

Notes

An earlier version of this chapter appeared in *Cognitive Psychology* 8 (1975): 382-439. Reprinted with permission from Academic Press and Eleanor Rosch and Carolyn Mervis.

This research was supported by grants to the first author (under her former name Eleanor Rosch Heider) by the National Science Foundation (GB-38245X), by the Grant Foundation, and by the National Institutes of Mental Health (1 R01 MH24316-01). We wish to thank David Johnson, Joseph Romeo, Ross Quigley, R. Scott Miller, Steve Frank, Alina Furnow, and Louise Jones for help with testing and analysis of the data. We also wish to thank Ed Smith, Ed Shoben, and Lance Rips for permission to refer to the multidimensional scaling study of superordinate categories, which was performed jointly with them. Carolyn Mervis is now at Cornell University. She was supported by an NSF Predoctoral Fellowship during the research.

1. The larger study was performed in collaboration with E. E. Smith, E. J. Shoben and L. J. Rips of Stanford University. Half of the subjects were tested at the University of California, Berkeley, half at Stanford University. The multidimensional scaling was performed entirely at Stanford.

Chapter 3

The Psychology of Intuition

Edward J. Wisniewski

What role should intuition play in attempting to understand how humans think and behave in the world? In this chapter, I argue that intuition should play a rather limited role in achieving this goal. Based on research from cognitive and social psychology and observations about evolution, I describe the dangers of appealing to intuition as the sole methodology for obtaining knowledge. I begin by describing several salient examples of fairly basic intuitions that seem obviously true, but that are contradicted by the results of psychological studies. Such examples are not rare and the fact that strong intuitions sometimes turn out to be wrong should make one very cautious about trusting any intuition (see Dawes 1988; Gilovich 1991; Kahneman, Slovic, and Tversky 1982; Nisbett and Ross 1980; and Stanovich 1996 for many examples and extensive discussion).

Importantly, I then suggest that intuition, uninformed by psychological research, is inherently limited as a method that researchers can use to understand how people think and behave. Specifically, people do not always have conscious access to their thought processes, and they are sometimes unaware of factors that influence their thoughts and behavior. However, intuition is fundamentally a *conscious* enterprise. Therefore, it is not possible to form intuitions about some aspects of thought and behavior by directly and consciously observing the processes and factors that give rise to them. Therefore, the lay person, as well as the researcher who relies solely on intuitions (the *intuition researcher*), is deprived of an important source of information for developing accurate intuitions about thought and behavior. Consistent with these observations, I present psychological research showing that lay people and intuition researchers can have poor intuitions about thought and behavior.

To develop accurate explanations of thought and behavior, standard methods of psychological research are necessary in order to indirectly infer that which is consciously unobservable. Furthermore, I will suggest that researchers who study behavior and thought within an experimental framework develop *better* intuitions about these phenomena than those of intuition researchers or lay people who do not study these phenomena within such a framework. The intuitions are better in the sense that they are more likely to be correct when subjected to experimental testing. Before I begin, let me very briefly explain the experimental approach that is used by psychologists to study thought and behavior. Unknown to many lay people, this approach has been applied to virtually every question

asked about human thought and behavior. It also characterizes all of the experiments that I discuss in this chapter.

The Experimental Approach

Consider a hypothetical situation in which there are two people who are *identical*—they are both named Marcella, they are identical-looking, they have identical kidneys, they think, perceive, and feel in an identical manner, they have identical experiences that happen in identical places and at identical times, and so on. Suppose that you are interested in whether a chemical compound causes people to be happier than they otherwise would be. You could design an experiment to test your hypothesis: the first Marcella takes pills that contain the compound and the second Marcella takes placebos—pills that do not contain the compound but that are otherwise identical. Assume that neither of the Marcellas can tell the difference between these pills nor knows about the hypothesis.

The Marcellas take their pills (at identical times and places and in identical ways) and a week later you notice a marked change in the behavior and feelings of the first Marcella—she smiles and reports feeling happier than the second Marcella, and so on. Furthermore, you have accurately perceived the behavior of each Marcella and acted identically toward each Marcella during the test. Also assume that smiling and reports of feeling happy accurately describe what happiness is all about. Then, it is logically the case that your chemical compound causes an increase in happiness. In this experiment, there are two conditions that are identical in every aspect but one—the variable of interest (i.e., the presence versus absence of a chemical compound). Thus, the chemical compound must have caused the difference in behavior between the two Marcellas.

Clearly, it is impossible to conduct an experiment like the one described above, as no two people are identical. However, in psychology one is able to conduct experiments in which it is likely that the only *systematic* difference between conditions is the variable of interest. To illustrate, consider how the psychologist would go about testing the hypothesis described above. The psychologist would randomly and independently assign subjects to either a condition in which they took pills that contained the compound or a condition in which they took the placebos. Of course, a subject in one group will not be identical to a subject in the other group, and they will vary along many dimensions—some of which are related to happiness (e.g., perhaps self-esteem, job security, level of stress, sex life, and so on). However, because a subject is equally likely to be assigned to one or the other condition, it is equally likely that the various values of these happiness variables will be assigned to one or the other condition. Therefore, before the pills are taken, it is unlikely that there will be a systematic difference between the two groups in terms of happiness. Thus, if the experimenter subsequently finds that the group that takes the pills containing the compound is happier than the placebo group, then the experimenter can conclude that it is likely (though not logically the case) that the compound causes happiness. This simple but powerful technique of random assignment does not equate

the groups on every happiness variable but rather minimizes the likelihood that there initially will be a systematic difference between the groups in terms of happiness. Other techniques are employed that do equate variables that can influence the results. For example, a single experimenter might be employed rather than separate experimenters for each condition to avoid the possibility that differences between happiness in the groups was due to differences in how different experimenters treated the groups.

Of course, this sketch of experimental methodology provides a very simple picture of the process. There are many other concerns that must be faced in designing and conducting a good experiment and in evaluating its results. For example, a finding might be due simply to chance, a result obtained from studying college sophomores may or may not generalize to other subject populations, subjects may think that they have figured out the experimenter's hypothesis and intentionally act in way that supports (or undermines) the hypothesis, an experimenter may interpret data in a biased fashion, a variable may not truly measure what it purports to measure (e.g., is "number of times that children touch toy guns" an accurate measure of violence?), and so on. Furthermore, because of ethical reasons, it is either difficult to study some things by using the experimental approach (e.g., causes of violence) or not possible to directly investigate them (e.g., whether cigarette smoking causes cancer in people). These concerns (and many others) are legitimate but by and large, they are wrinkles that are often ironed out (for discussion on how the wrinkles are ironed out see Mitchell and Jolley 1992; Underwood 1966). I now turn to some counterintuitive findings that have been uncovered by this experimental approach.

Strong but Wrong Intuitions

Consider the following statements that are meant to describe the typical human being who has normal cognitive functioning:

1. Given a choice, most people prefer less pain to more pain.
2. A person first comprehends an idea and then decides whether or not it is true.
3. Given two options (call them A and B), if a person *prefers* option A over option B, then that person should also *reject* option B in favor of option A.

I assume (although it is only a strong intuition on my part) that virtually everyone would confidently endorse these statements (including philosophers and psychologists). It also seems fairly easy to envision each statement acting as a basic premise in a philosophical analysis. For example, one might find statement 1 in a discussion of ethics, statement 2 in a discussion of language, and statement 3 in a discussion of folk psychology. As alluring as these statements

may be, they are contradicted by a number of findings from psychological experiments.

Pain

Recently, Kahneman and his colleagues have investigated how people evaluate past episodes of pain. In one study, Kahneman, Fredrickson, Schreiber, and Redelmeir (1993) exposed subjects to aversive experiences involving the immersion of their hands into cold water. On one trial of the experiment, subjects placed a hand into a tub of cold water (14.1° C [57.4° F]) for 60 seconds (the short trial). On another trial, these subjects placed their other hand into the same tub of cold water for 90 seconds (the long trial). In this case, the cold water remained at 14.1° C for 60 seconds (as in the first trial) but was then gradually increased from 14.1° C to 15.2° C (59.4° F) over the additional 30 seconds. (The order of the short and long trials and their assignment to the dominant or non-dominant hand were counterbalanced across subjects.) During a trial, subjects used their other hand to adjust a potentiometer, which recorded their subjective sense of discomfort on a scale from 0 to 14. The average discomfort rating was almost identical for the short trial and the first 60 seconds of the long trial (a rating a little over 8). For the additional 30 seconds of the long trial, the gradual increase in water temperature caused a drop in discomfort, but the average rating (5.7) still indicated that subjects were experiencing a fair amount of discomfort.

Interestingly, after completing the trials, subjects were asked to pick which of the two cold-water trials they would repeat—69 percent of the subjects chose the *longer* trial—thus preferring more pain to less pain. Kahneman et al. (1993) explain the finding by providing evidence that memory for a painful event is based on its worst and final moments and is relatively unaffected by duration. Even though the longer trial involved more pain, discomfort was less at the end of this trial than the shorter trial (with the peak levels of discomfort being the same in the two trials). Kahneman et al. (1993) also cite studies showing that this characterization of memory for pain is true for a number of other situations as well—including exposure to loud drilling noises and highly aversive films involving amputations and colonoscopies.

Comprehension and Truth Assessment

Gilbert and his colleagues have addressed the relationship between people's understanding of an idea and their judgments of the truth of that idea. Gilbert (1991) describes contrasting views of this relationship by two early philosophers: According to Descartes, people first comprehend an idea and then are free to judge or assess its truthfulness. Thus, there is a clear distinction between comprehension and assessment. Gilbert (1991) also suggests that among philosophers and psychologists the Cartesian view is the consensus view of how people understand ideas. In contrast, Spinoza suggested that comprehension and assessment were part of the same psychological act, and that in order to comprehend an idea peo-

ple initially *had* to accept it. Of course, they were free to unaccept that idea, but unacceptance was a secondary, deliberate revision of an initially accepted belief.

Gilbert, Krull, and Malone (1990) tested Spinoza's intuition in several studies. For example, in one experiment they had subjects learn some facts about an imaginary animal called a glark. They were then shown a series of propositions about glarks (e.g., Glarks have white fur) and were asked either to judge whether a proposition was true or false (based on what they had learned) or on a few occasions, to *just read and comprehend* a proposition (and thus not assess its veracity). More specifically, after learning facts about glarks, a proposition was presented on a computer screen either preceded by the question, "Is the following sentence TRUE?" or by the command, "Speed read the following sentence."

The sentences were presented in such a way that subjects assessed the veracity of a proposition under one of three conditions: (a) after having recently assessed the veracity of that same proposition (the assessment-then-assessment condition), (b) after having recently read (comprehended) that proposition (the comprehension-then-assessment condition), or (c) after not having recently seen that proposition (the assessment only condition).

Interestingly, after comprehending a proposition (that happened to be true), subjects were more likely to assess that proposition as true compared to the case in which they had previously not seen that proposition. Even more striking, after comprehending a proposition (that happened to be false), subjects were also more likely to assess that proposition as true compared to the situation in which they had not previously seen that proposition. In other words, subjects were more likely to consider a proposition as true if they had just recently read the proposition than if they had not.

As a side note, the astute reader might argue that in the comprehension-then-assessment condition, subjects' greater tendency to assess a proposition as true may have simply been due to prior exposure of that proposition rather than to involuntary acceptance of that proposition as true. That is, mere exposure to any proposition increases one's tendency to judge it as true (whether or not it is). However, other findings undermine this alternative explanation. Recall that in one condition, subjects had prior exposure to true and false propositions that they were asked to assess (rather than just comprehend). In these cases, subjects should also be more likely to judge a previously seen proposition as true (whether or not it is). However, subjects were *less likely* to assess a false proposition as true compared to the situation in which they had not previously seen that proposition.

Gilbert et al. (1990) explain their results (and those from further studies that they conducted) by suggesting that the understanding of an idea is inextricably bound to initially accepting that idea as true. When subjects are asked to just comprehend a proposition they cannot help but initially assume that the proposition is true. In cognitive psychology, such a process would be called automatic (Shiffrin and Schneider 1977). Thus, when subsequently asked to explicitly assess the truth of that proposition, subjects will be more accurate at assessing the propositions that are true but less accurate at assessing those that are false.

Gilbert (1991) cites a number of other studies from social, developmental, and cognitive psychology whose results converge on this view.

Preferring versus Rejecting

In a series of studies, Shafir (1993) examined people's choices between options that were framed either in terms of preferring or rejecting. In one study, subjects were presented with a description about Parent A and one about Parent B. They were asked to imagine that they were serving on the jury for an only-child sole-custody case following a fairly messy divorce. Half of the subjects were asked which parent they would *prefer* to award sole custody of the child. The other half were asked which parent they would *deny* (i.e., *reject*) sole custody of the child. Parent A had the following characteristics: average income, average health, average working hours, reasonable rapport with the child, and a relatively stable social life. Parent B was described as having above-average income, a very close relationship with the child, an extremely active social life, lots of work-related travel, and minor health problems.

Interestingly, the subjects who were asked which parent they preferred to award sole custody *and* the subjects who were asked which parent they would deny sole custody more often chose Parent B. That is, subjects were both more willing to award and deny sole custody of the child to the same parent. Shafir (1993) explains this result (and those of further studies that he conducted) by noting that options sometimes contain both highly positive and negative features. Framing a choice in terms of preference causes people to more heavily weight highly positive features, whereas framing a choice in terms of rejection causes them to more heavily weight highly negative features. Thus, it is possible to both prefer and reject the same option. In the descriptions of the parents above, note that Parent B has both highly desirable and undesirable features with respect to child raising, whereas Parent A has no striking positive or negative features.

There are many other studies that show such inconsistencies in choice. For example, in choosing between gambling options, people sometimes select a bet A over a bet B but are willing to sell bet B for more than bet A (Lichtenstein and Slovic 1973). As another example, the extent to which physicians (and patients) prefer one kind of cancer treatment versus another depends on whether the treatment options are framed in terms of lives lost versus lives saved (McNeil et al. 1986). These results, like those of Shafir's, are quite counterintuitive.

To summarize, I have described some examples of strong intuitions about thought and behavior that are actually wrong. Strong but wrong intuitions are not rare. An interesting question is why people do not have the right intuitions about these examples. Perhaps people entertain both the right and wrong intuitions but believe that the "wrong" intuitions are much more plausible. However, another possibility is that people never even consider the right intuitions in the first place. The intuitions are not considered because they involve influences of which people are not consciously aware. For example, when people make judgments about pain they may simply be unaware that sometimes they are affected by the peak and final moments of pain but not affected by its duration.

Thus, it is difficult for a person to have an intuition about which of these factors affect their recollections of pain.

Unconscious Processes

There has been a long history of debate about unconscious processing in psychology. In modern psychology, the importance of the unconscious was championed by Freud, then largely banished from psychology by the behaviorists only to re-emerge as a basic tenet in cognitive psychology, where it was taken as self-evident that much of thinking was unconscious. Certainly, many basic cognitive processes are largely unconscious (e.g., object recognition, depth perception, understanding a sentence, motor tasks, and so on). These processes often happen extremely fast. In fact, our corresponding phenomenological experience is a sense that there are no processes operating at all. For example, when I look into my living room, I almost instantaneously see objects laid out in depth—chairs, a table in front of a wall and windows, and so forth.

Language understanding and production are paradigmatic cases of basic cognitive processes that are largely unconscious. In producing and understanding linguistic expressions, people largely rely on tacit knowledge that is unconsciously available to them. That speakers possess such knowledge is evidenced by their command of the language—but they are not aware of using it and are rarely able to articulate it. For example, in the sentence "The pain that Marcella feels is unpleasant" the complementizer "that" is optional but in the sentence, "The dog that bit Marcella ran away," it is not. Any native speaker of English knows these distinctions but cannot explain why they know them.

Importantly, because such knowledge is not consciously accessible, one cannot rely solely on intuition to determine how people understand and produce language. To illustrate the difficulties of (just) using intuitions to draw conclusions about language understanding and use, consider some views from philosophy on the comprehension of nonliteral (i.e., figurative) speech and the use of count and mass syntax. In a paper that has greatly influenced psycholinguistic research, Grice (1975) argued that language use is a cooperative enterprise between a speaker and a listener in which the speaker adheres to certain conversational maxims in communicating ideas to the listener. For example, speakers typically do not say what they believe to be false or that for which they lack adequate evidence (the maxim of quality, Grice 1975: 46). Importantly, on this account, speakers sometimes deliberately violate these maxims and listeners recognize the violations so as to infer that the speaker intended something else. For example, suppose that you and Marcella are sitting in the kitchen on a summer day, a turkey broiling away in the oven, and the windows closed. Both of you are sweating. If Marcella were to say, "I'm hot" you would probably get up and open a window. According to Grice, Marcella's statement violates the conversational maxim of being informative since it is obvious to both you and Marcella that it is hot. You recognize the violation and assume that Marcella is making a request (or indirect speech act).

This intuitively compelling view of language use and understanding provides

a straightforward account of how people comprehend nonliteral speech, often called the *standard pragmatic view*. When a speaker produces an utterance, the listener derives its literal meaning and determines whether it makes sense in the present context. If not, the listener assumes that the speaker has deliberately violated a conversational maxim and attempts to determine a nonliteral interpretation of the utterance. For example, upon hearing the nominal metaphor "that dinner was a roller coaster" the listener determines that it does not make sense literally and that the speaker intended to violate a conversational maxim (e.g., quality). The listener then derives a nonliteral interpretation such as "that dinner involved a series of courses which alternated between tasting awful and tasting delicious." The philosopher Searle (1993) details just such an account of nonliteral language understanding, but it also has been proposed by psychologists (Clark and Lucy 1975) and cognitive psychologists (Gagne and Shoben, 1997).

However, although conversational maxims are an important aspect of language use, there is much evidence against the standard pragmatic account of how people understand figurative language (e.g., see Gibbs 1989, 1994; Glucksberg, Gildea, and Bookin 1982; Glucksberg 1991; Keysar 1989; Rumelhart 1993; Wisniewski and Love, forthcoming). For example, the standard pragmatic view predicts that people should be slower to interpret figurative than nonfigurative language. However, given an appropriate discourse context, people do not take longer to understand figurative utterances, such as "billboards are warts on the landscape" or "the ham sandwich left without paying" (see Gibbs 1994 for a review). As another example, the standard pragmatic view implies that people seek nonliteral meanings if and only if the literal meaning makes no sense in context (Glucksberg 1991). However, Keysar (1989) found that the time to judge the truth of a statement that literally made sense in a particular context was facilitated or inhibited by whether the context *also implied* that its metaphorical meaning was true or false (see Glucksberg et al. 1982 for a related finding).

To summarize, although the standard pragmatic account of how people understand figurative language is intuitively appealing, it is not correct. (Just exactly how people understand figurative language is not resolved, but see, e.g., Glucksberg and Keysar 1990 and Gentner 1989 for some proposals.) Of course, the work of Searle and Grice does show that conscious introspection about how we think can suggest plausible accounts of the processes that are involved. However, such intuitions were not based on consciously inspecting these processes.

As a second example of intuitions about language use, consider the distinction between count and mass nouns made in many languages. A number of syntactic constructions apply to count nouns but not (usually) to mass nouns, and vice versa. For example, count nouns in English typically can follow determiners such as "a" and "another" in their singular form (e.g., "a pig" but not "a water") and numerals in their plural form (e.g., "three pigs" but not "three waters"). On the other hand, singular mass nouns typically can follow determiners that pick out an indeterminate quantity of what the noun refers to (e.g., "too much beer" but not "too much computer").

In English, count/mass usage is quite varied and flexible. The distinction is made across many different kinds of nouns, including object and substance terms (e.g., "a dog" but not "a soap"), superordinates that refer to broad categories of

perceptually diverse things (e.g., "an animal" but not "a furniture"), and abstract nouns (e.g., "an idea" but not "an evidence"). Also, many nouns can have a role as either a count or a mass noun (e.g., "a candy" as well as "too much candy"). In addition, similar aspects of reality are nevertheless distinguished with count/mass usage. For example, Gleason (1969) notes that beans is a count noun whereas rice is a mass noun and that this grammatical distinction is present even though the nouns appear to name things that seem similar (i.e., small, edible things that almost always co-occur). Also, some pairs of count/mass nouns appear coreferential, such as fuzz versus cops, clothing versus garments, snow versus snowflakes (see Ware 1979 for discussion). That is, there is a count/mass distinction between some pairs of nouns even though intuitively they refer to similar things.

As a result of these observations, a number of linguists (e.g., Bloomfield 1933; Gleason 1969; McCawley 1975) and philosophers (e.g., Quine 1989; Ware 1979) have struggled with their intuitions to find some compelling reason that explains why this distinction is made across such diverse domains that have little in common and why both count and mass terms are apparently used to refer to the same aspects of reality. Unable to find such a reason, they have questioned whether there is a systematic relationship between count/mass grammar and conceptualization, suggesting that the relationship is arbitrary, unprincipled, or idiosyncratic.

However, other researchers argue that count/mass syntax marks a general distinction between cognitively individuated and unindividuated entities (e.g., Bloom 1994a, 1994b; Imai and Gentner 1994; Jackendoff 1991; Langacker 1987; Wierzbicka 1988; Wisniewski, Imai, and Casey 1996). According to this principle, speakers conceptualize the referents of count nouns as distinct, countable, individuated things and those of mass nouns as nondistinct, uncountable, unindividuated things.

Importantly, research suggests that there are a variety of factors that either enhance or compromise individuality that depend on the type of entity involved. Thus, the relationship between count/mass syntax and conceptualization is not arbitrary, unprincipled, or idiosyncratic. For example, whether people use mass or count syntax to refer to a set of objects depends on the degree to which they are spatially contiguous. In several unpublished studies, Lyman Casey and I presented subjects with pairs of object sets. In one set, the objects were very close together but in the other set they were farther apart. Subjects were then presented with a fictitious phrase such as "Here is some chorb" (mass syntax) or "Here are some chorbs" (count syntax) and asked to select the set that is best characterized by the phrase. Subjects typically selected the mass syntax phrases for sets of objects that are close together and the count syntax phrases for sets of objects that are farther apart. Whereas spatial contiguity often affects whether *objects* are individuated, temporal contiguity affects whether *sounds* are individuated. Bloom (1994b) showed that people preferred a plural count noun to label a sound occurring over discrete intervals, implying that they construed the sound as a number of distinct, temporally bounded, and separate individuals. However, people preferred a mass noun to label a sound occurring over a long, continuous period of time, implying that they construed the sound as a temporally unbounded, unindividuated entity.

Note that in both of these examples, subjects were most likely unaware of why they preferred to label one stimulus with count syntax and another with mass syntax. However, by carefully manipulating certain variables hypothesized to affect such choices (and preventing others from systematically varying), the experimenters determined which aspects of the world map onto the count/mass distinction.

In psychology, there is a growing appreciation of the ubiquity of unconscious influences across all aspects of cognition. Below I briefly describe unconscious influences on two other aspects of cognition—explanations for one's own behavior and beliefs about whether one's memory is of an event that actually happened.

Explaining Behavior and Feelings

A number of studies in social psychology suggest that people sometimes explain their behavior or feelings by using their intuitions to decide which factors of the situation caused their behavior. If their intuitions do not suggest plausible causes, then people believe that the reason for their behavior was random or arbitrary. As a result, people sometimes erroneously attribute their behavior to intuitively plausible causes that nevertheless did not influence their behavior, or they ignore intuitively implausible factors that nevertheless did influence their behavior. Thus, people sometimes fail to understand why they do things or feel a particular way. Of course, this view makes sense given that people do not always have conscious access to their thought processes.

To illustrate this phenomenon, consider several studies from social psychology. Nisbett and Wilson (1977) had subjects view a brief documentary on the plight of the Jewish poor in large cities. In one condition, subjects viewed the film while a power saw made noises in the hall outside. In another condition, subjects viewed the film without distraction. Then, subjects rated the film along several dimensions: how interesting they thought the film was, how much they thought other people would be affected by it, and how sympathetic they found the main character to be. After making the ratings the experimenter apologized to those who saw the film under the noisy condition and asked the subjects to indicate whether any of their ratings were affected by the noise. The presence of the noise did not affect subjects' ratings compared to the group who viewed the film without the noise. However, a majority of subjects in the noise condition erroneously reported that the noise had lowered at least one of their ratings.

As another example, Lewicki (1985) conducted a study that typifies the non-conscious influences of recently activated but irrelevant information on judgments. Subjects participating in an experiment entered a room where they were briefly interviewed by an experimenter before proceeding to the main part of the experiment. The experimenter asked the subject three questions, the last one being "What is your birth order?" The subjects, aged 18-19, typically did not understand the question and asked the experimenter exactly what she meant. In the slightly irritated condition, the experimenter replied in a slightly irritated way: "Don't you really know the meaning of birth order?!" The experimenter then explained the meaning and took down the subject's birth order. In the neutral condi-

tion, the experimenter explained the meaning of birth order after being asked and took down the subjects' birth order, leaving out the irritated reply. Half of the subjects were randomly assigned to each of these conditions.

After finishing this short interview, each subject went to another room where the main part of the experiment was to take place. In the other room there were two new experimenters sitting at separate tables: one looked somewhat similar to the first experimenter who had done the short interview (she had short hair and glasses like the first experimenter) and the other was dissimilar (she had long hair and did not wear glasses). The subject was to approach an experimenter who currently was not busy in order to start the experiment. However, both experimenters were busy writing things so a subject had to make a choice on some other basis. Subjects in the slightly irritated condition were four times as likely to choose the new experimenter who did not look like the first experimenter. After performing a task that involved choosing the most interesting irregular polygons, subjects filled out a questionnaire concerning their feelings during the experiment ostensibly designed to help the experimenters better understand the subjects in general and to make them more comfortable during the experiments. There were many questions that involved rank ordering numerous possibilities pertaining to all the phases and details of the experiment. Located toward the end of the questionnaire was a question that asked subjects why they chose the experimenter that they did. Most subjects ranked "my choice was completely random" as their first choice. Subjects also rated how much they liked the first experimenter—subjects in the slightly irritated condition did not rate the first experimenter any less likable than did subjects in the neutral condition.

The results of these studies suggest that people did not explain their thoughts, feelings, or behavior by consciously examining the cognitive processes that produced them. Rather, they evaluated the context in which the thoughts, behavior, or feelings occurred and attempted to construct an explanation that implicated factors that were intuitively plausible causes. In the Nisbett and Wilson (1977) study, subjects recalled that they viewed the film under a noisy condition and assumed quite plausibly that it affected their ratings of the film (although it did not actually affect their ratings). In the Lewicki (1985) study, there was no plausible, compelling factor that should affect a subject's choice of one or the other experimenter. Thus, a subject plausibly assumed that they had arbitrarily selected one of the experimenters.

Although people's intuitions about the causes of their thoughts and behavior can be accurate (Nisbett and Wilson 1977; Smith and Miller 1978), many studies suggest that a person's intuitions either fail to implicate the correct cause or implicate an erroneous cause (see Nisbett and Wilson 1977; Wilson and Stone 1985; Wilson and Brekke 1994 for reviews). Thus, by itself, intuition is a poor tool for understanding how people think and behave.

The Origin of Memories

Another interesting issue that has received much attention concerns intuitive judgments about the source or origin of memories and beliefs (see Johnson, Hashtroudi, and Lindsay 1993 for a review). For example, it is often crucial to

distinguish whether your memory of an event is something that actually happened in the past or something that you imagined in the past. That is, is the source of this memory a real or imagined event? This question is particularly relevant with respect to evaluating the validity of claims about repressed childhood memories (Loftus 1993). As another example, which is especially relevant in academics, it is often important to know whether you were the one who originally thought of the bright idea or whether the source was a colleague who told you that bright idea.

Baddeley (1990) describes a particularly notable example of confusing the source of a memory, which involved the psychologist Donald Thomson. As told by Baddeley, Thomson had conducted studies showing that eyewitness identification is likely to be strongly influenced by the clothes that the criminal was wearing at the time of the crime. He found that innocent people are likely to be picked out of a line-up if they are wearing clothes that resemble those worn by the criminal. Thomson appeared on a television program that discussed this issue. A few weeks later he was picked up by the police and subsequently picked out of a line-up by a woman who claimed that she had been raped by him. By coincidence, Thomson's appearance on the television program coincided with the time of the rape and the woman happened to be watching the show. She was correct in recognizing Thomson's face but had confused the source of the memory.

A fair amount of research suggests that the source of a memory or belief is not explicitly stored with that memory. Instead, people examine their memories and use certain processes to decide the source of that memory. Sometimes the decisions about the source are based on deliberative conscious processes, such as attributing a statement that you remember to your friend Marcella because she would be the only who would say that sort of thing, so she must have said it. In these cases, people might have *good* intuitions about how they determine a source of a memory.

However, this research also suggests that the accuracy of identifying the source depends on many factors of which people are largely unaware. One factor is the amount of perceptual and contextual information associated with the memory. Real events generally contain more of such information than imagined events, and people use this cue in remembering a past event as real or imagined. A second factor has to do with the cognitive operations associated with encoding the memory. Confusions between real and imagined events are related to the degree of cognitive processing involved in the event. For example, in a study by Finke, Johnson, and Shyi (1988), people were asked to rate the complexity of geometric forms, which included half forms. Whenever they saw a half form they were asked to "mentally complete" the figure in order to rate its complexity. From previous research, it is known that objects that are symmetrical along the vertical axis are easier to mentally complete than those symmetrical about the horizontal axis. After doing this task, subjects were shown examples of all the completed geometric forms, and they were asked to distinguish those figures they had actually seen completed from those they had imagined as completed. Subjects were more likely to say that they had imagined as completed the half forms that were symmetrical about the horizontal axis. The implication of this finding is that when something is easily imagined it will be harder to remember at a

later time whether it was imagined or actually occurred, and that people use this cue to determine whether they experienced something real or imagined.

Yet another factor that influences the identity of a source is whether the attributes associated with the source are uniquely characteristic of that source. For example, Johnson et al. (1993) describe a study in which subjects imagined themselves saying some words and heard the experimenter saying other words. Under this condition, they were relatively good at distinguishing whether they had imagined saying a word or whether they had heard the experimenter say the word. However, in another condition, subjects imagined themselves saying the words in the experimenter's voice. Later, they had much more difficulty deciding whether they imagined saying a word or whether the experimenter said it. Interestingly, there are also developmental differences in identifying the source of a memory. A child six years of age has trouble remembering whether she actually touched her nose or whether she only imagined touching her nose, although she can distinguish whether she touched her nose or whether another person touched his nose.

Evolution

At some level of explanation, a theory of mind and thinking must take seriously the evolutionary history of the human brain. Our current mind is grounded in mechanisms that developed over the course of evolution and that were built by refining or adding to existing mechanisms instead of optimally designing a new mechanism from scratch that perfectly solved a problem. That is, mind design was a compromise between preserving what was there and developing what was needed. Furthermore, the mechanisms of our current mind are largely beneficial to surviving in a past world that perhaps does not resemble our current world in at least some important respects. The mechanisms that were useful in that past world are useful in the current world, but the extent of their usefulness may vary if for no other reason than that the current world is different from the world in which those mechanisms emerged.

These observations have very important implications for the use of intuition in studying thought and behavior. People's intuitions are products of the world as we know it today and not of the world in which our minds evolved. In this light, it is useful to revisit Gilbert's findings that comprehension of an idea is intertwined with an initial assumption that the idea is true. Gilbert (1991) speculates that cognition is an evolutionary outgrowth of perception. The idea is that for the most part, our percepts accurately reflect aspects of reality. Because our percepts usually mirror reality, the perceptual system unconsciously and automatically assumes that they are true and often for good reason. In earlier times, a person contemplating the veracity of their percepts might be eaten by an approaching tiger.

Gilbert's claim then is that the evolution of thought resulted in a brain that is built upon existing mechanisms that served us quite well. More specifically, our belief mechanism has a comprehension and assessment component that is analogous to perceptual processes. In Gilbert's own words, "people believe in the

ideas they comprehend as quickly and automatically as they believe in the objects they see." At the same time our belief system has mechanisms that allow for subsequent unacceptance of ideas. Within an evolutionary framework, Gilbert's claims and supporting findings make sense. However, outside of this framework they may seem counterintuitive.

An Important Moral

It is clear that researchers must be very careful about relying on their intuitions in formulating theories of thought and behavior. Experimental methods are absolutely essential for determining the validity of such intuitions. Perhaps less obvious is that familiarity and direct experience with psychological research may lead to better intuitions about thinking and behavior. Cognitive and social psychologists, like other scientists, attempt to develop theories that explain a wide range of phenomena and that predict new phenomena. Consider a psychologist who has developed a theory that explains a number of surprising and counterintuitive findings. The psychologist will use the theory to make predictions about new phenomena that will be intuitive to the psychologist (because they follow from the theory) but that are likely to be surprising and counterintuitive to researchers and lay people who do not know about the theory. To the extent that the theory is a good one (by the usual scientific criteria), the psychologist's intuitions are more likely to turn out right than those of people who are not familiar with the theory and the phenomena it explains.

Chapter 4

Philosophical Intuitions and Cognitive Mechanisms

Eldar Shafir

Intuition occupies a central role in philosophical theorizing. Some of the most poignant and memorable passages in philosophical writings have relied on examples whose appeal to intuition can make compelling a theory that until then seemed obscure. The appeal to intuition can be observed in domains ranging from metaphysics and epistemology, to ethics and the philosophy of mind. In what follows, I shall be unabashedly descriptive in my treatment of intuitions. I shall focus on systematic and well-documented aspects of the psychology that underlies people's intuitions; I shall ignore questions such as whether there are moral facts, or facts about rationality, and whether we may have intuitive, perceptual, or other privileged access to such facts. This chapter will consider the systematic ways in which intuitions shift as a result of supposedly inconsequential manipulations, and the implications this might have for the stability and significance of philosophical theorizing.

A descriptive account of the psychology that underlies people's attitudes and intuitions should be given serious consideration, even by scholars mostly concerned with normative or prescriptive theory. The compelling nature of normative theory notwithstanding, most scholars of human behavior tend to endorse theories that they consider psychologically feasible. Even those who suppose an exceedingly high degree of rationality or morality on the part of individuals have typically regarded their assumptions to be plausible, if somewhat idealized. Unwilling to deny the relevance of human nature, these theorists adopt a naive account of mental life that, if approximately correct, could yield behaviors largely consistent with those dictated by normative theory. Requirements of deductive closure or unbounded memory, for example, are obviously unrealistic about us and thus not part of the assumptions that most people make. Likewise, moral principles are taken seriously to the extent that the creatures to which they are applied are assumed to be able to follow them. Many errors of reasoning, inconsistencies in choice, failures of self-control, and moral transgressions, to name a few, are considered interesting, if not embarrassing, precisely because there is the feeling that one could, and should, have done better.

The descriptive approach is based on empirical observation and experimental studies of behavior. The evidence indicates that people's sentiments and preferences exhibit patterns that are often at odds with intuitive assumptions, and em-