a patient has suffered a myocardial infarction. Judgment modes are qualitatively distinct means by which people arrive at their judgments. Contrast, say, a myocardial infarction diagnosis based solely on the physician’s personal clinical experience with another derived from a validated formula applied to signs and symptoms on a checklist. This entry explains why judgment modes matter in medicine. It also describes major judgment modes that are especially useful to distinguish in medical practice. And, as appropriate, the entry further indicates specific practical implications of such distinctions.

Why Judgment Modes Matter

Judgments are important in medicine because their accuracy imposes a ceiling on the quality of the decisions they inform. That ceiling in turn sets bounds on the patient’s well-being. A patient with severe chest muscle strain who is misdiagnosed as having had a myocardial infarction will be treated as a heart attack victim. This inappropriate treatment would be invasive and risky as well as needlessly expensive. Naturally, any physician or medical practice would like to minimize inaccurate judgments, be they diagnoses, prognoses, efficacy opinions, or any of the other myriad assessments that are required throughout every day in every clinic. Achieving that aim requires a deep understanding of precisely where those judgments originate. Such understanding makes it clearer how mistakes can occur and therefore what is sensible in efforts to prevent, correct, or compensate for them. If one actually misunderstands how particular medical judgments are achieved, then the resulting attempts to improve those assessments could easily backfire, making things worse. Assuming that judgments originate in procedures—that is, modes—that are fundamentally different from how they actually are generated is misunderstanding in the extreme.

A Judgment Mode Tree

Studies have shown that, as in most practical arenas, the judgments that support people’s medical decisions can arise from sources as different from one another as apples and oranges, and hence the term judgment modes aptly describes those sources.
Figure 1 shows a judgment mode tree. (The numbers on the nodes facilitate discussion.) This hierarchy is a taxonomy of major judgment modes, organized in a particular way. Specifically, there are reasons to expect the various modes to be invoked in roughly the order of a path from "northwest" to "southeast." For a given judgment problem, modes to the left and top of the tree are likely to be attempted before ones to the right and the bottom. It is important to bear in mind, however, that in a single decision episode, several different modes, applied one after another, easily might contribute to the judgment ultimately rendered.

**Individual Versus Collective Modes**

The first level of the tree distinguishes individual from collective modes. Individual modes (1) are those in which one person (e.g., the attending physician or, perhaps, a specialist whom the physician consults) provides the judgment in question. In contrast, in collective modes (2), the judgment is
supplied by several individuals working collaboratively in some manner (e.g., three physicians concurring to reach consensus on an especially challenging diagnosis). Individual modes normally have priority over collective modes if for no other reason than that people often work alone, sometimes by necessity. In addition, collective judgment is generally slower and more expensive, materially and emotionally; it entails higher process costs, such as the time and goodwill used up working through disagreements (e.g., between two physicians with opposing opinions about the true cause of a patient’s complaint).

Individual Modes

The first major distinction among individual judgment modes concerns whether judgments are adopted (1.1) or are, instead, self-generated (1.2). Self-generated judgments are produced by the person who introduces them into deliberations for the decision problem at hand, whereas adopted judgments are acquired from some other source. Incidentally, it is common in decision scholarship to refer to the person who creates judgments for a given situation as the “judge.” Consider, for instance, a physician deciding whether to apply standard heart attack treatment measures to a patient complaining of chest pain. If the physician makes the required diagnosis by herself, it is self-generated; she is the judge as well as the decider. But if she relies solely on the opinion of a colleague, the diagnosis is adopted.

Adoption Modes

Judgment via adoption is ubiquitous. It would be impossible for people to create personally all the judgments they need to make their decisions; they must depend on other sources. Sometimes those sources are humans, other people (1.1.1), as when a physician acquires the opinion of a pathologist who has examined a tissue sample. But at other times the sources are devices (1.1.2) of various kinds, such as computer programs that render probability assessments for potential diagnoses given presented signs and symptoms.

A decider who contemplates relying on adopted judgments ideally should resolve two practical concerns. The first is accuracy: How accurate would the adopted judgment be? For instance, how much more accurate than his own assessments should the physician expect the diagnoses of specialist Dr. Smith to be? The second concern is cost: How much would have to be paid—materially and otherwise—for the adopted judgment, and does the promised accuracy improvement outweigh the greater expense? For example, does the potentially improved accuracy of a new computerized diagnostic procedure more than offset the extra time and money it requires? There is reason to believe that both accuracy and cost concerns are often overlooked and that, when they are considered, they are not thought through adequately. Consider evidence that people’s conclusions about others’ expertise are strongly affected by factors (such as speech patterns) that easily can have little or nothing to do with objective accuracy indicators. Or take the fact that people often fail to ask questions about the functioning of judgment devices that have significant bearing on their appropriate use, such as questions about the information items the devices take into account and the items they ignore.

Self-Generation Modes

Under many conditions, although not all (e.g., where their inexperience is obvious), people tend to be overconfident about the quality of their own judgment. When that occurs, they should be expected to eschew adopted judgments for self-generated ones more than they should. And there are four main varieties of self-generation modes that might be pursued, each with its own special cases.

Precedence. The essence of precedence modes (1.2.1) is that the judge uses past occurrences in similar situations to inform the judgment needed presently. The most basic form this mode takes can be labeled expectancy learning (1.2.1.1). The core idea is the following: Consider some event of concern, say, a possible case of asthma. Furthermore, imagine that, in Clinical Setting 1, asthma occurs about 5% of the time, while in Setting 2, the rate is about 12%. The judge simply observes a large number of randomly presented cases (asthma vs. no asthma) in Setting 1 and does the same in Setting 2. There is no request to do anything like count, memorize, or even pay close attention to what is observed. Studies have shown that, in due
course, the judge will induce the fact that the asthma rates are different in the two contexts, as reflected in the judge's behavior.

Suppose the judge is told that, in a given setting, he will receive $100 if the next case that comes along is an asthma case. The judge is then allowed to choose the setting in which to exercise this opportunity. The judge will almost certainly pick Setting 2, where the asthma rate is 12%, rather than Setting 1, where that statistic is 5%; his expectancy for asthma is stronger in the former situation. Note that the judge was not asked to express a judgment explicitly. However, such a comparative likelihood judgment is implicit: "Asthma is more likely in Setting 2 than in Setting 1." If the judge were asked to state a probability judgment that the next patient in one of the settings will have asthma, there is no guarantee that that judgment would match the observed rate precisely. But the judgment would almost certainly be higher for Setting 2. The Setting 1 versus Setting 2 scenario is contrived, but the underlying principle is not. It is generalizable to more realistic circumstances, such as that in which "Setting 1" is replaced by "Patients like Ms. Jones."

The remaining precedence modes all entail some deliberate attempt to remember previous occurrences. In the recall contents mode (1.2.1.2), the judge tries to recall specific earlier instances when events similar to the one being considered presently actually happened. These are used to estimate, for example, the relative frequency of such past occurrences, which is then taken as the required probability judgment. Thus, suppose that, through such recall efforts, a physician estimates that, of the patients like Ms. Jones whom he has seen in the past, about 10% had asthma. He therefore concludes that there is a 10% chance that Ms. Jones has that condition, too.

The recall availability heuristic (1.2.1.3) is subtly but significantly different from the recall contents mode just described. In the latter, the judge uses the substance of what is remembered. Suppose that Ms. Jones's physician brings to mind 10 former patients similar to Ms. Jones and that one of them had asthma. That 1:10 ratio would yield the physician's probability judgment of asthma for Ms. Jones. In contrast, the inference in the recall availability approach is indirect. Ms. Jones's physician might try to bring to mind perhaps only a single case in which a patient similar to Ms. Jones had asthma. The physician next makes an assessment of how easy it was to recall that case, its "availability" for recall. The physician then invokes the key assumption underlying the recall availability heuristic: The easier it is to recall a particular kind of event, the more often that event must have occurred in the past. This assumption then justifies inferring that events similar to easy-to-recall past exemplars are highly likely to occur now also.

The records mode (1.2.1.4) is the "objective" or "scientific" variant of the precedence mode. The judge does not depend on fallible memory for a perhaps limited number of personally observed past cases. Instead, the judge calls upon reliable records of large, representative (if not exhaustive) samples of such cases. Thus, Ms. Jones's physician might consult an extensive database of valid records of patients who resemble Ms. Jones in relevant ways. If 13% of them had asthma, then Ms. Jones's doctor will take 13% as her probability of having asthma, too. This "evidence-based" approach is sometimes called an "actuarial" method, since it is basically the same as that used by insurance companies in arriving at the probability judgments that they use for setting premiums.

The key ideas underlying precedence modes are compelling. Nevertheless, those modes entail risks, too. One is nonstationarity, which essentially says that current tendencies are fundamentally different from those in the past. For example, shifts in environmental conditions might mean that the true asthma incidence changes substantially over time. Various psychological phenomena can compromise the adequacy of the nonobjective precedence modes also. For instance, recall is subject to a host of context influences, such as primacy and recency effects, whereby the earliest and the most recently observed cases, respectively, are especially likely to be remembered. The ease with which past instances can be brought to mind is also influenced by numerous other factors that have little to do with how often those instances actually occurred, such as their vividness (e.g., the breast cancer death of a prominent person).

Event Theories. The defining feature of event theory modes (1.2.2) is that, in some fashion or another, the judge draws on a theory of how the event in question literally comes about in nature. In
some cases, those theories are personal ones (1.2.2.1), sometimes described as "naive" in that typically they are buttressed not by rigorous scholarship but instead by plausible lay intuitions. Two examples of widely accepted theories in early medical history that almost certainly began as personal theories are illustrative. One is the "doctrine of signatures," according to which, supposedly, one could predict the efficacy of medicines derived from a plant by its resemblance to the organ of concern, for instance, the liverwort for treating disorders of the liver. Another is the miasma or "bad air" theory for explaining (and therefore predicting) diseases such as cholera.

The principle underlying judgment according to scenario availability (1.2.2.2) is similar to that deployed in the recall availability heuristic: An event is judged likely to occur to the extent that a scenario giving rise to it is easy to imagine. Thus, a physician's prognosis for a diabetes patient might be driven by how easily he can envision the patient adhering to a recommended treatment regimen. The operative principle is plausible but also perhaps overly self-generous: "If I personally cannot easily imagine how something can occur, that means it probably can't occur." One reason for having less than complete faith in this principle is that people have a hard time anticipating and even understanding the full range of complicated interactions that occur among the forces at play in many real-life scenarios, including ones in medicine.

Validated event theories (1.2.2.3) are the engines that drive modern scientific medicine. Such a theory provides an account for how a particular condition, such as an infectious disease (e.g., acute bronchitis), arises and how it progresses over time, affecting particular organs in specified ways and in specified order. That sequence directly guides prognostic judgment. And backward reasoning from particular signs and symptoms helps narrow down differential diagnoses.

Judges turn to event theories relatively early for a given case because humans have a natural need to understand the world, not merely to predict events accurately. One hazard of event theories, though, is that they likely give short shrift to uncertainty. After all, if one firmly believes that an event is the end product of a specific A → B → C → ... sequence of occurrences, there is little or no room for uncertainty.

Matching. The key feature of judgment via matching modes (1.2.3) is that, during the course of the judgment process, the judge matches one or more features of the event in question with some "reference," such as a prototype. Some algorithm-like diagnostic procedures provide good illustrations of the checkoff (1.2.3.1) variant of matching. Each possibility in a differential diagnosis is implicitly defined by a prototypical case consisting of several signs and symptoms, including, perhaps, test results. In a fashion similar to the parlor game "20 Questions," the diagnostician successively eliminates possible diagnoses whose prototypes require features missing from the case at hand until, eventually, only one diagnosis is left. In similarity (1.2.3.2) versions of matching, multiple characteristics of the given case are compared with corresponding features of the prototype more holistically, yielding an assessment of the degree of overall similarity between the case and the prototype. The judgment rests directly on that similarity assessment. For example, even if a given patient's profile does not exactly fit the classic pneumonia victim prototype, if the similarity is strong enough, pneumonia is the diagnosis rendered. Judgment according to the representativeness heuristic is illustrative.

Judgment via checkoffs in principle should be less demanding than judgment per similarity. However, an advantage of the latter is that it more readily acknowledges uncertainty. Unfortunately, that mode can also yield judgments that violate key formal principles such as Bayes's theorem.

Association Models. Judgment according to association models (1.2.4) seeks to exploit presumed statistical associations between the events in question and easily observed facts. For instance, "risk factors" such as hypertension are often used to sharpen stroke predictions, since strokes are thought to be (and are), especially common for people with high blood pressure.

In personal (1.2.4.1) forms of association model modes, the judge relies on personal intuitions and reasoning to derive a judgment. In special cases that are comprehensive (1.2.4.1.1), the judge attempts to exploit the predictiveness of at least two different facts in reaching judgments (e.g., hypertension and family history for stroke judgments). The anchoring and adjustment heuristic is a commonly discussed,
simple illustration, whereby the judge uses some general facts (e.g., a local incidence rate) to provide an initial ballpark judgment and then more case-specific facts (e.g., test results) to move that initial assessment in appropriate directions. Intuitive approximations of the logic of linear regression are other, perhaps more sophisticated, examples. In limited (1.2.4.1.2) forms of personal association modeling, the judge relies on just a single fact, perhaps deliberately flouting the principle that bringing to bear greater amounts of information can never reduce the statistical predictability of any event. Fast-and-frugal heuristics such as the "take the best rule" are illustrative. In that procedure, a physician would make a diagnosis solely on the basis of the one sign or symptom thought to be most predictive, ignoring all the rest. Limited association model procedures can sometimes outperform comprehensive ones. For instance, in attempting to synthesize the predictive value of multiple considerations, the judge might fail to recognize that, if some of those considerations are strongly correlated with one another, this redundancy should greatly affect how those facts are used. The burdens of juggling several facts at the same time can also reduce the judge's reliability, thereby undercutting judgment accuracy even further.

Formal (1.2.4.2) versions of the association models mode entail attempts to do the best job of "objectively" taking advantage of associations between facts and events of interest, almost always with the aid of computers. Statistical (1.2.4.2.1) variants include well-known linear regression-type procedures, such as discriminant function analysis, and also Bayesian updating routines. Simulation (1.2.4.2.2) variants take several forms. Some involve computer programs that are intended to mimic the routines that a recognized expert human judge (e.g., diagnostician) uses in actual practice. Artificial neural networks, another popular variety, are programs intended to imitate the means by which human neural networks are thought to perform mental tasks such as classifying stimuli into various categories.

It is noteworthy that, although most formal modes rely on objective association measures derived from formal records, personal modes do not. Instead, they depend on judges' own opinions about how various facts are correlated with the events at issue. And research has shown that the processes by which people arrive at their relationship beliefs are vulnerable to several forces that can introduce biases. For example, people's intuitions about how data should be used to draw relationship conclusions are often significantly different from standard statistical rules such as likelihood ratios. This clearly imposes bounds on how accurate judgments resting on personal association models can be.

Another Perspective: Deliberative Versus Nondeliberative Modes. All the various self-generation modes have been described as if they were executed consciously and purposefully. This is a reasonable assumption for "public" modes such as those resting on records and statistical models. But it is not always reasonable for modes carried out in the heads of individual judges. In fact, there is considerable evidence for another important mode distinction that underlies those displayed explicitly in the mode tree—the distinction between what may be called deliberative and nondeliberative modes.

Deliberative modes, sometimes said to involve System 2 thinking, are characterized by features such as control, effort, and awareness. In contrast, when a person applies nondeliberative modes, sometimes associated with terms such as System 1, automaticity, and intuition, the judgment process is often initiated on its own and cannot be stopped; once it begins, it requires virtually no mental capacity or effort, and it can function outside the person's awareness. In fact, the judge is likely to be unable to accurately describe how judgments are reached. Some modes listed in the mode tree, such as expectancy learning, might be nondeliberative virtually always. Others, however, are likely to transition from deliberative to nondeliberative as a result of experience. Thus, after a physician has made hundreds of diagnoses on the basis of a personal association model for a condition such as pneumonia, she is likely to lose awareness of how she arrives at those diagnoses.

The deliberative/nondeliberative distinction highlights the practical significance of modes. There have long been attempts to improve people's judgments by merely educating them about the existence of various biases. Such efforts have seldom succeeded. There is reason to believe that this is partly because the awareness approach presumes
that people are making their judgments on the basis of deliberative processes when they actually rest on ones that are nondeliberative, and hence beyond people's personal control. Markedly different strategies are therefore required to achieve improvements.

**Collective Modes**

There are two kinds of collective modes (2), distinguished by how the differing opinions of multiple judges are synthesized into a final, collaborative judgment. (Note that the judges do not have to be humans; some or all of them might be devices such as computer programs.) *Interactive* (2.1) synthesis basically entails discussion in some form, as when several clinicians hold a meeting and reach consensus about the diagnosis for a difficult case. In contrast, in *mechanical* (2.2) synthesis, the participating judges work independently and report their judgments individually. These assessments are then combined with a formula of some kind. For example, the recovery-time predictions of three different physicians might be simply averaged to yield a composite prediction that is reported to the patient.

There are many reasons to expect collective modes to yield more accurate judgments than individual modes (“Two heads are better than one,” ...). But there is also reason to suspect that that potential goes unrealized in many interactive situations. For instance, people often choose to discuss things that they know in common rather than to introduce into conversation information or expertise that they hold uniquely. On the other hand, even simple averaging of multiple individual judgments typically yields composite judgments that are markedly (and surprisingly) more accurate than those of any one person.

*J. Frank Yates and Andrea Angott*

**See also** Differential Diagnosis; Heuristics; Intuition Versus Analysis; Judgment; Teaching Diagnostic Clinical Reasoning

**Further Readings**


