Unbelieving the Unbelievable
Some Problems in the Rejection of False Information

Daniel T. Gilbert
University of Texas at Austin
Douglas S. Krull
University of Missouri at Columbia
Patrick S. Malone
University of Texas at Austin

ABSTRACT

Spinoza suggested that all information is accepted during comprehension
and that false information is then unaccepted. Subjects were presented
with true and false linguistic propositions and, on some trials, their
processing of that information was interrupted. As Spinoza's model
predicted, interruption increased the likelihood that subjects would
consider false propositions true but not vice versa (Study 1). This was so
even when the proposition was iconic and when its veracity was revealed
before its comprehension (Study 2). In fact, merely comprehending a false
proposition increased the likelihood that subjects would later consider it
true (Study 3). The results suggest that both true and false information are
initially represented as true and that people are not easily able to alter this
method of representation. Results are discussed in terms of contemporary
research on attribution, lie detection, hypothesis testing, and attitude
change.

> 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. – Descartes (1644/1984b, p. 205)

Rene Descartes was right about so many things that he surely deserved to be wrong about something: How people come to believe certain ideas and disbelieve others may be the something about which he was mistaken. Descartes insisted that ideas are initially represented in the mind without reference to their veracity. Thus, upon hearing the utterance "Armadillos may be lured from a thicket with soft cheese," Descartes suggested that the listener's mind simply held that proposition *in aequilibrio,* and only later submitted the proposition to a rational analysis by which it was determined to be true or false (Descartes, 1641/1984a ; see also Bennett, 1984 , pp. 159—167; Nuchelmans, 1983). In the centuries that followed, many psychologists (e.g., Bain, 1859 ; James, 1890) and philosophers (e.g., Reid, 1764/1895 ; Russell, 1921) found cause to question the Cartesian distinction between the mere representation and subsequent assessment of ideas. Nonetheless, modern psychology continues to embrace the twin Cartesian notions that (a) ideas may be mentally represented without any reference to their veracity, and (b) the acceptance and rejection of ideas are the dual outcomes of a single, controllable assessment process.

What are the alternatives to the Cartesian model? If people do not "merely comprehend" ideas before they decide to believe or disbelieve them, then what do they do instead? Baruch Spinoza (1677/1982, pp. 96—101) suggested that all ideas are accepted (i.e., represented in the mind as true) *prior* to a rational analysis of their veracity, and that some ideas are subsequently *unaccepted* (i.e., rerepresented as false). Thus, upon hearing that a sliver of Brie may be used to woo a reluctant armadillo, Spinoza suggested that a listener must momentarily accept this assertion before he or she can reject it as absurd. In short, the mental representation of a proposition or idea always has a truth value associated with it, and by default this value is *true.* This default value remains unaltered when the idea is subsequently assessed to be true, but is changed when the idea is subsequently assessed to be false. [1]

Libraries of the Mind

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 is somehow altered or "tagged" to indicate its truth value–otherwise, people would have to reassess the validity of their knowledge each time they deployed it. Spinoza and Descartes seem to have agreed with this assumption, but disagreed about the precise nature of the tagging system itself. A familiar metaphor may serve to illustrate the key elements of their division.

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. A first method would be to paste a red tag on each volume of fiction and a blue tag on each volume of

nonfiction; a second method would be to paste a tag on each volume of fiction and to leave the nonfiction untagged. Although each method would allow a librarian to discriminate easily between the two types of book, the second method has both a unique advantage and disadvantage. The red—blue system requires that every volume in the library be tagged, and thus demands a great deal more time and effort than does the tagged—untagged system (which requires only the tagging of a few volumes). On the other hand, the efficiency of the tagged—untagged system trades on its accuracy: For example, when a new, untagged volume of fiction arrives on the library shelves, it may be mistaken for nonfiction before it is read.

In a sense, 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), was read (assessed), and was then tagged (rerepresented) as either fiction (false) or nonfiction (true). Because new books (unassessed ideas) lacked a tag, they could not be identified as fiction or nonfiction until they had been read. Such unread books were "merely represented" in the library.

Spinoza felt that the mind was more like a library that used a tagged—untagged system. Books were first represented and then their contents assessed, but, because of the particular tagging system used, a new book that appeared without a tag looked exactly like (and thus was treated as) a work of nonfiction. In Spinoza's library, a book's spine *always* announced its contents, though sometimes erroneously. No book could be "merely represented" in the library because the absence of a tag was just as informative about the content of the book as was the presence of a tag. 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 given a special tag (cf. Wegner, Coulton, & Wenzlaff, 1985 ; see also Clark & Chase, 1972 , 1974 ; Gough, 1965 , 1966).

It is difficult to know which of these models best describes the human mind. Although the Cartesian and Spinozan systems are mechanically distinct, they produce the same conclusions under ideal conditions. For example, if one stood outside a library window and challenged the librarian's knowledge of famous books ("Can you tell me about *Civilization and Its Discontents* without reading it?"), the librarian's response ("That book is nonfiction") would not enable one to determine whether the library used a Cartesian or Spinozan tagging system. In other words, if the Spinozan and Cartesian procedures (shown in Figure 1) were allowed to run to completion, undisturbed, they would produce identical products, and thus these products would not be informative about the nature of the systems that produced them.

Nonetheless, if one could sneak a new book (e.g., *War of the Worlds*) onto the library's shelves and somehow prevent the librarian from assessing its contents and tagging its spine, then the librarian's response to an inquiry about that 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 shout through the window, "I don't know what sort of book this is. Come back tomorrow after it's been read and tagged." If, however, the library used the

tagged—untagged Spinozan system, then the librarian would mistakenly yell, "That one is nonfiction too!"

Study 1: The Hopi Language Experiment

This belabored metaphor suggests that the two procedures shown in Figure 1 will produce identical products when run to completion, but very different products when truncated (cf. the "principle of graceful degradation" in Norman & Bobrow, 1975 ; see also Ben Zur & Breznitz, 1981). The Spinozan hypothesis asserts that rejecting an idea requires the extra step of unaccepting or "tagging" a mental representation as false. If this is so, then people should initially accept both true and false ideas upon comprehension but, when the processing of the idea is interrupted, should not be able to go on and unaccept (or tag) false ideas. As such, interruption should cause Spinozan systems to mistake false ideas for true ones, but not vice versa. The Cartesian model, on the other hand, considers acceptance and rejection to be alternative consequences of a single assessment process, and therefore predicts no asymmetry between true and false ideas as a result of interruption. Study 1 was an initial attempt to test these competing predictions.

Method Overview

In the context of a language-learning experiment, subjects were presented with novel propositions on a computer screen. On most trials, subjects were subsequently informed that the preceding proposition was either true or false. On some of these trials, subjects' processing of the proposition was interrupted by having them quickly perform an unrelated task (namely, pushing a button in response to a tone). Finally, subjects were presented with the original propositions (in question form) and were asked to determine whether they were true or false.

Subjects

Thirty-five female students at the University of Texas participated to fulfill a requirement in their introductory psychology course. Only native speakers of English were eligible to participate.

Instructions

Subjects were invited to participate in an experiment on "how people learn languages when they visit foreign countries." Upon arriving at the laboratory, subjects were escorted to a private room, where they remained for the duration of the experiment. Subjects were seated in front of a microcomputer and given oral instructions. A male experimenter explained that the experiment was designed to simulate the learning of a foreign language as it occurred in a natural (rather than an academic) environment. He noted that an important difference between these two environments was that in the natural environment listeners often hear new words and must guess their meaning from sentential context; later, listeners may find that their guesses were right or wrong, or they may receive no feedback at all.

The learning task.

Subjects were told that to simulate this feature of natural language learning, the computer would present 28 propositions, one at a time, of the form *An X is a Y,* where *X* was a Hopi Indian noun and *Y* was its English equivalent. In fact, the ostensible Hopi noun was always a nonsense word. Subjects were told that shortly after the computer presented the proposition, it would print either the signal word *true* or *false* on the screen to indicate that the preceding proposition had been either accurate or inaccurate. Subjects were told that at the end of the session their knowledge of the Hopi vocabulary would be tested. We used these "nonsense propositions" (rather than real propositions such as *potatoes are grown in Idaho*) simply to ensure that subjects would use the signal word, and not their prior knowledge, to evaluate the truthfulness of the proposition.

Subjects were warned that on some trials no signal word would appear at all. Such trials were included as fillers for two reasons. First, pretesting showed that the inclusion of such trials made it more difficult for subjects to keep track of how many instances of each signal word they had seen and thus inhibited the tendency to respond *true* and *false* to approximately equal numbers of test items. Second, the inclusion of such filler trials enhanced the credibility of the cover story.

The interruption task.

Subjects were told that "some research suggests that the speed of response is related to the ability to learn language." As such, subjects were told that during the experiment a tone would occasionally sound, and that when this happened they were to press a response button as quickly as possible. Ostensibly, this would allow the experimenter to measure their general speed of response (which the computer did, in fact, record) and to correlate this measure with their ability to learn the Hopi vocabulary. In actuality, this task was designed to interrupt subjects' processing of the proposition that was currently being presented on the computer screen.

Procedure

The procedure for Study 1 is depicted in Table 1 . Each subject saw 28 propositions of the form *An X is a Y,* where *X* was a nonsense word (e.g., *twyrin*) and *Y* was a common English noun (e.g., *doctor*). Each proposition appeared for 8 s, and the order of presentation was held constant for all subjects (e.g., the proposition *A twyrin is a doctor* always appeared in the 21st position). The proposition was followed by a blank screen for either 2 or 10 s. [2]For one third of the subjects, the first blank screen was followed for 3 s by the signal word *true;* for one third it was followed by the signal word *false;* and for one third it was merely followed by another blank screen. The identity of the signal word was completely counterbalanced across these three orders so that, ultimately, every one of the 28 propositions was followed by one of the two signal words or by a blank screen.

Of the 28 total trials, the first and last 5 were included as buffers against primacy and recency effects. This left 18 trials, 12 of which were critical trials (6 trials followed by the

signal word *true* and 6 followed by the signal word *false*). On 4 of these 12 trials a 500-Hz tone sounded 750 ms after the presentation of the signal word, and lasted for 390 ms. Of these 4 interrupted critical trials, 2 were followed by the signal word *true* and 2 by the signal word *false.*

Our strategic use of novel propositions may strike some as curious. Recall that our three objectives in this experiment were (a) to allow subjects to comprehend propositions, (b) to control their assessment of these propositions, and (c) to interrupt that assessment in certain cases. Had the propositions referred to our subjects' real-world knowledge (e.g., *pelicans are mammals*), the second (and therefore the third) objective could not have been met.

The Identification Test

Subjects were given practice responding to the tone and were then allowed to go through the vocabulary-learning procedure described above. At the end of the last trial, subjects were given an identification test. Each of the 12 critical propositions was rephrased in the form of a question (e.g., *Is a twyrin a doctor?*), and these questions were printed in random order, one at a time, on the computer screen. In addition, 9 foils (i.e., questions about propositions that were not seen during the learning phase) were presented randomly and intermittently. Subjects were allowed up to 9 s to answer each question by pressing one of four response buttons labeled (respectively) *true, false, no information,* and *never seen.* Subjects were instructed to press the *true* button if the proposition was true, the *false* button if the proposition was false, the *no information* button if they had seen the proposition but had received no information about its veracity, and the *never seen* button if they had never seen the proposition. The computer recorded the subject's response to each item on the identification test; unbeknownst to the subject, the computer also recorded the speed of that response.

Results Identification of Propositions

Subjects read a set of propositions (some of which were interrupted) and later were asked to identify these propositions as true or false. We predicted that interruption would not affect the subject's ability to identify correctly the true propositions; that is, the baseline (uninterrupted) rate of correct identification of true propositions should have been unchanged by interruption. On the other hand, we expected interruption to increase subjects' tendency to misidentify false propositions as true; the baseline (uninterrupted) rate of correct identification of false propositions should have been decremented by interruption.

The proportion of propositions that each subject correctly identified was submitted to a 2 (Signal Word: true or false) × 2 (Interruption: yes or no) within-subjects analysis of variance (ANOVA). [3]This analysis revealed a main effect of signal word, F (1, 32) = 3.86, p = .058, which was qualified by the predicted Signal Word × Interruption interaction, F (1, 32) = 5.30, p = .028. As the left side of Figure 2 indicates, interruption had no effect on the correct identification of true propositions (55% when uninterrupted

vs. 58% when interrupted), but did significantly reduce correct identification of false propositions (55% when uninterrupted vs. 35% when interrupted). The fact that interruption did not affect the correct identification of true propositions is as important as the fact that interruption did affect correct identification of false propositions. Indeed, if correct identification of true propositions had increased with interruption, the results would merely indicate that interruption increases the probability that people will respond *true* to any query.

A similar ANOVA was performed on the proportion of true—false reversals (i.e., true propositions that were misidentified as false, and vice versa). The analysis revealed only the expected Signal Word × Interruption interaction, $F(1, 32) = 3.81$, $p = .06$. As the right side of Figure 2 indicates, when false propositions were uninterrupted, they were misidentified as true (21%) just as often as true propositions were misidentified as false (22%); however, when interrupted, false propositions were misidentified as true (33%) more often than true propositions were misidentified as false (17%). In short, interruption caused an increase in the percentage of false propositions that were misidentified as true, $t(1, 32) = 1.89$, $p = .067$, but not vice versa, $t < 1$.

Latency to Identify Propositions

The Spinozan hypothesis suggests that the processing of a false proposition requires more time and cognitive capacity than does the processing of a true proposition, simply because false propositions must be accepted and then unaccepted (see also Carpenter & Just, 1975 ; Clark & Chase, 1972 , 1974 ; Gough, 1965 , 1966 ; Trabasso, Rollins, & Shaughnessey, 1971). As such, one might expect subjects to respond more slowly to a probe that occurred after a false, rather than a true, proposition. An analysis of the time it took subjects to respond to the interrupting tone revealed such a tendency: Subjects responded somewhat more slowly to interruptions that occurred immediately after the presentation of a false proposition ($M = 577$ ms) than to those that occurred after the presentation of a true proposition, $M = 537$ ms, $t(1, 32) = 1.96$, $p = .058$.

The Problem of Guessing Bias

As described thus far, the data are consistent with the Spinozan hypothesis. Unfortunately, they are also consistent with two alternative explanations that turn on the notion of guessing bias. Although both of these alternatives suggest that subjects may simply have shown a tendency to guess *true* when presented with false items whose processing was previously interrupted, each posits a unique mediator of this bias.

The forgetting hypothesis.

The first alternative explanation is that interruption kept subjects from remembering false propositions (i.e., these items were never entered into long-term memory), and when later confronted with false items that they did not recognize, subjects tended to guess *true*. Subjects' memory for foil items (i.e., items that were not presented during the learning phase, but that were presented during the identification phase) belies this explanation. Of

the true and false items presented, subjects mistook 9% for unseen foils when the processing of those items was interrupted, and 8% for unseen foils when processing was not interrupted, $t < 1$. In other words, interruption did not increase the probability that subjects would mistake items they had seen for items they had not seen. Indeed, when subjects misidentified the unseen foil items, they mistook 8% of them for false propositions and only 3% of them for true propositions, $t(1, 32) = 2.30$, $p = .028$. In other words, when presented with items that could not have been in memory (i.e., the unseen foil items), subjects actually tended to guess *false* significantly more often than they guessed *true*.

The forgetting hypothesis states that (a) interruption prevents false items from being entered into long-term memory and that (b) when presented with such unrecognizable or "forgotten" items, subjects tend to guess *true*. Both of these suppositions are clearly ruled out in the present case.

The uncertainty hypothesis.

The second (and more reasonable) alternative is a variant of the first: Interruption may cause subjects to be uncertain of a false item's truth value (rather than causing them to forget having seen the item entirely), and when subjects experience such uncertainty, they tend to guess *true*. Did interruption raise subjects' uncertainty? One way to measure subjective uncertainty is simply to ask subjects how much confidence they have in their response, but this direct approach is not without serious problems (e.g., Kahneman & Tversky, 1982 ; Nisbett & Wilson, 1977). A more direct and unobtrusive method for assessing a person's certainty about a judgment is to measure his or her hesitation in making that judgment (e.g., see Collins & Quillian, 1969 ; Hyman, 1953 ; Rosch, 1975 ; Snodgrass, 1975 , p. 37). In the present case, this would mean measuring the time it took subjects to identify items as true and false. If subjects were uncertain about a particular identification, they should have been slower to make it.

The uncertainty hypothesis suggests that subjects should have shown greater overall hesitation when identifying previously interrupted items than when identifying previously uninterrupted items. This was clearly not the case. Subjects did *not* take more time to identify true or false items as true or false when the processing of those items had earlier been interrupted ($M = 4,182$ ms) than when the processing of those items had not been interrupted ($M = 4,887$ ms). In other words, subjects did not display greater hesitancy when identifying interrupted than uninterrupted items.

The uncertainty hypothesis suggests that we also examine the hesitation associated with the particular misidentification of false items as true; as Figure 2 shows, this is the critical error that drives our statistical effect. Did interruption cause subjects to feel more uncertain when confronted with a false item (and thus cause them merely to guess that such items were true)? In a word–no. Subjects did not show greater hesitation when they misidentified previously interrupted false items as true ($M = 3,405$ ms) than when they misidentified previously uninterrupted false items as true ($M = 3, 985$ ms). The uncertainty hypothesis states that (a) interruption during processing of an item increases

the subject's subsequent uncertainty about the truth value of that item and (b) when the subject feels uncertain, she tends to guess *true*. According to our hesitation data (which address the problem of uncertainty more directly than do verbal reports), interruption did not increase the subject's uncertainty about the truth value of the item. In short, subjects did not hesitate to make the quintessentially Spinozan mistake of misidentifying interrupted false propositions as true.

Discussion

No one can enter the library of the mind, and thus one may only deduce its holdings from the librarian's reports. Although one must always be cautious when inferring the nature of mental representation from behavioral responses ([Braitenberg, 1984](#) ; [Lloyd, 1989](#) , pp. 3—11), the results of Study 1 are consistent with the idea that people initially represent false information as true. It is not clear, however, whether people are compelled to do so by the nature of the cognitive mechanisms they use or whether they choose to do so because of the nature of the tasks they attempt. Societies place a premium on candor, and it seems likely that the majority of information that people encounter, assess, and remember is, in fact, true. (Indeed, the Spinozan tagging system is advantageous mainly to the extent that a library contains more nonfiction than fiction). Thus, it may be that people generalize from ordinary experience and consciously *assume* that all ideas are true unless otherwise noted. In other words, the initial belief in the truthfulness of information may be a flexible, heuristic assumption.

If people choose, but are not compelled, to represent as true the propositions offered them, then this heuristic assumption should be modifiable. For example, if people find themselves in situations in which they expect to receive false information (e.g., a propagandist speech), then they should be able to alter the default value assigned to incoming information. Individuals in such a position should adopt a "skeptic's set"; that is, they should choose initially to represent ideas as false, and then to recode a select few of them as true. Interruption should cause such skeptics to mistake true propositions for false ones, but not vice versa–precisely the opposite of the effect seen in Study 1. In the language of our metaphor, libraries whose holdings are primarily fiction may wish to modify the typical Spinozan procedure: They may wish to paste colored tags on their nonfiction and leave their fiction untagged. Workers in such a library should mistake untagged books for works of fiction.

The Spinozan position denies the possibility of a "skeptical reversal" of this sort. In its strongest form, the hypothesis states that people must initially represent ideas as true and only later rerepresent some of them as false, and they must conduct business this way regardless of the temporary circumstances in which they find themselves. If people can take changing contexts into account and consciously alter the default value assigned to incoming ideas, then this piece of the Spinozan position is plainly wrong.

Study 2: The Phony Man Experiment

Study 2 was an attempt to determine whether, under the appropriate conditions, people were able initially to represent a proposition as false and then alter that representation when the proposition turned out to be true. In addition, we attempted to extend the results of Study 1 beyond the linguistic domain by using iconic propositions (i.e., facial expressions of emotion rather than vocabulary definitions).

Method Overview

In the context of a lie-detection experiment, female subjects were presented with smiling male faces (i.e., icons of the proposition "I am happy") on a video screen. Subjects were informed either before or after each presentation that the smiling face was expressing either true or false happiness (i.e., that the iconic proposition was true or false). On some trials, subjects' processing of the face was interrupted by having them quickly perform an unrelated tone-discrimination task. Finally, subjects were once again presented with the original faces and asked to determine whether each was expressing true or false happiness.

Subjects

Twenty female students at the University of Texas participated to fulfill a requirement in their introductory psychology course. Only native speakers of English were eligible to participate.

Instructions

Female subjects were invited to participate in an experiment on "learning to detect emotional deception in men." Upon arriving at the laboratory, subjects were escorted to a private room, where they remained for the duration of the experiment. Subjects were seated in front of a video monitor and given written instructions, which explained that the experiment was designed to discover what kinds of information women found useful in determining the sincerity of a man's emotional expressions. Subjects were told that they would participate in both a learning phase (in which they would learn how to detect deceit in men's faces) and an application phase (in which they would apply their newly acquired ability to detect deceit). In fact, there was to be no application phase.

Cover story.

Subjects were told that earlier in the semester, several male volunteers had undergone a mood induction manipulation in which half were made sad and half were made happy. Immediately thereafter, each man had ostensibly been asked to smile while being recorded on videotape. Subjects were told that in the learning phase they would see the faces of 16 of these smiling men and would learn whether each was expressing true or false happiness. Supposedly, this procedure would enable subjects to "learn to identify the subtle facial differences" between sincere and insincere men.

The videotape.

Subjects were told that during the learning phase they would see 16 smiling male faces, one at a time, on the video monitor. Subjects were told that shortly before or after each face appeared, the signal word *true* or *false* would appear on the screen to indicate that the man's happiness had been either true (i.e., he had been made happy before being asked to smile) or false (i.e., he had been made sad before being asked to smile). Subjects were told that the experimenter was interested in discovering whether prior knowledge of a man's sincerity helped women recognize deceit, and thus the signal word would sometimes precede the face and sometimes follow it. Subjects were told that at the end of the session their memory for the sincerity of each face would be tested.

The interruption task.

Subjects were told that at any time during the experiment one of two tones might sound. Subjects were instructed to press a response button marked *high* if the tone was high-pitched and a response button marked *low* if the tone was low-pitched. Subjects were told that the facial lie-detection and tone-discrimination tasks were controlled by different cerebral hemispheres; as such, the experimenter ostensibly did not expect one task to affect the other and had included the tone-discrimination task merely to prove this hypothesis. In actuality, of course, the tone-discrimination task was designed to interrupt subjects' processing of the face and signal word that were being presented on the video screen on that trial. The response buttons were not connected to a computer in this experiment, and thus response times were not recorded.

Procedure

The procedure for Study 2 is depicted in Table 2 . Eight different videotapes were constructed to counterbalance the order of stimulus presentation (shown vertically in Table 2 as Videotapes A—H). Each subject saw one videotape of 16 smiling male faces. Each trial began with a screen that lasted 3 s and read *Face X* (where *X* was replaced on each trial by the Numbers 1—16), followed immediately by a blank screen that lasted 2 s. Following the blank screen either (a) a smiling male face appeared for 3 s, followed by a signal word (*true* or *false*) that appeared for 2 s, or (b) a signal word appeared for 2 s, followed by a face that appeared for 3 s. The order of faces was held constant (e.g., Bernard's face always appeared first). For half of the subjects, a particular face was preceded or followed by the signal word *true* , and for the remaining subjects it was preceded or followed by the signal word *false*. As Table 2 shows, the identity (true or false) and position (before or after) of the signal word were completely counterbalanced across the eight different videotapes, so that each of the faces was preceded and followed by both of the possible signal words an equal number of times.

Of the 16 total trials, Trials 1, 2, 14, 15, and 16 were included as buffers against primacy and recency effects. On 4 of the remaining 11 trials, either a 1,200-Hz or 200-Hz tone sounded immediately after the face or signal word (whichever came last) had been presented. The tone lasted for 665 ms. Of these four interrupted trials, one presented the signal word *true* after the face, one presented the signal word *true* before the face, one

presented the signal word *false* after the face, and one presented the signal word *false* before the face.

The Identification Test

Subjects were allowed to familiarize themselves with the tone-discrimination task. Subjects then watched a videotape of the smiling faces, as described previously. At the end of the last trial, subjects were immediately shown a new videotape that contained 8 of the faces seen earlier. Subjects were told that only a subset of the 16 original faces would be shown during the test, and thus "you should not be alarmed if a majority, or even all, of the test faces are expressing either true or false emotions." Only 8 of the original 16 faces (Numbers 3, 4, 6, 7, 9, 10, 12, and 13; shown in uppercase letters in Table 2) were shown during the test. We hoped that this would make subjects feel free to report that all, some, or none of the faces had been expressing true happiness. As Table 2 shows, the 4 faces whose processing had been interrupted were included among the 8 test faces. Each of the 8 test faces was shown (in random order) for 3 s, and subjects were allowed 5 s to determine whether the face was expressing true or false happiness.

Results

Subjects saw a series of smiling faces and later were asked to determine whether a subset of these faces had been expressing true or false happiness. Analogous to Study 1, we predicted that interruption would have no effect on correct identification of true faces; however, we expected interruption to reduce correct identification of false faces. In addition, we were interested in learning whether this asymmetry would be eradicated, reversed, or unaffected when the sincerity of the face was known before (rather than after) its presentation.

Omissions of Data

The data from two subjects were discarded before analysis. In one case the equipment failed during the experimental session, and in the other case the subject expressed strong suspicion about the cover story.

Face Identification

The number of faces correctly identified was submitted to a 2 (Signal Word: true or false) × 2 (Interruption: yes or no) × 2 (Position of Signal Word: before face or after face) within-subjects ANOVA , which revealed only the predicted Interruption × Signal Word interaction, F (1, 17) = 4.50, p < .05. This interaction indicates that interruption had no effect on correct identification of true faces (60% when uninterrupted vs. 65% when interrupted), but significantly reduced correct identification of false faces(65% when uninterrupted vs. 35% when interrupted). Interruption caused subjects to mistake false faces for true ones, but not vice versa. In the language of the Spinozan hypothesis, subjects seem initially to have represented each face as expressing true happiness, and then attempted to alter that representation when the face was found to be expressing false

happiness. Apparently, the tone-discrimination task impaired those recoding attempts. It is interesting, too, that the effects seen in Study 1 were replicated here with iconic propositions, which suggests that the representation of meaningful assertions as true is not limited to language-based assertions. [4]

It is important to note that the third-order interaction did not approach significance, $F < 1$. This indicates that the position of the signal word had no influence on the pattern of results described above. Subjects misidentified false faces as true when their processing of those faces was interrupted, and they were just as likely to do so when they learned about the face's insincerity before seeing the face as when they learned about the face's insincerity after seeing the face. In short, knowing ahead of time that information would be false apparently did not enable subjects to adopt a skeptic's set and represent the information initially as false. This suggests that the initial coding of ideas as true may (as Spinoza suggested) be an operation that is not readily amenable to voluntary control.

Study 3: The Furry Glark Experiment

Study 2 suggests that people may be unable to prevent themselves from representing as true that which they comprehend. In Study 3 we examined further the effects of attempts to "merely comprehend" on one's subsequent assessment of information. If the mental representation of information that is merely comprehended "looks like" the mental representation of information that is actually believed, then people should be unable to tell the two apart. Metaphorically, the Spinozan librarian who is asked merely to shelve an unread book should later be unable to distinguish between this unread volume and one of nonfiction. A Cartesian librarian, of course, should have no problem distinguishing between read and unread books.

In Study 3, subjects learned about the morphology and habits of an imaginary animal called a glark. Subjects were then shown a series of propositions about glarks (e.g., *Glarks have white fur*) and were asked to assess each as either true or false. On a few occasions, subjects were asked merely to read and comprehend the proposition a few moments before they were then asked to read and assess it. Spinoza's hypothesis predicted that "merely comprehending" a proposition would create a representation of that proposition as true, and that when subsequently asked to assess a proposition that they had earlier been asked to "merely comprehend," subjects would access this representation and judge the proposition true–even when it was demonstrably false.

Method Overview

Subjects learned a set of facts about an imaginary animal called a glark. Subjects were then shown a set of propositions and were asked either (a) to assess the proposition as true or false or (b) to read the proposition as quickly as possible. Propositions were presented repeatedly such that subjects were sometimes asked to assess the veracity of propositions that they had either read or assessed just a few moments earlier.

Subjects

Twenty-three male and 7 female students at the University of Texas participated to fulfill a requirement in their introductory psychology course. Only native speakers of English were eligible to participate.

Procedure

Subjects were greeted by a male experimenter who ushered them to an individual cubicle, where they remained for the duration of the experiment. Each cubicle contained a microcomputer with keys labeled *yes* and *no,* a video camera mounted at head-level and pointed at the subject's face, and several mysterious electronic devices, adorned with switches and colored lights, which were ostensibly connected to the video camera. Subjects were told that as part of an experiment on learning and reading, the computer would present them with a series of facts about glarks, and that after they had learned these facts the computer would ask them to answer some questions about glarks. Subjects were told that the camera and electronic devices were components of an eye-tracking device that would record their eye movements throughout the experiment (cf. [Gilbert, Krull, & Pelham, 1988](#)). In fact, the equipment was entirely inert. The purpose of this deception will be explained shortly.

The learning phase.

The experimenter pretended to calibrate the eye-tracking device and then left the room while the computer delivered written instructions to the subject. These instructions explained that during the initial learning phase, the subject would be presented with a series of facts about an imaginary animal called a glark, and that the subject's initial task was to learn these facts. The instructions also explained that during a subsequent testing phase, subjects would see a series of propositions about glarks and would be asked to make some judgments about them.

During the learning phase, 20 propositions of 5 to 11 words (e.g., *Glarks are covered with long white fur*) were presented on the computer screen, one at a time, in a random order. All propositions were affirmative sentences that began with the word *Glarks* and ended with a descriptive phrase (e.g., *Glarks live in large social groups*). Each proposition appeared on a single line at the center of the screen for 4 s and was followed by a blank screen for 2 s. During the learning phase, each proposition was displayed on three separate occasions, and all subjects indicated that this had provided more than ample opportunity to learn the propositions.

Testing phase.

Following the learning phase, the computer presented each subject with 80 test propositions. The presentation of each proposition was preceded by 1,500 ms by either the signal phrase "Is the following sentence TRUE ?" or by the signal phrase " SPEED READ the following sentence." Subjects were instructed that when they saw the signal phrase TRUE , they should read the proposition that followed and assess its veracity. If the proposition was true, they should press the key marked *yes* , and if the proposition

was false, they should press the key marked *no.* We will refer to these as the *assessment trials.* The importance of responding both quickly and accurately on assessment trials was stressed.

Subjects were also told that when they saw the signal phrase READ, they should simply read the proposition that followed as quickly as possible. After reading the proposition, they should press the space bar on the computer keyboard to indicate that they had finished reading. We will refer to these as the *comprehension trials.* Subjects were told that on comprehension trials, their reading speed was being measured by the computer. Ostensibly, these data would merely be used as baseline covariates for analyses of the assessment trials. The bogus eyetracking device was included so that subjects would feel compelled to read these propositions and would not simply ignore them. The importance of rapid responding was stressed for the comprehension trials. After the subject responded on either an assessment trial or a comprehension trial, the proposition was erased from the screen. A new trial began 500 ms later.

Proposition types.

During the testing phase, subjects encountered 15 true propositions (e.g., *Glarks have white fur*) and 15 false propositions (e.g., *Glarks have brown fur*). Three precautions were taken. First, true propositions presented during the testing phase were never worded identically to the corresponding propositions presented during the learning phase. This was done so that subjects could not respond to propositions presented during the testing phase on the basis of the proposition's familiar syntax. Second, 10 meaningless propositions (e.g., *Glarks have tired fur*) were presented during the testing phase and were always preceded by the READ signal phrase. This was done to keep subjects alert, introduce variability into the procedure, and minimize the possibility that subjects would guess the hypothesis. Third, the computer recorded subjects' response times so that we could be sure that they had followed instructions to read the propositions preceded by the READ signal phrase and to respond quickly.

Proposition order.

Each of 40 propositions (10 meaningless, 15 true, and 15 false) was presented twice during the testing phase. Subjects assessed the veracity of the true and false propositions either (a) after having previously assessed the veracity of those same propositions or (b) after having read those same propositions quickly. Propositions were arranged such that the subject's second exposure to each meaningful proposition occurred exactly three trials (approximately 10 s) after his or her first exposure.

Each subject experienced a total of 40 comprehension trials and 40 assessment trials during the learning phase. Of the 15 true propositions, 5 were comprehended (i.e., preceded by the signal phrase READ) on the first exposure and assessed (i.e., preceded by the signal phrase TRUE) on the second exposure, 5 were assessed on the first exposure and comprehended on the second exposure, and 5 were assessed on both

exposures. For each subject, each true proposition was randomly assigned to one of these three orders. The same procedure was used for the 15 false propositions.

Counterbalances.

Each subject saw a different randomly selected set of true and false statements; whereas one subject may have seen a true proposition about the color of a glark's fur, another subject may have seen a false proposition about the color of a glark's fur. No subject ever saw both a true and false proposition about fur color (or any other property of glarks). Finally, it is worth noting that despite all of these elaborate counterbalances, trials could not be completely counterbalanced without asking subjects to do the impossible task of assessing the veracity of meaningless propositions. As such, the 10 meaningless propositions were always comprehended on both exposures.

Summary

The comprehension of a proposition creates a mental representation of that proposition. According to Spinoza's hypothesis, this representation cannot be distinguished from the representation of a proposition that has been assessed to be true. Thus, when subjects are asked to assess a false proposition that they have previously comprehended, they should find that they have two competing representations of that proposition: One (created during the comprehension trial) should incorrectly represent the proposition as true, and one (created during the assessment trial) should correctly represent the proposition as false. On occasion, subjects would be expected to resolve this conflict in favor of the incorrect (in this case, the true) representation. Similarly, when subjects are asked to assess a true proposition that they have previously comprehended, they should find that they have complementary representations of that proposition–one of which was created during the comprehension trial and fortuitously represents the proposition as true. According to Descartes, people represent true and false propositions identically during comprehension; as such, Descartes's hypothesis makes no specific predictions about the effect of mere comprehension on later assessment, but is incapable of predicting that the effect will be asymmetrical.

Results Omissions of Data

One male subject was omitted from all analyses for failure to follow instructions. In addition, the data were trimmed by eliminating trials on which a subject's response time was more than three standard deviations from the mean for its trial type (i.e., assessment trials vs. comprehension trials). This resulted in the omission of 101 of the 2,400 observations.

Assessment of Propositions Effects of prior comprehension.

We predicted that merely comprehending a proposition would increase the likelihood that subjects would subsequently assess that proposition to be true. That is, prior comprehension should increase subjects' correct assessment of true propositions but

decrease their correct assessment of false propositions. The proportion of meaningful propositions that subjects correctly assessed was submitted to a 2 (Proposition's Veracity: true or false) × 2 (Trial Type: comprehension-then-assessment or assessment-only) ANOVA, which revealed an uninteresting main effect of veracity, $F(1, 28) = 21.11$, $p < .001$, as well as the predicted Veracity × Trial Type interaction, $F(1, 28) = 4.35$, $p < .05$. As a comparison of the first and third columns in Table 3 shows, prior comprehension increased subjects' correct assessment of true propositions from 91.6% to 95.9%, but decreased their correct assessment of false propositions from 83.1% to 75.5%. In other words, subjects were more likely to consider a proposition to be true if they had read it quickly just a few seconds before than if they had not.

Effects of prior assessment.

The effects of prior comprehension on subsequent assessment are clear. But what were the effects of prior assessment? A 2 (Proposition's Veracity: true or false) × 2 (Trial Type: assessment-then-assessment or assessment-only) ANOVA revealed main effects of both veracity, $F(1, 28) = 5.32$, $p < .03$, and trial type, $F(1, 28) = 7.61$, $p < .01$, but a nonsignificant interaction of the two, $F(1, 28) = 2.09$, $p = .16$. As a comparison of the first and second columns in Table 3 suggests, prior assessment did increase the likelihood that subjects would make a correct assessment upon subsequent exposure to the proposition, and this may even have been slightly more pronounced for false than for true propositions ($p = .16$). The important point, however, is that the tendency for prior assessments to improve subsequent assessments of false propositions stands in striking contrast to the tendency for prior comprehensions to debilitate subsequent assessments of false propositions. Clearly, prior exposure to a false proposition does not, in and of itself, debilitate subsequent assessments of that proposition; rather, debilitation occurs only when prior exposure does not allow an opportunity for assessment. [5]

Latency to Comprehend and Assess

We asked subjects to assess propositions in response to one signal phrase, and merely to comprehend them in response to another. The excellent performance of subjects on the assessment task clearly demonstrates that they obeyed the former instruction, but what evidence is there that they obeyed the latter? Indeed, the devil's advocate might suggest that on comprehension trials (a) subjects actually did assess the propositions that they were instructed merely to read, or (b) subjects did not even read the propositions that they were instructed to read. Both of these potential difficulties are belied by the data.

Could subjects have assessed propositions that they were instructed merely to comprehend? It seems unlikely. After all, prior assessment of a proposition affected subsequent assessment in an entirely different manner (i.e., it improved assessment of false propositions) than did prior comprehension of the same proposition (i.e., it debilitated assessments of false propositions). In addition, an analysis of response times indicates a striking difference in the amount of mental work that subjects performed on each kind of trial. Subjects responded much more rapidly on comprehension trials ($M = 1,336$ ms) than they did on assessment trials ($M = 1,899$ ms), $F(1, 28) = 65.9$, $p < .001$.

As one would expect, subjects did more mental work when asked to assess a proposition's veracity than when asked merely to comprehend it. Moreover, subjects responded more rapidly to true than to false propositions when asked to assess them (M s = 1,842 ms and 1,946 ms, respectively), F (1, 28) = 5.18, p < .04, but responded equally rapidly to true and false propositions when asked merely to comprehend them (M s = 1,322 ms and 1,350 ms, respectively), F < 1. In short, assessment was protracted and was affected by a proposition's veracity, whereas comprehension was rapid and was unaffected by a proposition's veracity. These facts strongly suggest that veracity was not being assessed on comprehension trials.

What of the opposite concern? How can we be sure that subjects did any mental work at all on comprehension trials? The answer is evident: If subjects had not even read the propositions presented on comprehension trials, then why would their experience of these trials significantly alter their subsequent assessments of the propositions in precisely the way that the Spinozan hypothesis predicts? Indeed, if subjects had ignored the propositions presented on comprehension trials, then the means in the first and third columns of Table 3 should be identical. It seems clear that subjects did do mental work on the comprehension trials, but that they did not do as much mental work as they did on the assessment trials. In other words, all evidence indicates that subjects did exactly what they were asked to do.

General Discussion

Seeing Is (Like) Believing

For many centuries, philosophers have wondered whether the having and holding of ideas are psychologically separable operations, and for just as many centuries, ordinary folk have considered this a perfectly stupid question. Clearly, one experiences belief as though one were capable of entertaining ideas before endorsing them. Upon hearing a friend claim that dachshunds enjoy sleeping on fresh asphalt, one feels as though judgment should be suspended while the pertinent facts about dachshunds are located and analyzed. In the case of abstract beliefs, the Cartesian procedure seems to describe the phenomenology of believing much better than does the Spinozan.

Yet, even ordinary folk will admit that the Spinozan procedure seems a more adequate description of concrete, perceptual experience than does the Cartesian. Upon seeing a dachshund lounging in the roadway, one immediately swerves one's car, as though the representation of this object was believed long before a rational analysis could inform one's actions. People find visual illusions such as the Mueller—Lyer lines fascinating precisely because such illusions require them to consciously reject a proposition ("One line is longer than the other") whose acceptance seems an uncontrollable consequence of its representation (see Bever, 1986).

These are precisely the sorts of experiences that Spinoza was considering when he concluded that all ideas are accepted upon comprehension. The Scottish philosopher, Thomas Reid (1764/1895, p. 209) , made this point most explicitly:

> 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* [italics added].

Indeed, the automatic acceptance of perceptual representations would seem evolutionarily prudent. An organism that questioned, analyzed, and logically assessed the validity of its every percept would probably find it next to impossible to use those percepts with enough expediency to survive. It is in the very nature of the perception of physical reality that the acceptance of, or belief in, a representation should precede the rational assessment of that representation's validity (see Fodor, 1983 ; Goldman, 1986).

The notion of acceptance upon comprehension, then, is really not so foreign a concept after all; motorists on a desert highway generally believe their eyes ("Watch out for the oil in the road!") before they are able to unbelieve them ("It's just a mirage", see Bargh, 1989). If the comprehension of concrete reality proceeds in this way, then why not the comprehension of abstract reality as well? The information communicated to one person by another may not be so very different from the information that a person receives directly through the senses, and some theorists have even considered the testimony of others to be a kind of "vicarious observation" (Quine & Ullian, 1978 , p. 51; see also Miller, 1981). In this sense, the utterance "There's a charging rhinoceros!" is similar to the direct observation of a charging rhinoceros itself, and thus it does not seem unreasonable to suspect that there would be important similarities in the way these two stimuli are processed; for example, both may be accepted immediately upon comprehension and then unaccepted only later as a lie or an illusion. Such functional similarities would be especially likely if the cognitive system that processed utterances had evolved from the perceptual system that processed rhinoceri (cf. Paivio, 1971 , 1986).

This discussion suggests that encountering an idea may be very much like encountering an object–a notion strongly implied by the concrete perceptual metaphors that ordinary folk use to describe their comprehension of abstract ideas (e.g., "I *see* your point" and "*Look* at it my way"; Lakoff & Johnson, 1980 , p. 48). If abstract ideas are, like physical objects, represented as true before their validity can be rationally assessed, then the Cartesian experience of comprehending without believing may be little more than an introspective illusion (cf. Nisbett & Wilson, 1977). Certainly, the results of the present experiments fit the Spinozan position better than they do the Cartesian, as does a variety of other psychological evidence–to which we now turn.

A Variety of Other Psychological Evidence Psycholinguistics

Decades of psycholinguistic research have shown that people are generally quicker to assess the validity of true than false affirmative sentences (e.g., Carpenter & Just, 1975 ; Gough, 1965 , 1966 ; Trabasso et al., 1971). Clark and Chase (1972 , 1974) have offered a rule-based model that accounts for this fact. The model assumes that when people

process assertions, they "start with the truth index set to true" and only later assess the actual correspondence between the assertion and its referents. If the assertion is later determined to be true, then "the truth index is left alone," whereas if the assertion is determined to be false, then "the truth index is changed to its opposite" (Clark & Clark, 1977 , p. 103). It is important to note that without an assumption about the initial setting of the hypothetical truth index, the model would not correctly predict the speed of subjects' responses to a variety of sentence types (or, for that matter, the speed with which subjects in Study 1 responded to true and false propositions). In other words, the Spinozan hypothesis (i.e., that false information, but not true information, requires rerepresentation) is latent in this important psycholinguistic model.

Interpersonal Communication

Several social psychological literatures also offer evidence that may be interpreted as supporting the Spinozan hypothesis. Researchers in the areas of attribution, persuasion, and lie detection have independently noted that people are particularly prone to believe that which others tell them (for respective reviews, see Jones, 1979 ; Petty & Cacioppo, 1986 ; Zuckerman, DePaulo, & Rosenthal, 1981). This seems to be the case whether others offer general assertions about the world ("Republicans are boring") or self-descriptive claims ("I am of the opinion that Republicans are boring"), and some investigators have even described this tendency as the most "fundamental" of all phenomena in person perception (Nisbett & Ross, 1980 ; Ross, 1977).

What is so interesting about these instances of inappropriate belief is that, in general, each is exacerbated by cognitive load and interruption (Baron, Baron, & Miller, 1973 ; Gilbert, 1989 ; Gilbert, Krull, & Pelham, 1988 ; Gilbert & Osborne, 1989 ; Gilbert, Pelham, & Krull, 1988 ; Keating & Brock, 1974 ; Osterhouse & Brock, 1970 ; Petty & Cacioppo, 1986). Whether people listen to a persuasive communication, attempt to detect a prevaricator, or merely read an autobiographical account, they tend to believe more readily the information with which they are presented when their processing of that information is interrupted by a competing task. This fact has generated a variety of theoretical explanations, but one could easily interpret all of these phenomena in Spinozan terms: Beleaguered believers are unable to conduct a full-scale assessment of the propositions that they automatically accept upon comprehension, and thus they continue to regard them as true.

Hypothesis Testing

The Spinozan hypothesis also predicts that because acceptance accompanies comprehension, people should find it quite difficult to suspend judgment; that is, they should have trouble merely entertaining ideas. In effect, when a Spinozan system possesses neither evidence for nor against a possibility, it should occasionally be inclined to treat that possibility as if it were probably true.

An extensive body of work attests to the fact that when people are asked to gather information about a mere possibility, they tend to seek confirmatory information (e.g.,

Snyder & Swann, 1978 ; Wason & Johnson-Laird, 1972 ; see Klayman & Ha, 1987 , for a review). Such a strategy is, of course, appropriate only if one already has reason to be inclined toward the hypothesis one is testing (see Trope & Bassock, 1982). If one has reason to suspect that a young woman is an extrovert, for example, then it makes sense to seek evidence about the extent of her extroversion ("Would you describe yourself as merely outgoing or as wholly outrageous?") and avoid useless questions about qualities that she is unlikely to possess ("Would you describe yourself as mildly reticent or pathologically shy?"). In short, when one believes a hypothesis, then confirmatory evidence is (at least subjectively) more informative than disconfirmatory evidence.

Just as the Spinozan hypothesis would predict, Swann and Giuliano (1987, p. 522) found "direct support for the notion that *simply entertaining a belief* elevates the perceived informativeness of evidence that may confirm that belief" [italics added]. The Spinozan hypothesis suggests that this occurs because in the course of entertaining a hypothesis subjects will embrace it, and thereby raise the perceived diagnosticity of confirmatory evidence. Thus, the tendency to conduct confirmatory searches may not reflect an inability to gather information as much as it reflects an inability to entertain or "merely represent" hypotheses. It is interesting to note that the most effective way to prevent subjects from conducting a confirmatory search is to have them merely entertain the hypothesis' logical opposite (Lord, Lepper, & Preston, 1984)–a remedy that might have been recommended by Spinoza himself!

Coda

One of the dilemmas of mental life is that people need to know *of* things that are untrue, and yet need to know that these things *are* untrue. In the course of a single day, everyone is exposed to a variety of deceptive communications, ill-conceived opinions, and erroneous facts, many of which they must comprehend, remember, and yet somehow manage not to believe. To forget that *the moon is made of green cheese* is to lose a precious piece of one's childhood, but to act as though one believes this assertion is to forego the prospect of meaningful adult relationships. A ubiquitous paradox for natural thinking systems is that they must possess, but must not deploy, a wide range of false information.

In theory, this would seem a rather simple task: Mental systems could keep false information from guiding their decisions simply by assessing each piece of information they encounter, and then coding that information as true or as false in the first place. Such was Descartes's view of the human mind. Spinozan systems, however, do not conduct business in this way. Rather, they easily accept all information before it is assessed, and then laboriously recode the information that is subsequently found to be false. On occasion, of course, such attempts to recode false information will fail, and when this happens, a Spinozan system will find itself believing what it should not. This method of initially representing ideas as true may be economical and it may be adaptive, but any system that uses it will err on the side of belief more often than doubt. That human beings are, in fact, more gullible than they are suspicious should probably "be counted among the first and most common notions that are innate in us."

References

Arkes, H. R., Hackett, C. & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making, 2*, 81-94.

Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 241-252.

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

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.

Begg, I., Armour, V. & Kerr, T. (1985). On believing what we remember. *Canadian Journal of Behavioural Science, 17*, 199-214.

Ben Zur, H. & Breznitz, S. J. (1981). The effects of time pressure on risky choice behavior. *Acta Psychologica, 47*, 89-104.

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

Bever, T. G. (1986). The aesthetic basis for cognitive structures.(In M. Brand & R. M. Harnish (Eds.), *The representation of knowledge and belief* (pp. 314—356). Tucson: University of Arizona Press.)

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology.* (Cambridge, MA: MIT Press)

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

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

Clark, H. H. & Chase, W. G. (1974). Perceptual coding strategies in the formation and verification of descriptions. *Memory and Cognition, 2*, 101-111.

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

Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240-247.

Descartes, R. (1984a). 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. (1984b). 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))

Fodor, J. A. (1983). *The modularity of mind.* (Cambridge, MA: MIT Press)

Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process.(In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189—211). New York: Guilford.)

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.

Goldman, A. I. (1986). *Epistemology and cognition.* (Cambridge, MA: Harvard University Press)

Gough, P. B. (1965). Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior, 4*, 107-111.

Gough, P. B. (1966). The verification of sentences: The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior, 5*, 492-496.

Hasher, L., Goldstein, D. & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior, 16*, 107-112.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45*, 188-196.

James, W. (1890). *The principles of psychology* (Vol. 2).(New York: Henry Holt & Co)

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist, 34*, 107-117.

Kahneman, D. & Tversky, A. (1982). Variants of uncertainty. *Cognition, 11*, 143-157.

Keating, J. P. & Brock, T. C. (1974). Acceptance of persuasion and the inhibition of counterargumentation under various distraction tasks. *Journal of Experimental Social Psychology, 10*, 301-309.

Klayman, J. & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis-testing. *Psychological Review, 94*, 211-228.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by.* (Chicago: University of Chicago Press)

Lloyd, D. (1989). *Simple minds.* (Cambridge, MA: MIT Press)

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.

Miller, G. A. (1981). *Language and speech.* (San Francisco: Freeman)

Nisbett, R. E. & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* (Englewood Cliffs, NJ: Prentice-Hall)

Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

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 Publishing Co)

Osterhouse, R. A. & Brock, T. C. (1970). Distraction increases yielding to propoganda by inhibiting counterarguing. *Journal of Personality and Social Psychology, 15*, 344-358.

Paivio, A. (1971). *Imagery and verbal process.* (New York: Holt, Rinehart, & Winston)

Paivio, A. (1986). *Mental representations: A dual coding approach.* (Oxford, England: Oxford University Press)

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). New York: Academic Press.)

Quine, W. V. & Ullian, J. S. (1978). *The web of belief* (2nd ed.).(New York: Random)

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, & Co. (Original work published 1764))

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192-233.

Ross, L. (1977). The intuitive psychologist and his shortcomings.(In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173—220). New York: Academic Press.)

Russell, B. (1921). *The analysis of mind.* (New York: Macmillan)

Schwartz, M. (1982). Repetition and rated truth value of statements. *American Journal of Psychology, 95*, 393-407.

Snodgrass, J. G. (1975). Psychophysics.(In B. Scharf (Ed.), *Experimental sensory psychology* (pp. 17—68). Glenview, IL: Foresman.)

Snyder, M. & Swann, W. B. (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)

Swann, W. B. & 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.

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.

Wason, P. C. & Johnson-Laird, P. N. (1972). *The psychology of reasoning.* (Cambridge, MA: Harvard University Press)

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.

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). New York: Academic Press.)

1

The 49th proposition of Spinoza's *Ethics* (1677/1982) states that "Will and intellect are one and the same thing," by which Spinoza meant that having and believing a mental representation are identical acts (see Bennett, 1984). It is important to recognize that Spinoza's thesis does not mean that people will continue to believe that which they comprehend, but only that they must initially do so for an instant.

2

Subjects were told that this variation was due to differences in the time it took the computer to locate a proposition in its internal dictionary. In fact, we were interested in seeing if the length of delay would affect the dependent measures. It did not, and therefore is not discussed further.

3

The analysis was performed on the proportion (rather than absolute number) of propositions that were correctly identified. Each subject saw twice as many uninterrupted as interrupted critical propositions, and thus an analysis of the absolute number of correctly identified critical propositions would reveal a spurious effect of interruption. In addition, the data from two subjects who failed to respond to one or more of the interrupting tones were discarded before analysis.

4

Another explanation of these results is that subjects merely guessed, and that these guesses were based on the striking sincerity of the smiles (e.g., "I can't recall if this smiling fellow was sincere or not, but he looks so sincere now that I guess he must have been a sincere one"). Two blind judges rated the smiling faces on a scale that ranged from *very insincere* (1) to *very sincere* (9). The mean rating of 4.9 was slightly below the midpoint of the scale, suggesting that the smiling faces were not strikingly sincere.

5

Aspects of these findings are reminiscent of the "referential validity effect" (Hasher, Goldstein, & Toppino, 1977), wherein repeated exposure to a statement increases the likelihood that the statement will ultimately be assessed as true. However, we found that repeated exposure increased the likelihood that a false statement would be assessed as true only when the first exposure precluded assessment; when the first exposure demanded assessment, subsequent exposure actually decreased the likelihood that a false statement would be assessed as true. The reasons for the discrepancy between our findings and others may be that (a) the delay between exposures in other experiments is typically several weeks, whereas in our study it was about 10 s (cf. Schwartz, 1982); and (b) other experiments have typically used statements that range from mildly plausible to mildly implausible, whereas we used statements that subjects learned were absolutely false just minutes before they assessed them (see Arkes, Hackett, & Boehm, 1989 ; Bacon, 1979 ; Begg, Armour, & Kerr, 1985).

Table 1

Three Orders of Stimulus Presentation Used in Study 1

Trial	Proposition	Order 1	2	3	Trial	Proposition	Order 1	2	3
1.	A ghoren is a jug	T	B	F	15.	*A suffa is a cloud*	T	B	F
2.	A monishna is a star	F	T	B	16.	A walive is a bear	B	F	T
3.	A cirell is a tree	B	F	T	17.	A tecrill is a mouse	F	T	B
4	A tarka is a wolf	T	F	B	18.	A basol is a fisherman	F	T	B
5	A dinca is a flame	B	T	F	19.	A casin is a rope	B	F	T
6.	*A polay is a stream*	B	F	T	20.	A nasli is a snake	T	B	F
7	A tica is a fox	T	B	F	21.	*A twyrin is a doctor*	T	F	B
8.	A bilicar is a spear	F	T	B	22.	A bandi is a raccoon	B	T	F
9.	A korrom is a mountain	T	F	B	23.	*A dalith is a root*	F	B	T
10.	*A curira is a necklace*	B	T	F	24	A tiloom is a cup	F	B	T
11.	A waihas is a fish	F	B	T	25.	A gafin is a pinecone	F	T	B
12.	A rotan is a hunter	T	B	F	26.	A hib is a canoe	B	F	T
13.	*A wika is a deer*	F	T	B	27.	A trica is a weasel	T	B	F
14.	A rirg is a valley	B	F	T	28.	A neseti is a bee	T	F	B

Note. T = true; F = false; B = blank screen; entries in italics were interrupted trials.

Table 2

Eight Orders of Stimulus Presentations Used in Study 2

	Face no. (trial no.)															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Videotape																
A	tb	fa	*FB*	TA	ta	TB	FB	fa	*FA*	*TA*	tb	FA	*TB*	fb	ta	fb
B	tb	fa	*TB*	FB	ta	TA	TB	fa	FB	*FA*	tb	*TA*	FA	fb	ta	fb
C	tb	fa	FA	*TB*	ta	*FB*	TA	fa	TB	FB	tb	*FA*	*TA*	fb	ta	fb
D	tb	fa	*TA*	FA	ta	*TB*	FB	fa	TA	TB	tb	FB	*FA*	fb	ta	fb
E	tb	fa	*FA*	*TA*	ta	FA	*TB*	fa	*FB*	TA	tb	TB	FB	fb	ta	fb
F	tb	fa	FB	*FA*	ta	*TA*	FA	fa	*TB*	*FB*	tb	TA	TB	fb	ta	fb
G	tb	fa	TB	FB	ta	*FA*	*TA*	fa	FA	*TB*	tb	*FB*	TA	fb	ta	fb
H	tb	fa	TA	TB	ta	FB	*FA*	fa	*TA*	FA	tb	*TB*	*FB*	fb	ta	fb

Note. F = false; T = true; A = signal word presented after face; B = signal word presented before face; entries in lowercase were filler trials; entries in uppercase were critical trials; entries in italics were interrupted trials.

Figure 1. What the Cartesian and Spinozan procedures do with false propositions.

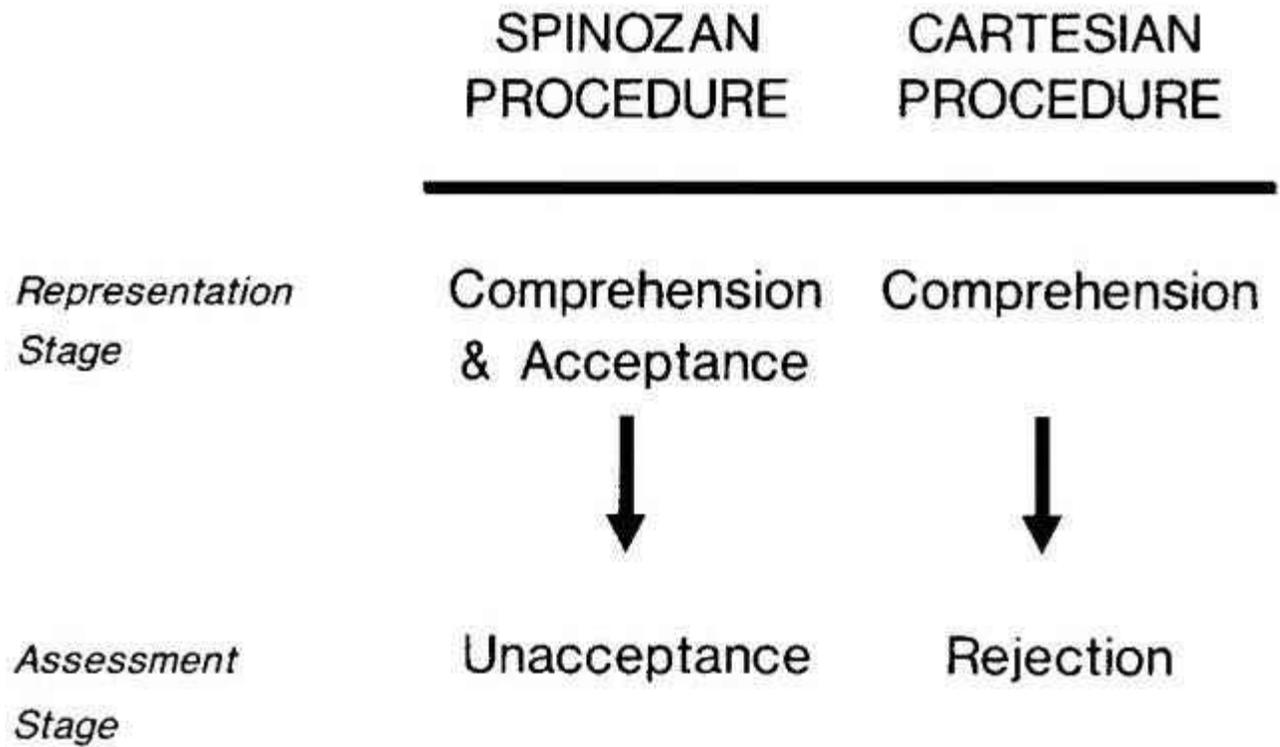|	SPINOZAN PROCEDURE	CARTESIAN PROCEDURE
Representation Stage	Comprehension & Acceptance	Comprehension
	↓	↓
Assessment Stage	Unacceptance	Rejection

Figure 2. Identifications and misidentifications of propositions in Study 1. (T-as-T denotes true propositions that were identified as true, T-as-F denotes true propositions that were identified as false, and so on.)