strong tendency for persons to communicate information that they believe to be accurate (Clark, 1984; Grice, 1975) and for persons to assume that others will do the same (Swann, Giuliano, & Wegner, 1982). As Dennett (1981) noted: "The faculty of communication would not gain ground in evolution unless it was by and large the faculty of transmitting true beliefs" (p. 18). Thus, just as perceptual systems enable timely action by capitalizing on the fact that most percepts are faithful, cognitive systems may achieve similar efficiency by capitalizing on the fact that most propositions are true. One might even argue that the savings of time and energy outweighs the intellectual deficits of inaccurate beliefs (Swann, 1984).

In the scheme of things, however, the laws of light reflectance should prove a great deal more reliable than the convention of truth telling, and thus a cognitive system (unlike a perceptual system) would have to develop some extra machinery for rejecting the representations it has accepted. If one were designing a cognitive system from scratch, then a perfectly rational Cartesian assessor might well be the optimal sort of machinery to build. But nature does not start from scratch; rather, she is an inveterate jury rigger who rarely invents a new mechanism to do splendidly what an old mechanism can be modified to do tolerably well. The Spinozan belief procedure may be just such a jury-rigged mechanism—a cognitive procedure that retains some features of the perceptual procedures from which it emerged (acceptance upon comprehension) and adds some new features of its own (subse-

quent unacceptance). Is the cognitive system Spinozan because it arose from the perceptual system? No one can say. But the commonalities are certainly suggestive. It may not be merely coincidental that the language of visual perception is so easily used to describe the comprehension of ideas, if you *see* the point (Lakoff & Johnson, 1980, p. 48).

Summary and Reflection

René Descartes did not think that God should be responsible for human folly, and so he reasoned that people must exercise free will when they choose to accept an erroneous idea. Baruch Spinoza, a Jewish scholar who was excommunicated for his trouble, was not so charitable toward the designer of the human mind. Spinoza argued that comprehending an idea did entail accepting that idea, however briefly. "Will and intellect are one and the same thing," he wrote, and thus, "I deny that a man makes no affirmation in so far as he has a perception" (1677/1982, pp. 97 and 99). Although Descartes's assumptions about the symmetry of acceptance and rejection and the disunity of comprehension and belief have silently dominated scientific thinking about these issues, psychological evidence suggests that Spinoza's hypotheses may have been closer to the truth. Findings from a multitude of research literatures converge on a single point: People are credulous creatures who find it very easy to believe and very difficult to doubt. In fact, believing is so easy, and perhaps so inevitable, that it may be more like involuntary comprehension than it is like rational assessment. As Reid (1764/1895) wrote:

> When I perceive a tree before me, my faculty of seeing gives me not only a notion or simple apprehension of the tree, but a belief of its existence, and of its figure, distance, and magnitude; and this judgment or belief is not got by comparing ideas, it is included in the very nature of the perception. (p. 209)

Such mandatory and immediate affirmation may make evolutionary sense when the mental representation concerns only the physical presence and properties of a stone, an animal, or a tree. The human eye is an exceptionally reliable instrument and it would be expensive, even foolhardy, to regularly question what it tells us. Nonetheless, the tendency to accept our mental representations of things before we assess them may spill over into the process by which we comprehend ideas as well. Just as we are compelled to believe immediately that which we see, so too may we be compelled to believe that which we hear and say. In either case, our belief may last for only a moment before it is unraveled by reason: Is that really water in the road ahead, or is it just a mirage? Does she really love me, or did she merely say so? If we comprehend words of love as we see the gloss of dark water, then we may thinkingly reject the reality of either. But only after we have first accepted them as so.

REFERENCES

Anderson, C. A. (1982). Inoculation and counter-explanation: Debiasing techniques in the perseverance of social theories. *Social Cognition, 1,* 126–139.

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology, 39,* 1037–1049.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 7, pp. 47–89). San Diego, CA: Academic Press.

Bain, A. (1859). *The emotions and the will.* London: Longmans, Green.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford.

Baron, R. S., Baron, P., & Miller, N. (1973). The relation between distraction and persuasion. *Psychological Bulletin, 80,* 310–323.

Bem, D. J. (1966). Inducing belief in false confessions. *Journal of Personality and Social Psychology, 3,* 707–710.

Bennett, J. (1984). *A study of Spinoza's Ethics.* Indianapolis, IN: Hackett.

Bjork, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 217–235). New York: Wiley.

Bloom, L. (1970). *Language development: Form and function in emerging grammars.* Cambridge, MA: MIT Press.

Bogdan, R. J. (1986). The importance of belief. In R. J. Bogdan (Ed.), *Belief: Form, content, and function* (pp. 1–16). Oxford, England: Clarendon Press.

Bok, S. (1978). *Lying.* New York: Random House.

Bransford, J. D., & Johnson, M. K. (1973). Consideration of some problems in comprehension. In W. G. Chase (Ed.), *Visual information processing* (pp. 383–438). San Diego, CA: Academic Press.

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review, 82,* 45–73.

Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility of children's memory: Psycholegal implications. *Journal of Experimental Psychology: General, 116,* 38–49.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (Vol. 5, pp. 3–39). Hillsdale, NJ: Erlbaum.

Clark, H. H. (1984). Language use and language users. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2, pp. 179–231). Reading, MA: Addison-Wesley.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3,* 472–517.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language.* New York: Harcourt, Brace, Jovanovich.

Dennett, D. C. (1981). *Brainstorms: Philosophical essays on mind and psychology.* Cambridge, MA: MIT Press.

Descartes, R. (1984). Fourth meditation. In J. Cottingham, R. Stoothoff, & D. Murdoch (Eds. and Trans.), *The philosophical writings of Descartes* (Vol. 2, pp. 37–43). Cambridge, England: Cambridge University Press. (Original work published 1641)

Descartes, R. (1984). Principles of philosophy. In J. Cottingham, R. Stoothoff, & D. Murdoch (Eds. and Trans.), *The philosophical writings of Descartes* (Vol. 1, pp. 193–291). Cambridge, England: Cambridge University Press. (Original work published 1644)

Dodd, D. H., & Bradshaw, J. M. (1980). Leading questions and memory: Pragmatic considerations. *Journal of Verbal Learning and Verbal Behavior, 21,* 207–219.

Donne, J. (1930). Satyre III. In J. Hayward (Ed.), *John Donne: Complete poetry and selected prose* (pp. 127–130). New York: Random House. (Original work published 1635)

Dowty, D. R., Wall, R. E., & Peters, S. (1981). *Introduction to Montague semantics.* Boston: Reidel.

Festinger, L., & Maccoby, N. (1964). On resistance to persuasive communications. *Journal of Abnormal and Social Psychology, 68,* 359–366.

Flavell, J. H. (1985). The development of children's knowledge about the appearance–reality distinction. *American Psychologist, 41,* 418–425.

Fleming, J., & Arrowood, A. J. (1979). Information processing and the perseverance of discredited self-perceptions. *Personality and Social Psychology Bulletin, 5,* 201–205.
Fodor, J. A. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.
Garfield, J. L. (1987). (Ed.). *Modularity in knowledge-representation and natural-language understanding.* Cambridge, MA: MIT Press.
Gazzaniga, M. S. (1985). *The social brain.* New York: Basic Books.
Geach, P. T. (1980). *Logic matters.* Berkeley: University of California Press.
Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189–211). New York: Guilford.
Gilbert, D. T., & Jones, E. E. (1986). Perceiver-induced constraint: Interpretations of self-generated reality. *Journal of Personality and Social Psychology, 50,* 269–280.
Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology, 59,* 601–613.
Gilbert, D. T., Krull, D. S., & Pelham, B. W. (1988). Of thoughts unspoken: Social inference and the self-regulation of behavior. *Journal of Personality and Social Psychology, 55,* 685–694.
Gilbert, D. T., & Osborne, R. E. (1989). Thinking backward: Some curable and incurable consequences of cognitive busyness. *Journal of Personality and Social Psychology, 57,* 940–949.
Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology, 54,* 733–740.
Gough, P. B. (1965). Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior, 4,* 107–111.
Greenberg, J. H. (1966). *Language universals.* The Hague, The Netherlands: Mouton.
Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol 3. Speech acts* (pp. 41–58). New York: Seminar Press.
Horn, L. R. (1989). *A natural history of negation.* Chicago: University of Chicago Press.
James, W. (1890). *The principles of psychology* (Vol. 2). New York: Holt.
Johnson-Laird, P. N. (1988). How is meaning mentally represented? In U. Eco, M. Santambrogio, & P. Viloli (Eds.), *Meaning and mental representations* (pp. 99–118). Bloomington: Indiana University Press.
Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist, 34,* 107–117.
Just, M. A., & Carpenter, P. A. (Eds.). (1977). *Cognitive processes in comprehension.* Hillsdale, NJ: Erlbaum.
Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.
Kiesler, C. A., & Kiesler, S. B. (1964). Role of forewarning in persuasive communications. *Journal of Abnormal and Social Psychology, 8,* 547–549.
Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis-testing. *Psychological Review, 94,* 211–228.
Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.
Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.
Lehrer, K. (1983). Belief, acceptance, and cognition. In H. Parret (Ed.), *On believing: Epistemological and semiotic approaches* (pp. 172–183). Berlin, FRG: Walter de Gruyter.
Lifton, R. J. (1961). *Thought reform and the psychology of totalism.* New York: Norton.
Loftus, E. F., & Davies, G. M. (1984). Distortions in the memory of children. *Journal of Social Issues, 40,* 51–67.
Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology, 47,* 1231–1243.
Lundholm, H. (1936). *The psychology of belief.* Durham, NC: Duke University Press.
Marr, D. (1982). *Vision.* San Francisco: Freeman.
Messerschmidt, R. (1933). The suggestibility of boys and girls between the ages of six and sixteen. *Journal of Genetic Psychology, 43,* 422–437.
McDougall, W. (1923). *Outline of psychology.* New York: Scribner.
Minsky, M. (1985). *The society of mind.* New York: Simon & Schuster.
Newell, A., & Simon, H. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.
Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44–64.
Nuchelmans, G. (1983). *Judgment and proposition: From Descartes to Kant.* Amsterdam: North-Holland.
Pea, R. D. (1980). The development of negation in early child language. In D. R. Olson (Ed.), *The social foundations of language and thought* (pp. 156–186). New York: Norton.
Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). San Diego, CA: Academic Press.
Petty, R. E., Wells, G. L., & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology, 34,* 874–884.
Piaget, J. (1962). *Play, dreams, and imitation in childhood* (C. Gattegno & F. M. Hodgson, Trans.). New York: Norton.
Price, H. H. (1969). *Belief.* London: Allen & Unwin.
Quine, W. V., & Ullian, J. S. (1978). *The web of belief* (2nd ed.). New York: Random House.
Reid, T. (1895). Inquiry into the human mind. In W. Hamilton, (Ed.), *The philosophical works of Thomas Reid* (Vol. 1, 8th ed., pp. 95–214). London: Longmans, Green. (Original work published 1764)
Rips, L. J., & Marcus, S. L. (1977). Suppositions and the analysis of conditional sentences. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 185–220). Hillsdale, NJ: Erlbaum.
Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). San Diego, CA: Academic Press.
Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attribution processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32,* 880–892.
Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 7–57). Cambridge, MA: MIT Press.
Russell, B. (1919). *Introduction to mathematical philosophy.* London: Allen & Unwin.
Russell, B. (1921). *The analysis of mind.* New York: McMillan.
Russell, B. (1948). *Human knowledge, its scope and limits.* New York: Simon & Schuster.
Schul, Y., & Burnstein, E. (1985). When discounting fails: Conditions under which individuals use discredited information in making a judgment. *Journal of Personality and Social Psychology, 49,* 894–903.
Snyder, M., & Swann, W. B., Jr. (1978). Hypothesis testing processes in social interaction. *Journal of Personality and Social Psychology, 36,* 1202–1212.
Spinoza, B. (1982). *The Ethics and selected letters* (S. Feldman, Ed. and S. Shirley, Trans.). Indianapolis, IN: Hackett. (Original work published 1677)
Stout, G. F. (1909a). *Analytic psychology* (3rd ed., Vol. 1). New York: Macmillan.
Stout, G. F. (1909b). *Analytic psychology* (3rd ed., Vol. 2). New York: Macmillan.
Swann, W. B., Jr. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review, 91,* 457–477.
Swann, W. B., Jr., & Giuliano, T. (1987). Confirmatory search strategies in social interaction: How, when, why, and with what consequences. *Journal of Social and Clinical Psychology, 5,* 511–524.
Swann, W. B., Jr., Giuliano, T., & Wegner, D. M. (1982). Where leading questions can lead: The power of conjecture in social interaction. *Journal of Personality and Social Psychology, 42,* 1025–1035.
Tarski, A. (1969). Truth and proof. *Scientific American, 220,* 63–77.
Tesniere, L. (1959). *Elements de syntaxe structurale* [Elements of structural syntax]. Paris: Klincksieck.

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading post-event information. *Memory & Cognition, 14,* 329–338.

Trabasso, T., Rollins, H., & Shaughnessey, E. (1971). Storage and verification stages in processing concepts. *Cognitive psychology, 2,* 239–289.

Trope, Y., & Bassock, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology, 43,* 22–34.

Tversky, A., & Kahneman, D. (1974). Judgments under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Valins, S. (1972). Persistent effects of information about internal reactions: Ineffectiveness of debriefing. In H. London & R. E. Nisbett (Eds.), *Thought and feeling: The cognitive alteration of feeling states* (pp. 116–124). Chicago: Aldine.

Walster, E., Berscheid, E., Abrahams, D., & Aronson, E. (1967). Effectiveness of debriefing following deception experiments. *Journal of Personality and Social Psychology, 6,* 371–380.

Wegner, D. M., Coulton, G., & Wenzlaff, R. (1985). The transparency of denial: Briefing in the debriefing paradigm. *Journal of Personality and Social Psychology, 49,* 338–346.

Wegner, D. M., Wenzlaff, R., Kerker, R. M., & Beattie, A. E. (1981). Incrimination through innuendo: Can media questions become public answers? *Journal of Personality and Social Psychology, 40,* 822–832.

Wyer, R. S., & Budesheim, T. L. (1987). Person memory and judgments: The impact of information that one is told to disregard. *Journal of Personality and Social Psychology, 53,* 14–29.

Wyer, R. S., & Unverzagt, W. H. (1985). The effects of instructions to disregard information on its subsequent recall and use in making judgments. *Journal of Personality and Social Psychology, 48,* 533–549.

Zuckerman, M., Depaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–59). San Diego, CA: Academic Press.